Clinical Outcomes of Pediatric Intestinal Failure Management after Implementation of a Multidisciplinary Intestinal Rehabilitation Program

by

Carol Oliveira

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Institute of Health Policy, Management and Evaluation
University of Toronto

© Copyright by Carol Oliveira 2016
Clinical Outcomes of Pediatric Intestinal Failure Management after Implementation of a Multidisciplinary Intestinal Rehabilitation Program

Carol Oliveira

Doctor of Philosophy

Institute of Health Policy, Management and Evaluation
University of Toronto

2016

Abstract

The objective of this thesis was to evaluate the effect of implementation of a multidisciplinary intestinal rehabilitation program (IRP) on clinical outcomes of pediatric intestinal failure (IF). This was attained by two comprehensive reviews and three related research projects that contribute both clinical and methodological knowledge.

Pediatric IF is a complex medical condition. Advances have been made clinically in recent years, but research in the field has suffered from methodological challenges and poor quality. Many of these were outlined and solutions suggested in this thesis.

The first project compared clinical outcomes of three eras of pediatric IF-management using univariate categorical and time-to-event analysis. While IRP-introduction led to an increased transplantation rate and improved clinical outcomes, further progress was made and the
natural course of the disease changed with IRP-maturation and adoption of novel treatment strategies.

The second project quantified the continuous change over time of pediatric IF over a 15-year time period and tested systematically introduced treatments using novel methods. Multivariate time series analysis was applied to assess the relationship between new treatment options and change in clinical outcomes. The study showed that IRP-introduction and omega-3 lipid emulsions independently decreased disease-specific mortality from liver failure and sepsis when adjusted for disease severity, while serial transverse enteroplasty and ethanol locks did not impact mortality significantly.

With improved mortality of pediatric IF, assessment of morbidity became imperative. After having identified a severe lack in methodology to measure neonatal and infant health-related quality of life (NIHRQOL) in a systematic literature review, the conceptual meaning of NIHRQOL was explored in the final project. It concluded that NIHRQOL is a multidimensional, multilayered and interconnected concept, in which the child’s needs are fundamental factors, and the caregiver’s and society’s ability to meet those needs characterize the interdependence between the child and its caregiving environment.

In summary, the complexity of pediatric IF and related methodological challenges were identified. Methodological solutions were described and illustrated. Clinical advancements of pediatric IF were quantified, and IRP-introduction and omega-3 lipid emulsions recognized as strongest contributors. Finally, the meaning of NIHRQOL was explored and organized in a conceptual framework in this thesis.
Acknowledgments

I am grateful to a number of people who provided me with support, guidance and encouragement, without which this thesis would not have been possible. My thanks go:

To my supervisor, Dr. Paul Wales, for his insight, guidance and mentorship;

To my thesis committee, Drs. Ahmed Bayoumi, Wendy Ungar, Jeffrey Hoch and Yaron Avitzur, for their guidance and inspiration;

To the surgeons of the Division of General and Thoracic Surgery at the Hospital for Sick Children, Drs. Jacob Langer, Annie Fecteau, Sharifa Himidan, Priscilla Chiu, Georges Azzie, Ted Gerstle, Peter Kim and Agostino Pierro for their support, mentorship and dedication to my academic development;

To the Group for Improvement of Intestinal Function and Treatment (GIFT), Karen Steinberg, Christina Kosar, Glenda Courtney-Martin, Alain Rogers, Stephanie So, Catherine Patterson, Dr. Anna Gold, Carina Silva, Megan Carricato, Kathryn Chambers, Andrea Carpenter, Debra Harrison and Jane Darch for their help with any questions and requests I had, and specially to Julia Maxwell, for her assistance with the focus groups.

To Nicole de Silva, for her important help and support with my thesis and non-thesis projects;

To the administrative staff of the Division of General and Thoracic Surgery at the Hospital for Sick Children, in particular Janice Bowers, Catherine Day, Anne Apong, Jennifer Ionson, Shannon Hannah and Tanya Kovaleva, for all their help with my administrative questions and organizational requests;

To Dr. Sanja Stanojevic, for helping me understand time series analysis.

To the director of Bioethics at the Hospital for Sick Children, Dr. Randi Zlotnik, for her insight into neonatal and infant quality of life;
To the director of the library at the Hospital for Sick Children, Elizabeth Uleryk, for her assistance with literature searches;

To the program director of the Division of General Surgery at the University of Toronto, Dr. Najma Ahmed, for her invaluable support;

To the surgeons of the Department of Pediatric Surgery at the University Children’s Hospital Zurich, Switzerland, in particular Drs. Martin Meuli and Clemens Schiestl, who supported my wish to pursue graduate studies at the University of Toronto, committed to my academic endeavors and provided me with mentorship;

To the Child Health Evaluative Sciences (CHES) Program at the Hospital for Sick Children, for providing a nurturing environment that cultivates academic productivity and success;

To the Institute of Health Policy, Management and Evaluation (IHPME), in particular past and present program directors, Drs. Ahmed Bayoumi, Sharon Dell and Robert Fowler, and program coordinators, Amber Gertzbein and Jennifer James, for their guidance and administrative support throughout my PhD studies;

To the Swiss National Science Foundation, the Stiefel-Zangger Foundation and the Walter and Gertrud Siegenthaler Foundation of the University of Zurich, and the Research Training Center at the Hospital For Sick Children for their financial support;

And finally to my family and friends, for their continuous support and encouragement, without which I could not have attained this achievement.
# Table of Contents

Abstract .......................................................................................................................... ii
Acknowledgments ........................................................................................................ iv
Table of Contents ......................................................................................................... vi
List of Tables ................................................................................................................ xiii
List of Figures ............................................................................................................... xiv
List of Appendices ....................................................................................................... xv
List of Abbreviations ................................................................................................... xvi
Thesis Overview .......................................................................................................... xxi

Chapter 1 ..................................................................................................................... 1

1 Overview of Pediatric Intestinal Failure ................................................................ 1
   1.1 Definition ........................................................................................................... 1
   1.2 History ............................................................................................................. 1
   1.3 Epidemiology ................................................................................................... 4
   1.4 Etiology ........................................................................................................... 5
   1.5 Pathophysiology .............................................................................................. 6
       1.5.1 Jejunal Resection .................................................................................... 6
       1.5.2 Ileal Resection ....................................................................................... 7
       1.5.3 Colonic Resection .................................................................................. 9
       1.5.4 Adaptation ............................................................................................ 9
   1.6 Clinical Management ...................................................................................... 11
       1.6.1 Multidisciplinary Intestinal Rehabilitation ........................................... 11
       1.6.2 Medical Management ............................................................................ 12
           1.6.2.1 Parenteral Nutrition ................................................................... 12
           1.6.2.2 Enteral Nutrition ......................................................................... 15
1.6.2.3 Hormone Therapy .......................................................... 16
1.6.2.4 Other Medication Frequently Used........................................... 18
1.6.3 Surgical Treatment.............................................................. 18
1.6.3.1 Bowel Preserving Initial Surgery ......................................... 19
1.6.3.2 Procedures To Improve Intestinal Function ............................. 20
1.6.3.3 Procedures To Improve Intestinal Motility ............................... 20
1.6.3.4 Intestinal Lengthening Procedures ....................................... 21
1.6.3.5 Transplantation .................................................................. 23
1.6.3.6 Other Procedures Frequently Performed In Patients With Intestinal Failure .......................................................... 25
1.7 Outcomes ................................................................................. 26
1.7.1 Complete Weaning From Parenteral Nutrition ............................... 26
1.7.2 Morbidity ............................................................................... 26
1.7.2.1 Comorbidity ........................................................................ 27
1.7.2.2 Complications From Intestinal Failure ................................... 27
1.7.2.2.1 Intestinal Failure Associated Liver Disease ......................... 27
1.7.2.2.2 Sepsis ........................................................................... 30
1.7.3 Quality Of Life ........................................................................ 32
1.7.4 Mortality ................................................................................ 34
1.7.4.1 Disease-Specific Mortality ..................................................... 35
1.7.4.2 Other Mortality ................................................................... 37
1.7.5 Cost-Effectiveness & Economic Burden ...................................... 37

Chapter 2 ....................................................................................... 39
2 Methodological Challenges of Pediatric Intestinal Failure Research .......... 39
  2.1 Study Design ........................................................................... 39
  2.1.1 Randomized Controlled Clinical Trials .................................... 40
  2.1.2 Alternative Study Designs ...................................................... 45
  2.1.2.1 Recruitment & Randomization ............................................. 45
  2.1.2.2 Outcome Measurement & Analysis ...................................... 47
  2.1.2.3 Observational Study Designs .............................................. 48
3 Intestinal Failure Outcomes of a Matured Multidisciplinary Intestinal Rehabilitation Program

3.1 Cover Page
3.1.1 Affiliations
3.1.2 Corresponding Author
3.1.3 Abbreviations
3.1.4 Key Words
3.1.5 Funding Sources
3.1.6 Financial Disclosure
3.1.7 Conflict Of Interest
3.1.8 Clinical Trial Registry
3.1.9 Manuscript Writing

3.2 Abstract

3.3 Introduction

3.4 Patients & Methods
3.4.1 Study Design
3.4.2 Study Sample
3.4.3 Data Collection
3.4.4 Data Analysis

3.5 Results
3.5.1 Patient Characteristics
3.5.2 Intestinal Anatomy
3.5.3 Outcomes
3.5.3.1 Nutritional Outcomes
3.5.3.2 Outcomes Related to Central Venous Catheters
3.5.3.3 Outcomes Related to Liver Disease
3.5.3.4 Mortality

3.6 Discussion
3.6.1 Limitations
3.6.2 Future Research

3.7 Conclusion
Chapter 4

4  Change of Outcome in Pediatric Intestinal Failure – The Use of Time Series Analysis to Assess the Evolution of an Intestinal Rehabilitation Program

4.1  Cover Page

4.1.1  Affiliations
4.1.2  Corresponding Author
4.1.3  Abbreviations
4.1.4  Key Words
4.1.5  Funding Source
4.1.6  Financial Disclosure
4.1.7  Meeting Presentation

4.2  Precis

4.3  Abstract

4.4  Introduction

4.5  Methods

4.5.1  Study Design
4.5.2  Study Sample
4.5.3  Data Collection
4.5.4  Data Analysis

4.6  Results

4.6.1  Patient Characteristics
4.6.2  Clinical Outcomes

4.7  Discussion

Chapter 5

5  Health-Related Quality of Life in Neonates and Infants: A Conceptual Framework

5.1  Preamble

5.2  Cover Page

5.2.1  Affiliations
5.2.2  Corresponding author
6.4.1 Implications for Clinicians ................................................................. 164
6.4.2 Implications for Policy-Makers .......................................................... 166
6.4.3 Implications for Researchers .............................................................. 167
6.4.4 Implications for Patients & Families .................................................. 168

6.5 Future Research Directions ..................................................................... 169
6.6 Conclusion .............................................................................................. 172

References .................................................................................................... 174

Appendices ..................................................................................................... 222

Appendix 4-A: Linear Trend Of Disease-Specific Mortality ............................ 222
Appendix 4-B: Linear Trend Of Parenteral Nutrition Weaning & Advanced Liver Disease ................................................................. 223
Appendix 4-C: Linear Trend Of Central Venous Catheter & Transplantation Related Outcomes ................................................................. 224
Appendix 5-A: Guideline Used By The Moderator ......................................... 225
Appendix 5-B: Terminology & Interconnectivity .......................................... 226
Appendix 5-C: Direct Quotes Of Participants ................................................ 228
Appendix 5-D: Levels, Domains & Attributes of Infant and Neonatal HRQOL Concept. .......................................................................................... 232
List of Tables

Table 3-1: Patient Characteristics ................................................................. 101
Table 3-2: Intestinal Anatomy ........................................................................ 102
Table 4-1: Patient Characteristics .................................................................. 122
Table 4-2: Clinical Outcomes .......................................................................... 123
Table 5-1: Participant Characteristics .............................................................. 141
Table 5-2: Most Important Factors to Participants .......................................... 142
List of Figures

Figure 3-1: Nutritional Outcomes ........................................................................................................... 104

Figure 3-2: Outcomes Related to Central Venous Catheters and Liver Disease ......................... 107

Figure 3-3: Mortality ................................................................................................................................. 109

Figure 4-1: Disease-Specific Mortality .................................................................................................. 124

Figure 4-2: Parenteral Nutrition Weaning & Advanced Liver Disease ............................................. 126

Figure 4-3: Central Venous Catheter & Transplantation Related Outcomes .................................. 128

Figure 5-1: Levels and Domains of Infant and Neonatal HRQOL Concept .................................... 144

Figure 5-2: Attributes of Infant and Neonatal HRQOL Concept ....................................................... 146
List of Appendices

Appendix 4-A: Linear Trend Of Disease-Specific Mortality.......................................................... 222

Appendix 4-B: Linear Trend Of Parenteral Nutrition Weaning & Advanced Liver Disease
.................................................................................................................................................. 223

Appendix 4-C: Linear Trend Of Central Venous Catheter & Transplantation Related
Outcomes........................................................................................................................................ 224

Appendix 5-A: Guideline Used By The Moderator......................................................................... 225

Appendix 5-B: Terminology & Interconnectivity.......................................................................... 226

Appendix 5-C: Direct Quotes Of Participants............................................................................... 228

Appendix 5-D: Levels, Domains & Attributes of Infant and Neonatal HRQOL Concept ... 232
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1AT</td>
<td>α-1 Antitrypsin</td>
</tr>
<tr>
<td>AC</td>
<td>Autocorrelation</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike criterion</td>
</tr>
<tr>
<td>AIRS</td>
<td>Autologous intestinal reconstruction surgery</td>
</tr>
<tr>
<td>ALT</td>
<td>Alanine transaminase</td>
</tr>
<tr>
<td>ARIMA</td>
<td>Autoregressive integrative moving average models</td>
</tr>
<tr>
<td>ASPEN</td>
<td>American Society for Parenteral and Enteral Nutrition</td>
</tr>
<tr>
<td>AST</td>
<td>Aspartate aminotransferase</td>
</tr>
<tr>
<td>BIC</td>
<td>Bayesian information criterion</td>
</tr>
<tr>
<td>BUN</td>
<td>Blood urea nitrogen</td>
</tr>
<tr>
<td>CAPS</td>
<td>Canadian Association of Pediatric Surgeons</td>
</tr>
<tr>
<td>CBC</td>
<td>Canadian Broadcasting Corporation</td>
</tr>
<tr>
<td>CCK</td>
<td>Cholecystokinin</td>
</tr>
<tr>
<td>CHQ</td>
<td>Child Health Questionnaire</td>
</tr>
<tr>
<td>CHQ PF-50</td>
<td>Child Health Questionnaire Parent Form</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CIPO</td>
<td>Chronic intestinal pseudo-obstruction</td>
</tr>
<tr>
<td>COMET</td>
<td>Core Outcome Measures for Effectiveness Trials initiative</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>CONSORT</td>
<td>Consolidated Standards of Reporting Trials</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CVC</td>
<td>Central venous catheter</td>
</tr>
<tr>
<td>CRBSI</td>
<td>Catheter-related bloodstream infection</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>ESPEN</td>
<td>European Society for Clinical Nutrition and Metabolism</td>
</tr>
<tr>
<td>FXR</td>
<td>Farnesoid X receptor</td>
</tr>
<tr>
<td>FGF-19</td>
<td>Fibroblast growth factor 19</td>
</tr>
<tr>
<td>G</td>
<td>Gastrostomy</td>
</tr>
<tr>
<td>GGT</td>
<td>Gamma-glutamyl transferase</td>
</tr>
<tr>
<td>GJ</td>
<td>Gastrojejunal</td>
</tr>
<tr>
<td>GIFT</td>
<td>Group for Improvement of Intestinal Function and Treatment</td>
</tr>
<tr>
<td>GIP</td>
<td>Gastric inhibitory polypeptide</td>
</tr>
<tr>
<td>GH</td>
<td>Growth hormone</td>
</tr>
<tr>
<td>GLP-1</td>
<td>Glucagon-like peptide 1</td>
</tr>
<tr>
<td>GLP-2</td>
<td>Glucagon-like peptide 2</td>
</tr>
<tr>
<td>GVHD</td>
<td>Graft-versus-host disease</td>
</tr>
<tr>
<td>HCP</td>
<td>Healthcare professional</td>
</tr>
<tr>
<td>HRQOL</td>
<td>Health-related quality of life</td>
</tr>
<tr>
<td>ICU</td>
<td>Intensive care unit</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>IDSA</td>
<td>Infectious Diseases Society of America</td>
</tr>
<tr>
<td>IF</td>
<td>Intestinal failure</td>
</tr>
<tr>
<td>IFALD</td>
<td>Intestinal failure associated liver disease</td>
</tr>
<tr>
<td>IGF-1</td>
<td>Insulin-like growth factor 1</td>
</tr>
<tr>
<td>INR</td>
<td>International normalized ration</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>IRP</td>
<td>Intestinal Rehabilitation Program</td>
</tr>
<tr>
<td>ITx</td>
<td>Isolated intestine graft transplantation</td>
</tr>
<tr>
<td>IVC</td>
<td>Inferior vena cava</td>
</tr>
<tr>
<td>KGF</td>
<td>Keratinocyte growth factor</td>
</tr>
<tr>
<td>LILP</td>
<td>Longitudinal intestinal lengthening procedure</td>
</tr>
<tr>
<td>LITx</td>
<td>Combined liver-intestine graft transplantation</td>
</tr>
<tr>
<td>LTx</td>
<td>Isolated liver graft transplantation</td>
</tr>
<tr>
<td>MELD</td>
<td>Model for End-Stage Liver Disease score</td>
</tr>
<tr>
<td>MVTx</td>
<td>Multivisceral graft transplantation</td>
</tr>
<tr>
<td>MVA</td>
<td>Microvillus atrophy</td>
</tr>
<tr>
<td>MVID</td>
<td>Microvillus inclusion disease</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>ND</td>
<td>Nasoduodenal</td>
</tr>
<tr>
<td>NEC</td>
<td>Necrotizing enterocolitis</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>NG</td>
<td>Nasogastric</td>
</tr>
<tr>
<td>NICU</td>
<td>Neonatal intensive care unit</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institute of Health</td>
</tr>
<tr>
<td>NIHRQOL</td>
<td>Neonatal and infants health-related quality of life</td>
</tr>
<tr>
<td>NIPS</td>
<td>Neonatal Infant Pain Scale</td>
</tr>
<tr>
<td>NJ</td>
<td>Nasojejunal</td>
</tr>
<tr>
<td>NNT</td>
<td>Number needed to treat</td>
</tr>
<tr>
<td>NORD</td>
<td>National Organization for Rare Disease</td>
</tr>
<tr>
<td>ω-3LCPUFA</td>
<td>Omega-3 long chain polyunsaturated fatty acid</td>
</tr>
<tr>
<td>ω-6LCPUFA</td>
<td>Omega-6 long chain polyunsaturated fatty acid</td>
</tr>
<tr>
<td>PAC</td>
<td>Partial autocorrelation</td>
</tr>
<tr>
<td>PAIN</td>
<td>Pain Assessment in Neonates Scale</td>
</tr>
<tr>
<td>PELD</td>
<td>Pediatric End-Stage Liver Disease score</td>
</tr>
<tr>
<td>PICC</td>
<td>Peripherally inserted central catheters</td>
</tr>
<tr>
<td>PN</td>
<td>Parenteral Nutrition</td>
</tr>
<tr>
<td>PNAC</td>
<td>Parenteral nutrition associated cholestasis</td>
</tr>
<tr>
<td>PNALD</td>
<td>Parenteral nutrition associated liver disease</td>
</tr>
<tr>
<td>PTLD</td>
<td>Post-transplant lymphoproliferative disorder</td>
</tr>
<tr>
<td>PYY</td>
<td>Peptide YY</td>
</tr>
<tr>
<td>QOL</td>
<td>Quality of life</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized controlled trial</td>
</tr>
<tr>
<td>SBS</td>
<td>Short bowel syndrome</td>
</tr>
<tr>
<td>SDGC</td>
<td>Study Center of the German Surgical Society</td>
</tr>
<tr>
<td>STEP</td>
<td>Serial transverse enteroplasty</td>
</tr>
<tr>
<td>STROBE</td>
<td>Strengthening the Reporting of Observational studies in Epidemiology</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>TPN</td>
<td>Total parenteral nutrition</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Thesis Overview

The objective of this thesis was to evaluate the change in clinical outcomes of pediatric intestinal failure (IF) that has been observed clinically over the past 10 years. The impact of implementation of a multidisciplinary intestinal rehabilitation program (IRP) and the effects of individual treatment options on clinical outcomes was assessed. With disease-specific mortality decreasing continuously, valid methodology for morbidity measurement became imperative.

Chapter 1 – 2 review topics relevant to the thesis projects:

Chapter 1 provides an overview over pediatric intestinal failure and its evolving management.

Chapter 2 discusses methodological challenges in pediatric IF research and outlines the rationale for the methodology chosen in the thesis projects.

Chapter 3 – 5 describe the three thesis projects including the systematic literature review on quality of life measurement in infants and neonates designed to meet the thesis objectives:

Chapter 3 represents the comparison of intestinal failure management eras. Early IRP experience is compared to a historic control and to a more evolved IRP including novel treatment options.

Chapter 4 uses time series analysis to evaluate change over time of clinical outcomes of pediatric IF. It quantifies the effect of a multidisciplinary IRP and individual treatment options adjusting for confounder and time-related issues such as delay of cause and effect and local trends.

Chapter 5 suggests a conceptual framework of neonatal and infant HRQOL based on focus groups and interviews with caregivers and healthcare professionals of neonates or infants with severe illness.
Chapter 6 synthesizes the results of all thesis projects, examines limitations and outlines future research directions.
1 Overview of Pediatric Intestinal Failure

1.1 Definition

Intestinal Failure (IF) is defined as the body’s inability to meet its energy and nutritional needs through the gastrointestinal tract\(^1\). More specifically in children, it is expressed as “insufficient functional intestinal mass needed to adequately digest and absorb nutrients and fluids required for appropriate growth and development”\(^{11}\). Pediatric patients with IF depend on parenteral nutrition (PN) for a prolonged period of time, which is associated with high morbidity and mortality\(^2\). The Canadian Association of Pediatric Surgeons adopted functional as well as anatomical diagnostic criteria consisting of PN-dependence for more than 42 days after bowel resection or functional loss, and/or less than 25% of expected age-specific remaining small bowel length\(^3\). This dual definition allows application to anatomical and functional IF-etiologies and has been used throughout the thesis.

1.2 History

The first successful intestinal resection was reported in 1880 and the direct relationship between remaining bowel length and long-term survival was described shortly thereafter\(^4,5\). In 1935, Haymond et al. reported near to normal survival for less than 1/3 of bowel reduction and certain mortality for >50% of bowel resection\(^5,6\). In the early 1960s, nomenclature was introduced to describe this condition more specifically\(^7,8\). Resection of more than 1/3 of intestine was considered massive intestinal resection, leading to long-term features of malabsorption including diarrhea, steatorrhea, weight loss, anemia, hypocalcaemia, hypomagnesaemia and hypoproteinaemia; it was called Massive Bowel Resection Syndrome or Short Bowel Syndrome (SBS)\(^7,8\). By 1967, several cases of adult SBS were reported and the first case of SBS in a child was described by Leonard et al.\(^7,10\).
Before the introduction of safe PN, mortality was the predominant outcome for any type of IF. Although PN has been available since the 1930s, it had not been established in a clinical setting until the 1960s\textsuperscript{11}. Over many years, the side effects of PN prohibited systematic, clinical application. This changed in 1962, when Arvid Wretlind presented the first non-toxic total parenteral nutrition (TPN) program, which at the time consisted of a glucose and fat solution soon to be known as IntraLipid\textsuperscript{8,11}. The results were published in 1964, making them available to a global audience\textsuperscript{12,13}. In 1968, Dudrick et al. showed that catheter placement in the superior vena cava was a safe approach allowing long-term PN-support\textsuperscript{14}. By 1972, “all-in-one” TPN-solutions consisting of water, energy from lipids and carbohydrates, amino acids, vitamins and trace elements, became available and were applied in cancer patients\textsuperscript{15}. Since the mid-1970s, success with PN-administration at home has been reported in adults and children\textsuperscript{16-19}. In parallel to advancements in PN, enteral nutrition was also evolving, introducing naso-jejunal and jejunostomy feeding approaches and developing elemental diet formulas leading to low fecal output in collaboration with the National Aeronautics and Space Administration (NASA)\textsuperscript{20,21}. Although the advances in PN were more dramatic, enteral nutrition remained a cheaper, safe and effective treatment of IF.

With the advances in solid organ transplantation by Alexis Carrel in the early 20th century, another treatment option for IF had been born. After Peter Medawar noticed the particular role of the immune system in the transplantation process, Murray and coworkers performed the first successful transplantation of a kidney in identical twins in 1954\textsuperscript{22,23}. Lillehi and Starzl performed first intestinal and multivisceral transplantations in dogs, leading to 25 years of intestinal transplantations in humans without survival\textsuperscript{24}. In 1987, the first successful multivisceral transplantation was performed by the Pittsburgh group, with early mortality due to lymphoma\textsuperscript{25}. The first long-term success of small bowel transplant with enteral autonomy was reported by Grant et al. in 1990\textsuperscript{26}. Finally, with the introduction of a potent immunosuppressant tacrolimus the fate of intestinal transplantations changed and it finally evolved into a successful and widely applied treatment option for IF\textsuperscript{27}.

In the early 1960s, non-transplant surgical solutions for SBS were developed with limited success\textsuperscript{7,9}. Autologous intestinal reconstruction surgery (AIRS) aimed at prolonging transit time such as reverses small intestinal interposition, recirculating small bowel loops, vagotomy, pyloroplasty, intestinal valves and gastric pouches were starting to be performed
in humans after being evaluated in animal models. Unforeseen complications from these procedures such as bacterial overgrowth resulted in limited utility and a quest for alternatives. In the 1980s, the focus was on procedures intending to increase functional intestinal absorptive surface such as intestinal plication and tapering with or without intestinal lengthening for patients with dilated bowel segments. Plication consists of folding the bowel wall in on itself at the antimesenteric border and then securing it with a longitudinal suture line leading to preservation of the entire mucosal surface and reduction of bowel diameter. Tapering enteroplasty on the other hand narrows the intestinal lumen by removing the antimesenteric section of the bowel and closing it with a longitudinal suture line. In patients with short residual bowel length, intestinal surface preservation may be important and a combination with intestinal lengthening ideal. A first intestinal lengthening procedure, the longitudinal intestinal lengthening procedure (LILP), was proposed by Bianchi et al. in 1980. Thirty years later, Kim et al. reported an alternative lengthening procedure, the serial transverse enteroplasty (STEP), that was technically simpler and allowed repeated performance. The LILP and STEP both take advantage of the intestinal dilation to increase bowel length and improve motility while preserving bowel surface.

A major advance in IF-management and outcome occurred with the introduction of multidisciplinary intestinal rehabilitation programs (IRPs) towards the end of the millennium. With a dedicated healthcare team, continuity in care was increased and longitudinal outcome registries became feasible. This led to a rapid evolution of new treatment options and significant optimization of IF-management. Although it was known since 1971 that there was an association between PN and the development of intestinal failure associated liver disease (IFALD), the particular role of lipids in IFALD pathogenesis was recognized by Colomb et al. in 2000. With IFALD representing the major cause of morbidity and mortality of IF, efforts were made to prevent progression by optimizing PN-protocols and IF-management. Most significantly, alternative lipid strategies such as lipid minimization or substitution of omega-6 long chain polyunsaturated fatty acid (ω-6LCPUFA) rich soybean-based lipid emulsions in PN with fish oil containing substantial omega-3 long chain polyunsaturated fatty acids (ω-3LCPUFA) led to a fundamental change in the management of IF in most recent years.
The clinical picture of IF has changed repeatedly over the last 50 years. Frontiers were pushed relentlessly in many different directions with the objective to continuously improve morbidity and mortality of IF, with an increasing emphasis on long-term outcomes in most recent years.

1.3 Epidemiology

The incidence and prevalence of pediatric IF are highly dependent on the definitions used and are therefore difficult to determine. In addition, prevalence is influenced by the provided healthcare, as for example with more efficient healthcare mortality of a disease may be reduced; thereby increasing the number of patients with a chronic course and the disease prevalence\textsuperscript{2,59}. The most frequent category of IF, short bowel syndrome (SBS), has been recorded as a rare disease in the American National Organization for Rare Disease (NORD).

Pediatric IF-prevalence and incidence have been extremely variable in the literature as a result of the heterogeneous definitions used and the rapidly evolving prenatal and neonatal management leading to altered outcomes\textsuperscript{2,59-61}.

Intestinal failure point prevalence has been estimated as 12.7 patients per 1 million children in the United Kingdom (UK)\textsuperscript{62}. In Germany the period prevalence is being described as 170 patients per year with a population of 81.8 million\textsuperscript{61}. The incidence density of IF has been portrayed as 85.7 patients per 100,000 live births in tertiary healthcare centers in Italy\textsuperscript{60}.

The incidence density of SBS has been reported as 24.5 patients per 100,000 live births for a Canadian tertiary healthcare center and the Pediatric Intestinal Failure Consortium stated a cumulative incidence of 3 – 5 per 100,000 live births per year\textsuperscript{2,3,63-65}.

The prevalence of children on home-PN due to SBS were projected as 2 patients per 1 million population in the UK in 1990, which is similar to the annual prevalence in Spain in 2005\textsuperscript{66,67}.
1.4 Etiology

Intestinal failure etiologies differ between adults and children. In adults, approximately 40% suffer IF secondary to cancer (SBS), while the remainder of cases are caused by Crohn’s disease, mesenteric ischemia, radiation enteritis, severe motility disorders and congenital abnormalities. In children, IF may be categorized into three etiological groups: structural loss of intestine (SBS), functional loss due to primary motility disorders, and structural enterocyte defects causing mucosal enteropathy. Some conditions may include aspects of two or all three etiological groups, however, as seen in gastroschisis.

Short bowel syndrome represents the most frequent IF-etioloogy in children. It is most frequently caused by congenital loss (e.g. congenital short bowel, intestinal atresias) or surgical resection of small intestine (e.g. necrotizing enterocolitis (NEC)) during the neonatal period. Other less frequent conditions leading to SBS are volvulus, Crohn’s disease, trauma, malignancies, intussusception, intestinal ischemia and spontaneous perforation. Short bowel syndrome may further be divided following to the anatomy of the intestinal loss: 1) small bowel resection with small bowel anastomosis and intact colon; 2) small bowel and partial colon resection resulting in enterocolonic anastomosis; 3) small bowel loss with high output jejunostomy. Type 2 anatomy (enterocolonic anastomosis with loss of the ileum) is the most common anatomical subtype.

Functional loss of small intestine from severe motility disorders and structural enterocyte defects is considerably less frequent accounting for approximately 10% of pediatric IF-cases. Severe motility disorders include chronic intestinal pseudo-obstruction (CIPO) and long-segment Hirschsprung’s disease/aganglionosis, while structural enterocyte defects include microvillus inclusion disease (MVID)/microvillus atrophy (MVA), intestinal epithelial dysplasia/Tufting enteropathy, intestinal lymphangiectasia and autoimmune enteropathy. As these patients suffer from a constitutive disorder of intestinal epithelial cells, they traditionally received intestinal transplantation, as their potential for enteral autonomy is absent.
1.5 Pathophysiology

After extensive small bowel loss, symptoms of malabsorption arise including diarrhea, steatorrhea, dehydration and abdominal distension and failure to thrive. Nutritional deficiencies may lead to anemia, edema, bleeding disorders, metabolic bone defects and neurologic manifestations. The degree of malabsorption depends on the location and extent of intestinal resection, as well as, the amount of intestinal surface and site-specific transport process loss. The ability of the intestine to adapt to the new circumstances ultimately determines whether enteral tolerance will be re-established.

1.5.1 Jejunal Resection

The jejunum represents the primary location of digestion and nutrient absorption in the human body. A jejunal resection will therefore lead to decreased absorption of most nutrients. The jejunum itself has a moderate ability to adapt to the new conditions after a resection and does that primarily by changing transport and enzyme activity.\(^76\). Interestingly, the ileum will eventually adapt its structure and function, and compensate for the jejunal loss. While the outcome after <75% of jejunum resection is generally good, the management is much more challenging when <25% of jejunum is remaining. Patients with a proximal jejunostomy typically have rapid transit leaving insufficient time for mixture of biliary and pancreatic secretions, digestion and enterocyte contact time with chyme, resulting in malabsorption and severe fluid and electrolyte loss.\(^35\) The epithelium of the jejunum allows free movement of water and salt across the mucosa to maintain an intraluminal sodium concentration between 70-90 mmol. This is in contrast to the ileum where the tight junctions between epithelial cells are “tighter” resulting in a net absorption of water and salt. Gastrin, cholecystokinin (CCK), secretin, gastric inhibitory polypeptide (GIP) and motilin secretion are typically preserved in IF-patients, but extensive intestinal resection will cause an imbalance of the fine-tuned regulatory secretion and motility mechanisms affecting the function of the residual bowel and aggravate malabsorption. Absorption of calcium, magnesium, phosphorus, iron and water-soluble vitamins takes place in the duodenum and proximal jejunum predominantly. Most patients with IF have a preserved duodenum and some jejunum, but usually have malabsorption of calcium and magnesium from precipitation with long-chain fatty acids in
from fat malabsorption. Along with fat malabsorption there may be malabsorption of fat-soluble vitamins A, D, E and K that require continuous supplementation.

1.5.2 Ileal Resection

The absorption of vitamin B12 and bile acids is unique to the ileum. Bile acids play important roles in bile flow regulation, intestinal lipid, nutrient and vitamin absorption, and as signaling molecules and inflammatory agents in lipid, glucose and energy metabolism. As part of the enterohepatic circulation, bile acids regulate bile acid synthesis and whole-body lipid homeostasis. Primary bile acids (i.e. cholic acid and chenodeoxycholic acid) are synthesized in the liver from cholesterol to replace bile acids lost with the stool. Then they are conjugated with taurine and glycine, secreted into the intestine where they solubilize lipids in micelles for absorption, re-absorbed in the terminal ileum and finally returned to the liver via the portal venous system. The bile acid synthesis process is down-regulated by cholic acid, up-regulated by cholesterol and inhibited by fibroblast growth factor 19 (FGF-19) released by the ileum in response to bile acid re-absorption. In the ileum, absorbed bile acids bind to FXR, which stimulates FGF-19 transcription. In the liver, bile acids activate the nuclear farnesoid X factor (FXR) and down-regulate de novo bile acid synthesis via a FGF-19 pathway. After ileal resection the enterohepatic circulation will be deranged, resulting in bile acid malabsorption. As bile acids reach the colon, bacterial deconjugation and dehydroxylation will produce secondary bile acids (i.e. deoxycholate and lithoholate). These will stimulate electrolyte and water secretion, resulting in diarrhea. Disruption of the enterohepatic circulation will also increase the relative cholesterol saturation due to loss of bile salts, facilitating the development of cholelithiasis.

Removal of >60cm of ileum will result in the need for lifelong vitamin B12 replacement in adults. After ileal resection, colorrheic diarrhea is observed where unabsorbed bile acids and long-chain fatty acids reach the colon and cause electrolyte and fluid loss. Prolonged lack of bile acid reabsorption leads to depletion of the bile acid pool and ultimately minimal micelle formation, that results in fat malabsorption, fat-soluble vitamin deficiency and steatorrhea. This is typically observed when >100cm of ileum is removed. The length of ileal resection causing Vitamin B12 and progression from colorrheic diarrhea to steaorrhea in children is not known. Patients with ileal resection and bile acid malabsorption are at a higher
risk for gallstones as bile salts maintain bile soluble. Fat malabsorption on the other hand promotes oxalate kidney stone formation as free fatty acids bind with calcium reducing the availability of calcium to complex with oxalate and therefore lead to higher oxalate absorption\textsuperscript{82}. As the ileum reabsorbs a large proportion of the fluid secreted by the jejunum during the digestive process, substantial ileal resection will lead to fluid and electrolyte loss\textsuperscript{83,84}.

Ileal resection is frequently combined with removal of the ileocecal valve that may favor development of bacterial overgrowth by colonization of the small intestine with colonic bacteria. This aggravates malabsorption as absorptive surface area is lost due to chronic mucosal inflammation. In addition, these refluxed bacteria consume vitamin B12 magnifying vitamin B12 deficiency and they deconjugate bile acids fostering bile acid depletion and steatorrhea\textsuperscript{85}. Bacterial overgrowth is associated with progressive intestinal dilatation and dysmotility that may fuel further bacterial overgrowth and severe complications such as bacterial translocation, liver injury, D-lactic acidosis, arthritis and colitis\textsuperscript{85-87}. The independent effect of the ileocecal valve on bacterial overgrowth, small intestinal transit time and complete weaning from PN remains unclear, however\textsuperscript{88}.

Glucagon-like protein 1 (GLP-1), Glucagon-like protein 2 (GLP-2), neurotensin and peptide YY (PYY) production may be significantly reduced with the loss of the ileum and proximal colon, resulting in rapid gastric emptying and intestinal transit. These hormones respond to intraluminal carbohydrates and fat by delaying gastric emptying, inhibiting gastric acid secretion and slowing intestinal transit\textsuperscript{88,89}. This mechanism, the “ileal brake”, has been shown to promote intestinal growth\textsuperscript{35}. Glucagon-like protein 2 is produced by the enteroendocrine L-cells in the ileum and colon. It is stimulated by luminal nutrients, and works through still unclear secondary messengers to facilitate intestinal adaptation. It enhances mucosal hyperplasia, decreases epithelial permeability, slows transit and delays gastric emptying.

With the limited ability of the jejunum to overtake ileal functions, the adaptability after ileal resection is poor and the prognosis fair\textsuperscript{35}. 
1.5.3 Colonic Resection

The colon plays an important role in water, electrolyte and short-chain fatty acid absorption. The presence of the colon in patients with extensive small bowl loss is beneficial, as it has a capacity for electrolyte and water and energy absorption, and augmented energy absorption from creation of short-chain fatty acids from metabolism of carbohydrates and fiber from colonic anaerobic bacteria\(^{90-92}\). In addition, the right colon is a site of GLP-2 production that has an intestinotrophic affect on the whole intestinal tract. The colon undergoes adaptation after small bowel resection by increasing the number of enterocytes and intestinal hormones such as PYY\(^{93}\).

Colon preservation is associated with a higher probability of complete weaning from PN after small bowel resection in adults\(^{94-96}\). In contrast, the outcome of infants and children with SBS is not dependent on colon presence and this seems less dramatic due to the tremendous small intestine growth potential that infants possess\(^{87,97,98}\).

1.5.4 Adaptation

The intestinal adaptation process involves structural and functional changes in the residual small bowel after extensive small bowel resection to compensate and improve nutrient and fluid absorptive capacity of the remnant bowel\(^{99,100}\). Adaptation starts within 48h after bowel resection and may require 2-3 years to complete\(^{50,101}\).

On a microscopic level, structural adaptation to intestinal resection consists of enhanced crypt cell proliferation in the small bowel leading to an increased density of enterocytes, villus hyperplasia, amplified local angiogenesis with increased tissue oxygenation and blood flow, and hypertrophy of the longitudinal and circular smooth muscular layers\(^{99,102-106}\). Gross alterations include bowel dilatation and elongation with the objective to enlarge the absorptive surface area\(^{50,107}\).

Functional changes take place after intestinal resection attempting to compensate for the sudden decrease of digestive and absorptive capacity\(^{50}\). The gene expression pattern is altered leading to increased expression of transporter proteins involved in nutrient, electrolyte, and water absorption\(^{107-111}\). Enterocytes mature more rapidly expressing digestive enzymes and
amino acid transporters earlier. Reduced intestinal transit time allows longer contact time between the nutrients and the absorptive mucosa. These alterations ultimately result is an increase of the digestive and absorptive capacity of the residual bowel.

A variety of factors influence the adaptation process including patient and anatomic characteristics, route of nutrition delivery, and hormones and growth factors. In the neonatal period, the intestine grows rapidly, making neonates more likely to achieve full enteral autonomy after extensive resection. Adaptive capacity is greater after jejunal resection as the ileum may adopt many of the morphologic features of the jejunum and the microscopic adaptation processes are more pronounced in the ileum. Residual bowel length is correlated with the amount of adaptation taking place and probability to achieve enteral autonomy. Enteral nutrition is crucial for stimulation of the adaptive response as GLP-2 is released by the presence of luminal nutrients. Ileal anatomy (extent of ileal resection, presence of ileocecal valve, continuity of small bowel with colon) and general nutritional status are other factors associated with successful PN-weaning. Hormones and growth factors play a key role in the adaptation process, making them logical candidates as pharmacological therapeutic agents to enhance enteral tolerance. Growth hormone (GH) alone or in combination with glutamine may improve energy absorption while being administered, but the clinical value and long-term safety remains unclear. Glucagon-like peptide 2 (GLP-2) has been shown to support the adaptation process by enhancing structural and functional changes, and it is approved for clinical use in adults. Other trophic proteins with promising preclinical results are epidermal growth factor (EGF), insulin-like growth factor 1 (IGF-1) and hepatocyte growth factor, with EGF having encouraging preliminary results for a pediatric population.

Functional aspects of adaptation may be measured in 72-hour metabolic balance studies such as fecal fat for fat absorption, α-1 Antitrypsin (A1AT) clearance for protein absorption and Serum levels of D-xylose for carbohydrate absorption. Plasma citrulline, a nonessential amino acid produced by small intestinal enterocytes has been promoted as a marker for residual muscular surface area and structural adaptation.
1.6 Clinical Management

The goal of IF-management is to optimize residual intestinal absorptive capacity and adaptation to regain enteral autonomy, while preventing or slowing the progression of other complications of chronic parenteral nutritional support. Medical and surgical treatment options have arisen in recent years with the objective to reduce morbidity and mortality from IF. These efforts emphasize preservation of functional bowel, and prevention of progressive IFALD and sepsis.

Clinical management of SBS takes place in stages starting with the use of PN\(^3\). Enteral nutrition is advanced until it eventually replaces PN. The goal is full enteral nutritional support, ideally by the oral route.

1.6.1 Multidisciplinary Intestinal Rehabilitation

Pediatric patient with IF are a complex, severely ill population. A dedicated, multidisciplinary program has shown to be beneficial\(^44,46,49,165\). The multidisciplinary IRPs typically consist of a gastroenterologist, a surgeon, nurse practitioners, dieticians, pharmacists, occupational/physical therapists and include social and psychological support\(^35,43,44,46,49,165-175\). Such coordinated and multidisciplinary care enables continuity, improved communication between specialties and institutions and with caregivers, more coherent and integrated planning, close monitoring of medication and nutritional strategies, early recognition and treatment of complications, better understanding of the natural disease, continuous scientific evaluation and quality assessment\(^175\).

The subjective perception that inception of IRPs led to improved clinical outcomes was quantified in a meta-analysis consisting of three pre-post studies\(^44,46,49,165\). Introduction of IRPs led to a significant increase of survival from IF and overall survival, each by 22%\(^44\). The systematic literature review further supported the hypothesis that IRPs are associated with reduced septic complications and development of IFALD, increased intestinal adaptation and more frequent intestinal lengthening procedures and transplantation\(^44\). It remains unclear whether the beneficial IRP effect is inherent to the multidisciplinary
approach itself or whether the improved outcomes are the result of new treatment options introduced with the IRP.

1.6.2 Medical Management

Immediately after intestinal resection there is a phase of high gastric output due to intestinal ileus for 2-5 days, followed by perfuse diarrhea and massive electrolyte loss\textsuperscript{35}. During this phase fluid and electrolyte replacement predominates. Parenteral nutrition offers the only nutritional support possible, and close clinical and laboratory evaluation is indispensable. Gastric and ostomy losses need to be replaced uncompromisingly. Enteral nutrition should be started as soon as the patient becomes stable to promote intestinal adaptation. With progression of the adaptation process, enteral feeding may be increased and PN weaned gradually. Early oral stimulation is crucial to prevent oral aversion. Hormones and trophic factors are currently being evaluated for their effectiveness in promoting intestinal adaptation.

1.6.2.1 Parenteral Nutrition

Parenteral nutrition is the cornerstone of IF-management and has been available to the pediatric population since the late 1960s\textsuperscript{14}. Since PN administration differs between adults and children, this chapter will focus on the pediatric population due to its relevance to the thesis research projects.

Parenteral nutrition is indicated for many medical and surgical conditions where the patient cannot be fed enterally. The intestinal tract may be functioning partially or not at all, making PN or total PN (TPN) necessary. Medical conditions include chronic diarrhea, inflammatory bowel disease, NEC, multi-organ failure, severe malnutrition, congenital motility disorders and malabsorptive syndromes, and prematurity, while surgical indications involve congenital anomalies and gastrointestinal surgery. Trauma including burns may make PN necessary as well. The role of PN is fundamental in IF, such that PN dependence is part of most definitions of IF.

Central venous access, defined as catheter tip placed in the proximal vena cava or right atrium, is required for long-term PN-administration due to high osmolarity of the PN-solution
infused to patients with significant or complete PN-dependence. Central venous access is associated with a lower risk for thrombus formation and vascular intimal damage experienced in patients with IF, as it allows instantaneous dilution of hypertonic PN-solution\textsuperscript{35,176}. The central venous catheters (CVC) are typically tunneled with the tip advanced into the superior vena cava (peripherally inserted central catheters (PICC), tunneled cuffed central catheters) or implanted under the skin (subcutaneous portacath). Since significant morbidity and mortality of IF may arise from CVCs, meticulous catheter insertion and management is crucial\textsuperscript{176,177}. Complications from CVC insertion include pneumothorax, hemothorax, arrhythmias, arterial puncture and primary malposition or malfunction of the catheter\textsuperscript{35}. Later complications of CVCs consist of catheter dislodgement and migration, catheter-related bloodstream infections (CRBSI) and line and venous thrombosis\textsuperscript{35,177}. Numerous risk factors and preventative features were tested, and the European Society for Clinical Nutrition and Metabolism (ESPEN) and the American Society for Parenteral and Enteral Nutrition (ASPEN) published guidelines for evidence-based CVC-management and PN-administration\textsuperscript{177,178}.

Prescription of PN is based on anthropometrics and clinical condition. Fluids and electrolytes are typically provided together with the PN-solutions and based on the results of a comprehensive nutritional assessment that includes baseline biochemical laboratory tests\textsuperscript{179,180}. Infused fluids need to account for insensible losses and other losses such as from stomas or fistulas. Fluid restriction may be required in patients with comorbid cardiac or renal disease. The baseline caloric requirements may be calculated following the ASPEN guidelines and range from 90-120 kcal/kg/day in a preterm neonate to 30-50 kcal/kg/day in a child aged 12-18 years\textsuperscript{181}. This approximate energy requirement may need to be adjusted for a number of patient factors such as sepsis or mechanical ventilation. Electrolytes, trace elements and vitamins require separate prescription and close monitoring as imbalances and deficiencies are frequent in patients with IF.

Targets for protein intake range between 3-4 g/kg/day for a preterm neonate and 0.8-1.5 g/kg/day for an adolescent\textsuperscript{181}. Crystalline amino acid solutions frequently used in PN-solutions provide 4kcal/g. Amino acids are typically started at 1-2 g/kg/day and advanced by 1-4 g/kg/day increments, which a higher starting dose being beneficial in premature neonates\textsuperscript{182}. Protein intake monitoring consists of blood urea nitrogen (BUN), pre-albumin...
and C-reactive protein, and in the case of liver failure, serum ammonia measurements. Severity of illness, stress factors such as sepsis, renal and hepatic function and metabolic disorders may influence protein requirements. Special pediatric amino acid solutions are available that contain a higher concentration of essential amino acids and have been associated with decreased risk of IFALD.

Intravenous lipid emulsions provide 20-50% of the energy provided by PN. Standard initial dose of lipids is 1 g/kg/day, that may be advanced to 3 g/kg/day for neonates and 1-2 g/kg/days for older children. The caloric value of lipids is calculated using the factor 9 kcal/g. Energy from lipids should not be less than 4% to avoid essential fatty acid deficiencies and not exceed 50% to prevent ketosis. Measuring serum triglyceride concentration monitors intravenous lipid intake. Hypertriglyceridemia may be influenced by common conditions observed in IF such as sepsis, acute stress response reaction, renal/hepatic dysfunction and by medication. Reducing the infusion time or infusing lipids on alternating days of the week may improve persistent hypertriglyceridemia. Soybean-based lipid emulsions (e.g. IntraLipid®) are used most frequently in PN-solutions, as they are rich in essential fatty acid providing omega-6 fatty acids. There is increasing evidence however, that high doses of soy-based lipid emulsions are associated with increased inflammation and liver injury leading to IFALD. Newer fish-oil based lipid emulsions (e.g. Omegaven®, SMOFLipid®) have a less inflammatory profile and seem promising to prevent progression and reversal of IFALD in children. Another tactic to reduce IFALD is by minimizing the lipid intake to 1g/kg/day or less, but concerns regarding long-term growth and neurodevelopment are a concern with such calorie restriction. Since Omegaven® is rich in Omega-3 LCPUFAs, exclusive nutrition with these emulsions may lead to essential fatty acid deficiencies. The composite lipid emulsion, SMOFLipid®, has recently been evaluated for its ability to prevent IFALD progression. The product consists of soybean oil (30%), medium chain triglycerides (30%), olive oil (25%), and fish oil (15%), and has an omega-6 to omega-3 ratio of 2.5:1. This composition is closer to the composition of breast milk than the traditionally used IntraLipid®. The hepatoprotective effect and improvement of the fatty acid profile observed in association with fish-oil based and compound lipid emulsions is explained by the anti-inflammatory cytokine pattern resulting from increased omega-3 lipids.
Glucose is the main energy source in the PN-solution and provides 40-60% of total calories at a caloric concentration of 3.4 kcal/g. The target carbohydrate dose is calculated based on the energy needs that are not met with fat or protein. Glucose should be initiated in an incremental fashion (starting at 5 mg/kg/min, increasing by 2-5 mg/kg/min per day) respecting the infusion rate that is age-appropriate. Blood glucose levels should be checked regularly to avoid hyperglycemia and glycosuria, and ultimately prevent hepatic steatosis.

Stable IF-patients with an adequate home environment and support may be candidates for home PN. Home PN enables nutritional support in a more natural environment, which promotes the child’s development and participation in social activities. With appropriate caregiver training, home PN is associated with less frequent CRBSI. Home PN should be managed by an IRP as strict monitoring, regular clinic visits and effective communication between the family and healthcare professions are crucial for successful home PN administration. A home nurse with significant experience with PN solutions, delivery systems and clinical management of IF-patients is another requirement. Quality of life has significantly increased with the availability of home PN programs.

1.6.2.2 Enteral Nutrition

Once the patient is stable enteral feeding should be initiated as early as possible, as it supports the adaptation process and is associated with favorable outcomes in the pediatric population. Continuous or bolus enteral feeding protocols were developed. While continuous enteral feeding allows for gradual increase of the enteral diet, which has been associated with reduced emesis, increased carrier protein saturation and decreased thermal energy losses, arguments can be made for bolus feeds as well. Clinical experience with bolus feeds were positive as they represent a more physiologic pattern of enteral nutrition, are developmentally appropriate, allow attempts at oral feeding better and provide more freedom to the patient (non-published data). Enteral diet advancement may be monitored by assessing stool output volume, or pH levels indicative for carbohydrate malabsorption in patients with functional colon. Appropriate hydration is important during this process, as mild dehydration has been associated with increased emesis and reduced absorption. The route of enteral feeding is based on the patient’s needs and typically starts with a nasogastric (NG) tube. Transpyloric nasoduodenal (ND) or nasojejunal (NJ)
feeding may be helpful for delayed gastric emptying, gastroesophageal reflux and patients at risk for aspiration. Patients with long-term enteral feeding may benefit from a gastrostomy (G) tube or gastrojejunal (GJ) tube, as they are less prone to dislodgement. It is unclear what the best formula is for enteral feeding of infants and children. Initially, breast milk or semi-elemental formula is recommended to benefit from the advantages of breast milk, such as EGF or GLP-2 in colostrum, or the hypoallergenicity of hydrosolate formula, as children with SBS are at an increased risk of developing protein allergies. Generally infants tolerate hydrosolate formulas better, as they have a lower osmotic load and higher fat content, but data is limited making the formula selection primarily following the clinical response. Older children do not require protein hydrosolate and more complex nutrients are required to stimulate intestinal adaptation. Approximately 40% of the caloric intake should be from lipids as they have a low osmotic value, high energy content and motility reducing effect. Particularly long-chain fatty acids have a strong adaptation stimulating effect. To prevent oral aversion, patients should be started on minimal oral feeds as soon as possible and even if this provides only minimal caloric content. Particularly infants and children are at risk to develop oral aversion, which may delay the transition to full oral feeds. As enteral feeding is tolerated, PN is weaned progressively. Fluid balance, electrolyte levels and nutritional status is followed closely during this weeklong to yearlong PN-weaning process. While macronutrients are typically well absorbed, attention must be given to micronutrient deficiencies and appropriate replacement provided.

Dietary supplements and trophic factors may be used to support the adaptation process and boost enteral tolerance. Soluble and insoluble fibers may be helpful as they slow down intestinal transit time and gastric emptying, and may be metabolized into rapidly absorbed short-chain fatty acids.

1.6.2.3 Hormone Therapy

Hormonal therapy involving GLP-2, GH, insulin-like growth hormone 1 (IGF-1), epidermal growth factor (EGF) and keratinocyte growth factor (KGF) has been advancing in most recent years. Overall, to date, their benefit has been somewhat disappointing with most gains being reduced parenteral fluid volume requirements. The results are reducible meaning therapy cannot be discontinued. The role in children is yet to be determined. Perhaps since
these treatments will be applied to children who are still undergoing active gut development, the results will be more robust. Long-term concerns exist and are related to neoplastic risk and cost.

Glucagon-like protein 2 seems to be a key factor in the adaptation process and it has a particular role in postnatal intestinal development. It is released by in the terminal ileum in response to direct contact with long-chain fatty acids and stimulates intestinal crypt-cell proliferation by increasing IGF-1 production by the pericryptal fibroblasts leading to augmented villus height and surface area, thereby increasing absorptive capacity\textsuperscript{136,139,218-222}. In addition, GLP-2 stimulates the EGF-receptor pathway and chylomicron formation in the mucosa and slows down proximal motility\textsuperscript{223,224}. Glucagon-like protein 2-stimulation seems to be modulated by enteric neuronal pathways\textsuperscript{225}. Clinical trials with GLP-2 have shown promising results in animals and adults\textsuperscript{130,135-144,226}. Effectiveness and safety of GLP-2 treatment in children with SBS have not been reported to this date, although results are anticipated soon.

Using the same pathway, exogenous IGF-1 has a strong impact on intestinal mucosa. Systemic therapy with IGF-1 for IF is risky in most cases however, because of the non-specificity of IGF-1 effects especially in a growing child\textsuperscript{222,227,228}. It may have a therapeutic value in a child with primary or secondary GH or IGF-1 deficiency in addition to IF, where exogenous GH or IGF-1 would treat both conditions concurrently\textsuperscript{229,230}.

Growth hormone likely affects the GLP-2 pathway by increasing IGF-1 systemically thereby increasing intestinal mass. Growth hormone with or without glutamine has led to variable results during exposure in children and adults with SBS, long-term efficiency and safety have not been determined yet\textsuperscript{107,125-134,231-235}. Localized therapy delivering GH or IGF-1 directly and exclusively to the intestine would be most useful and may increase the therapeutic value of these two hormones in the management of IF\textsuperscript{35}. In a recent Cochrane review, any benefits of GH were reversed after cessation of therapy\textsuperscript{125}.

Epidermal growth factor has been shown to be involved in the intestinal post-resection response and its ligands and receptors to be influenced by GLP-2\textsuperscript{224,236,237}. It is present in breast milk and saliva, and has a role in antenatal gut development. Enteral EGF seems to
improving active glucose absorption and reducing bacterial translocation in infants, but clinical experience with enteral or intravenous EGF is extremely limited\textsuperscript{35,146}.

The intestinal protective effect of KGF used in oncology to prevent postchemotherapy mucositis has been hypothesized to be useful in IF-management\textsuperscript{238-241}. Keratinocyte growth factor is procured by the peri-cryptal myofibroblasts, stimulates cell division and defines mucosal cell phenotype\textsuperscript{239-241}. Clinical studies have returned results indicating exogenous KGF’s primarily effecting the colon and leaving the clinical utility of KGF in IF undefined\textsuperscript{240}.

1.6.2.4 Other Medication Frequently Used

During the early phase after intestinal resection, transiently increased gastric acid secretion and serum gastrin levels may be controlled with intravenous H2 blockers or proton pump inhibitors\textsuperscript{200,242}. Hypergastrinemia after massive gut resection lasts for 6-12 months.

Motility medications to promote contraction can be very helpful and include agents such as domperidone, metoclopramide and cisapride. Antisecretory agents to reduce large stool volumes include loperamide, clonidine and octreotide, a somatostatin analogue, and have been shown to be safe and effective in prolonging transit time, reducing daily fecal volume, and diminishing fluid and electrolyte loss\textsuperscript{35,243-245}. Since side effects of these medications include an increased risk for bacterial overgrowth, they must be used carefully\textsuperscript{35,246}. While these short-term effects may be desirable, long-term effects remain unclear and side effects such as inhibition of intestinal regeneration, development of cholelithiasis, decreased hepatobiliary secretion, sphincter Oddi dysfunction and impairment of pancreatic exocrine function make somatostatin a non-routine medication in IF-management\textsuperscript{247-249}.

1.6.3 Surgical Treatment

Most cases of IF are the consequence of major intestinal surgery typically during the neonatal period in infants born prematurely and with comorbidities. At this time point numerous decisions must be made, many of them with long-term consequences.
1.6.3.1 Bowel Preserving Initial Surgery

The goal of the primary intestinal surgery is to preserve as much functional bowel as possible, while weighing the risk to life of metabolic derangement due to ischemia or sepsis. The majority of patients requiring intestinal surgery will not result in SBS, but some may end up with insufficient functional bowel required for growth and survival. The residual bowel length is best estimated by calculating the percentage of present bowel to the expected bowel by age \(^{98,250}\). Infants with less than 10% of remaining bowel have an increased risk of mortality and a significantly lower probability to achieve independence from PN\(^{98}\).

In addition to structural bowel remnant, the functional status must be considered when predicting outcomes. This is of particular importance in conditions where motility and/or absorptive capacity are affected as in gastroschisis or even after ischemic events. There is no standardized tool available to judge functionality of residual bowel and predict outcome, however.

Recent reports of successful management of patients with very short bowel after catastrophic bowel loss, as for example in midgut volvulus or NEC, more surgeons offer potentially life-saving nearly complete small bowel resection, although this comes with high costs in morbidity and long-term mortality\(^{251-253}\). This decision remains an ethical challenge during the initial phase of IF-management.

Judgment of viability of the intestine is another challenging aspect during the initial surgical intervention. Techniques that are useful in adults, such as Doppler to assess mesenteric vessels, may not be applicable to pediatric patients. In general, intestine that has borderline viability is left in the abdomen. A valuable option in patients with extensive bowel pathology is the second-look operation, where the bowel is reassessed 24-48 hours later, hopefully revealing a clearer picture\(^{254-256}\).

The decision to create a diverting stoma at the primary operation or perform an anastomosis is a judgment call and dependent on patient stability. Intuitively, bowel continuity permits more aggressive, early enteral feeding that stimulates intestinal adaptation. In addition, stomas can be associated with significant morbidity\(^{199,257,258}\). Stomas remain the safest choice and patients at a higher risk for anastomotic failure or very short residual bowel length may
benefit from temporary diversion\textsuperscript{35,255,259,260}. Early restoring surgery is favored and may take place once an infant is stable and \(>2\text{kg}\) of weight, typically 4-6 weeks after primary surgery\textsuperscript{35}.

1.6.3.2 **Procedures To Improve Intestinal Function**

Patients with abdominal pathology are at a higher risk of developing intestinal stenosis and obstruction leading to proximal bowel dilation, dysmotility and bacteria overgrowth\textsuperscript{50}. Such mechanical obstruction may be improved with lysis of adhesive bands, stricturoplasty for short residual bowel or segmental resection for patients with adequate bowel length\textsuperscript{50,261}.

Patients with stomas may benefit from reestablishing intestinal continuity as early as possible, as it increases absorptive surface area and allows colonic absorption providing up to 5-10\% of calories if the colon is preserved\textsuperscript{262}. In addition, exposure to luminal nutrients, gastrointestinal secretions and humoral factors prevents mucosal atrophy caused by diverted intestine and it supports the adaptation process\textsuperscript{262-267}. Disadvantages of stoma closure consist of diarrhea resulting in perineal irritation, and a higher risk for oxalate nephrolithiasis in patients with colon\textsuperscript{268,269}.

1.6.3.3 **Procedures To Improve Intestinal Motility**

Intestinal tapering and plication may offer a surgical solution for symptoms from dysmotile intestine secondary to dilation. It is important to discriminate between dysmotile bowel segments and obstructed bowel with proximal bowel dilation, as mechanical obstruction may be solved with release of adhesions, stricturoplasty or resection of the obstructive segment, while tapering or plication improves motility and decreases bacterial overgrowth by narrowing the intestinal lumen\textsuperscript{50}. Tapering consists of partial resection of the anti-mesenteric border of the dilated bowel, which comes with an inevitable loss of some absorptive surface area. Plication reduces the bowel diameter without bowel wall resection by folding the dilated bowel longitudinally and securing it with a suture line\textsuperscript{270}. Disadvantages of this approach include bowel obstruction from the folded bowel segment and suture line failures\textsuperscript{268}. 
Antiperistaltic small intestine interposition segments represent a surgical treatment option to slow intestinal transit time\textsuperscript{50,271-273}. The technique is mostly used in adults and its favorable response does not come without complications such as complete bowel obstruction\textsuperscript{271-273}. Since neonates and infants have a greater potential for intestinal growth and there is a lack of sustained benefit of this procedure in this population, it is not widely used in pediatric IF\textsuperscript{50}.

A similar surgical method to reduce intestinal transit time is isoperistaltic interposition of a colon segment into the small intestine. The main advantage over antiperistaltic small intestine segments is regarded to be the lower risk for intestinal obstruction\textsuperscript{50,274}. The experience with this procedure is small and the results in pediatric patients variable, so that it is not a first line treatment option in IF\textsuperscript{50,275,276}.

With the understanding that the ileocecal valve affects intestinal transit time and bacterial reflux from the colon into the ileum, surgical methods were developed to create intestinal valves expecting similar effects\textsuperscript{50,277}. Intestinal valves are basically a short intussusception, where intestine is telescoped into the distal bowel and secured with seromuscular sutures\textsuperscript{278}. The obstructing effect of the valve promotes adaptation and induces proximal bowel dilatation that may be used for an intestinal lengthening procedure\textsuperscript{268,279}. One limitation of the procedure is that it may only be performed where sufficient residual bowel is available as 8 cm are required to perform the valve construction\textsuperscript{50,273,274}.

### 1.6.3.4 Intestinal Lengthening Procedures

Bowel dilatation is part of the adaptation process and represents an effort of the bowel to compensate for the lost absorptive surface area. There are a number of surgical procedures that utilize this additional surface area and taper the bowel at the same time leading to improved function and motility.

The first intestinal lengthening procedure was developed by Bianchi et al. in 1980\textsuperscript{39}. The longitudinal intestinal lengthening procedure (LILP) profits from the dual blood supply of the small intestine that is hemi-circumferential. The bowel and mesentery are divided longitudinally between the two blood supply networks and the two newly created bowel segments are then anastomosed end-to-end resulting in a bowel segment that is twice as long and half as wide\textsuperscript{50,273}. This procedure is indicated in patients with TPN and very short, dilated
residual bowel. There is significant experience with this procedure in adults and children, and overall it leads to complete PN-weaning in approximately 71% of survivors\textsuperscript{280-283}. Complications from LILP include adhesive bowel obstruction, anastomotic strictures or leaks, and recurrent bowel dilatation\textsuperscript{50,273,283}. Patients with recurrent bowel dilation may become candidates for other intestinal lengthening procedures such as the Kimura or STEP.

The technically challenging aspect of separating the mesentery and handling its delicate vasculature, its dependency on a mobile and intact mesentery, and the impracticality of repeated LILP procedure may be regarded as arguments against this procedure.

The Kimura procedure, or isolated bowel segments, is an intestinal lengthening procedure that may be used when the mesentery is not compatible with LILP\textsuperscript{284}. The procedure consists of two parts starting with attaching the anti-mesenteric border of the dilated bowel segment to an intra-abdominal host organ, such as the abdominal wall or the liver, that has had its peritoneum removed\textsuperscript{50,285}. After collateral blood supply has developed from the host organs to the anti-mesenteric surface of the bowel, the bowel is then divided horizontally, and the new bowel segment interposed and anastomosed end-to-end resulting in a lengthened and tapered bowel\textsuperscript{50,285}. Disadvantages of this procedure include the difficulty to access the abdominal cavity due its non-anatomical placement of the bowel, and theoretically, the lack of nutrient blood to flow through the portal system. It is essentially never performed clinically. The experience with this procedure is limited, making it a second-line treatment option\textsuperscript{50}.

The STEP procedure has been added to the intestinal lengthening procedures since its inception in 2003\textsuperscript{40,41}. The procedure is regarded technically less challenging than the LILP, may lead to more than double of bowel length and may be performed primarily or secondarily after LILP or STEP\textsuperscript{50}. The procedure starts with removing all abdominal adhesions and marking the anti-mesenteric border with a surgical pen. A stapler is then applied from alternating sides of the intestine through small mesenteric windows resulting in a zigzag channel. As with LILP and Kimura procedure, the bowel is lengthened and tapered, but in contrast to the LILP, the vessels of the mesentery are not in danger in STEP, making it a safer option. Long-term outcomes after STEP are positive, with approximately 50% of all patients and 75% of surviving patients achieving PN-independence\textsuperscript{101}. 
Sequential bowel lengthening procedures after intentional bowel dilation with intestinal valves or after primary lengthening procedures in patients with refractory SBS have been performed successfully. Some patients may also require transplantation after initial intestinal lengthening procedure and this has been reported as 26% after LILP and 16% after STEP.

1.6.3.5 Transplantation

Transplantation may offer a valuable treatment option for patients with IF and small potential for intestinal adaptation. In general, indication for transplantation is given when the probability of survival with a transplant is larger than the probability of survival with continuing non-transplant therapy. This prediction is difficult however, due to the heterogeneous nature of IF, variable IF-management among different healthcare centers, fast treatment evolution, and small sample size and retrospective design of studies. Optimal timing for referral is therefore vague.

Nevertheless, criteria for referral are required to guide clinical decision-making, as transplantation is an important treatment option for IF-patients when conservative measures fail. Multidisciplinary intestinal rehabilitation teams have noted that early referral for transplant assessment is a crucial factor contributing to improved IF-outcomes. Consensus groups have defined pediatric criteria for referral for transplant evaluation. The most recent consensus statement consists of the following referral criteria: massive small bowel resection, severely diseased bowel and unacceptable morbidity, continuing prognostic or diagnostic uncertainty, microvillus inclusion disease or intestinal epithelial dysplasia, persistent hyperbilirubinemia >6 mg/dl (102.6 µmol/L), thrombosis of 2 of 4 upper body central veins and request of the patient or family.

The multidisciplinary transplant team decides on whether the criteria for transplantation are fulfilled, deliberates over the type of graft needed and assesses comorbidities. At this time point imaging performed and factors relevant to the surgery are evaluated including abdominal vasculature and anatomy, surgical history and central venous access. There are a number of different transplantation grafts available for IF-patients and their suitability depends on the clinical IF-features: isolated intestine graft transplantation (ITx), combined
liver-intestine graft transplantation (LITx), multivisceral graft transplantation (MVTx) or isolated liver graft transplantation (LTx)\textsuperscript{291}. Isolated intestinal transplantation is ideal for a patient who suffers from permanent IF without significant IFALD. In ITx the proximal donor jejunum is anastomosed to the recipient’s jejunum and an end ileostomy is placed to monitor the graft visually and with serial biopsies. Arterial blood supply is obtained from the aorta and venous drainage is into the inferior vena cava (IVC) or the portal system\textsuperscript{291}. Combined liver-intestine transplantation is indicated for IF with IFALD. In LITx the donor’s liver, duodenum and pancreas is transplanted together with the intestine, leaving the recipient with two duodenums and two pancreases\textsuperscript{291}. This technique has established itself because it avoids donor portal dissection and biliary anastomosis, which was necessary with the preceding technique where the duodenum and pancreas of the donor were removed\textsuperscript{26,35,292}. The recipient jejunum is anastomosed end-to-side to the donor’s jejunum. Arterial inflow is typically derived from the infra-renal aorta and venous outflow runs through a piggybacked anastomosis to the supra-hepatic IVC\textsuperscript{26,35,292}. Multivisceral graft transplantation is indicated when there is a need to replace the stomach together with the duodenopancreatic complex or where there is a significant size discrepancy between the donor and the recipient, which may be the case in small infants. In MVTs, removal of more recipient’s organs and absence of double organs generates more space\textsuperscript{35}. The proximal anastomosis is done between the two gastric fundi and pyloroplasty is performed, as the stomach lacks vagal innervation\textsuperscript{35}. The vascular anastomoses are similar to the ones in LITx. For patients with a high probability to achieve enteral autonomy and IFALD, LTx is the preferred option as portal hypertension may be a contributing factor to the inability to wean from PN completely\textsuperscript{35,293}. Traditionally, for a patient to be a candidate for isolated liver transplantation, the residual intestine needs to show adaptive potential. Patients need to tolerate at least 50% of their nutrition enterally for this option to be considered.

Availability of organs is limited and scoring systems are used to determine the urgency of each case. For children <12 years old the Pediatric End-Stage Liver Disease (PELD) score is used in North America. It calculates the 3-month mortality risk based on bilirubin, albumin, and international normalized ration (INR) values, and a growth impairment factor for infants <1 year of age, and accounts for the need of a simultaneous intestinal transplant. Isolated intestinal transplantations are solely based on the waiting time. Waitlist mortality and post-
transplant mortality have improved significantly over the past 10 years\textsuperscript{291,294}. In 2009, a pooled experience with 2000 adult and pediatric transplantations showed a 1-year and 3-year survival rate of 80\% and 65\%, respectively. In contrast to these improved short-term outcomes, long-term 10-year survival remained at 45\%. High-volume centers seem to have a beneficial impact on the outcome achieving a 5-year survival rate of 77\%\textsuperscript{291,296}. There seems to be an increasing trend for patient being re-transplanted after organ loss, with one high-volume center reporting 8.1\% of pediatric intestinal transplants being re-transplantations with a 5-year survival rate of 71\%\textsuperscript{291,296}. Principal cause of death among transplanted patients is sepsis in 50\%\textsuperscript{290,295}. Other complications potentially leading to organ loss or mortality include surgical complications such as bleeding, thrombosis, anastomotic leaks, bowel perforation and loss of domain, organ rejection, viral infections, post-transplant lymphoproliferative disorder (PTLD) and graft-versus-host disease (GVHD)\textsuperscript{35}.

It has been perceived clinically that in recent years the need for transplantation in patients with IF decreased with an increasing number of patients not being waitlisted for transplantation after assessment and those on the waitlist to be delisted secondary to improved clinical status. The international intestinal transplantation registry has shown a steady decrease in bowel transplants in patients aged 0-18 years since 2008. This is thought to be related to improved patient outcomes with multidisciplinary intestinal rehabilitation teams.

1.6.3.6 Other Procedures Frequently Performed In Patients With Intestinal Failure

Patients with IF typically have a rich surgical history with numerous surgical procedures. Besides the initial intestinal resection procedure, autologous bowel reconstruction and transplantation, these patients require interventions to place central venous catheters, create and reverse stomas, and insert feeding enterostomies (e.g. gastrostomies). At times cholecystectomy is required for PN induced cholelithiasis that can occur in 25\% of patients. Liver biopsy is commonly performed at the time of autologous bowel reconstruction and also after transplantation. Endoscopy is frequently performed to inspect the gastrointestinal structure and diagnose any local complications and assess for rejection in setting of bowel transplantation.
1.7 Outcomes

Intestinal failure is associated with high morbidity and mortality. Overall mortality is traditionally between 25 – 40%, but the introduction of IRPs has had a significant impact on improving this mortality\textsuperscript{46,49,98,199,297-299}.

The ideal outcome of IF is complete weaning from PN and achievement of enteral autonomy without undergoing transplantation. Although individual risk factors for poor outcome of IF have been determined, the results of individual studies have been inconsistent and may have changed over time\textsuperscript{98}. Also, no outcome predicting score or severity of illness measurement tool has been developed. Similarly, health-related quality of life (HRQOL) measurement tools for the typical age group of IF, which is the neonatal and infant age period, have not been developed, nor have disease-specific HRQOL instruments been developed for older IF-populations. Comprehensive and valid morbidity measurement is therefore nearly impossible for this population at this time point.

1.7.1 Complete Weaning From Parenteral Nutrition

Of children with IF, 40% to 85% eventually wean off PN with SBS-patients having a higher probability when compared to those with motility disorders or structural enterocyte defects\textsuperscript{43,60,98,300}. The presence of an ileocecal valve and residual bowel length of ≥10% of expected length are associated with increased probability of 11.8 times and 3.9 times to wean completely off PN\textsuperscript{98}. Serum citrulline, a marker for enterocyte mass, has been shown to be predictive for PN-weaning\textsuperscript{301}. Serum citrulline levels of ≥19 µmol/L have a sensitivity of 94% and specificity of 67% of predicting pediatric patients with IF who will be able to be weaned off PN\textsuperscript{301}.

1.7.2 Morbidity

Comorbidities and complications are very frequent in pediatric IF, and significant efforts were made with the introduction of IRPs to improve this morbidity and impact on quality of life.
1.7.2.1 Comorbidity

The great majority of IF-patients are born prematurely\(^3\text{-}^70\). Their poorer survival rate is not necessarily related to IF, as their potential for intestinal adaptation is greater, but to comorbidities associated with premature birth\(^70\). Comorbidities are a significant factor in the decision-making process of these patients, as they may not allow extensive surgical interventions or they may support the decision for withdrawal of care\(^70\).

Comorbidities account for a significant part of mortality ranging from 20% to 33% in a cohort before introduction of IRP and in one representing an early IRP experience\(^49\). With significant improvement of IF-management, this proportion may have increased recently.

1.7.2.2 Complications From Intestinal Failure

The most significant complications of IF are related to long-term need of a CVC and exposure to PN. Intestinal failure associated liver disease and repeated sepsis represent limiting factors during the adaptation process with potentially fatal outcome. Recent advances in IF-management such as novel lipid strategies or ethanol locks have aided in reducing the progression of IFALD and incidence of sepsis.

1.7.2.2.1 Intestinal Failure Associated Liver Disease

Although introduction of safe PN signified a major advance in the management of IF, it was rapidly understood that long-term exposure led to cholestatic liver disease or IFALD\(^48\). Intestinal failure associated liver disease represents a clinical spectrum ranging from mild cholestasis to hepatic fibrosis and liver failure. After 1 to 4 weeks of PN, biochemical characteristics of IFALD such as elevated conjugated serum bilirubin and gamma-glutamyl transferase (GGT) become evident\(^302\text{-}^303\). Children are more susceptible to IFALD than adults and demonstrate a predominantly cholestatic picture with comparatively rapid progression into advanced liver disease\(^304\text{-}^305\). Adults with IFALD experience steatosis primarily and present initially with elevated transaminases, making jaundice a late and concerning sign\(^306\text{-}^307\). In adults and children, the development of biliary sludge and cholestasis is related to the duration of PN administration and both processes are probably part of the pathogenesis of IFALD\(^308\text{-}^311\). With IFALD progression other features of liver failure such as synthetic
dysfunction and hypersplenism resulting in coagulopathy and thrombocytopenia become evident\textsuperscript{35}. Serum conjugated bilirubin levels have been shown to have 94% sensitivity and 87% specificity for end-stage liver disease in children\textsuperscript{312}. Although some patients may develop progressive fibrosis or cirrhosis in the absence of biochemical changes, the most practical biomarker used for tracking IFALD clinically has remained conjugated bilirubin.

The epidemiology of IFALD is difficult to determine as a variety of different definitions are being used\textsuperscript{3,65,185,306,307,313-315}. Intestinal failure associated liver disease is defined based on laboratory values, but these may vary between studies\textsuperscript{35}. Most recently, definitions used are based on serum conjugated bilirubin levels only, which is in contrast to earlier definitions which use constant elevation of various liver enzymes instead\textsuperscript{3,65,185,306,307,314}. The definition will alter the incidence of IFALD in the sense that it is increased if multiple liver tests are taken into account and it is decreased if serum conjugated bilirubin is used exclusively\textsuperscript{35,313,315}. The reported incidence of IFALD in infants on long-term PN ranges from 40% to 60%, and in adults on home PN it lies between 15% and 40%\textsuperscript{35,304}. End-stage liver disease develops in approximately 25% of pediatric patients with SBS and 15% in adult patients on home PN\textsuperscript{49,306}. End-stage liver disease in children results frequently in mortality due to limited availability of size-appropriate allografts\textsuperscript{35}.

Risk factors for IFALD and progression into end-stage liver failure have been identified. Children born prematurely and with a birth weight of <500 grams have the greatest risk to develop IFALD\textsuperscript{65,302,316}. Preterm infants with SBS on PN for >3 months have a risk for IFALD of up to 90%\textsuperscript{35,317}. Etiology of IF plays another important role, with intractable diarrhea of infancy, gastrochisis and jejunal atresia being associated with an increased risk for IFALD\textsuperscript{173,316}. In adults, residual small bowel length, comorbid liver disease, female gender and absence of colon in continuity have been associated with elevated risk for IFALD\textsuperscript{185,307,313,318}. Besides demographic risk factors, PN-associated and host factors have been found to contribute to IFALD development\textsuperscript{35}. Hypercaloric feeds with excess macronutrients and a delicate interplay of amino acids, lipids, minerals and trace elements have been suggested to be involved in IFALD pathogenesis\textsuperscript{35,319-322}. Host factors favoring IFALD development include immaturity of hepatic pathways from prematurity, and IF itself, which is associated with absence of enteral nutrition, deranged enterohepatic circulation, intestinal stasis and recurrent episodes of sepsis\textsuperscript{35,65,199,305,323-327}. Immature hepatic pathways
in premature neonates favor cholestasis. Animal models indicate that this is related to altered expression of nuclear receptors (e.g. farsenoid X receptor (FXR)) and their effect on expression of bile acid synthesis and transport. In addition, toxic bile acids may accumulate and cause secondary oxidant injury due to a deficiency in antioxidants and other protective molecules (e.g. glutathione) of the neonatal liver.

Based on the known IFALD pathomechanisms, a variety of treatments have been developed. Nutritional strategies include avoidance of overfeeding of parenteral macronutrients and calories, minimizing certain trace minerals such as manganese and chromium, and aggressive enteral feeding. Restoration of gastrointestinal continuity and autologous bowel reconstructive procedures have been suggested to positively affect IFALD progression. For patients with expected prolonged PN dependence, cycling infusions of PN may be beneficial as they result in lower insulin levels, less hepatomegaly and improved biochemistry. The interrupted PN administration allows for free fatty acid mobilization from body stores, which is thought to protect the liver by promoting more efficient energy utilization and reducing hepatic stress. Lipids seem to play a key role in IFALD development making them particularly interesting for therapeutic approaches. It has been shown that temporary cessation or reduction of traditional soybean-based parenteral lipids improves or resolves hyperbilirubinemia and thrombocytopenia. This may due to the reduced exposure to omega-6 fatty acids or to the reduced exposure to phytosterols present in soybean emulsions. Concerns regarding long-term neurodevelopment have been raised with lipid reduction.

Changing the composition of parenteral lipids has been established as a therapeutic option to reverse IFALD. Replacement of soy-based lipid emulsions (e.g. IntraLipid®) with lipid emulsions derived from fish oil containing substantial amounts of ω-3LCPUFA (e.g. Omegaven®) have been shown to resolve IFALD in 45% to 63% of pediatric IF cases with IFALD. Another alternative lipid emulsion consisting of 30% soybean oil, 30% medium-chained triglycerides, 25% olive oil and 15% fish oil (SMOFLipid®) are currently being tested for their efficacy in treating IFALD as well. The therapeutic effect is hypothesized to be secondary to a less inflammatory and liver injury causing profile of newer fish-oil based lipid emulsions (e.g. Omegaven®, SMOFLipid®) than high doses of soy-based lipid emulsions, which may prevent progression and reversal of IFALD in children.
All alternative lipid emulsions are currently being used in investigatory settings only\textsuperscript{35,190,335,336}.

Pharmaceutical treatment options for IFALD are limited. Ursodeoxycholic acid, a bile acid, has been given orally to treat cholestatic liver disease in adults, but weak evidence of therapeutic effect in children and side effects have limited its applicability in pediatric patients with SBS\textsuperscript{35,333,337-343}. In patients with no ileum it may worsen symptomatology and increase diarrhea\textsuperscript{35}. The therapeutic effect of ursodeoxycholic acid is not fully understood, but hypothesized to be by correcting bile acid deficiency, improving bile flow, displacing cytotoxic bile acids, or providing immunomodulatory protection\textsuperscript{333,344}. Other pharmaceutical options that are being used despite lack of high-quality evidence include N-acetylcysteine and choline supplementation\textsuperscript{345,346}. Controlled studies have shown that cholecystokinin and tauroursodeoxycholic acid are not effective as IFALD treatment\textsuperscript{343,347,348}.

The ultimate solution for patients with IFALD refractory to other treatments is transplantation. Limited availability of size-appropriate grafts and intermediate outcomes after intestinal transplantation makes ITx, LITx and MVTx a less popular therapeutic option\textsuperscript{3,49,291,294,305,349,350}. In patients with a high probability to achieve enteral autonomy, LTx may represent a preferable treatment\textsuperscript{351,352}.

1.7.2.2.2 Sepsis

As sepsis is a major contributor to IFALD development and mortality, prevention of CRBSI and bacterial overgrowth leading to bacterial translocation is therefore crucial\textsuperscript{35,49,86}. Sepsis represents the most frequent and serious complication from CVCs required for PN\textsuperscript{353}. Medical care for sepsis may include admission to the ICU for a mean length of 6 days and up to 21 days of admission to the hospital\textsuperscript{35,354}. Mortality from sepsis is estimated as 12\% to 35\% per infection\textsuperscript{35,354}. Patients with IF are at a particular risk for CRBSI and approximately 50\% of admissions of patients on home PN is due to a CVC-associated infection\textsuperscript{47}. Non-immunosuppressed patients have a reported CRBSI rate of 0.5-2 infections per 1000 catheter-days, while immunosuppressed patients including hematology and oncology patients have 2-3 infections per 10000 catheter-days\textsuperscript{355-358}. Children with SBS have a considerably higher CRBSI rate of 2.1 to 9 infections per 1000 catheter-days\textsuperscript{47,359-363}.
Fever and glucose intolerance in patients are suggestive for sepsis. Paired blood culture sampling from the hub or tip of the suspected CVC and a peripheral vein will confirm the diagnosis \(^{364}\). Treatment should follow evidence-based guidelines such as the Infectious Diseases Society of America (IDSA) guidelines \(^{364}\). In general, IF-patients should be started on empirical combination antibiotics and possibly anti-fungal therapy \(^{35,364}\). Depending on the severity of illness, catheter type and causative microbiological agent, the CVC may have to be replaced \(^{35,364-368}\).

Catheter-related bloodstream infections may originate from the insertion site, the catheter hub or via the bloodstream \(^{35}\). The catheter hub has been considered the most common source of CRBSIs, but infections may be secondary to bacteria around the insertion site or seeding from other locations as for example the bowel \(^{357,369-371}\). Preventative measures directed towards all three sources for infection have been tested. Prevention of CRBSI is crucial in pediatric patients with IF, as central venous access sites are limited and CRBSI increases IFALD development, diminished HRQOL drastically and is potentially fatal. Many CVC handling techniques have been optimized, which have led to improvement in complication rates \(^{35,177}\). Proper catheter selection, routine hygiene measures for dressing changes and CVC accessing, and dressing products with a chlorhexidine-impregnated sponge have been shown to be beneficial \(^{35,177,178,372,373}\). Putting a formal protocol for catheter insertion and care into place has been shown to reduce CRBSI rates from 11.3 to 1.6 per 10000 catheter days \(^{369}\). While tunneling of catheters has turned out to not be beneficial, polyurethane and silicone catheters appear to be superior materials reducing the CRBSI rate \(^{120,370}\). Antimicrobial-coated catheter devices have been evaluated widely creating evidence that externally impregnated CVCs with chlorhexidine, silver sulfadiazine or minocycline and rifampin lead to reduced CRBSI-rates \(^{35,374,375}\). In children with IF their use is limited due to concerns regarding development of antibiotic resistance and unavailability of such devices for small infants \(^{35}\). Systemic antibiotic CRBSI therapy is successful in 60% to 91% and failure is usually due to incomplete eradication of organisms in the biofilm \(^{376}\). Antibiotic-locks have been reported to prevent and treat CRBSI by filling the CVC with higher antibiotic concentrations and for a prolonged period of time to facilitate biofilm penetration \(^{364,377-379}\). A similar technique, which avoids antibiotic resistance development, is ethanol-lock therapy \(^{380-389}\). A dramatic CRBSI rate reduction of 81% has been reported in a meta-analysis including pediatric IF-patients.
exclusively\textsuperscript{381}. This dramatic CRBSI rate reduction is of particular importance for pediatric patients with IF and limited central venous access sites. Advantages of this technique include that the organism and its sensitivity does not have to be identified, multi-drug-resistant organisms are sensitive to ethanol and the use of broad-spectrum antibiotics and associated resistance development may be reduced\textsuperscript{35,379}. Ethanol-lock protocols using 70\% ethanol solution and dwelling time of \( \geq 2 \) to \( \geq 4h \) per day have been applied in children with IF\textsuperscript{360,381-384}. The ideal ethanol-lock protocol has still to be defined and a potentially increased risk for thrombotic events has to be evaluated in controlled trials. Other complications suggested include an impact on catheter integrity, compatibility with heparin or citrate solutions and a theoretical impact on neurodevelopment of pediatric patients when the ethanol lock solution is flushed\textsuperscript{35,360,390-392}. Prevention of bacterial translocation by preventing and treating small intestinal bacterial overgrowth and optimizing enteral motility may influence CRBSI development, which illustrates the complexity of IF-patients.

1.7.3 Quality Of Life

Mortality and morbidity for pediatric IF have decreased significantly over the past decade, making the question of how this improvement has affected HRQOL in this population increasingly relevant. While generic HRQOL measurement tools for adults and older children have been developed and validated, disease-specific instruments for IF and those developed specifically for the most relevant age-range in pediatric IF, neonates and infants, are nearly absent. Moreover, a valid conceptual framework for describing HRQOL in neonates and infants is lacking. Due to this lack of validated methodology, a deficiency in knowledge and evidence around pediatric IF HRQOL is apparent. The crucial role of high-quality HRQOL data in patient- and family-centered medical care, as well as resource distribution and medical decision-making further illustrates the potential problems arising from the absence of valid HRQOL measurement tools and constructs.

Another difficultly related to HRQOL measurement in pediatric IF is that the disease consists of a variety of health states ranging from patients admitted to the ICU with a CRBSI and liver failure to a stable patient on home-PN. Furthermore, the age range of pediatric IF typically covers infancy to early school age, which includes major developmental changes, so that the meaning of HRQOL may transform along with the age. To truly understand HRQOL of these
patients, HRQOL has to be quantified for different health states and ages. In addition, HRQOL should take into account that pediatric IF-patients are frequently premature neonates, for which HRQOL may consist of different attributes than term born neonates. As with any other pediatric condition, pediatric IF-patients are part of a family and society, which may furthermore influence HRQOL and may have to be accounted for when measuring HRQOL in this population. Finally, HRQOL measurement for IF needs to rely on proxy measurements in most instances, as the majority of affected patients are too young to indicate their perception of HRQOL.

Despite the challenges and limited methodology, efforts have been made to measure HRQOL in pediatric patients with IF using generic measurement tools. A large, multi-center and cross-sectional study reported HRQOL measures for children on home PN using age-specific, generic French HRQOL instruments, while multiple studies qualified HRQOL in children after ITx and LTx. Over all age groups, pediatric patients on home PN are reported to have a high, parent-rated HRQOL, which is similar to healthy reference populations. The same study stated that home PN characteristics such as duration, PN start in early childhood, weekly number of infusions and hospitalization frequency did not affect HRQOL, while the presence of a stoma reduced HRQOL significantly. One reason for this may be that general HRQOL measurement tools were used, which might not be sufficiently sensitive to capture HRQOL-affecting factors specifically related to IF. One thesis publication described compromised HRQOL of children with IF on home PN or after ITx when compared to the general public, but found no significant difference between home PN and after ITx management. Intestinal failure patients aged 10-18 years who received ITx judged their HRQOL similar to healthy controls when using the Child Health Questionnaire (CHQ). There seems to be a difference between parent-reported HRQOL of pediatric IF-patients after ITx and LTx. Four out of 12 domains of the CHQ and 3 out of 5 domains of the PedsQL were rated lower for ITx patients when compared to those who underwent LTx. Several studies assessed HRQOL of pediatric patients after LTx and concluded that HRQOL improved with LTx, HRQOL after LTx is lower than in healthy children, and greater HRQOL when compared to other chronic diseases.

Since the children’s HRQOL is tightly connected to the family’s function, a more recent study also incorporated family function to their assessment of HRQOL after pediatric LTx
using the Child Health Questionnaire Parent Form (CHQ PF-50)\textsuperscript{395}. Mothers’ of children on home PN tend have lower HRQOL when compared to the father, while siblings of affected patients have similar HRQOL as the patient\textsuperscript{399}. Several studies observed that HRQOL of pediatric IF-patients reported by parents are lower than when quantified by the patients themselves, and health care professionals are described to judge the patient’s HRQOL even lower\textsuperscript{35,394,399}.

1.7.4 Mortality

Overall mortality of pediatric children with IF is high and has been reported to range between 6.4\% and 40\% depending on patient and anatomical characteristics, treating healthcare center and IF-management approach\textsuperscript{44,49,98,199,297,298}. Clinically, a decrease of mortality rates has been perceived since the early 2000s, while published survival rates are still variable and being reported as 70-90\%\textsuperscript{46,49,165,168,175}. Studies quantifying the change of mortality over time and identifying treatment options affecting mortality are missing at this point in time.

Short bowel syndrome has an incidence of 24.5 neonates per 100,000 live births with a case fatality rate of 37.5\%, translating into 1.4\% of deaths among children <4 years of age, of which most deaths are related to hepatic disease and sepsis\textsuperscript{3,44,294,305,315,349}. Approximately 2/3 of patients with IF will develop IFALD and 25\% traditionally progress into liver failure\textsuperscript{46,49,165,168,175,333}. Children with IF require central catheters for long-term PN, putting them at a particular risk for CRBSI, as they are receiving PN over the central line, are immunocompromised due to their suboptimal nourishment state, and bacterial translocation and overgrowth due to intestinal pathology. Sepsis has been described to be responsible of 20\%-33\% of deaths of neonates with IF\textsuperscript{49}. In addition to the direct mortality from sepsis, frequent CRBSIs promote IFALD development and subsequent mortality from liver failure\textsuperscript{401,402}. As IFALD and sepsis cause the majority of fatalities in pediatric IF and are a direct consequence of the disease itself, mortality from liver disease and sepsis has been summarized as disease-specific mortality. Attempts to reduce IF mortality and impact on quality of life in pediatric patients are therefore directed towards preventing or treating IFALD, inhibiting progression into fulminant liver failure and reduction of septic episodes\textsuperscript{333}. 
Intestinal failure patients are frequently premature neonates, who suffer from diverse co-morbidities causing additional morbidity and mortality\textsuperscript{2,49}. While new treatment options affecting disease-specific mortality from IF, i.e. from liver failure and sepsis, a significant change in overall mortality may not have been detectable in previous studies. The cause of mortality may have shifted from disease-specific complications to such related to prematurity and other congenital malformations. Studies analyzing this clinically perceived development are warranted.

1.7.4.1 Disease-Specific Mortality

Mortality from liver disease has been reported to account for 60% and 33% of all mortality of pediatric IF, while sepsis is causing 20% and 33% of deaths in children with SBS, with and without an IRP in place, respectively\textsuperscript{49}. Disease-specific mortality has been observed to continuously decrease over the past decade and IRPs are believed to play a major role in this development\textsuperscript{44,49}. It remains unclear what aspects of the IRP led to the change, as evaluation of individual treatment options that became available along with the IRP has not been performed\textsuperscript{44,46,49,165,403}. Novel lipid strategies, transplantations, STEP and ethanol locks are hypothesized to most directly impact disease-specific mortality by decreasing the risk for progressive liver disease and CRBSI\textsuperscript{44,50,58,101,190,191,291,336,351,360,381-384,401}.

Novel lipid strategies, such as outlined above, are hypothesized to halt and reverse liver disease progression by changing the inflammatory cytokine pattern\textsuperscript{58,187-189,191}. High-quality studies, namely randomized controlled trials (RCTs), are extremely rare, making generalizable conclusions on the effect of novel lipid strategies impossible\textsuperscript{333}. In addition, these strategies may vary, ranging from lipid minimization to different types of lipid emulsions (e.g. SMOFLipid®, Omegaven®), which creates heterogeneity and makes summarizing studies more challenging\textsuperscript{333}.

Transplantation represents the ultimate solution for IF complicated by liver failure, but it is limited by graft availability and long-term complications. Survival after transplantation has improved significantly between the first experimental transplants to the now standardized and refined protocols\textsuperscript{291}. The indication for transplants seems to be crucial however, as survival benefit is only shown for IF-patients with complications of long-term PN, and not when
transplant is performed for HRQOL or high-risk disease such as severe motility disorders and structural enterocyte defects\textsuperscript{291,404}.

First studies describing long-term follow-up after STEP have shown that full intestinal adaptation may take up to 5 years after the procedure to complete\textsuperscript{101}. How and to which extent this effect translates into decreased disease-specific mortality is unclear at this time point. Hypothetically it is expected to affect disease-specific mortality by supporting patients with non-progressing intestinal adaptation to overcome this arrest and thereby affect late disease-specific IF-mortality.

Ethanol locks have only recently been introduced, but seem to have a promising effect on IF-patients with frequent CRBSI\textsuperscript{35,360,390-392}. While it has not been possible to translate the effect of ethanol locks into disease-specific mortality reduction in children with IF, central catheters represent the lifeline of these patients. While CRBSIs cause disease-specific mortality directly at a rate of 12\%-35\% per infection, it further impacts this outcome by promoting IFALD development and using up limited vascular access, ultimately leading to transplantation\textsuperscript{35,354}.

The independent effect of an IRP and its maturation on disease-specific mortality is unknown\textsuperscript{44,46,49,165,403}. While clinically there is an obvious benefit of IRPs, no study has quantified this using disease-specific mortality. Overall mortality seems not a sufficiently specific outcome to capture this effect in a small comparative study\textsuperscript{49}. On the other hand, two comparative studies have shown survival improvement, disease-specific and overall, with IRP introduction\textsuperscript{44,46,165}. Further research is required to fortify this evidence and characterize the effect of IRP-maturation over time.

The limited evidence of effective IF-management is primarily due to the small sample size associated with a rare disease such as pediatric IF. Research implementing specific methodology and elaborate statistical analysis is required to determine the independent effect of treatment options and modifiable patient characteristics on disease-specific mortality.

In addition, pediatric patients with IF have become more complex and severely ill in more recent years, possibly contributing to the limited benefit in mortality measured in outcome studies. With increasing treatment options and improved clinical course, the threshold for
withdrawal of care may have decreased, making more premature and co-morbid patients with shorter residual bowel length part of the IF-cohort.

1.7.4.2 Other Mortality

As pediatric patients with IF are frequently prematurely born, the majority of non-IF-related mortality is secondary to complications from prematurity. A decade ago, approximately 13% of deaths in pediatric IF-cohorts were caused by non-IF related conditions, such as cardiorespiratory and neurological. Recently this proportion has been perceived as having increased significantly, but this observation has not been quantified in clinical studies thus far.

Recent data shows that neonatal mortality at the NICU is caused by severe prematurity and low birth weight in 14% of the cases. The second and third most common cause of death in this population is IF-related with 12% being sepsis and 11% acquired bowel disease. With these causes being addressed and minimized with optimization of IF-management, other causes of death become more dominant. Such other causes include lung hypoplasia in 9.5%, intraventricular or intracranial hemorrhage in 9.4%, respiratory distress syndrome in 8%, lethal anomaly in 7.6%, hypoxic–ischemic encephalopathy in 6.1%, genetic syndromes in 5%, major heart defects in 3.4%, bronchopulmonary dysplasia in 2.8%, hemorrhagic shock or profound anemia in 2.2%, pulmonary hemorrhage in 2%, renal failure in 2%, congenital diaphragmatic hernia in 1.7%, air leak syndrome in 0.6% and pulmonary hypertension in 0.6%.

Similarly to IF-specific mortality, other mortality has been affected by recent improvements in medical care allowing more complex and severely ill neonates and infants to be born and survive initially. When comparing mortality using historic control groups and unadjusted analysis, caution must be exercised when generalizing conclusions, as the comparison groups may be inherently different in terms of severity of illness and patient characteristics.

1.7.5 Cost-Effectiveness & Economic Burden

No comprehensive cost-effectiveness or cost-utility analysis has been performed to assess interventions for children with IF. This may be explained by the fact that it is a rare disease
making appropriate methodology challenging. In addition, the need for cost evaluation of IF-management may only have developed in recent years with an increasing number of treatment options becoming available and improved clinical outcomes leading to chronic clinical courses and more complex patients incurring increasing costs. Even though, the need for economic evaluation of IF-management has become prominent, a lack of methodology and suboptimal clinical evidence make valid analyses nearly impossible. For example, a cost-effectiveness analysis without generalizable and valid data on effectiveness of treatment options to use in the statistical model will produce limited results. In addition, a cost-utility analysis of IF-management is not possible at this time point, as methodology for HRQOL and utility measurement in infants and neonates, which compose the majority of IF-cases has yet to be developed.

The necessity to have an understanding of costs related to IF-management fostered two cost-of-illness studies of pediatric IF-management in America and the Netherlands. The economic burden of pediatric IF in North America during the first year of care has been estimated to be 505,250 US$ for 2005. Most significant factors contributing to this value were ICU admission, surgical interventions and readmissions. After the first year, annual costs were mainly due to home care ranging between US$250,000 and US$300,000. A slight increase of annual costs over time after the first year of care reflects the increasing rate of long-term PN complications requiring hospital admissions. The mean 5-year costs for a child with IF in the United States is estimated at 1.6 million (+/- 1 million), while the total costs in the Netherlands for 2006 was reported as €269,700 (US$355,195).

Although cost-of-illness studies allow estimation of the financial burden of care, they do not relate the costs to clinical outcomes, which is crucial in a disease with recent improvement of clinical outcomes such as pediatric IF. Results therefore fail to justify resource allocation and guide policy-making as they ignore the benefit of the increased spending or investment completely. In a time of exponentially increasing health care costs and financial crisis, an urgency to develop valid methodology resulting in generalizable cost-effectiveness analyses for IF and other neonatal and infant diseases is clear.
Chapter 2

2 Methodological Challenges of Pediatric Intestinal Failure Research

In a time of evidence-based medicine, medical decisions and health resource distribution should be guided by results from methodologically sound research and based on the highest possible evidence. While this assures high quality and cost-effective medical care, this standard is hard to achieve for many medical conditions due to methodological challenges encountered, which are related directly to the nature of the disease. Pediatric IF represents one of these conditions. It is a rare, surgical disease, prevalent in a vulnerable population, employs heterogeneous definitions and involves poorly developed measurement tools. It has a high morbidity and mortality, and evolves rapidly over time. It is important to appreciate the methodological challenges of such a condition to understand the reasons for the limited quality of evidence available and guide efforts towards improving these methodological obstacles. The objective of this chapter is to highlight the methodological challenges of pediatric IF-research and propose solutions based on literature, case studies and reasoning. Issues arising when selecting the study design, choosing and measuring outcomes, analyzing data, recruiting study participants, obtaining consent, generalizing results, attempting economic evaluation, publishing the study and assessing study quality will be discussed in the context of pediatric IF.

2.1 Study Design

Highly controlled and artificial study designs such as the randomized controlled clinical trial are the closest ones to experimental science in epidemiology. They are the gold standard for identifying disease etiology and causality between exposures and outcomes, which are sought in the majority of clinical studies. Even with this study design, rarely one study alone will provide sufficient evidence to support generalizable conclusions, as one high-quality study alone may only provide results true to the context of that particular population sample and ensure internal validity. Multiple studies reporting similar results under different circumstances are required to deem findings generalizable. In pediatric IF-research,
randomized controlled clinical trials are frequently not feasible due to numerous methodological difficulties as outlined below, leaving observation of reality and modeling of processes the main option of Epidemiology to discover relationships among variables. All study designs come with uncertainty about how well the study results represent the truth. To account for uncertainty, Epidemiology uses levels of precision, risk factors, confounding factors, a variety of study designs and pooling of evidence to base conclusions upon. In rare diseases research, such as in IF, uncertainty is more prevailing, making rigorous methodology, high-quality data and bias control crucial for valid results.

2.1.1 Randomized Controlled Clinical Trials

In a clinical trial, three basic requirements must be met: first, the trial must be able to answer an important clinical question; second, it must be based on rigorous methodology to produce valid results; and third, it must not harm participants. In IF-research, all three criteria must be fulfilled, but the second one will pose the largest challenge. A small number of available participants increases the risk of invalid results due to non-generalizable results or biased conclusions. A small study is also more likely to fail to demonstrate an effect when there truly is one present (i.e. insufficient power), and more prone to variability leading to estimates with low precision. On the other hand, a small study may have the advantages of representing a homogenous group with similar baseline characteristics, and of having close relationships with investigators leading to a higher willingness to participate and level of compliance.

Clinical trials must follow several basic concepts to produce valid results. The study has to examine an unbiased sample of the patient population of interest to produce results that are externally valid and generalizable. Inclusion and exclusion criteria have to be clearly stated. The study sample should then be compared to a concurrent internal control group representing a placebo treatment, reference treatment, no intervention or a dose-comparison treatment. If everything else is unchanged except for the exposure, it is appropriate to conclude that the intervention resulted in the difference detected between both groups. Intestinal failure research frequently derives study cohorts from registries. Methodological advantages of such a selection include having consecutive and homogenous participants, and maximizing available participants. Disadvantages may be related to a highly
specific population interfering with external validity and generalizability. In many instances of IF-research, an internal control group is not available, so that non-comparative study designs or external control groups may have to be utilized. External control groups, including historical/retrospective controls, are necessary as insufficient participants are available making powerful clinical trials too lengthy or it would be unethical to use internal controls because of presence of questionable equipoise. Although internal validity may be threatened when using external controls, for many aspects in pediatric IF-research they are the only comparative groups available, thereby risking that there may be other variables besides the treatment exposure that are different between the comparative groups. One way to address this is to account for confounders in multivariable analyses. The downside to this approach is that adjustments can only be made for known and measurable confounders, and a large sample size is required for this statistical bias control method. Another way to address selection bias when using external controls is matching patients according to individual co-variables or propensity scores. The limitation here is similar, as this approach can only adjust for known confounders and requires a relatively large pool of available patients to achieve specific matching.

Despite methodological attempts to deal with selection bias, it has not been possible to determine which treatment options independently contributed to the change in morbidity and mortality of pediatric IF over the past 15 years. Similarly, efficacy of individual IF-treatment options has not been demonstrated for all therapies believed to have contributed to the improved clinical outcomes. How these treatment options were introduced into pediatric IF-care helps one to understand the deficiency of randomized controlled clinical trials for many of these therapies.

The leading reason for developing multi-disciplinary IRPs was that there was a clinical perception that pediatric IF patients were experiencing excessive morbidity and mortality. These complex patients with multisystem comorbidities needed coordinated, comprehensive care with a team approach. Once IRPs were in place, it was unethical to conduct trials, exposing part of the patients to the historic regimen that was believed to be insufficient, to actually prove the effects of the IRP. In retrospect, one could argue for a trial prior to generalized IRP-implementation exposing part of the patients to the new regimen and prospectively measure the IRP effect. But in a rare disease as pediatric IF, the time required
conducting such a trial with sufficient power would have been too lengthy given the clinical urgency for a change in clinical management of these patients. Also, it would have been nearly impossible to expose a selection of patients blindly to two different clinical managements in one institution for administrative and care delivery reasons. While testing the role of IRPs in RCTs is impracticable, patients managed in such a program may be easier to recruit for RCTs of other specific treatments as IF-patients are now organized and followed closely.

For other novel treatment strategies in pediatric IF, such as parenteral omega-3 lipid emulsions, STEP and ethanol locks, trials could have been conducted and were attempted. The role of omega-3 lipids was evaluated for a multitude of medical conditions, but it was only in the late 1990s, that they were assessed in conjunction with IFALD. Several animal models reported promising results so that in 2006 first results of enteral and parenteral omega-3 supplementation in children suffering from IFALD were reported. Attempts were made to prove efficacy of omega-3 lipids in comparative studies, but methodological difficulties mainly related to the rarity of the disease and retrospective design resulted in suboptimal studies with a high risk of bias. As an example, Puder et al. in a retrospective cohort study, compared the effect of fish oil high in omega-3 lipids (1g/kg/day) to standard PN-lipids (Intralipid®, 1-4g/kg/day) in infants with a predicted PN-duration of >30 days and 2 consecutive direct bilirubin values of >2mg/dL using a historic control group (34.2µmol/L) Efficacy was tested by comparing the natural logarithm of the conjugated bilirubin trend differences between the two groups using generalized estimating equations models and adjusting for the baseline bilirubin value. This study illustrates multiple methodological difficulties and the associated risk of introducing bias: 1) comparison to a historic control group that may be significantly different to the exposure group (e.g. historical patient populations exposed to different general medical management such as an IRP), 2) introduction of important confounders related to differential lipid dose exposure (the control group received conventional lipid dosing at 1-4g/kg/day), enteral nutrition delivery, and septic complications, 3) nonspecific inclusion and exclusion criteria (e.g. introduction of subjectivity by using estimates for PN-requirement), 4) overly conservative IFALD-endpoint (conjugated bilirubin >34.2µmol/L vs. >50 or >100 µmol/L) making this a relatively healthy population, 5) no adjustment for selection bias in the design or analysis, and difficulty
interpreting the data due to multiple transformations used. Efficacy of omega-3 lipid supplementation was then shown in a multi-center randomized controlled clinical trial by Diamond et al. and reported at a conference in 2013\textsuperscript{335}. Just as importantly, they were able to show that a high-quality randomized controlled clinical trial in a pediatric IF-setting was feasible with substantial collaborative effort between centers and rigorous application of Clinical Epidemiology knowledge. The multi-center study design enhances generalizability, but still more studies are necessary to generate the highest level of evidence for this treatment and regulatory agencies to approve its use.

Other autologous bowel lengthening procedures have existed since 1980 and employed globally prior to the development of the STEP procedure in 2003\textsuperscript{36,39,41,280,284,285}. Once a successful STEP was performed in a human, it was a logical step to utilize this procedure in patients who plateaued in their ability to wean from parenteral nutrition. The STEP procedure became attractive, as it is technically easier to perform than other bowel lengthening procedures. Shortly after its first description, STEP was implemented at our institution\textsuperscript{40,41,101,156}. The theoretical benefits of offering this surgical option to pediatric IF-patients was described shortly thereafter, and the short and long-term outcomes reported, but no randomized controlled clinical trial testing efficacy of STEP was ever conducted\textsuperscript{101,156,253,286,410}. Reasons for this include challenges from the small number of available candidates, surgical nature of the intervention, questionable equipoise, the surgeon’s preferences and expertise, the surgical learning curve and blinding difficulties. In addition, the lower-risk exposure together with the clinical urgency, availability of descriptive studies, liberal regulations for introduction of new surgical interventions and devices, absence of financial gain for an industry sponsor, and widespread implementation by surgeons may contribute to the perception that a randomized controlled clinical trial may not be necessary, or if deemed necessary, that resources should first be allocated to more urgent questions.

The method by which ethanol locks were introduced to pediatric patients with IF is another example that nicely illustrates how new therapies are introduced to a new population without high-quality research. Central venous catheters were developed in the late 1970s and quickly adopted in many specialties for their numerous benefits as they provide a reliable access for administration of chemotherapy, blood products, PN, intravenous fluids and antibiotics, and
allow painless collection of blood samples\textsuperscript{385,419,420}. As these devices come with a risk of catheter related bloodstream infections (CRBSI), methods were tested to prevent such septic episodes originating from the CVC. Catheter lock techniques using local and systemic antibiotics, alone or in combination with thrombolytic and anticoagulant agents were described\textsuperscript{378,421-427}. Although ethanol locks have been used over several years as well, a retrospective comparison of ethanol locks vs. systemic antibiotics was performed only in 2003\textsuperscript{385}. The study concluded that ethanol locks were a safe and effective way to treat CRBSI even in small children\textsuperscript{385}. Although the authors also recommended a randomized controlled clinical trial to support their findings, no such study has been conducted since then. Instead four retrospective pre-post studies were published and summarized in a meta-analysis\textsuperscript{360,381-384}. The meta-analysis applied rigorous methods, identified two flaws in the included studies using the Newcastle-Ottawa risk of bias assessment tool and deemed the failings (i.e. not demonstrating the absence of the outcome of interest at the beginning of the study and no description of excluded patients) of minor importance\textsuperscript{381}. They acknowledged the risk of bias from using retrospective, non-randomized studies, and the increase of internal validity from using pre-post study designs. This evidence was sufficient for the intervention to be included in new guidelines and institutional drug formularies. Research of ethanol locks in the prevention of CRBSI in pediatric patients with IF jumped from pre-clinical studies to small, retrospective phase II studies without ever having been tested in a formal randomized controlled clinical trial. From a strict methodological perspective, it may appear irresponsible to expose more patients to this intervention with the lack of high-level evidence for efficacy and safety of this treatment, while the high risk of morbidity and mortality of the medical condition itself may outweigh, making it the lesser of two evils. Despite the methodological difficulties and clinical urgency, efforts should and are being made to conduct a formal randomized controlled clinical trial to show efficacy and safety of ethanol locks in preventing CRBSI in pediatric IF-patients.

Although it is important to understand which specific novel treatments have most impacted IF-outcomes over the past 15 years, until now, no study has been able to determine what changes in clinical management ultimately resulted in the decreased morbidity and mortality of this disease. While randomized controlled clinical trials are able to evaluate efficacy, more pragmatic approaches are necessary to determine effectiveness of treatment and when the
nature of the disease makes it impossible to follow the gold standard of a randomized controlled clinical trial, alternative approaches may become justified and invaluable to generate scientific evidence.

2.1.2 Alternative Study Designs

Although randomized controlled clinical trials deliver the highest level of scientific evidence, it is important to be aware that results from well-conducted cohort or case-control studies may generate similar estimates of treatment effects\(^{428}\). As the risk of bias decreases as the effect size increases, randomized controlled clinical trials may not be necessary for highly effective, life-saving treatments\(^{429}\). A cutoff of a 5- to 10-fold improvement has been suggested to determine the need for randomized controlled clinical trials\(^{429}\). The role of randomized controlled clinical trials in surgical research has been under review. It has been proposed that for skill-dependent interventions the traditional pyramid hierarchical model of evidence quality is not generalizable to surgical research and should be abandoned in favor of a circular model, that accounts for the conflict between internal validity (i.e. methodological rigor to avoid bias) and generalizability (i.e. relevance for a wider surgical community)\(^{430-432}\). Solomon et al. showed that only 40% of surgical research questions are amendable to randomized controlled trials, even under ideal clinical circumstances\(^{433,434}\). The authors concluded that the most important factors precluding randomized controlled clinical trials in surgical research were the rarity of the condition, lack of community equipoise, methodological issues, and patient and surgical preferences\(^{433}\). While alternative, ‘quasi-experimental’ designs are more easily applicable to surgical research questions, it is crucial that they follow similar quality criteria as the ones of randomized controlled clinical trials, including the use of standard definitions, and the detailed description of inclusion and exclusion criteria, valid methods and correct analysis. Specific quality criteria for non-randomized clinical studies were developed, as for example the Newcastle-Ottawa Scale, and should be used to guide clinical research using observational study designs\(^{435}\).

2.1.2.1 Recruitment & Randomization

A few valuable methodological approaches may be considered for research involving a rare and surgically treated disease, ranging from methods modifying single aspects of randomized controlled clinical trials to those employing alternative study designs. When maximization of
the available sample size is desired, adaptive randomization may be considered in the context of a randomized controlled clinical trial. It adjusts the allocation probabilities according to baseline covariates or efficacy response, thereby increasing internal validity. The covariate adaptive approach permits balancing important patient characteristics and prognostic factors, while the response-adaptive design allows more patients to benefit from the superior intervention. This is of particular value to surgical studies where equipoise may be controversial. Another beneficial design in rare diseases research is the group sequential adaptive trial design, where the required sample size decreases by recruiting participants in stages and using interim analyses. Specific data from each interim analysis, such as effect size and random variation, is used to calculate required sample size for the next stage. Flexible study designs have created some controversy with followers of the sequential analysis, as they may endanger validity, integrity and at times efficacy. On the other hand, they enhance feasibility and flexibility. In a research area like pediatric IF and in cases where methodological challenges may not be overcome with standard methodological approaches, a study using an alternative design carefully and clearly acknowledging limitations may be more valuable than no study at all. One good example from the general pediatric surgery research area is the recent multi-center Ravitch versus Nuss procedure for pectus excavatum multi-center randomized controlled clinical trial where there was no institutional equipoise (personal communication, currently unpublished). Instead of randomizing participants to surgeons, they were randomized to institutions with either an expertise in the Ravitch or the Nuss procedure, to avoid learning curve effects for individual surgeons and questionable equipoise at individual institutions.

Closest to a randomized controlled clinical trial is the prospective parallel-group non-randomized trial. In this study design a cohort of patients that undergoes a novel procedure is followed prospectively and compared to another group that receives the standard treatment. This study design should fulfill all requirements of a randomized controlled clinical trial apart from the randomization and blinding. It is important to evenly distribute the baseline characteristics between the two groups to avoid selection bias, by for example using propensity score matching (see also Chapter 2.4.1). Further bias adjustment may be made in the statistical analysis of the data by using multivariate analysis using covariates or propensity scores (see also Chapter 2.4.1). The benefit of such a trial is that it simplifies the
process of randomization and blinding, which may be problematic in a surgical setting. The difficulty of this study design in the setting of pediatric IF-research is that it is extremely difficult to find internal controls to be followed prospectively due to the rarity of the condition making such a trial too lengthy, particularly when a larger sample is required to allow for valid baseline characteristic adjustment using multivariate analysis.

2.1.2.2 Outcome Measurement & Analysis

Another way to address small sample sizes is to design a study with repeated measurements, that allows for between-subject and within-subject comparisons, increasing the data available and allowing optimal internal validity. Advanced statistical analyses accounting for intra-subject correlations of repeated measurements are available to maximize the utility of such a study design (e.g. generalized estimating equations, hierarchical linear or nonlinear models). One important, but unassuming type of repeated measurement design is the clinical trial based on repeated outcome measurements under different exposures of 1 participant, i.e. N-of-1 study design. Although these statistical methods offer valuable solutions to issues frequently found in pediatric IF-research, it is difficult to identify any published studies in this field using such methodology. One exception is the recently presented N-of-1 study of our group on the effect of enteral omega-3 lipids on controlling small bowel hemorrhage from ulcers along the staple lines after a STEP procedure requiring repeated blood transfusions. Interestingly, every time the patient stopped taking oral omega-3 lipids the bleeding re-occurred leading to low hemoglobin values requiring blood transfusions. Time series analysis was able to statistically quantify this association.

Interrupted time series is an excellent method to monitor treatment response over time before and after the introduction of a novel treatment. This study design is suitable for changing indications for the treatment over time and learning curve effects, both of which may be addressed by adjusting for confounders in a multivariate model. In IF-research, this study design revealed itself to be valuable as limitations encountered when comparing three treatment groups, as described in Chapter 3, were addressed by analyzing an interrupted time series using multivariate autoregressive integrative moving average models (ARIMA) models. This permitted answering the question of what kind of new treatment option contributed significantly and independently to the improved morbidity and mortality of
pediatric IF over the past 15 years (Chapter 4)\textsuperscript{,442,443}. Changing the scope slightly and assessing data longitudinally instead of categorically produced important scientific evidence, which has not been attainable previously due to methodological difficulties inherent to pediatric IF-research.

2.1.2.3 Observational Study Designs

The value of well-designed observational studies such as cohort, case-control or cross-sectional studies should not be underestimated as they have been contributing to scientific evidence significantly, especially in the surgical field and for rare diseases. Cohort studies compare groups of patients exposed to different treatments, either retrospectively or prospectively. They are able to gather data on the sequence of events, examine multiple outcomes for a given exposure and provide information on factors contributing to the outcome of interest. They are particularly useful for rare exposures and allow calculation of outcome rates in exposed and unexposed individuals over time (i.e. incidence, relative risk calculations). On the other hand, they require a large number of subjects to test rare exposures and are susceptible to selection bias. Prospective cohort studies may be expensive to conduct and require long durations of follow-up, making them susceptible to loss to follow-up and withdrawals. Retrospective cohort studies are more prone to recall and information bias, and permit less control over assessed variables. Understanding these limitations allows the researchers to employ methods to address some of these shortcomings in advance. For example, maintaining periodic contact with the participants may minimize loss to follow-up, or multivariate analysis allows for adjustment of known confounders. Variants of cohort studies such as nested case-control, nested case-cohort, multiple cohort studies and external controls have their strengths and weaknesses, and may be considered as well when designing a non-randomized clinical study. Despite the challenge the rarity of pediatric IF brings, cohort studies have been completed successfully in this setting\textsuperscript{,65}. A cohort of neonates was drawn from a tertiary care center ICU population and used to assess characteristics and outcomes of neonatal patients undergoing laparotomy and developing SBS, or not. Sampling patients from a population with a higher IF-incidence (i.e. neonatal patients undergoing a laparotomy) solved the methodological difficulty. While this approach was valuable for the study objectives of comparing neonates who develop SBS after a laparotomy and those who
do not, it sacrificed the ability of a cohort study to generate an incidence value of all the infants at risk.

Case-control studies are ideal to find risk factors of a certain outcome of interest. Two samples of patients with the outcome are identified, and the occurrence of the risk factor is assessed retrospectively for both groups. The advantage of case-control studies is that they are not expensive, are time efficient and only require a few subjects to provide valuable information. The downside is that they are not able to provide incidence, prevalence, attributable risk or risk difference values. They also only allow assessment of one outcome at a time, while cohort, cross-sectional and clinical trials permit testing of multiple outcomes. With their high efficiency, case-control studies bring a higher risk of bias introduction, as they are sampling study subjects from two different samples and measure predictor valuables retrospectively. Awareness of this risk and skillful application of appropriate measures to control for sampling and measurement bias, e.g. using matching of cases and controls for most important baseline characteristics, or using data that was recoded before outcome occurred, are crucial in this study design. This study design has been useful in pediatric IF-research, as it does not require a large sample size and the study is completed quickly, thereby helping dealing with the rarity, high mortality and fast-changing clinical picture of the disease. Case-crossover studies, also called pre-post studies, are a variant of the case-control design, where subjects with the outcome of interest are compared to their former selves. This study design holds high internal validity as it reduces random error and confounding by using cases as their own controls. On the other hand, it is only applicable under certain circumstances, making it less generally applicable. In the setting of pediatric IF, pre-post studies were used to show effectiveness of ethanol locks in CRBSI-prevention.

The cross-sectional study design is yet another approach accessible to pediatric IF-research. In a cross-sectional study, associations between factors are tested in a sample at one time point. This study design does not allow for incidence measurement, as is has no longitudinal time dimension. It does, however, permit estimation of prevalence. Cross-sectional studies are attractive because they are relatively quick to conduct, inexpensive, and they do not require any follow-up. On the other hand, they have difficulty establishing causal relationships and are impractical for rare diseases.
For a rare disease such as pediatric IF, case series may actually be suited better to describe factors of patients with a certain condition or outcome of interest. For example, a case series describing gestational age of patients with IF would likely reveal that the majority of them were born prematurely suggesting an association between prematurity and IF. Case series and case reports are found at the base of the evidence pyramid and often criticized for their limited contribution to scientific evidence. In surgery and rare diseases research, these study designs are frequently dominant and have introduced significant clinical advancement as they have provided novel findings and generated valuable hypotheses. Many new surgical techniques are often first described using case reports and introduced to a wider audience in case series. One example related to IF-research is the one of the STEP procedure that was first described in a case report in 2003 and then applied and described more widely in subsequent case series.

2.1.2.4 Meta-Analysis

Once two or more studies exist, the data may be pooled in a meta-analysis providing significant results where individual studies were insufficiently powered to answer the research question or the results were conflicting. Heterogeneity between included studies, such as different study population, minor differences in therapeutic regimen or the year of study conduct, is a fundamental issue in meta-analyses. There are methodological methods to address study variability, such as random-effects models or meta-regression models. To increase the power of a meta-analysis, analysis of pooled patient-level data instead of study-level data may be considered, if this more specific data is available. Although meta-analyses ideally consist of randomized controlled clinical trials, they may include well-designed non-randomized studies. A recent systematic comparison found that the quality of meta-analyses of well-designed non-randomized trials was as high as those based on randomized controlled clinical trials. Although data from case series and case reports may be included in meta-analyses by contributing patient-level data as well, one should be aware that the quality of a meta-analysis is highly dependent on the quality of the data that is included in a meta-analysis (i.e. “garbage in, garbage out” phenomenon). In IF-research several non-RCT meta-analyses have been conducted providing important scientific evidence. For example, Oliveira et al. showed that application of Ethanol locks in children with a CVC for PN-administration reduced the CRBSI-rate significantly. This study illustrates well how
pooled results from small pre-post studies led to high-quality scientific evidence. On different occasions, meta-analyses provided clinically important results in pediatric IF-research, although the pooled data did not primarily consist of classical randomized clinical trial data, but retrospective comparisons or pre-post studies. In general, it has been a challenge to pool IF-studies successfully due to conflicting disease and outcome definitions, varying demographics, countless management practices, inconsistent gestational age stratification and lack of (high-quality) control groups.

There are many different study designs available to the pediatric IF-investigator. As described above, the RCT may not be feasible in many cases and alternative methodologies are necessary to create new knowledge. Training and knowledge of clinical research methodology are required to skillfully overcome challenges originating from the rarity and surgical nature of this disease. At times, it may be more appropriate to leave an important research question unanswered for the time-being, while new methods are developed to avoid dissemination of erroneous results from poorly conducted studies. For instance, this author recently made this decision regarding an eagerly awaited cost-utility analysis of intestinal rehabilitation programs. With the advent of IRPs and increasing survival of IF-patients, morbidity of survivors with complex anatomy is unclear and associated costs of care are rising. It would therefore be of outmost importance to clinicians and policy makers to assess HRQOL of these patients as well as cost-utility of IRPs. As no valid HRQOL measurement tools for the population of interest are available at this time point, a conceptual framework for HRQOL in infants and neonates was developed to serve as guide for measurement tool development (Chapter 5).

2.2 Outcome Selection & Measurement

In research, meaningful outcomes must be chosen and be explicitly defined and measured in a standardized manner to answer questions with validity, accuracy and precision. Outcome selection before the study has been initiated is fundamental to avoid bias and ascertain study validity. The relevance and appropriateness of primary and secondary outcomes in pediatric IF-research will vary depending on the research question, but should also be assessed in the context of the scientific literature with the aim to add to available knowledge in the most meaningful manner.
Pediatric IF-research has struggled with inconsistent outcome selection, heterogeneous variable definitions, selective reporting and lack of valid measurement tools. Understanding the importance of this methodological aspect of clinical research will assist one in appreciating the difficulties encountered in pediatric IF-research when attempting to make general and clinically relevant conclusions.

2.2.1 Outcome Selection

The selection of outcomes is a crucial stage in study design. As the risk of false-positive conclusions increases with the number of outcomes measured per study, it is important to keep the number of measured outcomes as parsimonius as possible. It is necessary to pre-specify one primary outcome that is most specific to the research question. Relevant secondary outcomes are valuable as they may provide interesting inferences, but unless the study was primarily powered to answer these secondary outcomes, they should remain hypothesis generating.

For outcomes to provide valid results, they must measure what they are intended to measure, in an unbiased way and be relevant to the clinical disease. They should be precise, accurate, feasible, reliable, responsive and sensitive to change, measureable under limited time and money, and easily interpretable. Decisions appropriate for the study design must be made a priori about how, when and by whom the outcomes shall be measured and by what technique. It is common for these methodological decisions to be omitted from published articles, thus making it difficult to judge the appropriateness of selected outcomes. A study may, for example, not report on a certain relevant outcome because it was not available, it was of poor quality or it did not yield favorable results.

2.2.1.1 Heterogeneous Definitions

One methodological problem identified in pediatric IF-research is the non-standardized definition of outcome variables, with inconsistent outcome selection and reporting across different study groups. Interestingly, IF itself does not have a standardized definition. Some researchers define it based on anatomical criteria; while others consider it a functional diagnosis and others still, combine both anatomical and functional elements. In the early 1990s, Galea et al. chose to define SBS as a jejunoileal length of <25% of the mean expected
length by age plus 1 standard deviation, while Georgeson et al. recognized that bowel length alone does not account sufficiently for different disease etiologies and decided to define SBS as a complex syndrome of malabsorption associated with substantial shortening of the jejunoleum requiring total or supplemental PN for at least 3 months to achieve adequate growth and development\textsuperscript{445,446}. Unfortunately, heterogeneous disease definitions are still being used 20 years later. The Canadian Association of Pediatric Surgeons (CAPS) defines IF as PN-dependence for more than 42 days after bowel resection or functional loss, and/or less than 25\% of expected age-specific remaining small bowel length, but not all IF-managing centers in Canada use the same definition proposed by CAPS\textsuperscript{3,49,165}. Adding to the confusion, a large number of studies are investigating SBS patients only while others include all forms of IF in their study population such as mucosal enteropathies and dysmotility syndromes\textsuperscript{49,60,165,175,190}. At times, studies use the terms IF and SBS interchangeably, but the study cohort consists of SBS-patients only\textsuperscript{299,360}. Nonspecific IF-SBS terminology may be explained by the historical context, where IF was first described in SBS-patients, and by the fact that SBS is the most prevalent subtype of IF. At times it is appropriate to include the whole IF population for certain research questions, such as when assessing effectiveness of ethanol locks on CRBSI-prevention. Nevertheless, consistent and precise terminology utilization is favorable and should be applied whenever possible to optimize study validity.

Lack of standardized definitions for key outcome measures makes comparison of studies between institutions difficult. Cholestasis, residual bowel length, adaptation and mortality are a few important examples.

Cholestatic liver disease, the most common cause of death in pediatric IF, is a complication of chronic PN exposure. For years this condition was described as parenteral nutrition associated cholestasis (PNAC) or parenteral nutrition associated liver disease (PNALD). More recently, there has been an appreciation that the liver dysfunction experienced by these patients is multifactorial in nature and related to prematurity, lack of enteral feeding, recurrent sepsis and PN composition. The condition is now described as intestinal failure associated liver disease (IFALD). There is no standard definition for IFALD. By convention, most clinicians describe it in terms of serum conjugated bilirubin levels, but the cut off is not agreed upon. Some U.S. centers define IFALD at a serum conjugated bilirubin level of 34.2 µmol/L (17.1mg/dL). In Toronto, the definition has been a conjugated bilirubin level of
50 µmol/L since 2003. Other biochemistry has been proposed such as liver transaminase levels (aspartate aminotransferase (AST), alanine transaminase (ALT)), AST/platelet ratio, abnormal synthetic function (albumin or INR), but none have been validated in the IF population. The definition of advanced liver disease or liver failure is even more contentious. Composite clinical scores such as Child-Pugh or MELD/PELD scores have not been adopted. Liver biopsy showing histological fibrosis is not available in many patients and may not be ethical to obtain. Other clinical findings such as evidence of portal hypertension (esophageal varices, splenomegaly, thrombocytopenia) or need for transplant have been used as well.2,49,165,413,414

Residual bowel length is fundamental when describing patient with SBS, as it represents the remaining bowel that will undergo adaptation and provide the patient with nutrients to grow and survive. While residual bowel length generally refers to the remaining small bowel, it is important to specify whether it is in absolute terms or whether age-adjusted values are used. As the bowel grows dramatically between 30 weeks gestational age and 6 months postnatally, it is important to account for gestational age when calculating residual bowel length in neonatal patient with IF.69 Many studies cited Touloukian et al.’s article using 3 broad gestational groups ranging from 24 to 40 weeks and showing that the bowel of a deceased 28-week fetus is approximately 200 cm and increases to 300 cm at term.447 In 2009, Strujs et al. published normal values separately for small and large bowel length for premature infants from 24 gestational weeks to 5 years of age.250 They have shown that a term infant has approximately 200 cm of small bowel and 40-50 cm of large bowel. The outcome of a 28-week gestation infant with 30 cm of residual intestine may be significantly different from an older infant, as the bowel’s ability to adapt and grow is expected to be higher. For these reasons, availability of narrow normal values and age-adjusted reporting of residual bowel length is important. Studies reporting absolute bowel length without adjusting for gestational age are misleading. To account for gestational age, residual bowel may be reported as percentage of the original bowel length or as an adjusted value for expected bowel length.69

Development of enteral independence and mortality are both frequently considered primary outcomes to assess treatment effectiveness of IF-therapies.101,156,165,175,299 While most studies report on clinically observable PN-independence for a certain amount of time (e.g. 6 weeks, 2 months, 1 year), recently authors have suggested to use citrulline levels as a marker for
enterocyte mass and predictor of PN-weaning\textsuperscript{158-164,301}. Utilization of a biomarker for enteral autonomy may be justified by its more objective nature and standardized measurement, but on the other hand, clinical PN-independence over a certain period of time is more relevant to the patient as it represents the ultimate outcome of interest. Ensuring the patient maintained adequate growth after cessation of PN is paramount, but usually not reported in the form of anthropometrics.

Mortality was an important clinical outcome in pediatric IF-research. It is attractive as it is clinically meaningful, and easy to measure. Historically, these patients succumbed to liver failure or sepsis. As prematurity is common in any pediatric IF cohort, other causes of death such as respiratory, neurologic or cardiac are frequent. Describing both overall mortality, as well as, disease-specific mortality provides different information. Disease-specific mortality (liver failure or sepsis) helps to assess impact of therapies specific to IF management.

Clearly, heterogeneous terminology and selective outcome reporting interferes with validity, multi-centered study designs, pooled analysis and generalizability of pediatric IF-studies. It augments methodological challenges related to the rarity of the disease and sabotages any efforts to pool results and build a solid base of evidence. If the conducted pediatric IF-studies would use standardized terminology, expose the patients to similar treatments, and measure and report outcomes in a consistent way, available evidence could be pooled and stronger conclusions drawn, even from studies employing suboptimal designs such as observational studies. The heterogeneous terminology and outcomes are explained by its origins as a rare, surgical diagnosis, as it facilitates uncoordinated development of research groups. Regrettably, there has been little progress in developing consensus amongst pediatric IF-researchers\textsuperscript{138}. This may be related to inadequate infrastructure or personnel at some institutions to reliably measure certain outcomes or abstract data. Too commonly, investigators prefer to build on their previous research and apply the same definitions they have used in the past\textsuperscript{1}. Other factors contributing to failed terminology standardization in pediatric IF may be that on occasion, multiple consensus groups are conducted simultaneously and in an uncoordinated manner, resulting in multiple disease definitions at one given time. It is necessary to build awareness for this problem among pediatric IF-researchers, as the bias introduced with using non-standardized outcome measures and
Selective outcome reporting is significant. It disrupts academic efforts to create high-quality scientific evidence and ultimately denies patients optimal health care.

Perhaps pressure from publishers and international pediatric surgery and gastroenterology associations would help support a movement towards standardized terminology and outcome reporting. Disease definitions can be established in a standardized process resulting in an expert consensus statement that incorporates evidence-based knowledge and patient perspectives. It should be endorsed by respective professional associations, evaluated against other used terminology, validated externally, and be widely disseminated. A core outcome set for pediatric IF would be of great utility to help standardize outcome reporting. Core outcome sets are lists that catalogue important outcomes to be measured and reported in all pragmatic trials of a specific disease or condition. They are ideally based on consensus of key stakeholders such as patients, healthcare professionals and funding bodies. The Core Outcome Measures for Effectiveness Trials (COMET) initiative was launched in 2010 by researchers interested in the development and application of agreed standardized sets of outcomes for a variety of surgical trials, e.g. such involving esophageal, colorectal, and head and neck cancer, but no such sets have been developed for IF.

Another aspect of heterogeneity in pediatric IF is the enormous number of different exposures described by pediatric IF-research groups. A multi-center retrospective cohort study showed that pediatric patients with IF are exposed to a multitude of different medical treatments including different antibiotics, histamine-2 blockers, proton pump inhibitors, promotility and anti-motility agents, bile salt binding medications and anti-secretory agents, and variety of surgical procedures including small bowel resection, exploratory laparotomy, ostomy procedures, gastrostomy procedures, autologous bowel lengthening procedures including LILP and STEP, tapering procedures and other unspecified interventions. To further illustrate the extent of this heterogeneity of treatments, the same cohort of 272 patients received 20 different formulas to initiate enteral feeds, and overall, 40 different enteral formulas were used in these patients. This is problematic particularly in the setting of retrospective studies using historic or external controls, where many different exposures may be present in an unequal way and thereby introduce bias to the study. Also, results from one case series may not be applicable to other groups as their management may differ significantly. Frequently, the details of these various treatments are not described in
publications, so that generalization of conclusions becomes extremely difficult. It is therefore important that clinical protocols and management are made available to the academic audience, by encouraging publishing journals to request this information from authors and providing space in their articles for such data.

2.2.1.2 Important Outcomes in Pediatric Intestinal Failure

In IF-research, there are several clinically relevant outcomes that have been measured frequently to assess treatment effectiveness in studies. They typically include some version of mortality and successful intestinal adaptation. Mortality has been reported as overall mortality or survival, mortality from the disease (e.g. mortality from sepsis and liver disease), mortality among patients with liver failure, mortality among those patients with liver failure, and mortality causes. Many pediatric IF-studies primarily report overall mortality and then indicate mortality causes separately. Disease-specific mortality appears to be a valid primary outcome in pediatric IF. It measures what it is supposed to measure, what is likely the worst possible clinical outcome and captures death from the two main IF-complications: mortality from a septic episode and mortality from terminal liver disease. Disease-specific mortality is a more specific mortality outcome than overall mortality, as it measures an outcome that was intended to be affected by treatment exposures, and excludes mortality that is inaccessible to IF-treatments, such as mortality caused by cardiorespiratory or neurological conditions. Since IF-patients are frequently prematurely born and suffer from several comorbidities, measuring overall mortality when evaluating the effect of IF-therapies is not appropriate as it dilutes and confounds the measurable effect of the IF-treatment of interest. Disease-specific mortality is also easily interpretable and measurable without causing additional costs, making it a suitable primary outcome. In comparison to other outcomes measuring treatment efficiency, such as for example early IFALD, disease-specific mortality may not be as sensitive to change, as new treatments may affect intermediate outcomes such as cholestasis or septic episodes before influencing disease-specific mortality. The conservative nature of this outcome, however, allows solid conclusions on the impact of treatment options tested, if statistically significant changes in disease-specific mortality are found.
Similar heterogeneity is present amongst the measurement of successful intestinal adaptation. There is no consensus on the time period of PN independence before an individual is considered to have achieved enteral autonomy. Some studies employ the date PN was ceased, while others use various time periods such as 2 weeks, 2 months or a year \cite{44,49,199}. In addition, the variables were measured using different scales. Enteral autonomy for example was measured with continuous values of total PN-days, binary values for 50% or 100% enteral calorie intake achievement, time to 50% or 100% enteral calorie intake, and percentage of days receiving enteral breast milk or formula \cite{44,49,199}. Enteral autonomy, whether it is measured in achieving complete PN-weaning or time to 100% enteral feeds, represents the best possible outcome and completion of the intestinal adaptation process. It is affected directly by IF-therapies as it represents the physiological resolution of the disease, and therefore a specific primary outcome to measure efficiency in this setting. Enteral autonomy is also meaningful in a clinical setting and to patients, making it relevant and easily interpretable. It is critical to remember; however, that a patient must demonstrate maintenance of weight or growth in the case of children to truly be considered free of parenteral support.

Secondary outcomes that are meaningful when assessing efficacy of therapies in pediatric IF are related to morbidity from IF, such as cholestasis, CRBSI and need for transplantation. Other important secondary outcomes in the setting of pediatric IF include bone health, micronutrient deficiencies and neurocognitive development. Health-related quality of life is another essential IF-outcome, as it quantifies overall morbidity. More recently there has also been an emphasis on including patient-reported outcomes, that may reflect effectiveness and quality of surgical treatment more accurately \cite{452,453}. In pediatric IF-research this aspect is crucial, as the disease affects patients and their entire family massively, but sadly this has been almost completely ignored in outcome studies and trials. This may be related to absence of validated measurement tools for the neonatal and infant population of pediatric IF, particularly for HRQOL outcomes. Nevertheless, simpler patient-related outcomes such as pain levels, or family-reported overall wellbeing of the affected child, are measurable outcomes that should be assessed more frequently in IF-studies. Clinical research in neonates and infants is complicated by the fact that all subjective clinical outcomes are dependent on proxy reporting. For example, pain cannot be quantified directly using a traditional 1-10 scale
frequently used in adults. More elaborate scales have been developed to assess pain in neonates and infants, such as the Neonatal Infant Pain Scale (NIPS) or the Pain Assessment in Neonates (PAIN) scale, which include physiological and behavioral aspects, but these scales are not able to account for the neonate’s/infant’s direct perception of pain. Another example is HRQOL. Health-related quality of life can only be measured in neonates and infants by asking proxies about their impression on the neonates’ or infant’s wellbeing. As the meaning of neonatal and infant HRQOL (NIHRQOL) remains undefined in the literature, age-specific measurement tools have not been developed, and concepts applicable to adults, such as the adaptation process and the role of resilience, are undefined for this age-group measurement of these NIHRQOL as secondary outcome remains challenging, despite the clinical importance.

Outcomes related to resource utilization, such as for example costs, are further outcomes that may be assessed in clinical studies, but have only rarely been the focus of pediatric IF-research. Such outcomes are important, particularly when put into relation to clinical effectiveness and quality of life, as they allow efficient and fair resource distribution and maximization of limited resources. Such outcomes may be evaluated in simple cost analyses, cost-benefit analyses, cost-utility analyses or cost-effectiveness analyses, depending on the study objective.

2.2.1.3 Alternative Outcome Types

Composite outcomes may be indicated at times when ≥2 measurable outcomes represent one clinical entity. This may be sometimes useful when both measurable outcomes occur infrequently, as for example in recent pediatric IF-cohorts death from liver failure and transplantation, both representing end-stage IFALD. Also, surrogate, early marker of the outcome may be chosen when the outcome of interest takes too much time to develop. Biomarkers have taken a fundamental role in pediatric IF diagnosis-making and disease monitoring, as well as, become valuable hard outcomes in pediatric IF-research as an indicator for biological processes, disease progression and pharmacological response to therapeutic intervention. Examples of biomarkers for IFALD include direct bilirubin, bile acids, GGT and INR. While direct bilirubin represents a particularly cheap, easily attainable, widely recognized and most practicable marker, elevated bile acids are the earliest indicator
of IFALD, but highly variable in the population. Although widely used, not all laboratories may be able to provide direct bilirubin values or may preferentially report total or indirect bilirubin values. The total bilirubin consists of an indirect, non-conjugated, water-insoluble component (indirect bilirubin); a direct, conjugated, water-soluble component (direct bilirubin); and another water-soluble component that is covalently bound to albumin (delta bilirubin). The direct bilirubin is further divided into two fractions, bilirubin monogluconide and bilirubin diglucuronide. In conventional methods, delta bilirubin is measured as part of conjugated bilirubin, but these details have to be clarified prior to pooling data from multiple centers. Although the concept of surrogate outcomes and biomarkers are important in IF-research, overly optimistic conclusions from results using surrogate markers may result in harm to patients and should therefore be used prudently. The issue with surrogate outcomes is that a change may not necessarily mean a difference of the hard outcome, because the surrogate biomarker may only be associated with the hard outcome statistically and not biologically, or because there is an alternative mechanism, where the treatment impacts the hard outcome that is independent from the surrogate outcome. The role of biomarkers and surrogate outcomes may therefore be mainly in early Phase I trials. For example, omega-3 lipids have been shown in studies to reverse IFALD by measuring direct bilirubin values as a surrogate for liver health. Whether this treatment actually affects the liver on a structural level is unclear and more studies are required to determine whether pathological findings of liver fibrosis are resolved with omega-3 lipid supplementation.

2.2.1.4 Changing Priorities

The management and measurement of outcomes of pediatric IF patients have undergone rapid transformation over the last 15 years. This has affected the relevance of outcomes and co-variables. For example, end-stage liver failure from chronic PN was a valid clinical outcome 5 years ago, but may be regarded as less important in current studies as the incidence of this complication has decreased significantly. Today early IFALD appears more meaningful as an outcome, as end-stage liver failure develops only rarely now and early IFALD plays a more significant role in the clinical decision-making process. This difficulty interferes with retrospective study designs, which are the predominant study designs in pediatric IF. Rapidly changing outcome definitions and relevance makes it difficult to find valid historic control groups as some variables of interest may not have been measured at all,
or may have been defined and measured differently among groups. In addition, measurable and immeasurable variables may have changed over time in this rapidly transforming clinical condition, thereby introducing possible confounders when using historic controls. In Chapter 4, a way to deal with rapidly changing, measurable confounders is described. It includes change over time assessment and adjustment of significant variables in a multivariate ARIMA model. Also prospective randomized clinical trials may be affected by the rapid clinical transformation in the setting of pediatric IF, as relevance of outcomes and co-variables may have changed over the time of enrollment, which may be long due to the rarity of the disease. By the time the trial is completed, the results may not be clinically meaningful anymore.

2.2.2 Outcome Measurement

Outcome selection is a crucial step in clinical research and may interfere significantly with scientific evidence formation and application of evidence-based medicine in a clinical setting. Even the best outcome selection will not produce valid results that are free of random error and bias, if they are not measured appropriately.

2.2.2.1 Precision

Variable measurement may be judged based on its ability to measure an outcome precisely, meaning in a reproducible, reliable and consistent way, and to be as free of random error as possible. The more precise measurements are performed in a study, the more powerful the study is to detect a significant difference between groups, as random error is kept low. Random error or chance variability may result from imprecision of the observer, the measuring instrument or variability of the subject itself. Precision may be measured by quantifying reproducibility of repeated measurements, either performed by the same observer or different observers, and if deemed necessary, there are methodological strategies to apply to reduce random error. These include standardization of measurement methods, training and certification of the observers, refinement of instruments, automating instruments and repeating measurements. When designing studies, investigators must decide on how rigorously these strategies should be pursued, accounting for the importance of the variable, feasibility and costs.
In pediatric IF-research, rarely does a study discuss precision of its measurements. The methodology section is typically very short not providing any information on how outcomes were measured and what strategies were employed to minimize random error. Using strategies to improve precision would be valuable in pediatric IF-research that struggles with methodological challenges from the rarity of the disease, and should be applied where feasible and affordable to increase study power. While standardization and observer training are easily applied, instrument refining, instrument automation and repeated measurement is frequently assumed without formal description. For example, data was fed prospectively into the clinical registry used for the two studies described in Chapter 3 and Chapter 4 by only one highly trained nurse practitioner, who is familiar with all the IF-patients at our institution. Clear definitions of outcome variables were used, such as disease-specific mortality that was defined as mortality caused either by a septic episode or liver disease. Most biochemical outcomes were measured and recorded in an automatic way according to standard operating protocols; such as for example weekly serum direct bilirubin levels or standard diagnosis of a CRBSI.

2.2.2.2 Accuracy

A variable may also be assessed by its accuracy, or in other words its ability to represent what it is intended to represent, which is fundamental for the variable’s validity. Accuracy is different and may not necessarily be linked to precision. Many times precision and accuracy go hand in hand however, as methods to improve one will often improve the other measurement feature as well. While accuracy is negatively correlated to systemic error, precision is correlated in the same way to random error. Three types of systematic error (bias) may affect accuracy: observer bias, instrument bias and subject bias. Accuracy is best assessed when compared to a gold standard measurement that is considered accurate, and expressed in mean difference for continuous scales, sensitivity and specificity for dichotomous scales and kappa for categorical scales. Closely related to accuracy is validity, which stands for the degree to what a variable measures what it is supposed to measure. Three different aspects of validity of a measurement may be assessed: content validity (measures all aspects of a phenomenon, including face validity), construct validity (measures what it is supposed to measure) and criterion validity (follows an anticipated behavior, including predictive validity). A valid measurement approach may be derived from the
literature, which builds on previous studies and makes the study more generalizable; however, if no validated instrument is available, a new measurement tool or protocol may be necessary. Strategies to increase accuracy are similar to such used to optimize precision. Standardizing measurement, observer training, instrument refinement and automation may be used for both measurement features. Accuracy may further be improved by making disguised measurements, calibrating the instrument and blinding the observer and subject. The investigator has to decide how vigorously these strategies should and can be pursued, based on anticipated impact of inaccuracy on study conclusions, feasibility and costs.

In the two studies described in Chapter 3 and Chapter 4, all measurements were performed by clinical indication, making tendencies of the subject such as treatment compliance less related to the research question. Occasionally, families were aware that a novel treatment was available and offered to them, such as omega-3 lipid emulsions or ethanol locks. This may affect accuracy of subjective outcomes, but is barely able to affect hard outcomes such as disease-specific mortality or serum direct bilirubin levels. Laboratory values do not allow for many adjustments, as the machine and protocols used by the lab determine measurement standardization, observer training, instrument refinement, automation and calibration, and thereby accuracy. Blinding was not feasible in the traditional sense for these two studies due to the interventional nature of the disease and exposure tested (see also Chapter 2.1.2). Families may not have been fully blinded, since they may have been aware of the new treatment options. This awareness is not highly significant in this context however, as the reported primary and secondary outcomes were not influenced directly by the families. One could still argue that awareness of novel treatment options increases hope in family and therefore may influence the decision to withdrawing care and thereby mortality, but whether such a mechanism is clinically meaningful is unclear. While the investigators were not blinded, the nurse practitioner entering the data into the registry was blinded in the sense that she was not aware what groups would be compared in future studies. The subjects (neonates and infants) themselves were completely blinded due to their cognitive developmental stage that did not allow them to understand new treatments or research questions. In this sense, they are rarely prone to subject bias, except when it is caused by non-blinded family or caregivers, such as a mother fearful of a new surgical procedure increasing the stress reaction
in the infant and possibly leading to immunocompromise and a higher risk of post-operative infections.

2.2.2.3 Other Features of Measurement

Other aspects of measurement should be appreciated as well when making decisions about variable measurement. The chosen measurement should be sensitive enough to detect clinically significant differences. It should be specific and represent only the characteristic of interest. An excellent example to illustrate this is disease-specific mortality versus overall mortality. Disease-specific mortality is more specific as it represents only mortality from the disease (IF) and does not include mortality from other, non-related causes. Measurements should also be appropriate for the study objectives. A study aiming to evaluate the effect of IRPs should define in advance what type of effect it is intended to measure, e.g. disease-specific mortality versus transplantation versus enteral autonomy. While transplants were a valuable and relevant measure of IRP efficiency 5 years ago, presently this outcome is less relevant because the frequency of transplantation has decreased secondary to improved intestinal rehabilitation. Measurements should also provide an adequate distribution of responses and not cluster around one extreme value. For example, a measurement instrument to measure stress levels in parents of infants with IF before and after IRP introduction requires a scale discriminating high levels of stress in a sensible way, as an overly simplistic scale from 0-5 may result in all parents scoring at the extremes of the scale. Finally, measurements should be objective without resulting in narrow-mindedness limiting the scope of the observations and ability to appreciate unexpected phenomena. Objective measurement of an outcome is often more straightforward with quantitative design, however, qualitative variables can often enrich the findings. For example, the study in Chapter 3 chose to evaluate effectiveness of IRP by assessing disease-specific mortality. More subjective outcome measurement of IRP effectiveness such as family satisfaction or stress levels would have added an interesting perspective to the study.

When deciding on how to measure outcomes of interest the value of efficiency and parsimony should be kept in mind. As much relevant data should be collected as efficiently as possible with the time and money available, while overly numerous outcome measurement should be avoided. A clinical IF-registry recording data that is clinically relevant without
causing additional tests and measurements illustrates this balance nicely, as it allows measurement and documentation of a large number of variables, without overwhelming the study subjects and research team, and causing any additional costs. The study design significantly influences the prospect of a study to optimize the measurement process, as not all strategies may be feasible. A retrospective study uses pre-selected, pre-defined and pre-measured variables, whereas prospective studies hold more freedom to optimize variable measurement. Suboptimal measurement; therefore, contributes to the lower scientific quality of retrospective studies by suboptimal power and validity.

2.3 Recruitment

In clinical research, an unbiased sample of subjects that ideally represents all subjects affected by the disease is selected and examined. Recruitment for clinical studies is an important and at times challenging step, as it has the potential to introduce bias, is influenced by the frequency of a disease and has to take into consideration legal and ethical aspects.

The majority of patients with IF are diagnosed early in life, typically as neonates or infants\textsuperscript{50,199,269,457}. Neonates and infants with IF exemplify a particularly challenging population as participants are rare and vulnerable, they cannot give consent, and the disease involves sensitive data. In addition, they have not been studied thoroughly before, leading to a lack of strong evidence supporting new research, and age-specific methodology has frequently not been developed (see also Chapter 2.2).

2.3.1 Sample Selection

In pediatric IF-studies, study samples most frequently consist of convenience samples. That is, participants who meet the inclusion criteria and are easily available to the investigator. Pediatric IF-patients are typically recruited from a single institution\textsuperscript{49,101,410}. This approach may help produce homogeneous samples and maximize recruitment of participants. Consecutive samples are especially valuable as they minimize selection bias. Appropriate sample selection aims to recruit a sample that adequately represents the target population and is large enough to answer the research question. In pediatric IF-research, where sample size is a challenge, collaboration with other institutions may become necessary. Although the argument may be made that involving multiple research teams, specialties and healthcare
centers may result in a less homogeneous sample and introduce bias, collaborative efforts between institutions have been shown to generate high-quality research and increase generalizability. Collaboration may also help with funding and manpower issues, and increase study efficiency, particularly when using a randomized controlled clinical study design. In pediatric IF-research, multicenter studies may be challenging, secondary to the absence of standardized outcome definitions and a multitude of clinical protocols used by the various intestinal rehabilitation programs (see also Chapter 2.2.1).

A study without a control group is not able to quantify the change of an outcome. Case reports and case series make a contribution, by describing observations in one or more patients. They leave the reader to judge the relationship between exposure and outcome. Case series tend to be more hypothesis generating. For obvious reasons, these study designs result in a weaker level of evidence than a study that includes a control group. Controls may represent a group of patients who receive placebo treatment, a reference treatment, no intervention or a dose comparison treatment. In pediatric IF-research, many studies do not include control groups and hence only provide descriptive results. The over abundance of case reports and series in the surgical literature has been criticized repeatedly and reasons contributing to this low level of evidence have been discussed (see also Chapter 2.1).

Selection of unbiased cases and controls is fundamental to ensure internal validity of a study. Ideally controls are identical to the exposed group except for the exposure of interest. Pediatric IF-studies that include a control group frequently use historic controls for convenience reasons. This may introduce selection bias, as historic controls may be systematically different from more contemporary patients. Particularly in a rapidly evolving condition such as pediatric IF where several medical advances for IF, as well as, other comorbidities (prematurity) may have changed over time. Historical patients; therefore, may not reflect current medical practice and introduce bias. For this reason, internal and contemporary controls are preferred.

Surgical research is particularly susceptible to selection bias as surgical indications may be subjective and governed by a surgeon’s judgment. Even when the surgical exposure is standardized, a surgeon decides on an individual basis whether an operation and thereby
enrollment in a study is indicated. This decision-process may be significantly different between individual surgeons. Distinction must be made between receiving an operation and requiring one. While indications for surgical intervention in pediatric IF are mostly clear, there may be significant variability in the type of procedure performed depending on knowledge and experience. These factors affect both internal and external validity.

Despite its importance, the recruitment process, including inclusion and exclusion criteria, is often not reported\textsuperscript{435,461-463}. For example, in the study by Opilla et al. examining ethanol lock therapy for prevention of CRBSI, the authors describe in detail how the first case was recruited, but then fail to state how the other 8 patients were enrolled. It is uncertain how “frequent” CRBSI was defined, whether the series consisted of consecutive cases, whether any cases where excluded from the analysis, and what inclusion and exclusion criteria were applied\textsuperscript{388}. If only limited information on recruitment is available, assessment of sampling bias, sample representativeness and appropriateness of the control group may be difficult or impossible. In the mentioned study, the control group was clear as the exposed patients served as their own internal controls thereby minimizing the risk of systematic differences between the comparative groups\textsuperscript{388}.

### 2.3.2 Sample Representativeness

When selecting patients for a study, it is important for a sample to be representative of the entire population to which the results will be applied. It follows that the more representative the study sample, the more relevant the results will be. The process of selecting a representative sample starts with the study design, where populations and sampling methods are chosen sensibly, and continues with delineation of clear inclusion and exclusion criteria. There is a balance between sample homogeneity to assist in conduct of a trial and generalizability to the total population.

In rare diseases research, it is paramount to maximize recruitment to optimize sample size. Non-participants, meaning the patients who are eligible but refuse to consent to a study may bias the study sample, as they tend to be systematically different from study participants. While in a randomized controlled clinical trial, participating subjects may be distributed randomly between the study groups, in observational studies non-participants are more likely to be dispersed unequally. In a randomized controlled clinical trial, early non-responders are
randomly distributed between groups and blind to the exposure. Late non-responders or dropouts should also be minimized in randomized controlled clinical trials by the tight protocol surveillance allowing for active intervention to avoid these situations. Due to the risk of selection bias the number of non-responders should be kept to a minimum. At times, a rate of 25% is regarded as a good achievement, although it may introduce bias and distort the results. Collection of data on non-responders assists in determining if nonparticipants represent a population who are systematically different than participants. The best method to address non-responders is to prevent them from happening.

A functional relationship and transparent communication between the patient, the treating team and the research group play a major role in successful patient enrollment, as contradictory information, power struggles, ineffective communication and failure to provide necessary support may lead to refusal and withdrawal of the patient from the study. Individual reasons why patients refuse to participate should be understood and all reasonable efforts made to minimize these hurdles. Employing various means of subject contact (mail, phone, email) is helpful. Subjects may be more likely to consent to a study if the study design minimizes invasive and uncomfortable procedures, information is clear to alleviate anxiety, patients feel they have access to therapy that is otherwise not available, and offers incentives such as reimbursement of transportation costs. If language barriers are present, using translators or translated questionnaires may be helpful as well. On the other hand, coercion, undue inducement, intimidation and exploitation must be avoided. Financial compensation for time, travels or inconvenience, or as incentive may be help recruit study participants, but requires rigorous control by institutional research ethics boards, especially in vulnerable populations such as poor, homeless, sick and pediatric patients. In rare diseases, financial compensation may be lucrative to patients, as they may be financially indigent due to high medical costs that may not be covered by health insurance plans.

Since IF patients suffer significant morbidity and many patients are managed at highly specialized centers where they develop close relationships with their care providers, the issue of non-participation is less of an obstacle. In many cases, families are quite keen to try novel therapies to improve their child’s outcome. It is therefore critical to ensure that protocols are well designed and ethical so these vulnerable patients are not exploited. As the research team is frequently also the treating team, patients may feel obliged to participate in studies or
worry about receiving suboptimal treatment if they refuse to participate. Patients with IF are also at risk of coercive influence from unrealistic hope as they are suffering from a chronic, life-threatening disease making them and their families feel as if there is little choice. National and international guidelines have been developed to guide investigators and research ethics boards in clinical research outlining special considerations for vulnerable populations, and suggesting solutions to balance the right and need of all populations to participate in research with proving sufficient protection from harm. Research should be designed to encourage participation of all groups to avoid development of knowledge that helps only a subset of the population while protecting everyone’s rights and welfare, by carefully considering study designs and anticipating complexity in seemingly benign situations. Theories to optimize the patient-clinician relationship have been developed and should be referred to when planning and conducting studies on rare diseases. Kavanaugh et al. proposed the use of Swanson’s middle-range theory of caring to optimize the caring relationships between health care professionals and vulnerable patients participating in sensitive research, of which pediatric IF is a representative. Five processes (maintaining belief, knowing, being with, doing for and enabling) with numerous sub-processes are presented to promote caring relationships and thereby improve subject recruitment and retention.

### 2.3.3 Rare Participants

While the definition of “Rare Diseases” differs between countries, there is an underlying understanding that a rare disease has a low prevalence, and treatments and research are perceived as suboptimal. Rare diseases, also known as orphan diseases, have been defined as diseases that affect less than 1 in 650-10,500 people. Approximately 10% of diseases are classified as rare and their cumulative prevalence is between 6%-8% making them a significant public health issue. Advances in rare disease research may not only be beneficial to the disease itself, but also reveal knowledge translatable to more common diseases. While the majority of developed countries have established a national Orphan Drug program, Canada remains one of the few that does not have a governmental plan to protect patients with rare diseases from exorbitant drug costs.
Clinical research of rare diseases poses obvious obstacles related to sample size and power. Small samples affect study design, statistical analysis and generalizability directly. Prospective studies on a disease with low incidence and/or prevalence may require many years to complete due to limited enrolment, over which time period medical care may change significantly and thereby introduce bias. Retrospective cohort studies as an alternative are associated with lower costs, shorter study period, and improved feasibility, but they are associated with well-known biases such as selection bias and information bias (see also Chapter 2.1). Statistical adjustments for bias, such as multivariable analysis for confounding factors, or multiple imputation for missing data, may be an option. Small studies with insufficient study power and statistically insignificant results, may lead to less frequent publication due to publication bias. Challenges of conducting research on a rare disease may also change over time. With decreased mortality of pediatric IF, the effect size may become smaller, meaning that more patients are required to detect a difference, and recruitment may become longer as events happen less frequently.

Despite these facts, studies with small sample size or retrospective design are frequently the only evidence available for clinicians to base medical decisions upon. Inevitably, medical care will be associated with a higher level of uncertainty, increased risks, poorer outcomes and slower medical advances. Justification of clinical and academic resource allocation to rare diseases may therefore be difficult.

Awareness of rare diseases among health care professionals and the public may be helpful in increasing the number of available participants. Awareness may be increased using conventional broadcast news outlets (e.g. articles on health topics of the Canadian Broadcasting Corporation (CBC)), as well as social media (email newsletters, chat room, Facebook posts (e.g. New England Journal of Medicine case postings) or Twitter). International online registries for rare conditions permit data collection without geographical limitations, thereby enhancing sample size and generalizability. Examples of such electronic disease registries are the Pleuropulmonary Blastoma Registry and the Toxic Shock in Pediatric Burns Registry. Particular attention to patient confidentiality is required in these circumstances. There have been national initiatives to support the process of making the diagnosis of rare diseases with telephone and online tools such as the Orphanet from the French national strategy for rare diseases. Study-related information may be
distributed using a study-specific website and patient recruitment may be increased by allowing direct recruitment through the website\textsuperscript{448,485}. The website should be appropriately indexed searchable with common Internet search engines (e.g. www.google.com). While this comes at little extra cost, the website may be disseminated on paid Internet advertisements\textsuperscript{486,487}. Limitations of web-based recruitment tools include confidentiality, lack of evidence for effectiveness, misdiagnosis, spam and overcrowding, travel-related to necessary in-person evaluations, selections bias of more wealthy, younger and higher educated participants, and limited knowledge on web-related ethics and consents\textsuperscript{448}. Patient recruitment may be further enhanced through rare diseases organizations or patient organizations\textsuperscript{448}. For example, in the United States the National Organization of Rare Disorders (www.rarediseases.org) offers posting of ongoing studies on their website for no cost. Orphanet (www.orpha.net) and EURODIS (www.eurodis.org) represent two other rare diseases organizations facilitating communication between researchers, patients and potential study participants\textsuperscript{448}. Another option to coordinate research is clinical research networks that include multiple clinical sites supported by one coordinating research center. Clinical research networks allow patient data pooling, establishment of comprehensive databases and recruitment standards, improvement of clinical trial feasibility, standardization and validation of disease definitions and outcomes, and interdisciplinary communication and mentorship\textsuperscript{448,488,489}. In rare diseases, where resources and patients are scarce, research collaboration and multi-institutional studies are of particular importance to minimize research repetition and competition, and maximize generalizability and efficiency\textsuperscript{448}. The academic funding system that traditionally has favored basic science research is only recently focusing on more tangible clinical research deliverables and promoting team grants.

Specialty clinics and clinical patient registries are a valuable source for patient recruitment, especially in rare diseases and without being dependent on research funding. Clinical and research patient registries are defined as an organized program for data collection, storage, retrieval and dissemination of clearly defined data points for a specific purpose (public health, health services research, health promotion, patient care, clinical research and public safety)\textsuperscript{408}. Patient registries may also be regarded as systematic data collection programs, which may be divided into three categories: disease/condition/syndrome registry, exposure registry or patient characteristics registry. Registries are not a new concept, as they have been
present for hundreds of years, such as for leprosy and tuberculosis.\textsuperscript{490-492} They are typically used as a first step to estimate prevalence and incidence, and create identity for a new (rare) disease. Registries assist in understanding the natural history of a disease, determine relevant patient outcomes and patient characteristics, and generate hypotheses. Specific data quality criteria should be fulfilled: data should be complete, valid and entered in a timely manner.\textsuperscript{493,494} Benefits of a rare disease registry are that it helps maximize sample size and identify subjects for recruitment to clinical trials, increases available data and decreases biased information, may be used for observational studies, may produce long-term outcomes, and may provide appropriate controls.

There are certain errors and biases associated with registry data that may threaten data accuracy. Random errors are unpredictable and may lead to inconsistency in repeated measures. Systematic errors may distort results from the true values, and may result in measurement (e.g. outcome definition), selection (e.g. diagnosis definition and enrolment into the registry) and information bias (e.g. interviewer bias, recall bias). It is important to understand whether these biases are differential or not, and if necessary, to use appropriate methods to control for confounding. Several registries related to pediatric IF exist.\textsuperscript{49,495,496} A challenging aspect of IF-research using registries is related to outcome measurement. As stated previously, there is a lack of consensus definitions for key outcome variables. In addition, as outcome has changed over the last 10-15 years, the outcomes deemed important have also changed. Certain variables may not have been collected in more historic patients and therefore cannot be assessed presently.

Lead-time bias may be encountered, as in more recent years, early survival of patients with severe IF has improved; thereby creating a population that is surviving in a sicker state with more complex surgical anatomy. When comparing these recent cases to historic controls, the results may be distorted, as the controls may reflect a healthier sub-population as sick patients were more likely to die early. To account for this bias, multivariable analysis adjusting for severity of illness can be performed (see also Chapter 4) as can patient matching or propensity scoring. Another issue related to using registry data is variation in follow up for different groups, leading to bias for assessment of long-term outcomes. Time event analysis or study design that ensures an adequate minimum follow-up are possible solutions.
2.3.4 Vulnerable Participants

Research participants may be considered vulnerable due to the group they belong to, the nature of the situation they are in, or the research itself. Pediatric IF-patients are considered a vulnerable population for several reasons, including their extremely young age, high risk for morbidity and mortality, and the clinical urgency.

Like comatose, mentally ill and fetal patients, neonates and infants represent a vulnerable population, as they are not capable of giving consent to receive treatment or participate in research. Substitute decision makers must be identified to obtain consent. Medical decisions must therefore be based on the parent’s or legal guardian’s values. Due to the severity of the illness and clinical urgency, obtaining consent from parents may be complicated and at times impossible, because of the emotional involvement of these substitute decision makers.

Neonates and infants with IF are also vulnerable because they are at a high risk of severe morbidity and mortality. Caregivers may feel pressured to participate in interventional studies due a perception of desperation. They may regard experimental treatments as the only hope for their child. Parents in these situations may also appreciate research participation as a way of telling their story and giving meaning to their burden. The clinical urgency of pediatric IF may put investigators and parents under pressure to develop new knowledge and find answers to many unanswered questions. This again may put neonates and infants at risk for being involved in risky research and exposed to poorly understood treatment options.

In addition, pediatric IF represents a sensitive condition, as IF is a life-threatening condition affecting deeply personal aspects of life. It is associated with potential child death, frequent comorbidities, congenital malformations and premature birth and sacred topics of motherhood that may be associated with parental conflict and feelings of guilt. These topics are intensely personal. Research in such conditions is delicate, as it requires close guidance to avoid patient and family harm and exploitation. With the psychological burden that comes along with pediatric IF, parents may feel overwhelmed and unable to consent to research studies. This phenomenon was observed in the third project of this PhD-thesis, where a family declined to participate in a focus group to generate a conceptual framework of neonatal and infant HRQOL as they found the topic too emotionally distressing (Chapter 5).
Guidelines have become necessary to protect vulnerable populations from harm\textsuperscript{497,504,505}. Three basic ethical principles applicable to research involving vulnerable populations include respect for persons, beneficence, and justice\textsuperscript{468,506}. More specific guidelines to protect neonates and infants in the setting of research were proposed by the U.S. National Institute of Health (NIH) Research. They state that preclinical studies must have been performed that allow estimation of the risk to the neonate/infants, informed consent must be obtained from the mother and if possible from the father, or the legal guardian, all participants must understand the informed consent process and research study, and the researchers cannot have a vital role in determining viability of the neonate/infant\textsuperscript{497,507,508}.

Obtaining consent from research subjects is one way to ascertain that a discussion of risk and harms related to the research between the investigator and participant has occurred. Patients who are unable to provide permission to give consent require special regulations. The permission to conduct research on a child must be obtained from a parent or legal guardian, and when developmentally appropriate the assent of a child must be obtained additionally. Research in children is typically limited to such involving minimal risk only. Research with more than minimal risk to the child may only be conducted if there is direct benefit to the child. A research ethics board may also decide to allow research projects where the risk increase is only minor and vital information about the child’s disorder or condition is obtained\textsuperscript{456}. The consent process must be voluntary, informed and ongoing\textsuperscript{509}. In very select situations, research ethics boards may deem research without obtaining informed consent from parents or legal guardians appropriate to help facilitate important research\textsuperscript{509}. Generally, if the research does not involve living human subjects, or if the research is based on secondary analyses of existing data exemption from consent may be requested\textsuperscript{456}. More specifically, the Canadian Panel on Research Ethics describes acceptable alterations to consent requirements in certain situations as follows: if the research involves no more than minimal risk to the participants; if the alteration to consent requirements is unlikely to adversely affect the welfare of participants; if it is impossible or impractical to carry out the research and to address the research question properly, given the research design, if the prior consent of participants is required; if, in the case of a proposed alteration, the precise nature and extent of any proposed alteration is defined; and if the plan to provide briefing (if any) which may also offer participants the possibility of refusing and/or withdrawing data and/or
human biological materials, shall be in accordance with debriefing in the context of alteration to consent requirements guidelines (Article 3.7)\textsuperscript{509}. Exceptional rigor has to be applied when assessing ethical aspects of research involving a vulnerable population such as pediatric IF-patients.

The complexity and importance of psychological dynamics including vulnerability of study subjects and sensitivity of data in the context of pediatric IF-research is frequently underestimated when studies are designed and rarely addressed in subsequent publications. Vulnerability of a population of interest poses a challenge to clinical research, as it may lead to compromise in methodological rigor, complicate patient recruitment, extend the research timeline and delay new knowledge development. Recognition of these issues may help address them at an early stage, resulting in smoother and more efficient clinical research.

2.4 Analysis

Statistical analysis and data interpretation is a crucial aspect of clinical research, as it scientifically synthesizes data and quantifies uncertainty\textsuperscript{510}. Regardless the quality of a research question or study design, credible conclusions can only be drawn if valid statistical methods are applied. This is true for any type of analysis, from the simplest descriptive approach to complex modeling. Before initiation of a clinical study, several decisions regarding the statistical analysis should be made: 1) selection of variables for ‘Table 1’ summarizing baseline characteristics, 2) specification of primary, secondary and subgroup analyses, 3) determination of analytic methods to determine whether the effect is real, 4) selection of methodological approach for missing data, and 5) establishment of clinical importance\textsuperscript{455}. Applicability and importance of these aspects may vary depending on the study design. In pediatric IF-research where many studies are observational, analytical options may be restricted. On the other hand, skillful application of statistical methods may address some of the methodological challenges related to the rarity of the disease.

2.4.1 Baseline Comparisons

The first table of a study should include all baseline characteristics that are likely to influence risk or responsiveness to a treatment. These factors are also important to understand the nature of the sample and will guide generalization of the results. Appropriate descriptive
Statistical methods should be used based on sample size and data distribution. This step of the statistical analysis is applicable to the majority of study designs. In case series, description of participants fulfills a major part of the study objectives, and if the sample size is large enough, this may be supplemented by testing for associations between patient characteristics and outcomes. Later is an important way of hypothesis generation in pediatric IF-research. For example, one case series of 6 patients suggested that exposure to fish oil reversed IFALD, which led to multiple comparative studies and a significant change in the clinical picture of pediatric IF.\(^\text{414}\)

In comparative studies, significance levels of baseline characteristics between groups may be reported, depending on the advantages and disadvantages of providing such information. The risk of reporting significance levels of baseline characteristics in small studies typically used in pediatric IF-research is that differences between groups may not be significant although they are clinically important, leading to false security in terms of balanced comparison groups and selection bias. In large trials, baseline differences may be significant without being clinically meaningful. Adjustment for significant baseline characteristics in a multivariate analysis is frequently sought, but rarely completed because of the small samples available in pediatric IF-research. Ways around this issue may be the application of propensity scores or inverse probability weights, or utilization of compound variables representing comorbidities, other risk factors or severity of illness in multivariate models. The later option is limited by the fact that many rare diseases do not have validated risk or severity of illness scores. The statistical reassurance that is sought with baseline assessment and adjusted multivariate analyses should always be sensible to the fact that tests and adjustments can only be made for known and measured confounders.

### 2.4.2 A Priori Decisions

In a comparative study, primary, secondary and subgroup analyses should be pre-determined and aim to answer the research question. It should be in keeping with the study objectives, outcome selection and measurement, and available study sample. Certain aspects of the analysis must be discussed and well-thought decisions made in terms of using one versus two-sided hypothesis testing, testing superiority versus non-inferiority, anticipated data distributions and approaches of model specification to produce valid results. Many of the
methods used when analyzing randomized clinical controlled trials with large samples may not be applicable in pediatric IF-research due to the rarity of the disease (see also Chapter 2.1 and Chapter 2.3.3). Whether a study’s statistical approach was determined before a study was initiated is rarely described in published studies, although this information is crucial to assess validity of a study. For instance, a research team could use different statistical approaches, but selectively report the one producing significant results and not the ones derived from a methodologically sounder analysis. Or it could perform as many analyses until one produces significant results, thereby increasing the risk of false–positive results (type I error). Such unethical and bias-introducing behavior is problematic, but published articles rarely provide sufficient information to judge whether the statistical approach was pre-determined. By addressing this analytical aspect in a publication, the study results become more credible and methodological quality of pediatric IF-research may increase by creating awareness among investigators and readers.

A sample size calculation should be done prior to study initiation, but it is rarely performed or reported in the pediatric IF-literature. This information is primarily important to guide recruitment and form conclusions, but may also help choosing the ideal statistical analysis and primary outcome of a study. Although sample size calculation is more relevant when designing an interventional trial, as the freedom exists to increase the sample size, change the primary outcome or comparison group prior to study initiation, it may still provide valuable information for the design of observational studies. It is also important because it helps maximize available resources by distributing them to feasible studies that have a higher probability of finding a meaningful result. In pediatric IF-research, where recruitment time of clinical trials may be impermissibly long due to the rarity of the disease, sample size calculation is crucial to avoid directing time and funding towards a study that will not be able to answer the research question and detect a true difference between groups.

2.4.3 Statistical Limitations of Small Trials

In pediatric IF-research, small sample sizes, use of external control groups and the frequent use of retrospective study design limit analytical methods. While it is feasible to determine the analytical approach before data are actually being analyzed, data availability may limit the analytical options. Every statistical comparison brings a risk of yielding false positive and
false negative results. An upper limit for the false positive rate is typically set at 5%, indicating that if a study is to be repeated 100 times, 5 of the studies may be false positive. The study should also be able to detect a difference, if there truly is one, that is, having the power to identify a difference between two groups. This ability is directly related to the sample size of the study and precision of outcome measurement. In IF-research, sample size is frequently a limiting factor leading to studies with insufficient power to detect a significant effect. One way to maximize the power of observational studies at the study design level is to carry out case-control studies.

Small samples may also cause analytical challenges as they limit the number of applicable statistical approaches. For example, a multivariable analysis may not be feasible for a sample of 30 patients due to over-specification of the model. While distribution issues from small samples used to be another statistical challenge, modern statistical software packages (e.g. STATA, SAS) include non-parametric options for a majority of statistical tests, making it easier to apply appropriate statistical methods to small samples. While the study samples may have improved slightly in pediatric IF-research with increased survival and more frequent multi-center studies, events have become less frequent with improved clinical outcomes, thereby shifting the cause of statistical challenges without significantly improving or changing them. For example, multivariable analysis may not be feasible for a rare binary outcome even with the sample size increased to 50, or a study may not be sufficiently powerful to detect a smaller effect size even with a few more patients in the sample.

Although small studies cannot provide as much scientific evidence as a large trial, it is the only feasible study design in many situations. It is important that the investigators acknowledge the lower level of evidence provided by small studies, openly discuss the study limitations, select appropriate statistical methods and conclude that the results do not provide confirmatory, but rather exploratory evidence. Strategic methodological decisions may be made to maximize power, such as choosing a matched study design, using continuous instead of discrete variables, or using strategies to maximize precision of outcome measurement is another approach to decrease variability, and thereby optimize conditions of small studies facilitating their ability to detect a true difference between groups (see also Chapter 2.2.2.1).
At times, changing the perspective and analyzing data longitudinally instead of categorically may be helpful when addressing previously encountered methodological challenges related to small samples (see also Chapter 2.1.2.2). Generalized estimating equation models and hierarchical linear or nonlinear models represent advanced statistical approaches that account for intra-subject correlations of repeated measurements and maximize the utility of such study designs\textsuperscript{408,438-440}. To account for time as a factor and analyze temporal trends, time-to-event analysis using a Cox-Proportional Hazards model with time as a co-variable may be considered as well. Interrupted time series analysis is a valuable statistical method to measure change over time, as it also allows adjustment for seasonality and time lag of exposure and effect. This method allowed quantification of individual effectiveness of four pediatric IF treatment options using a multivariate ARIMA model, as described in Chapter 4.

### 2.4.4 Alternative Statistical Approaches

Alternative statistical approaches should be considered for pediatric IF-studies, as they might provide valuable methods for small samples. Traditionally, statistical tests are done to test hypotheses. In this four-step process, statistics are used to determine the likelihood that a given hypothesis is true: first, a null hypothesis (H0) and an alternative hypothesis (H1) are defined, then a statistical test is used to test whether the null hypothesis is true. From the test statistic, the $p$-value is calculated indicating the probability that a test statistic is more extreme as the one observed if the null hypothesis is true. Finally, the calculated $p$-value is compared to a previously set significance level, alpha ($\alpha$), and the null hypothesis is rejected in favor of the alternative hypothesis, if the $p$-value is lesser than alpha. Alternative statistical designs that part from this standard paradigm include statistical prediction models and Bayesian statistics.

In statistical prediction models, a distribution of control measurements is created, that represents the expected normal range for a certain measurement. If experimental measurements are contained in the established prediction limits, it is concluded that the experimental intervention does not have a significant impact on the observed condition. If they exceed predicted values, then the experiment intervention has demonstrated an effect\textsuperscript{408}. This methodology is especially applicable if the number of potential endpoints is large and the number of available subjects is small\textsuperscript{408}. Prediction designs can be used parametrically or
non-parametrically. Involving multiple responses from each subject, by taking measurements at the same time or sequentially, may improve their performance. Statistical prediction models have been used by the National Aeronautics and Space Administration (NASA) to assess effectiveness of a program to prevent bone density loss in astronauts, that obviously represent a very small population. Pediatric IF impacts a small number of people as well, and outcomes of interest are relatively frequent, making this statistical approach a relevant alternative to classical hypothesis testing in this population.

Bayesian methods are particularly useful in small clinical trials, as they model knowledge and uncertainty more directly using probabilities. Bayesian statistics may be regarded as a more natural statistical approach representing more authentically how people think about probabilities in real life, taking into consideration prior experiences when assessing situations. In contrast to frequentist methods, Bayesian methods include prior knowledge and probabilities into the statistical analysis, instead of completely ignoring any previous experience, by determining prior probability distributions, i.e. priors. The prior probability distribution is combined with the probability distribution of the new data to yield the posterior distribution, which in turn is used for future inferences. While this allows maximization of knowledge, as it includes previous knowledge into the current analysis, and direct answering of questions on probabilities, it also introduces a certain level of subjectivity into the statistical analysis. In IF-research, Bayesian methods have been used to evaluate expert beliefs on the probability of IFALD with different lipid strategies. Guidelines for the use of Bayesian statistics have been published and should guide application.

Ranking and selection designs are other alternative statistical approaches that may be applied in small trials. These designs are helpful to decide which treatment of multiple available treatments is most effective. They regard larger responses to treatment as preferable and rank treatments according to the expected mean responses. As the type I error rate is kept at 50% in this approach, sample size requirements are less extensive. While this proves to be an advantage in small trials, investigators should be aware that this approach produces considerable lower level of scientific evidence.

While all described methods represent valuable methodological alternatives for IR-research, every possible effort should be made to achieve the highest methodological quality applying
standard methods, as this will facilitate combining results from multiple studies and creating strong scientific evidence. Communicating the pros and cons of alternative statistical methods in published articles is important, as it will facilitate quality assessment by readers and reviewers, many of which may not be familiar with these alternative approaches.

2.4.5 Missing Values

Missing values are frequently encountered, but often neglected in clinical research. It is tempting to treat missing values as if they were lost and ignore the particular observation. In rare diseases research where every observation counts, such an approach is unfortunate as it further decreases the sample size and study power. It also endangers validity and credibility of the study by possible introduction of attrition bias. The severity of this problem may be gauged by reporting the percentage of missing values in respect to known values. There is no consensus regarding the magnitude of missing data that is problematic. Suggested cutoffs range from 5% up to 20%, while other authors argue that the amount itself is not as relevant as the nature of the missing data and whether the study is still able to answer the research question.

In pediatric IF-research, missing values are addressed in publications inconsistently, leaving it unclear at times whether there were no missing values present in the dataset, or whether observations with missing values were excluded from the analysis. Randomized controlled clinical trials evaluating IF-patients appear to report on missing data more consistently. Interestingly, not all quality guidelines for reporting of clinical studies include aspects of missing data equally unambiguously. The Newcastle-Ottawa scale has been used repeatedly in pediatric IF-research to assess quality of non-randomized studies, but only asks for description of excluded patients. This leaves it unclear whether the described patients/observations were excluded due to missing values or other reasons and does not require any discussion on how missing values were dealt with. On the other hand, the Cochrane risk of bias tool that has been used in a meta-analysis of randomized controlled clinical trials on the efficiency of human growth hormone and glutamine for patients with SBS, explicitly asks for adequate handling of incomplete data.
While missing values can be minimized with good study design and data abstraction, complete absence of missing values may not always be possible. This is particularly true for retrospective studies where certain data values may not have been measured or documented. In pediatric IF-research, where trials may take a long time to complete due to the rarity of the disease, patients may be lost to follow-up creating incomplete observations. There are multiple statistical methods to deal with missing values. Missing values may occur completely at random (i.e. no systematic reason), at random (i.e. some information on the reasons for missingness is known and examinable using other collected variables, but it is not related to the outcome), or not at random (i.e. some information on the reasons for missingness is known and examinable using other collected variables, and it is related to the outcome). If data is not missing at random, such that there is a pattern to the missing values related to the outcome of interest, then missing data introduces attrition bias into the study.

While it is up to the investigator to judge the nature of the missing values, the reasoning should be available to the reader to judge the study’s validity. Unfortunately, this methodical aspect is rarely reported in non-randomized studies including those performed in the field of pediatric IF.

The simplest method of excluding observations with missing data (listwise deletion) may be tempting, but is inefficient and wasteful, as study power is decreased and valuable information is lost. If data is missing at random, it should therefore be included in the analysis. The two approaches that are most frequently recommended and accepted are multiple imputation and maximum likelihood. Both approaches have been shown to produce very similar results and one has not been regarded superior to the other. In maximum likelihood, values with the highest probability are assigned using probability density of the realized data, also called the likelihood function. Using this likelihood function provides estimates based on all available data, including incomplete observations. This approach may be suboptimal in pediatric IF-research, as simulation studies have shown that it struggles with small sample sizes providing biased results. In large studies, it is however the preferred method for dealing with missing data. In small studies, multiple imputation is preferred to deal with missing data and used to replace missing values with imputed values from models that use information from all available data to estimate a distribution of possible values. This process is repeated multiple times and the possible
distributions are combined. Unfortunately, these two methods are only infrequently used in pediatric IF-research.

Simpler approaches include mean substitution and regression-based methods, both of which are single imputation techniques. In mean substitution, the missing values are replaced by the mean of the observed values. This technique uses information only from the affected variable and thereby preserves the mean of a variable distribution, but it will distort other aspects of the variable such as the variance and median\textsuperscript{522}. Regression based methods replace a missing value by using observed values of the affected case. This technique is problematic when correlations and analysis of covariance are tested, because the percentage of variance explained (R\textsuperscript{2}) is assumed to be perfect and the estimates of variability are underestimated\textsuperscript{518,522}. Other statistical approaches that may be considered for missing data, include expectation maximation that provides a maximum-likelihood estimate of the covariance using all available data similar to above mentioned maximum likelihood. Management of missing data should be a conscious step of statistical planning, where strengths and limitations of each method are evaluated and the most appropriate approach selected.

Although these techniques offer valuable options to IF-researchers to deal with missing values, they are rarely applied. This may be due to technical challenges and conceptual difficulties. Regrettably, these techniques are still perceived by many investigators as methods to simply make up values. Appropriate application of these methods offers valuable help for some of the methodological issues encountered in this field.

2.4.6 Minimally Important Values

It is important to understand the concept of minimally important values, as results are only meaningful if they are significant to the patient in the clinical setting.

Treatment effect may be quantified by mathematically comparing the efficiency of different treatments, e.g. by calculating the absolute risk reduction. Whether the calculated differences are minimally important differences is up to the clinician and patient to decide. For clinicians, minimally important differences are typically expressed in terms of the number of patients that need to be treated to prevent one outcome (i.e. number needed to treat, NNT), or to cause
one harmful event (i.e. number needed to harm, NNH). Minimally important differences from
the patient perspective are crucial when assessing meaningfulness of treatment differences,
but less easily calculated. A patient may perceive risks and benefits differently than the
clinician. Adjusting the number needed to treat or to harm with a patient’s perception of risk
permits quantification of the minimal important difference from both the clinician’s and
patient’s perspective. This can be illustrated with an example where two treatments aiming to
prevent sepsis are compared. For an individual patient being admitted with 4 episodes of
recurrent sepsis that appears more severe than a typical patient, the true number needed to
treat for this patient could be represented as NNT/4. Even if the relevant NNT for the
clinician appears too large, the adjusted/weighted NNT may represent a clinically meaningful
effect.

In pediatric IF-research, calculation of the number needed to treat or harm are rarely
performed. In addition, weighting patient’s perceptions into the judgment of minimally
important values is nearly impossible due to the young age of the population. However,
similar to HRQOL measurement, proxy perceptions may be used to incorporate a patient’s
perspective on treatment risks and benefits. This aspect is regularly forgotten in pediatric IF-
studies, but should be achievable by including results from questionnaires that elucidate
caregiver perception of the treatment in the study.

Minimally important values and clinical meaningfulness are hardly discussed in pediatric IF-
studies, but they strengthen the conclusions and provide context. When provision of exact
numbers needed to treat or harm are not feasible, discussing the results in the clinical context
including the patient’s perspective may be sufficient to increase meaningfulness of research
results. An example are studies that include patient/family views of home PN, where the
possibility of having the child receive PN therapy at home rather than in hospital was
perceived as an important benefit of the treatment.68,173

2.4.7 Forming Conclusions

Forming valid conclusions based on study results is as important to the study validity as the
study design, outcome selection and measurement, and analysis. A sound knowledge of
methodology will guide the process of forming valid conclusions.
Although conclusions should not be exaggerated, inflated conclusions have been observed repeatedly in clinical trials\(^5^{23}\). It is tempting to focus on dramatic measures, such as a high relative risk reduction, and ignore less compelling results in a study\(^4^{55}\). There is increasing evidence that relative risk reductions create higher opinions about efficacy among physicians and health policy makers, than their corresponding absolute risk reduction or numbers needed to treat\(^5^{24},5^{25}\). Reporting guidelines, such as the CONSORT statement, attempt to counter this behavior by recommending reporting of summary results, as well as, effect sizes with precision estimates for every tested primary and secondary outcome\(^4^{55},4^{63}\). Conclusions reported in a publication should be based directly on the study results, and should be clearly distinguished from conclusions derived from inferences based on clinical experience.

Subgroup analysis as part of a secondary analysis is a valuable tool for hypothesis generation, but strong conclusions from such analyses should be kept to a minimum due to the risk of spurious results arising from multiple testing\(^5^{26}\). Statistical methods such as the Bonferroni correction may be considered to address this problem. Advantages and disadvantages of applying such corrections should be evaluated and decisions should be in line with the study objectives. While corrections for multiple testing decrease the risk of type I errors, they also increase the risk of type II errors\(^5^{27}\). Therefore, if a study aims to explain observations and generate hypotheses, correction methods may not be necessary. While this topic is controversial, awareness and discussion of risks associated with multiple testing in the pediatric IF-field will increase credibility of study conclusions. In addition, unless the sample size was determined based on a priori determination of subgroup analysis, it is common for such an analysis to lack the statistical power to show a meaningful difference. This phenomenon is very common in the IF literature.

Study results that fail to show statistical significance require reporting and interpretation as much as statistically significant results do. Excluding non-significant results from studies introduces reporting bias. Such biased exclusions are difficult to trace in individual studies, as authors may simply choose to omit these results in their manuscript. Biased dissemination of results is more obvious at the journal level, where publication bias leads to negative outcome studies being published less frequently\(^4^{77}-4^{79}\). These biases may be decreased by appropriate interpretation of indeterminate results and temperance of conclusions. Indeterminate results may not be truly negative, if the study power was not sufficient to detect the measured effect
size. This may occur even in well-designed clinical trials when the effect size used in the a priori sample size calculation may not be equal to the one actually measured in the study. Although the measured effect size may be smaller than the one used in the sample size calculated, it may still be clinically meaningful or meaningful to the patient. Truly negative results (statistically non-significant results of a sufficiently powered study with an effect size that is not meaningful) are valuable as well and require reporting and publication. In pediatric IF-research, where sample sizes are frequently small and even small effects may be meaningful to families in the context of a complex disease with high morbidity and mortality, thorough understanding of statistical concepts and careful interpretation of results are crucial to create appropriate conclusions. Reporting effect sizes and confidence intervals instead of \( p \)-values are particularly important as the confidence intervals provide the reader with a measure of both significance and precision. An increasing number of journals now request reporting of confidence intervals instead of \( p \)-values as uncertainty measure.

### 2.5 Dissemination

Study results only become beneficial to patients if they are generalizable, accessible to clinicians and implementable in clinical practice. The limitations of pediatric IF-research are related to the rarity of the condition and the surgical nature of the disease, that negatively impacts generalizability and quality of evidence.

#### 2.5.1 Generalizability

Generalizability of results, also called external validity, is an important aspect of clinical research as it represents the step that interferes with dissemination and translation of the results to other research and clinical settings. While internal validity refers to the methodological quality and ability of the study to answer the exact research question, external validity refers to the capacity to generalize the results of the study to other settings and patients. A study that is only internally invalid should not be attempted, whereas a study with limited external validity may still be justifiable if it presents a step forward of testing an idea at a reasonable price\(^{455}\). When designing a study, an investigator will make a trade-off between internal and external validity. Results from a study with low internal validity are not useful, even with a high external validity; therefore, determination of internal validity should always be a priority when designing studies.
The extent of generalizability depends on study objectives, the nature of the clinical disorder and financial resources. Studies aiming to prove efficacy will be conducted in a highly artificial setting that simulates an experiment and aims to have the highest internal validity. Particularly in surgical research, such study designs may not be feasible as discussed in Chapter 2.1.1. Studies with less stringent criteria are frequently more realistic. They typically aim to prove effectiveness of the treatment and are more generalizable, but may pay for this advantage with decreased internal validity and a higher risk for bias. In pediatric IF-research, many studies tend to be on this end of the spectrum. While the pragmatic approach of this research enhances generalizability, the fact that pediatric IF-research is typically conducted at highly specialized centers decreases generalizability. The rarity of the disease may affect generalizability in several ways. First, studies are typically small, which decreases generalizability of study results. Recruiting patients from registries and different healthcare centers may improve this problem. Second, the surgical nature of the disease with variation in surgical indications and technique, and differences in surgical skill and experience will diminish generalizability. Inconsistent definitions and outcome selection in pediatric IF-research also impacts generalizability as the definitions of one study may be different than those used by other groups, and outcomes meaningful to one group may not be relevant to another. The rapidly evolving field of pediatric IF, also affects generalizability as the fast change of clinical outcomes may result in results that are no longer relevant to a wide audience by the time of publication.

Financial resources also influence quality and generalizability of a study. Better funded studies can employ more staff, and can enroll more patients from multiple centers increasing representativeness of a sample. Not surprisingly, the amount of available funding also affects internal validity of a study. Lower quality research makes it more difficult to secure future funding. Elucidation of methodological challenges related to pediatric IF-research and explanation of proposed solutions in funding applications may help justify resource allocation to this field of research.

2.5.2 Quality of Evidence

In the mid-1990s, the standard of randomized controlled trials was inferior when the studies were performed by surgeons, reported in surgical journals and conducted to evaluate
effectiveness of surgical procedures\textsuperscript{460}. Despite the relevance of surgical research, methodology was poor and led to some pointed comments such as the one by Dr. Richard Horton, Editor-in-Chief of \textit{The Lancet}, who compared surgical research to comic opera and regarded half of the surgical research as misconceived\textsuperscript{432,459}. Since then surgeons have produced an increasing number of well-conducted randomized controlled clinical trials\textsuperscript{432}. More challenges need to be addressed in surgical, clinical research, to release the full potential of this important discipline and related sub-disciplines that impact human life, anatomy, physiology and tissue as directly as no other\textsuperscript{432}. Since the milestone, blinded randomized clinical trial on the role of streptomycin for pulmonary tuberculosis, this study design has been the methodological gold standard as it minimizes bias from known and unknown confounders\textsuperscript{529}. Since introduction of minimal reporting standards for clinical trials, such as the Consolidated Standards of Reporting Trials (CONSORT), overall methodological quality of randomized controlled trials has improved, but remains suboptimal in certain subdomains such as reporting of adverse events, phase II oncology studies and non-pharmacological treatment including surgery\textsuperscript{463,530,531}. In 2013, Adie et al. found that of 150 surgical randomized controlled clinical trials less than half reported on important methodological details such as sample size calculation (45\%), random sequence allocation (43\%), allocation concealment (45\%) or blinding (37\%)\textsuperscript{532}. In solid organ transplantation trials, only 47\% of the CONSORT items were reported\textsuperscript{533}. Adherence to non-pharmacological treatment extension of CONSORT guideline appears to be particularly low: of 45 surgical trials, 8 of the CONSORT items were reported in less than 30\%, of which 7 items were specific to the non-pharmacological extension\textsuperscript{462,534}. In addition, surgical intervention trials often fail to report on funding sources (36\%), complete statistical power calculations (45\%) and factors pertinent to generalizability\textsuperscript{535}. Surgical trials also seem to be especially prone to be discontinued (43\% vs. 27\% of medical trials, \(p=0.001\)) and remain unpublished\textsuperscript{536}. Most commonly reported reason for early study discontinuation is poor recruitment\textsuperscript{536,537}. Approximately 2/3 of the surgical trials that are completed result in publication with a median lag of 4.9 years\textsuperscript{537}. Trial discontinuation and non-publication result in waste of resources, hidden trials data and ethical concerns\textsuperscript{538}. Moreover, unfortunate choices related to control arms appear to be contributing to the poor quality and biased results of surgical trials (e.g. when controls used in a trial perform worse than other groups using the same technique)\textsuperscript{539}. These issues are hypothesized to be true for pediatric IF-research, but to
date no study has systematically assessed these aspects of quality in pediatric IF. This may represent an interesting research question, but such a study may not be practicable for another few years due to the small number of randomized clinical trials conducted in pediatric IF.

Why is it that methodology in surgical research is persistently worse than in medical studies? There are methodological issues that are more pronounced in the surgical field such as lack of equipoise, blinding, funding and methodological training, standardization and learning curve effects, and regulatory factors that may play a role in this phenomenon\textsuperscript{432}. For a trial to be ethically sound, the subject and the physician must have a sense that both treatment options are equally beneficial; however, many surgeons may have an opinion about what intervention is more effective, making equipoise impossible. A recent survey on beliefs of practicing surgeons showed that they did not recognize equipoise because of limited appreciation of the methodological weakness of nonrandomized studies, little understanding of pragmatic trial designs, and lack of appreciation in the value of RCTs for generating high-quality data to change clinical practice\textsuperscript{432,540}. An example in pediatric IF that highlights these issues is the choice of surgical procedure to lengthen and taper a shortened intestine. There are essentially two options to lengthen an intestine, the longitudinal intestinal lengthening procedure (LILP) and serial transverse enteroplasty (STEP). While there are theoretical benefits and disadvantages to LILP and STEP, there is no randomized controlled trial supporting either surgical approach over the other. Surgeons who perform this procedure, however, typically have a preference and expertise in one of these operations that influences their judgment. Moreover, patients may not consent to let chance decide which surgical procedure they will receive after having sought the specific expertise of a treating surgeon. Patients may also be influenced by promotion and marketing of a novel intervention by institutions, leading to high rates of dropout and crossover\textsuperscript{541}.

Efficacy of surgical procedures depends significantly on a surgeon’s skills, preferences, and experience. It is difficult to standardize surgical interventions; therefore, comparing surgical interventions in a randomized controlled clinical trial is problematic. The treatment effect will be confounded by variability in surgeon performance. In addition to inter-surgeon variability, there is also intra-surgeon variability as the surgeon may experience a learning curve for the tested surgical intervention. Alternatively, if trials only test surgical interventions performed in a highly standardized manner by highly experienced surgeons,
external validity will be compromised, as the results will not reflect the larger surgical community. In pediatric SBS, the primary intervention is typically performed by any kind of pediatric surgeons, while surgeons who have developed an expertise in the field frequently perform subsequent procedures, such as bowel lengthening procedures. This is especially true in tertiary healthcare centers, but may not be the case in less specialized healthcare centers, where no surgeons have acquired expertise in pediatric IF. Furthermore, in teaching hospitals, fellows that are still acquiring surgical skills may perform procedures in those patients as well. And all of these procedures may be performed slightly differently depending on the surgeons’ preference and anatomy of the patient. This variability in training, preference, expertise and anatomy undoubtedly introduces heterogeneity and potentially bias into the study, particularly in study designs that are more easily attainable in pediatric IF, such as non-randomized trials or retrospective studies.

In surgical research it is not possible to blind investigators, but outcome assessors should be blinded to avoid bias and with some logistical effort, blinding of participants should be feasible as well. For example, laparoscopic versus open cholecystectomy was assessed using similarly sized wound dressings to conceal the incisions from the participants. Sham surgery may be ideal as a control, but it is not ethical in humans. This strategy is used frequently in pre-clinical studies involving animals. In IF-research, blinding is more feasible for pharmacologic treatment such as for ethanol locks versus heparin locks to prevent CBRSI in patients receiving PN. The limitation; however has been the exorbitant cost of the blinding process; specifically supplying ethanol and heparin locks in identical syringes. In this particular example, the clinical urgency and the encouraging, albeit imperfect, evidence supporting ethanol locks to prevent CRBSI in IF-patients has made it difficult to justify spending scarce funding resources on a treatment that appears to be effective and of low risk. Funding for surgical research has been a continuous challenge. In the USA, the success rate of surgeons obtaining funding from the National Institutes of Health (NIH) was been declining relative to nonsurgical colleagues. This seems to be related to the lower number of applications coming from surgical research teams. Whether this development is paralleled in Canada has not been assessed thoroughly. Research of a surgical and rare disease such as pediatric IF, however, is frequently not supported by many of the Canadian funding sources due to methodological challenges and the fact that the illness may not impact
enough patients to justify the investment. For example, the National Organization for Rare Disorders provides research funding for very specific, rare conditions such as Creutzfeld-Jacob-Disease, but not for any relevant to pediatric IF\textsuperscript{544}. The disadvantage of surgical research in the funding competition has been recognized repeatedly and suggestions have been made for national funding bodies to acknowledge this systematic difficulty and actively support surgery-driven research\textsuperscript{545}. Contributing to the problem may be the fact that only a very small number of surgical investigators have formal training in research methodology\textsuperscript{432,546}. In addition, surgeons are trained in a traditional “master-pupil” model where expertise is based on a school or methods they have internalized and prefer to apply, making them reluctant to change their practice and adopt new, unfamiliar approaches for the sake of a clinical trial. Also, the nature of surgical practice requires creativity and at times improvisation depending on the clinical circumstances. Hence, the line between what constitutes innovation and research can be unclear.

There is an ethical obligation to protect subjects from being exposed to unnecessary or harmful medical or surgical interventions. On several occasions, surgical procedures have been performed that subsequently were found to be ineffective or harmful. Examples include extracranial-intracranial bypass for ischemic stroke, internal mammary artery ligation, Halsted radical mastectomy, kidney decapsulation for hypertension and uterine suspension\textsuperscript{432,547,548}. Significant differences in the effective size of randomized versus non-randomized clinical trials were found, making the validity of uncontrolled trials questionable\textsuperscript{549,550}. What appears a highly effective treatment in a non-randomized study may be found ineffective when subjected to a randomized trial. A relevant example in pediatric IF focuses on the role of cholecystokinin (CCK) for the treatment of IFALD in patients on chronic PN. Teitelbaum et al. found in a retrospective cohort study of neonates matched by age and PN duration that there was a lower proportion of cholestasis in the group that received CCK\textsuperscript{551}. When the question was subjected to a randomized controlled clinical trial in over 200 neonates randomized to CCK or placebo, no difference in the rate of IFALD was found\textsuperscript{347}. Consequently, Cholecystokinin as a therapy for IFALD has disappeared.

To improve the quality of evidence in rare diseases, collaboration between multiple research groups has been suggested repeatedly. Collaborative efforts between institutions have been shown to more likely generate high quality research, although more attention has to be paid
to methodological details to avoid threats to internal validity from heterogeneity in multi-center studies. Many countries have developed national institutions to facilitate planning, conducting, and analyzing randomized multicenter trials for surgical diseases, such as the Study Center of the German Surgical Society (SDGC), the UK Medical Research Council Hubs for Trials Methodology Research, and the American College of Surgeons Continuous Quality Improvement Surgical Research Committee. In Canada, the benefit of central coordination of multi-center research has been recognized in the literature, but no formal office has been established to take over this role. Collaboration between institutions, as well as medical disciplines increases efficiency and generalizability, but also helps with challenges related to funding and manpower. In pediatric IF-research, the benefits of collaboration with multiple medical disciplines has been recognized. Involving different types of researchers (e.g. clinical, basic, qualitative, quantitative researchers), as well as integrating multiple academic disciplines (e.g. Surgery, Medicine, Nursing, Social Work) and nonacademic individuals (e.g. families, engineers) into clinical research may help resolve real-world patient care issues, create more practice-based evidence and help close the gap between trial-derived knowledge and actual patient care.

In research of a rare disease, such as pediatric IF, compromises need to be made to attain the highest quality while still making the research endeavor feasible. Many of the methodological limitations have been described above. It is important to remember that the objective of clinical research is to guide best care delivery while causing the least harm to the patient. Researchers should be transparent and openly describe limitations and discuss potential impact on the results. The more detailed the description of the study (e.g. treatment indications, recruitment process including inclusion and exclusion criteria, baseline characteristics, surgeon’s expertise and preconceptions, assumptions, potential for bias introduction) the more reliable the results will be perceived and the more valuable the conclusions.

2.5.3 Publications & Guidelines

Quality of clinical research and reporting has been impacted by the introduction of guidelines outlining methodological quality criteria. Clinical trials, was one of the first study designs to undergo strict conduction and reporting regulations, that are summarized in the Consolidated
Standards of Reporting Trials (CONSORT) guidelines\(^{408,553,554}\). A similar guideline was developed for observational studies, including cohort, case-control and cross-sectional studies: the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) criteria\(^{555,556}\). Although the same methodological rigor should be applied to research of both common and rare diseases, the latter group faces more limitations. Problems are related to insufficient sample size, inappropriate controls, lack of long-term outcomes, and scarce and biased information (see also *Chapter 2.5.2*). The CONSORT statement developed a version directed towards trials assessing non-pharmacological treatments, although this has not been adopted thoroughly in the surgical literature\(^{457,529}\). In pediatric IF-research reporting guidelines have mainly been used in meta-analyses, e.g. Newcastle-Ottawa risk of bias tool\(^{44,125,381}\).

Discontinuation and non-publication of surgical trials represents a major source of waste in a specialty that is already struggling with insufficient funding and well-trained investigators. Reasons for trial discontinuation in surgery appear to primarily consist of recruitment difficulties\(^{537}\). Surgical trials with larger sample sizes (>100) have been associated with a higher likelihood of publication, while industry funding appears to reduce this likelihood\(^{537}\). Clinical investigators’ experiences support the perception that larger, positive and qualitatively higher studies are published more easily. It is hypothesized that due to the typical small sample size, low power and observational study design, pediatric IF-research is published in lower impact journals. Selective publication of positive studies has been recognized as a problem that undermines the validity of evidence-based medicine by introducing publication bias\(^{557}\). Publication bias may be assessed in meta-analyses, but the few meta-analyses that have been published in pediatric IF exclude this aspect of quality assessment\(^{44,125,381}\). The primary reason for this is mainly the small number of studies included in these meta-analyses, making it impossible to reliably quantify publication bias using statistical or graphic methods\(^{44,125,381}\).
Chapter 3

3 Intestinal Failure Outcomes of a Matured Multidisciplinary Intestinal Rehabilitation Program

This chapter is formatted for publication in the Journal of Pediatrics.

3.1 Cover Page

INTESTINAL FAILURE OUTCOMES OF A MATURED MULTIDISCIPLINARY INTESTINAL REHABILITATION PROGRAM

Short Title: Outcomes Of Matured Intestinal Rehabilitation

Carol Oliveira, MD\textsuperscript{1,2,3,4}, Nicole T. de Silva, MSc\textsuperscript{1,2}, Yaron Avitzur, MD\textsuperscript{1,5}, Ahmed M. Bayoumi, MD\textsuperscript{4,6}, Wendy J. Ungar, PhD\textsuperscript{3,4}, Jeffrey S. Hoch, PhD\textsuperscript{4,7}, Paul W. Wales, MD\textsuperscript{1,2,3,4}

3.1.1 Affiliations

\textsuperscript{1}Group for Improvement of Intestinal Function and Treatment Program (GIFT), The Hospital for Sick Children, Toronto, Canada

\textsuperscript{2}Division of General and Thoracic Surgery, The Hospital for Sick Children, Toronto, Canada
3 Child Health Evaluative Sciences (CHES) Program, The Hospital for Sick Children, Toronto, Canada

4 Institute of Health Policy, Management and Evaluation (IHPME), University of Toronto, Toronto, Canada

5 Division of Gastroenterology, Hepatology and Nutrition, The Hospital for Sick Children, Toronto, Canada

6 Centre for Research on Inner City Health, Li Ka Shing Knowledge Institute and Division of General Internal Medicine, St. Michael’s Hospital, Toronto, Canada

7 Centre for Excellence in Economic Analysis Research (CLEAR), St. Michael’s Hospital, Toronto, Canada

3.1.2 Corresponding Author

Paul W. Wales, BSc, MD, MSc, FRCSC, FACS
Neonatal and Pediatric Surgeon
Director, The Group for Improvement of Intestinal Function and Treatment (GIFT)
Division of General and Thoracic Surgery
Associate Scientist, Child Health Evaluative Sciences, Research Institute
The Hospital for Sick Children
Associate Professor, Department of Surgery, University of Toronto
Room 1526, 555 University Ave
Toronto, ON M5G 1X8, Canada
Tel: +1 416 813 7654 x201490
Fax: +1 416 813 7477
E-mail address: paul.wales@sickkids.ca

3.1.3 Abbreviations

GIFT – The Group for Improvement of Intestinal Function and Treatment; IFALD – intestinal failure associated liver disease; CRBSI – catheter-related bloodstream infection;
IRP – intestinal rehabilitation program; IF – intestinal failure; NICU – Neonatal Intensive Care Unit; PN – parenteral nutrition; SBS – Short Bowel Syndrome

3.1.4  Key Words

short bowel syndrome, parenteral nutrition, multidisciplinary care, gastroenterology, birth defects, fetus/newborn, infant

3.1.5  Funding Sources

The Swiss National Science Foundation, The Walter and Gertrud Siegenthaler Foundation and The University of Zurich, Switzerland; The Hospital For Sick Children, Toronto, Canada (Research Scholarships)

3.1.6  Financial Disclosure

All authors have no financial relationship relevant to this article to disclose.

3.1.7  Conflict Of Interest

All authors have no conflict of interest relevant to this article to disclose.

3.1.8  Clinical Trial Registry

Not applicable

3.1.9  Manuscript Writing

Dr. Carol Oliveira wrote the first draft of the manuscript. No honorarium, grant, or other form of payment was given to anyone to produce the manuscript.

3.2  Abstract

Objective: The paradigm of pediatric intestinal failure management changed with the introduction of multidisciplinary intestinal rehabilitation programs (IRP) and new medical
and surgical treatment options. The study objective was to examine practice change over time and quantify the effect of IRP-introduction and maturation on patient outcomes.

**Study Design:** A prospective cohort study was performed using data from a tertiary children’s hospital’s IRP-registry. Patients were divided into 3 eras: Pre-IRP (1998-2000, n=50), Early-IRP (2003-2005, n=43) and Late-IRP (2006-2009, n=58). Non-parametric and time-to-event analysis was performed.

**Results:** Median gestational age was 33-35 weeks, etiology and remaining small bowel was similar (75%-77%) amongst the groups. Transplantation and serial-transverse enteroplasty were predominant in Early-IRP, while their utilization decreased in Late-IRP where lipid management became more prominent. Independence from parenteral nutrition was achieved in 54%-70% after a follow-up of 657-309 days. Days in NICU decreased from a median (interquartile range) of 86 (95-145) to 59 (17-102) ($p=0.002$), but overall hospitalization remained unchanged. Sepsis/1000 catheter-days declined from 22 (11-35) to 12 (2-25) ($p=0.014$). Disease-specific mortality from liver disease and sepsis decreased from 26% to 3.5% ($p=0.004$).

**Conclusions:** Introduction of IRP initially increased the transplantation rate and improved outcomes. Further improvement and change of the natural course of the disease were observed with IRP-maturation and adoption of novel lipid strategies. No conclusions on causal inference of individual treatment options may be made. Observed effects represent the summary of a compound exposure of early and mature IRP including introduction of new medical and surgical treatment options.

### 3.3 Introduction

Pediatric intestinal failure (IF) is a devastating condition caused by inability to absorb sufficient nutrients to grow and survive, either due to anatomical or functional loss of intestine$^{2,3}$. These patients are dependent on parenteral nutrition (PN) for a prolonged period of time, which is associated with high morbidity, mortality and cost$^2$. Short bowel syndrome (SBS) represents the most frequent IF-type in children and is caused by neonatal or congenital loss of significant length of functional small bowel. Recent advances in medical
care have led to improved clinical outcomes in SBS-patients and possibly to an increased prevalence\textsuperscript{2,59}. Multidisciplinary intestinal rehabilitation (IRP) programs have mainly promoted these advances in SBS-management.

In 2002, we established a multidisciplinary IRP at our institution (The Group for Improvement of Intestinal Function and Treatment – GIFT) with the objective to improve clinical care of pediatric IF-patients. The early experience of GIFT (2003-2005) was previously evaluated and revealed a reduction in the catheter-related bloodstream infection (CRBSI) rate, reduction in mortality from liver failure from 22% to 11% and increase in the number of performed transplantations, when compared to historic controls\textsuperscript{49}. This effect was ascribed to improved communication between the involved medical specialties and streamlining of care, while the clinical course of the disease remained basically unaffected. With the maturation of the IRP and introduction of new therapeutic options, further improvement of clinical outcomes is hypothesized.

In this study, we aimed to evaluate the effect of a matured multidisciplinary intestinal rehabilitation program including new therapeutic options on clinical outcomes of pediatric SBS.

3.4 Patients & Methods

3.4.1 Study Design

A retrospective analysis of prospectively collected, observational data from a single study site was performed. Ethics board approval was obtained from The Hospital For Sick Children and the University of Toronto, in agreement with which no consents were obtained.

3.4.2 Study Sample

Consecutive patients diagnosed with SBS between 0 and 365 days of age were included in the analysis. To allow for a washout period, an historic control group (Pre-IRP) was selected including consecutive patients born between 1998 and 2000. The experimental group was exposed to IRP and selected based on the referral date. Early-IRP and Late-IRP consisted of consecutive patients referred between 01/2003 and 12/2005, and 01/2006 and 12/2009,
respectively. These time periods were selected based on IRP-introduction in 2003 and initial evaluation of clinical outcomes after the first 3 years of IRP-experience\textsuperscript{49,65}.

The IRP consists of a pediatric staff surgeon, gastroenterologist, nurse practitioner, physiotherapist, occupational therapist, nutritionist and social worker. Goals of the IRP consist of providing continuity in care (e.g. daily review of all patients with program director, outpatient multidisciplinary clinic), optimizing communication between medical disciplines (e.g. weekly IRP team meetings, earlier transplant listing), recruiting dedicated multidisciplinary staff to the IRP, providing continuous training to develop experts in intestinal failure and rehabilitation (e.g. journal clubs, informal teaching, clinical experience), improving parent education (e.g. central line teaching), facilitating outpatient care and providing outpatient support (e.g. home TPN, home care nurse, accessible IRP team), introducing and evaluating new treatment options for pediatric IF (e.g. introduction of novel lipid strategies, bowel lengthening procedures) and conducting high-quality research to improve evidence-based knowledge on this rare condition. With increased experience, these goals are achieved more persistently and extensively.

The SBS-criteria used by the GIFT-team for referral of new patients are primarily functional with dependence on parenteral nutrition for greater than 42 days after bowel resection or injury\textsuperscript{50}. For patients with less than 25% of expected age-specific small bowel length, referral is immediate. Typically patients are initially treated at the neonatal or pediatric intensive care unit and then transferred to a dedicated IRP unit once they are deemed stable enough for the ward. Patients with abdominal wall defects were eligible for this study sample; those with dysmotility or mucosal enteropathy were excluded.

\subsection*{3.4.3 Data Collection}

Data for this study were derived from the IRP patient database maintained by GIFT. A single experienced data coordinator performed data extraction and input. Time duration was calculated using statistical software and dates were extracted from patient records. Time 0 was defined as the date of primary surgery. Follow-up was continued until patients were discharged from IRP-care (n=60), died (n=45) or to study end (n=46; Pre-IRP: December 12, 2002; Early-IRP and Late-IRP: December 31, 2009).
Data on patient characteristics, surgical interventions, medical treatment, and outcomes related to nutrition, sepsis, IF-associated liver disease (IFALD), transplantation and mortality were extracted from the registry and analyzed. Cholestasis was defined as a conjugated bilirubin level >50µmol/L for at least 2 weeks without association with sepsis; and advanced liver dysfunction as conjugated bilirubin >100µmol/L for at least 2 weeks and not associated with sepsis. Catheter-related bloodstream infections were defined as a positive blood culture from the central venous catheter (CVC) not related to another source of bacteremia and reported per 1000 catheter days. Complete weaning from PN was defined as independence from PN for ≥2 consecutive weeks with adaptive weight gain. If PN had to be restarted during this period, the total PN duration was calculated adding the PN dependent episodes together. Length of hospitalization and admission to the intensive care unit (ICU) refers to the primary hospitalization only.

3.4.4 Data Analysis

Data were analyzed using Stata 11.2 software. Continuous variables were described using the median and interquartile range (IQR). Significance level was set at a two-sided alpha of 0.05. Kruskal-Wallis test was used to compare continuous variables among the three groups. Chi-Square and Fisher’s exact tests were applied for binary outcomes. Univariate time-to-event analysis was performed using a Log-Rank test. Multivariate analysis was attempted including application of data transformations and propensity scores, but failed to provide valid results, as the proportional hazard assumption was not met. Missing data were rare (<2% per outcome such as birth weight, home-PN and central catheter complication rate) and related to patients initially managed at external hospitals and subsequently transferred to the study site. Casewise deletion was implemented for these observations. Extensive sensitivity analysis was performed for exclusion of referred patients, early deaths and patients with interrupted PN-exposure (not reported, as no clinically meaningful impact on results).
3.5 Results

3.5.1 Patient Characteristics

Our IRP managed 43 and 58 SBS-patients who met the inclusion criteria between 2003 and 2005, and between 2006 and 2009, respectively. These patients were compared to 50 historic controls that were treated at our institution before the inception of GIFT between 1998 and 2000. Data on patient characteristics are outlined in Table 3-1.

<table>
<thead>
<tr>
<th>Table 3-1: Patient Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
</tr>
<tr>
<td>Gestational age [weeks]</td>
</tr>
<tr>
<td>Birth weight [grams]</td>
</tr>
<tr>
<td>Gender [male]</td>
</tr>
<tr>
<td>Age at primary surgery [days]</td>
</tr>
<tr>
<td>Etiology</td>
</tr>
<tr>
<td>Necrotizing enterocolitis</td>
</tr>
<tr>
<td>Abdominal wall defects</td>
</tr>
<tr>
<td>Meconium ileus</td>
</tr>
<tr>
<td>Small bowel atresia</td>
</tr>
<tr>
<td>Volvulus</td>
</tr>
<tr>
<td>Spontaneous perforation</td>
</tr>
<tr>
<td>Hirschsprung’s disease</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>ICU admission [days]</td>
</tr>
<tr>
<td>Primary Hospitalization [days]</td>
</tr>
<tr>
<td>Home parenteral nutrition [yes]</td>
</tr>
<tr>
<td>Follow-up [days]</td>
</tr>
</tbody>
</table>

Values represent medians and interquartile ranges or frequencies with percentages (%). Abbreviations: ICU – Intensive Care Unit, * – statistically significant

The three comparison groups were similar in terms of gestational age, birth weight, gender distribution and age at primary surgery. Median gestational age was 33-35 weeks (p=0.1) and
median birth weight ranged between 1740 and 2010 grams across the three groups \((p=0.3)\). Slightly more patients were male (53-63\%, \(p=0.6\)) and the age at primary surgery was 4-5 days \((p=0.9)\). The underlying etiologies of SBS were similar with Necrotizing Enterocolitis (NEC) in 34-38\%, abdominal wall defects in 18-26\% and small bowel atresia in 10-21\%.

Patients were in the Neonatal Intensive Care Unit (NICU) for a significantly shorter period of time in Late-IRP (Pre-IRP 86 (59-145), Early-IRP 91 (72-126), Late-IRP 59 (17-102) days; \(p_{\text{PreIRP:LateIRP}}=0.004\), \(p_{\text{EarlyIRP:LateIRP}}=0.001\)). Overall hospitalization was only slightly shorter in Pre-IRP (Pre-IRP 100 (60-182), Early-IRP 130 (89-194), Late-IRP 135 (89-187) days; \(p=0.3\)). Home-PN was used in 12\% during Pre-IRP, 19\% during Early-IRP and 27\% during Late-IRP \((p=0.1)\). Out of province referrals occurred in 5\% in Early-IRP and 17\% in Late-IRP. Follow-up time was insignificantly shorter for Late-IRP (Pre-IRP 657 (240-935), Early-IRP 448 (168-1030), Late-IRP 309 (106-581) days; \(p=0.1\)).

### 3.5.2 Intestinal Anatomy

Details on the intestinal anatomy of included patients are outlined in Table 3-2.

<table>
<thead>
<tr>
<th></th>
<th>Pre-IRP ((n = 50))</th>
<th>Early-IRP ((n = 43))</th>
<th>Late-IRP ((n = 58))</th>
<th>(p)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small bowel length [cm]</td>
<td>75 (40-106)</td>
<td>77 (59-115)</td>
<td>75 (40-120)</td>
<td>0.6</td>
</tr>
<tr>
<td>Small bowel length [% of expected]</td>
<td>81 (42-100)</td>
<td>85 (36-100)</td>
<td>62 (27-97)</td>
<td>0.3</td>
</tr>
<tr>
<td>Colon length [cm]</td>
<td>41 (29-45)</td>
<td>41 (25-46)</td>
<td>30 (22-40)</td>
<td>0.014*</td>
</tr>
<tr>
<td>Ileocecal valve resection [yes]</td>
<td>13 (26%)</td>
<td>8 (19%)</td>
<td>18 (31%)</td>
<td>0.4</td>
</tr>
<tr>
<td>Gastrostomy feeding tube [yes]</td>
<td>14 (28%)</td>
<td>20 (47%)</td>
<td>29 (50%)</td>
<td>0.053</td>
</tr>
<tr>
<td>STEP [yes]</td>
<td>0 (0%)</td>
<td>6 (14%)</td>
<td>5 (9%)</td>
<td>0.015*</td>
</tr>
</tbody>
</table>

Values represent medians and interquartile ranges or frequencies with percentages (%). *Abbreviations:* STEP - serial-transverse enteroplasty; \* - statistically significant.

After primary surgery remaining small bowel length was between 75cm and 77cm (Pre-IRP 75cm (40-106), Early-IRP 77cm (59-115), Late-IRP 75cm (40-120); \(p=0.6\)), while the percentage of expected age-adjusted small bowel length was slightly smaller in Late-IRP.
(62% (27-97)) when compared to Pre-IRP (81% (42-100)) and Early-IRP (85% (36-100); $p=0.3$). Colon length was significantly longer in Pre-IRP (41cm (29-45)) and Early-IRP (41cm (25-46)) than in Late-IRP (30cm (22-40); $p=0.014$). Ileocecal valve resection occurred less frequently in Early-IRP (19%) than in Pre-IRP (26%) and Late-IRP (31%; $p=0.4$).

A greater utilization of gastrostomy feeding tubes was observed over time with 28% in Pre-IRP, 47% in Early-IRP and 50% in Late-IRP ($p=0.053$). The first STEP at our institution was performed in 2003. During Early-IRP, STEP was performed on 14% of the patients, while during Late-IRP only 9% underwent this procedure ($p=0.015$) (Table 3-1).

3.5.3 Outcomes

3.5.3.1 Nutritional Outcomes

Nutritional outcomes improved over time (Figure 3-1). In the categorical comparison, oral tolerance of 50% of the required daily calories was achieved in 72% in Pre-IRP, 84% in Early-IRP and 85% in Late-IRP ($p=0.2$; relative risk (RR) Early-IRP vs. Pre-IRP 1.16, Late-IRP vs. Pre-IRP 1.17). Time-to-event analysis revealed that this was achieved later, without being statistically significant ($p=0.7$) (Figure 3-1: A). Half of the patients achieved 50% enteral tolerance at 89 days Pre-IRP, 88 days Early-IRP and 91 days Late-IRP.

In the categorical comparison, complete PN-weaning was achieved in 62%, 70% and 61% in Pre-IRP, Early-IRP and Late-IRP ($p=0.3$; RR Early-IRP vs. Pre-IRP 1.13, Late-IRP vs. Pre-IRP 0.86). When adjusting for censoring (mostly through loss to mortality) in the time-to-event analysis, it becomes evident that with the IRP-introduction and maturation, complete PN-weaning was achieved more frequently and it took longer to complete ($p=0.012$) (Figure 3-1: B).
Figure 3-1: Nutritional Outcomes

A) Kaplan-Meier failure estimates
50% Enteral Feeds

Kaplan-Meier failure estimates
Pre-IRP ----- Early-IRP
Late-IRP

B) Kaplan-Meier failure estimates
Complete Weaning

Kaplan-Meier failure estimates
Pre-IRP ----- Early-IRP
Late-IRP

\[ p_{Log-Rank} = 0.7 \]

\[ p_{Log-Rank} = 0.012^* \]
Overall the adaptation process took longer, but was completed more frequently with the maturation of IRP. A) Enteral tolerance of 50% of daily calories represents an early outcome of the adaptation process. This outcome was similar for the three comparison groups. B) The final outcome of the adaptation process is weaning from parenteral nutrition and enteral tolerance of 100% of daily calories. This process is significantly different for Late-IRP when compared to Pre-IRP and Early-IRP \((p=0.005, p=0.02)\). * – statistically significant

Half of the patients were weaned completely from PN at 114 days Pre-IRP, 120 days Early-IRP and 363 days Late-IRP.

### 3.5.3.2 Outcomes Related to Central Venous Catheters

Morbidity from long-term PN was evaluated by assessing CVC-complication rate, CRBSI-rate and IFALD-development (Figure 3-2). The number of CVCs per 1000 catheter-days per patient was significantly greater in Pre-IRP (32 (20-41)) when compared to Early-IRP (20 (16-34)) and Late-IRP (17 (9-29); \(p<0.001\)). Complications from CVCs included dislodgement, thrombosis, occlusion and breakage. The complication rate per 1000 catheter-days was 18 (16-23) Pre-IRP, 14 (7-32) Early-IR and 10 (4-17) Late-IRP \((p=0.09)\) (Figure 3-2: A).

The CRBSI-rate was significantly lower after IRP-introduction and maturation \((p=0.0138)\) (Figure 3-2: B). In Pre-IRP 22 (11-35) CRBSI per 1000 catheter-days occurred, which is 46% more than in Early-IRP (15 (5-39)) and 83% more than in Late-IRP (12 (2-25)). There was no difference between gram-positive, gram-negative and fungal infections between eras.

### 3.5.3.3 Outcomes Related to Liver Disease

Impact of IFALD was assessed by evaluating the cholestasis, advanced liver dysfunction and death from liver failure (Figure 3-2: C). After IRP-introduction, more patients developed cholestasis (RR Early-IRP vs. Pre-IRP 1.06, Late-IRP vs. Pre-IRP 1.09) and advanced liver disease (RR Early-IRP vs. Pre-IRP 1.09, Late-IRP vs. Pre-IRP 1.09), which is likely secondary to increased awareness of IFALD by treating physicians, while less patients died from progression into liver failure \((p=0.003; \text{RR Early-IRP vs. Pre-IRP 0.47, Late-IRP vs. Pre-IRP 0.17).} \)
Omega-3 intravenous lipid emulsions for patients with advanced liver dysfunction (conjugated bilirubin >100µmol/L) were introduced in late 2005. In 2006, 64% of the patients received Omega-3 intravenous lipid emulsions, while this percentage decreased to 35% at the end of Late-IRP. Early referral for transplantation assessment was persistent throughout the IRP-experience. While during Early-IRP 60-90% of the assessed patients were waitlisted for transplantation, this percentage decreased continuously to 10% at the end of Late-IRP (Figure 3-2: D).
Morbidity from central venous catheters decreased with the introduction and maturation of IRP. Elevated conjugated bilirubin levels were observed more frequently, while mortality from liver failure decreased. A) The rate of complications from central venous catheters decreased slightly from 18 to 14 to 10 complications/1000 catheter days. B) The rate of catheter-related bloodstream infections (CRBSI) decreased significantly from 22 to 15 to 12/1000 catheter days with the introduction and maturation of IRP. C) The development of IFALD changed with the introduction and maturation of IRP. The increase of cholestasis and advanced liver disease, and decrease of mortality from liver failure may reflect more aggressive parenteral nutrition administration, more complex patients and effective prevention of mortality from liver failure over time. D) While the number of patients being assessed for transplantation remained high, the number of patients on the waiting list and receiving a transplant decreased continuously with maturation of the IRP. * – statistically significant

During the IRP-period, one transplantation per year was performed in average with a peak of 5 in 2005 and a nadir of 0 in 2008. In contrast to Early-IRP, several patients were removed from the waiting list as their clinical status improved in Late-IRP.
3.5.3.4 Mortality

Mortality changed significantly with IRP-introduction and maturation (Figure 3-3). Disease-specific mortality (from liver failure and sepsis) decreased significantly with IRP-maturation ($p=0.04$) (Figure 3-3: A). In the categorical comparison, the proportion of patients dying from disease-specific complications decreased from 26% in Pre-IRP, to 19% in Early-IRP and 4% in Late-IRP ($p=0.004$; RR Early-IRP vs. Pre-IRP 0.72, Late-IRP vs. Pre-IRP 0.13).

The case-fatality rate for overall mortality decreased steadily from Pre-IRP 38% to Early-IRP 30% and Late-IRP 22% ($p=0.2$; RR Early-IRP vs. Pre-IRP 0.8, Late-IRP vs. Pre-IRP 0.6). Mortality cause changed significantly with IRP-maturation ($p=0.008$) (Figure 3-3: B).
Figure 3-3: Mortality

With the introduction and maturation of IRP the pattern of mortality changed significantly. A) Disease-specific mortality decreased significantly with maturation of IRP. B) The cause of mortality changed from SBS complications (i.e. parenteral-nutrition related liver disease, catheter-related bloodstream infection) being nearly 70% to 15%. * – statistically significant
While in Pre-IRP 69% of deaths were caused by liver failure or sepsis, only 39% of the mortality was disease-specific in Early-IRP. In Late-IRP 85% of deaths were secondary to comorbidities of prematurity such as cardio-respiratory, neurological, or renal origin.

3.6 Discussion

Pediatric SBS is a severe, chronic condition associated with high morbidity and mortality\textsuperscript{3,65}. Several multi-disciplinary IRP have been introduced across North America and Europe with the aim to improve clinical management and outcomes of this condition\textsuperscript{42,44,173,175,407,559-565}. In a recent systematic literature review of IRP, three studies were identified that evaluated the effectiveness of IRP comparing patient outcomes to a historic control\textsuperscript{44,46,49,165}. They revealed improved survival, decreased occurrence of liver failure, time to intestinal autonomy and more frequent transplantation\textsuperscript{44,46,49,165}. While these studies were able to show how the introduction of IRP affected the clinical outcome of pediatric SBS-patients, our study is the first to evaluate the change in outcome over time, as a program evolves. Maturation of a program may include improved process streaming, increased knowledge and application of new evidence, increased experience of the health care providers and introduction of new medical treatment options. We hypothesize that these factors together further improved the clinical outcome of pediatric SBS and ultimately led to a change in the course of the disease.

Patient characteristics of our cohort were similar among the three groups of comparison and when compared to other studies\textsuperscript{44,566}. The decreased NICU admission time and unchanged total length of hospitalization illustrates increased comfort and confidence on the GIFT inpatient unit to accept care of complex neonates from the NICU earlier.

As we have observed clinically and in our neonatal piglet model of short bowel syndrome, it is common for patients to achieve tolerance of 50% of calorie requirements enterally\textsuperscript{567,568}. The process of reaching 100% enteral calories and sustained PN weaning represents the major challenge and may take several years to complete. A major benefit of an IRP is that it can prevent or slow progression of complications that would otherwise shorten a patient’s life before enteral autonomy can be realized. Historically, the most common cause of death in the IF-population was liver failure\textsuperscript{49}. Protocols to reduce sepsis, alter PN composition (e.g. novel lipid management strategies) and optimize intestinal absorptive capacity (feeding strategies,
autologous bowel reconstruction techniques) help prevent development of hepatic failure providing more time for full adaptation to occur. Patients in Early-IRP appear to benefit from these protocols, as the surviving patients are adapting more frequently (Figure 3-2: B). Late-IRP patients survive more frequently; adding patients to this group that would have died in the earlier groups and require more time to adapt fully (Figure 3-2: B, Figure 3-3: A). As the adaptation process may take up to 5 years to complete, the follow-up time may not be sufficiently long to capture the improved complete PN weaning proportion for the Late-IRP group.

The introduction and maturation of IRP significantly affected rates of central venous catheter (CVC) related complications. This change illustrates the compound effects of IRP as no novel treatment options directed towards sepsis and complication prevention were introduced during this time of the program. We hypothesize that this effect is due to more experienced health care professionals handling CVCs and a healthier intestinal barrier secondary to a more aggressive enteral nutrition administration and more sophisticated utilization of motility agents. Other contributing factors may be a less immunosuppressed state of the patient due to less malnourishment and more frequent use of antibiotics to prevent bacterial overgrowth. Ethanol locks have been introduced successfully at our institution in 2009 and further improvement of the CRBSI and CVC-complication rate has been shown in a previous study.

The course of IFALD has clearly changed with the introduction and maturation of IRP. The early increase in the proportion of patients with IFALD may have been due to increased awareness of IFALD and possible detection bias. With maturation of IRP, the incidence of cholestasis and advanced liver failure decreased. Initially, the addition of novel lipid management strategies such as Omega-3 lipid emulsions helped salvage infants with advanced liver disease. Interestingly, however, the use of Omega-3 lipid emulsions has decreased over time and the proportion of patients developing progressive IFALD has continued to diminish. Other factors such as reduction in sepsis and earlier, more aggressive enteral feeding may have also had an impact. Similarly, the need for STEP and transplantations decreased over time indicating that more complex patients are surviving while progressing slower and requiring less surgical and medical treatment options more recently. It is hypothesized that this effect is primarily related to intangible effects of a
matured IRP that provides increased continuity in care and streamlining of processes, improved communication between medical disciplines, dedicated multidisciplinary staff with increasing medical expertise and experience, improved parent education, more skillful outpatient care and availability of outpatient support.

3.6.1 Limitations

As with every rare disease study, the sample size is frequently limited leading to analytic challenges such as lack of adequate statistical power to conduct multivariate analysis. Although different methods were explored in this analysis using transformation of variables and propensity scores, valid multivariate analysis was not feasible. Our prospective database, however, is one of the largest and with the longest follow-up reported in literature. Rarity of a disease and associated methodological limitations should not prevent outcome assessment of new treatment modalities as it provides unique and valid evidence when reported correctly and interpreted within its limitations.

Our comparison of three eras does not permit conclusions on the cause-effect relationship of individual factors, as a concurrent control group was not possible. In addition, treatment options were introduced and administered dynamically during one era. The aim of the study was therefore not to determine the role of individual treatment options, but to describe the development of our IRP and quantify improvements in clinical outcomes as the program evolved. The eras represent a compound exposure, that includes measurable treatment interventions such as alternative lipid emulsions and more intangible factors that occur when there is an experienced team managing these patients on a dedicated unit with dedicated resources and developed treatment guidelines. In our study, we were able to show that the introduction of our IRP initially led mainly to a significant increase in transplantation and STEP, while with maturation of the program, surgical interventions became a less crucial therapeutic option. With time, disease-specific mortality decreased significantly as did CVC-complication and sepsis rates.
3.6.2 Future Research

The care provided by a multidisciplinary IRP represents a complex construct of medical interventions and human factors. New therapies are introduced over time and outcomes evolve as programs mature. Future studies assessing the impact of specific treatment modalities are necessary. Most literature is retrospective in design and determination of therapeutic efficacy can be confounded by other treatments and a global improvement in medical care that occurs over time. Ideally, individual therapies would be evaluated with prospective, randomized controlled trials. If such a design is not feasible, longitudinal study design and time-series analysis, to capture change over time and determine specific factors influencing its development, may be a valuable option. In addition, a formal cost-effectiveness analysis should be performed to evaluate resource allocation in context of health utility. Finally, a validated severity of illness measurement tool for the pediatric IF-population would be helpful to permit more reliable risk adjustment between treatment groups and facilitate comparison of patients between studies.

3.7 Conclusion

Our study is the first to characterize the fluid change in outcome that is the new paradigm of infantile SBS. While IRP-introduction led to an increased transplantation rate and improved clinical outcomes, further progress was made and the natural course of the disease changed with IRP-maturation and adoption of novel treatment strategies. The described effects are secondary to a compound exposure to early and more mature IRP including the introduction of new medical and surgical treatment options.
Chapter 4

4 Change of Outcome in Pediatric Intestinal Failure – The Use of Time Series Analysis to Assess the Evolution of an Intestinal Rehabilitation Program

This chapter is formatted for publication in the *Journal of the American College of Surgeons*.

4.1 Cover Page

**CHANGE OF OUTCOME IN PEDIATRIC INTESTINAL FAILURE – THE USE OF TIME SERIES ANALYSIS TO ASSESS THE EVOLUTION OF AN INTESTINAL REHABILITATION PROGRAM**

Short Title: Pediatric Intestinal Rehabilitation Evolution

Carol Oliveira, MD\(^1,2,3,4\), Nicole T. de Silva, MSc\(^1,2\), Sanja Stanojevic, PhD\(^3,4,5\), Yaron Avitzur, MD\(^1,6\), Ahmed M. Bayoumi, MD, MSc\(^4,7\), Wendy J. Ungar, PhD\(^3,4\), Jeffrey S. Hoch, PhD\(^4,8\), Paul W. Wales, MD, MSc, FACS\(^1,2,3,4\)

4.1.1 Affiliations

\(^1\)Group for Improvement of Intestinal Function and Treatment Program (*GIFT*), The Hospital for Sick Children, Toronto, Canada
2Division of General and Thoracic Surgery, The Hospital for Sick Children, Toronto, Canada

3Child Health Evaluative Sciences (CHES) Program, The Hospital for Sick Children, Toronto, Canada

4Institute of Health Policy, Management and Evaluation (IHPME), University of Toronto, Toronto, Canada

5Division of Respiratory Medicine, The Hospital for Sick Children, Toronto, Canada

6Division of Gastroenterology, Hepatology and Nutrition, The Hospital for Sick Children, Toronto, Canada

7Centre for Research on Inner City Health, Li Ka Shing Knowledge Institute and Division of General Internal Medicine, St. Michael’s Hospital

8Centre for Excellence in Economic Analysis Research (CLEAR), St. Michael’s Hospital, Toronto, Canada

4.1.2 Corresponding Author

Paul W. Wales, BSc, MD, MSc, FRCSC, FACS
Neonatal and Pediatric Surgeon
Director, The Group for Improvement of Intestinal Function and Treatment (GIFT)
Division of General and Thoracic Surgery
Associate Scientist, Child Health Evaluative Sciences, Research Institute
The Hospital for Sick Children
Associate Professor, Department of Surgery, University of Toronto
Room 1526, 555 University Ave
Toronto, ON M5G 1X8, Canada
Tel: +1 416 813 7654 x201490
Fax: +1 416 813 7477
E-mail address: paul.wales@sickkids.ca
4.1.3 Abbreviations

IF – intestinal failure, IRP – intestinal rehabilitation program, PN – parenteral nutrition, 
CRBSI – catheter-related bloodstream infections, CVC – central venous catheter, ICU – 
intensive care unit, ARIMA - autoregressive integrative moving average models, IQR – 
interquartile range, CI – confidence interval

4.1.4 Key Words

short bowel syndrome, intestinal failure, intestinal rehabilitation, parenteral nutrition, 
multidisciplinary care, gastroenterology, bowel lengthening procedure, serial transverse 
enteroplasty, omega-3 lipids, neonate, infant

4.1.5 Funding Source

The Swiss National Science Foundation, The Walter and Gertrud Siegenthaler Foundation 
and The University of Zurich, Switzerland; and The Hospital For Sick Children, Toronto, 
Canada, have financially supported the first author with research scholarships during 
conduction of this research.

4.1.6 Financial Disclosure

All authors have no financial relationship relevant to this article to disclose.

4.1.7 Meeting Presentation

The study abstract has been presented at the annual meeting of the Canadian Association of 
Paediatric Surgeons (CAPS) in 2013.

4.2 Precis

Independent effects of new treatment options in the evolution of pediatric intestinal failure 
over 15 years are explored using time series analysis. Intestinal rehabilitation programs and 
omega-3 lipids decreased disease-specific mortality. Serial transverse enteroplasty and 
ethanol locks did not impact significantly.
4.3 Abstract

**Background:** The clinical picture of pediatric intestinal failure has changed over the past 15 years, while effectiveness evolving treatment options remains unclear. The study explores evolution in care and quantifies independent effects of new treatment options.

**Study Design:** Consecutive patients (n=196) with neonatal/infantile intestinal failure born between 07/1996 and 12/2011 were derived from an intestinal rehabilitation program (IRP) patient registry. Change over time was analyzed using multivariable Box-Jenkins-method based autoregressive integrative moving average models (ARIMA), robust linear regression and non-parametric trend analysis. Four systematically introduced treatment options (IRP, serial transverse enteroplasty, omega-3 lipid emulsions, ethanol locks) were evaluated. Analyses were adjusted for patient characteristics and disease severity. The primary outcome was disease-specific mortality from liver failure and sepsis. Secondary outcomes included parenteral nutrition weaning, transplantations, catheter complications and liver disease.

**Results:** Patient characteristics remained unchanged over time, except for decreasing small bowel length (-0.5%/quarter; 95% Confidence Interval (95%-CI) -0.85, -0.16) and intensive care unit admission time (-0.6 days/quarter; 95%-CI -1.03, -0.18). Disease-specific mortality diminished significantly over time (-0.02 deaths/quarter; 95%-CI -0.03, -0.01) by IRP and omega-3 lipids introduction (-0.6 deaths/quarter each, 95%-CI -1.23, -0.02 and -0.77, -0.45). Serial transverse enteroplasty and ethanol locks did not impact significantly. Parenteral nutrition weaning and transplantations remained unchanged, while catheter sepsis and complication rates decreased by 0.3 episodes/1000 catheter-days each (95%-CI -0.43, -0.2 and -0.45, -0.24).

**Conclusions:** Introduction of IRP and omega-3 lipids independently decreased disease-specific mortality. For the first time, time series analysis was applied to evaluate effectiveness of treatment options in intestinal rehabilitation.

4.4 Introduction

Pediatric intestinal failure (IF) is a devastating condition caused by the inability to absorb sufficient nutrients to grow and survive, either due to anatomical or functional loss of
The course and outcome has changed in the last 15 years, however, with the introduction of multidisciplinary intestinal rehabilitation programs (IRP), and new medical and surgical treatment options\textsuperscript{49,569}. Multidisciplinary IRPs have been shown to improve clinical outcomes by integrating expert management, improving communication, and enhancing continuity and comprehensiveness of care\textsuperscript{44}.

As previously reported, the early effects of IRP included significantly reduced mortality from liver failure, septic episodes, critical care unit admission length and an increased rate of transplantation\textsuperscript{49}. A systematic literature review reported in 2012 that with increased experience and maturation of the IRP, novel medical and surgical treatment options were implemented, that further improved the clinical outcomes of pediatric IF-patients\textsuperscript{44}. Due to small sample sizes of IF-studies, it was difficult to compare adjusted effectiveness of individual treatment options and no conclusions on causal inference could be made\textsuperscript{44}. Time series analysis represents a valuable methodological option that permits longitudinal analysis of change over time thereby overcoming previously encountered methodological limitations related to small samples. It includes time as a factor in the analysis, accounts for delay of exposure and effect, and allows confounder adjustment in multivariate models. Although this statistical method has been derived from Economics, it has been shown to be valid and useful in the medical setting\textsuperscript{442,443}.

The objective of the study was to evaluate changes in clinical outcome of pediatric IF over time, and to identify and quantify the independent effect of newly introduced treatment options using novel methodology to overcome limitations experienced in the literature.

### 4.5 Methods

#### 4.5.1 Study Design

A retrospective analysis of prospectively collected, observational data from a single study site was performed. Ethics board approval was obtained from The Hospital For Sick Children and the University of Toronto, in agreement with which no consents were obtained.
4.5.2 Study Sample

Study subjects were drawn from an institutional IRP-database. The criteria for referral of new patients to the IRP are primarily functional with dependence on parenteral nutrition (PN) for greater than 42 days after bowel resection or injury\textsuperscript{3}. For patients with less than 25\% of expected age-specific small bowel length, referral is immediate. Patients suffering from primary dysmotility or mucosal enteropathy were excluded from this study. Patients with abdominal wall defects were permitted.

Data from patients diagnosed with IF between 0 and 365 days of age were derived from a prospective IRP-patient registry and included in the analysis. Data from patients managed at our institution prior to IRP-inception were identified using an intensive care unit patient database and added to the IRP-registry retrospectively. The final database records extended from July 1996 to December 2011. Patients were categorized to quarterly time periods based on their date of birth. A fraction of these patients have been described in earlier studies\textsuperscript{49,65,569}.

4.5.3 Data Collection

A single, experienced data coordinator performed data extraction and input. Any time interval was calculated using statistical software (Stata 11.2, Microsoft Access) and dates were extracted from patient records. Time 0 was defined as the date of primary surgery. Follow-up was continued until patients were discharged from IRP-care, died or to study end (February 10, 2013). Missing observations were missing at random and imputed based on time and whether treatment options were available using multivariate regression models (1997/1, 1998/3, 2001/1, 2001/3-2002/3; representing 6.75\% of the dataset, single and multiple imputation using 50 imputations, based on poisson and linear regression)\textsuperscript{514,518,519,521}. All variables were missing similarly frequent; no data was missing on treatment availability.

Data on patient characteristics, anatomy and etiology, surgical interventions, medical treatment, and outcomes related to nutrition, sepsis, IF-associated liver disease, transplantation and mortality were measured and analyzed. The primary outcome was disease-specific mortality from liver failure and sepsis. Secondary outcomes included
parenteral nutrition weaning, transplantations, catheter complications and liver disease. Age-adjusted expected small bowel length was calculated using normal values published by Struijs et al.\textsuperscript{250}. Intestinal failure secondary to necrotizing enterocolitis, perforation or abdominal wall defect was regarded as inflammatory etiology. Clinical outcomes were defined as in previous studies by our group: cholestasis as conjugated bilirubin levels of \( >50\mu\text{mol} / \text{L} \) for at least 2 weeks without association with sepsis; advanced liver disease as conjugated bilirubin of \( >100\mu\text{mol} / \text{L} \) for at least 2 weeks and not associated with sepsis; catheter-related bloodstream infections (CRBSI) as a positive blood culture from the central venous catheter (CVC) not related to another source of bacteremia and reported per 1000 catheter-days; complete PN-weaning as independence from PN for \( \geq 2 \) consecutive weeks with adaptive weight gain and, if PN had to be restarted during this period, the total PN-duration was calculated adding PN-dependent episodes together\textsuperscript{49,569}. Length of hospitalization and admission to the intensive care unit (ICU) refers to the primary hospitalization only. The median per time interval was calculated for continuous variables, frequencies per time interval were analyzed for binary outcomes.

4.5.4 Data Analysis

Data were analyzed using Stata 11.2 software\textsuperscript{403}. Change over time of patient characteristics was evaluated using robust univariate linear regression models and non-parametric trend tests (Cuzick test)\textsuperscript{570}.

Change over time of clinical outcomes was analyzed using multivariable Box-Jenkins-method based autoregressive integrative moving average models (ARIMA), which allow accounting for delay of exposure and effect, trend changes, cycles and seasonal changes. As recent observations are regarded as more accurate, the time series was smoothed using exponentially weighted moving average models, which weigh recent observations more heavily and sensitive to systematic exposure regimen changes. Smoothing parameters (\( \alpha \)) were selected based on the smallest sum-of-squared forecast errors and the means of the series at time 0 (\( \gamma_0 \)) were regarded as the mean of the first half of the sample. Other \( \alpha \) and \( \gamma_0 \) for exponential smoothing were tested and double-exponential smoothing performed, but not reported, as it did not provide superior information. Guided by data distribution, secondary analyses included non-parametric trend tests (Cuzick-test) and robust univariate linear
regression models. Robust univariate linear regression coefficients, which indicate the linear change over time, and $p$-values are reported, while non-parametric trend test results were used for validation purposes and therefore not reported, unless significantly different. The linear change over time of clinical outcomes is depicted in supplemental figures *(Appendix 4-A, Appendix 4-B: A/B, Appendix 4-C: A-D)*.

Multivariable ARIMA-models were defined based on clinical reasoning. Four treatment options that were introduced systematically over time and the three most relevant patient characteristics were included in the final multivariable model. Assessed treatment options consisted of availability of IRP (since 2002/4), serial transverse enteroplasty (offered to patients plateauing in their PN-weaning process or showing complications from chronic PN since 2003/2), omega-3 lipid emulsions (administered to patients with advanced liver dysfunction since 2006/2) and ethanol locks (applying in patients with frequent CRBSI since 2009/2). The primary outcome of interest was disease-specific mortality (i.e. mortality from liver disease and sepsis), while secondary outcomes of interest included advanced liver disease, complete PN-weaning, CRBSI-rate, CVC-complication rate, number of transplant assessments and frequency of transplantations.

Pre-estimation of univariate ARIMA-models was based on stationarity (Augmented and Generalized Least Squares Dickey-Fuller tests, Phillips and Perron test) and autocorrelation (AC) and partial-autocorrelation (PAC) functions. Goodness-of-fit was assessed using Akaike (AIC) and Bayesian Information Criterion (BIC), and the Box-Jenkins principle of parsimony. Each listed co-variable was added to the model in a systematic way and their effect on the model recorded. The same steps were followed for a robust linear regression model and the results were compared to the ARIMA-values for validation purposes. The final ARIMA-models, which include all four treatment options and three patient characteristics, were then tested for goodness-of-fit, residual autocorrelation (Barlett’s white noise test, Portmanteau Q test), residual normality (visual assessment) and heteroskedasticity (visual assessment). If post-estimation revealed borderline results, the final ARIMA-model conclusions were verified using robust standard error estimation.
4.6 Results

4.6.1 Patient Characteristics

One-hundred-and-ninety-six patients with neonatal/infantile IF born between July 1, 1996, and December 31, 2011, were identified. The patient characteristics and their change over time are outlined in Table 4-1.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Median (IQR)</th>
<th>Quarterly change*</th>
<th>95%-CI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients per quarter</td>
<td>4 (2-5)</td>
<td>-0.02</td>
<td>-0.05, 0.01</td>
</tr>
<tr>
<td>Male patients per quarter</td>
<td>2 (1-2)</td>
<td>-0.01</td>
<td>-0.03, 0.01</td>
</tr>
<tr>
<td>Gestational age (weeks)</td>
<td>32.5 (30.5-35)</td>
<td>0.02</td>
<td>-0.03, 0.07</td>
</tr>
<tr>
<td>Birth weight (grams)</td>
<td>1846 (1315-2305)</td>
<td>5</td>
<td>-5, 15</td>
</tr>
<tr>
<td><strong>Expected small bowel length (%)</strong></td>
<td><strong>64 (53.5-89)</strong></td>
<td><strong>-0.5</strong></td>
<td><strong>-0.85, -0.16</strong></td>
</tr>
<tr>
<td>Ileocecal valve resections per quarter</td>
<td>1 (0-2)</td>
<td>0.001</td>
<td>-0.01, 0.02</td>
</tr>
<tr>
<td>Inflammatory etiologies per quarter #</td>
<td>2 (1-4)</td>
<td>-0.01</td>
<td>-0.03, 0.02</td>
</tr>
<tr>
<td>Gastrostomies per quarter</td>
<td>1 (1-2)</td>
<td>0.01</td>
<td>-0.004, 0.03</td>
</tr>
<tr>
<td>Home-PN per quarter</td>
<td>1 (0-1)</td>
<td>0.004</td>
<td>-0.01, 0.02</td>
</tr>
<tr>
<td>Hospitalization (days)</td>
<td>142 (104-181)</td>
<td>0.4</td>
<td>-0.35, 1.23</td>
</tr>
<tr>
<td>Admission at ICU (days)</td>
<td><strong>86.75 (59-99)</strong></td>
<td><strong>-0.6</strong></td>
<td><strong>-1.03, -0.18</strong></td>
</tr>
</tbody>
</table>

Legend: ICU – intensive care unit; Home-PN – Patients on home parenteral nutrition; IQR – interquartile range (25th – 75th percentile); 95%-CI – 95% confidence interval; **bold** – statistically significant; * – derived from univariate robust linear regression; # – inflammatory etiologies (necrotizing enterocolitis, perforation, abdominal wall defect)

The quarterly number of patients of 4 remained stable over time (inter-quartile range (IQR) 2-5; -0.02 patients/quarter; 95% confidence interval (95%-CI) -0.05, 0.01) and patient characteristics did not change significantly. In median 2 patients were males (IQR 1-2; -0.01; 95%-CI -0.03, 0.01). Gestational age was 32.5 weeks (IQR 30.5-35; 0.02; 95%-CI -0.03, 0.07) and birth weight 1846 grams (IQR 1315-2305; 5; 95%-CI -5, 15). Over a time period of 15 years, expected small bowel length decreased significantly by 0.5% per quarter (64.09%
in median, IQR 53.5-89; 95%-CI -0.85, -0.16). Each quarter, one patient had an ileocecal valve resection (IQR 0-2; 0.001; 95%-CI -0.01, 0.02) and two cases were secondary to an inflammatory pathology (IQR 1-4; -0.01; 95%-CI -0.03, 0.02). One patient each had a gastrostomy (IQR 1-2; 0.004; 95%-CI -0.004, 0.03) or was on home-PN (IQR 0-1; 0.004; 95%-CI -0.01, 0.02). Overall admission length did not change significantly over time (142 days in median; IQR 104-181; 0.4; 95%-CI 0.35, 1.23), while ICU admission length decreased by 0.6 days (86.75 days in median; IQR 59-99; 95%-CI 1.03, -0.18).

4.6.2 Clinical Outcomes

Change over time in clinical outcomes is summarized in Table 4-2.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Median (IQR)</th>
<th>Quarterly change*</th>
<th>95%-CI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease-specific mortality per quarter</td>
<td>0 (0-1)</td>
<td>-0.02</td>
<td>-0.03, -0.01</td>
</tr>
<tr>
<td>PN-weaning per quarter</td>
<td>2 (1-3)</td>
<td>-0.01</td>
<td>-0.03, 0.01</td>
</tr>
<tr>
<td>Advanced liver disease per quarter</td>
<td>1 (1-2)</td>
<td>-0.004</td>
<td>-0.02, 0.01</td>
</tr>
<tr>
<td>CRBSI-rate per 1000 catheter-days</td>
<td>17 (9.5-22.8)</td>
<td>-0.3</td>
<td>-0.43, -0.2</td>
</tr>
<tr>
<td>CVC-complication rate per 1000 catheter-day</td>
<td>16 (8.2-21.9)</td>
<td>-0.3</td>
<td>-0.45, -0.24</td>
</tr>
<tr>
<td>Transplant assessments per quarter</td>
<td>1 (0-2)</td>
<td>0.01</td>
<td>-0.01, 0.02</td>
</tr>
<tr>
<td>Transplantations per quarter</td>
<td>0 (0-0)</td>
<td>0.004</td>
<td>-0.003, 0.01*</td>
</tr>
</tbody>
</table>

Legend: STEP – serial-transverse enteroplasty; 95%-CI – 95% confidence interval; IQR – interquartile range (25th – 75th percentile); PN – parenteral nutrition; CRBSI – catheter-related bloodstream infection; CVC – central venous catheter; bold – statistically significant; * – derived from univariate robust linear regression; # – derived from parametric linear regression, as all weights went to zero in robust regression.

Mortality from liver disease or sepsis decreased significantly over time (-0.02 deaths/quarter; 95%-CI -0.03, -0.01) (Appendix 4-A). The median quarterly disease-specific mortality was 0 (IQR 0-1). Of the four treatment options tested, IRP-presence and omega-3 lipid availability independently and significantly affected disease-specific mortality (Figure 4-1).
Multidisciplinary rehabilitation and omega-3 lipids each reduced mortality by 0.6 deaths per quarter (95%-CI -1.23, -0.02; 95%-CI -0.77, -0.45). The only patient characteristic having an independent and significant impact on disease-specific mortality was inflammatory etiology, which increased it by 0.15 deaths per quarter (95%-CI 0.07, 0.22).

A median of 2 patients per quarter were weaned completely from PN, a rate which did not change significantly over time (-0.01 patient/quarter; 95%-CI -0.03, 0.01) (Appendix 4-B: A). In the multivariable ARIMA-model, omega-3 availability decreased PN-weaning by 0.7 per quarter (95-CI -1.29, -0.19), while other treatment options did not have a significant effect (Figure 4-2: A). Small bowel length and inflammatory etiology were positively correlated with PN-weaning. Each percentage of additional small bowel remaining was
associated with 0.02 patients per quarter who were weaned from PN (95%-CI 0.01, 0.03),
while inflammatory etiology enhanced this process causing 0.6 patients per quarter to become
fully adapted (95%-CI 0.41, 0.7).

Advanced liver disease did not change significantly over time (-0.004 patient/quarter; 95%-CI
-0.02, 0.01) (Appendix 4-B: B). A median of 1 patient per quarter developed advanced
liver disease (IQR 1-2). While none of the tested treatment options had a significant impact
on advanced liver disease development, inflammatory etiology increased the frequency of
advanced liver disease by 0.3 patients per quarter (95%-CI 0.14, 0.48) (Figure 4-2: B).
A) Complete weaning from parenteral nutrition remained unchanged over the 15 year time period.

B) Increased awareness led to an increase of advanced liver disease shortly after IRP-introduction. With IRP-maturation progression into advanced liver disease has been increasingly prevented.
Catheter-related bloodstream infections decreased over time (Appendix 4-C: A). The median CRBSI-rate of 17/1000 catheter-days per quarter decreased by 0.3/1000 catheter-days (95%-CI -0.43, -0.2) (Figure 4-3: A). Small bowel length was the only factor contributing independently (0.13/1000 catheter-days/quarter; 95%-CI 0.01, 0.25). Complications of CVCs such as breakage, occlusion or leakage per 1000 catheter-days decreased significantly over the 15-year time period (Appendix 4-C: B). The median CVC-complication rate was 16/1000 catheter-days, which decreased by 0.3/1000 catheter-days per quarter (95%-CI -0.45, -0.24). None of the treatment options were observed to independently cause this improvement (Figure 4-3: B). Residual small bowel length increased this rate by 0.13 complications/1000 catheter-days per quarter (95%-CI 0.05, 0.22).

Transplant assessment and transplantations performed did not change linearly over time (Appendix 4-3: C/D). The transplant team assessed a median of 1 patient per quarter, which led to 0 transplantations in average. Transplant assessments remained stable (0.01/quarter; 95%-CI -.01, 0.02), as did performed transplantations (0.004/quarter; 95%-CI -0.003, 0.01). Availability of ethanol locks diminished the need for transplant assessments by 1.6 patients/quarter (95%-CI -2.6, -0.57), while inflammatory etiology was associated with more frequent assessments (0.23 assessments/quarter; 95%-CI 0.04, 0.42) (Figure 4-3: C). None of the treatment options or patient characteristics significantly affected the number of transplantations performed (Figure 4-3: D).
A/B) Central venous catheter complications and CRBSI decreased by 0.3 episodes/1000 catheter-days per quarter.

C/D) Transplant assessments and transplantations followed a similar course as advanced liver disease. With increased awareness of liver disease and improved communication with the transplant team during the first four years of IRP, more patients were assessed and transplanted during this time period. Maturation of IRP led to prevention of advanced liver disease making transplant assessment and transplantation less necessary.

4.7 Discussion

Time series analysis was used to determine the role of individual treatment options in the changing paradigm of pediatric IF. While small sample sizes and distributional issues limited prior studies, this approach permits analysis of change over time using time intervals instead of comparing categories with each other. Robust linear regression models were used to determine the linear trend of clinical change over the 15-year time period, while smoothing
was applied to filter out the noise component in the series. Exponential weighted moving average smoothing infers individual observations (e.g. disease-specific mortality for 2000/2) from the entire available history and weights recent observations more heavily as they are considered to be more accurate\(^5\). Autoregressive integrative moving average models adjust for local trends, cycles, seasonality and random noise of a time series and therefore are more specific and sensitive to measure non-linear change over time, similar to the one observed in the change of outcome of pediatric IF, when compared to linear regression models. With 62 observations, this time series may be regarded as short when compared to time series used in Economics, the original field of ARIMA-models, where they are being used to analyze past market changes and predict future market behavior. However, a time series over 15 years is long in a medical setting particularly in pediatric IF\(^4\).\(^4\)\(^3\). Missing values are ubiquitous in quantitative research and biased analyses are expected if >10%-20% of observations are missing\(^5\).\(^1\)\(^3\),\(^1\)\(^5\). Simple, ad-hoc methods to deal with missing data such as list-wise or pairwise deletion, complete case analysis, mean substitution or single imputation have been deemed suboptimal as each may introduce serious bias\(^5\).\(^1\)\(^4\),\(^5\).\(^1\)\(^5\),\(^5\)\(^1\)\(^8\),\(^5\)\(^1\)\(^9\),\(^5\)\(^2\)\(^1\). The missing data in this study (6.75%) were therefore addressed using single and multiple imputation methods, as technically feasible. Multiple imputation has been validated and applied widely to replace missing values by using available information and preserving relationships in the dataset\(^5\).\(^1\)\(^4\),\(^5\).\(^1\)\(^5\),\(^5\)\(^1\)\(^8\),\(^5\)\(^1\)\(^9\),\(^5\)\(^2\)\(^1\). While the imputed data represent estimates of most probable values given the available information, the limitation of single imputation is that it ignores uncertainty and may therefore underestimate variance. As the complete dataset fulfilled the assumptions required for ARIMA-modeling, the models were stable and the results are in agreement with clinical perception of how pediatric IF changed over time. The application of this methodology is considered a valid and novel approach to evaluate change of pediatric IF and overcome limitations of previous studies, exploring the relationship between treatment options and clinical outcomes by accounting for timing and confounders\(^4\).\(^4\),\(^4\)\(^9\),\(^5\)\(^6\). It represents a pragmatic approach demonstrating effectiveness rather than efficacy, which typically is derived from a more artificial setting of randomized, controlled trials. Confounding factors were adjusted for using multivariable ARIMA-models as performed in our study and the results may be generalizable to any other highly specialized tertiary healthcare center in a developed country. This study overcomes numerous limitations of the current literature on
pediatric IF, as it is the first to perform multivariable analysis and generating pragmatic evidence on treatment effectiveness\textsuperscript{44,49,165,169,569}.

Earlier studies and clinical experience showed that disease-specific mortality has decreased dramatically\textsuperscript{44,49,569}. It was however unclear which treatment options (if any) caused this improvement. While it has been anticipated that availability of omega-3 lipid emulsions improved disease-specific mortality by preventing liver disease progression, the independent effect of IRP was less clear. These results show that dedication of a highly specialized team of experts to this rare pediatric condition has just as a strong impact on mortality prevention as have novel lipid strategies. While STEP and ethanol locks have been shown to support the PN-weaning process and decrease CRBSI-rates in case series and univariate analyses, the effect on disease-specific mortality has not been assessed previously\textsuperscript{101,381}. In this study, both treatment options were provided to a small proportion of patients who either suffer from a plateaued adaptation process or complications from prolonged PN, or present with frequent CRBSI. The fact that these patients represent a sicker subpopulation and the small number of patients who actually benefit from these treatment options, as well as an indirect effect on this outcome, may explain why their effect did not achieve statistical significance.

Over time patients were admitted to the ICU for a shorter time, despite shorter small bowel length, which is regarded as a disease severity marker. This change has been perceived clinically, but prior studies have not been able to measure this change due to methodological limitations\textsuperscript{569}. This study showed that despite more severe illness and reduced mortality, complete weaning of PN has remained unchanged over the 15-year time period (Figure 4-2: A). While prior to IRP-inception PN-weaning was limited by high mortality, patients of more recent years do not achieve complete PN-weaning because of increased severity of disease, shorter follow-up time and improved survival, which is facilitated by omega-3 lipids. This explains why omega-3 lipids were found to diminish complete PN-weaning in the multivariable analysis. With improved mortality, a larger proportion of patients survive with shorter small bowel over a longer period of time, forming a cohort of chronically ill patients with more complex anatomy and a longer adaptation process. This is in agreement with earlier results\textsuperscript{569}. 
Advanced liver disease seems to have increased with IRP-inception, but this is understood to be secondary to a higher awareness of the IRP compared to earlier IF-management. A similar course of an initial increase with IRP-inception and secondary decrease after 4 years of IRP was observed for transplant assessment and transplantations (Figure 4-2: B, Figure 4-3: C/D). This indicates successful prevention of complications from long-term PN, after an initial phase of increased awareness, earlier referral to the transplant team and more frequent transplantation. Interestingly, none of the four tested treatment options were found to significantly contribute to this effect. It is important to understand that omega-3 lipid emulsions are not expected to influence this improvement at our institution as they are only administered to patients with advanced liver disease (conjugated bilirubin >100µmol/L) with the goal to prevent further progression to full liver failure.

Central venous catheter complication and CRBSI-rates decreased significantly, and seems to be positively associated with short bowel length. This may be explained by the fact that patients with very short bowel may have died early and never had the opportunity to develop a CVC-complication or CRBSI. Further studies are required to clarify this association as well as evaluate how clinical outcome change of pediatric IF affects quality of life and cost-effectiveness.

Whether these results represent a simple association between treatment options and improved outcomes, or a true causal relationship is debatable. When following the Bradford Hill criteria to assess causal inference, the results may be regarded as reflecting causality for disease-specific mortality: first, the effect size was large and achieved statistical significance despite the small sample size; second, a decrease in mortality from liver disease and/or sepsis was observed by several different clinical and research groups; third, the study cohort was specific despite not being in an experimental setting, meaning that the cohort only included clearly defined participants with one medical condition managed all at the same medical center; fourth, there is a time-relationship between the exposure and the effect in the study design as well as the statistical analysis applied; fifth, a biological gradient between exposure and effect was not observed due to the nature of the tested exposures, however, a gradient response was observed for the maturation of IRP, that is, the more mature the IRP, the more pronounced the decrease in disease-specific mortality; sixth, the cause-effect relationship between the tested treatment options and disease-specific mortality is plausible; seventh, the
findings are believed to be coherent in different settings, although formal laboratory results are not available for any of the four tested treatment options due to limited applicability (e.g. the effect of STEP availability on disease-specific mortality is not assessable in a laboratory setting); eighth, availability of experimental evidence is limited due to methodological difficulties of a rare, complex, surgical pediatric condition; and finally, similar factors result in similar results, such as for example, novel lipid strategies using omega-3 lipids and such minimizing lipid dose exposure in PN result in improved liver function and ultimately in decreased disease-specific mortality. The associations found between the four treatment options and primary outcome of interest using time series analysis in this study may therefore likely represent a causal relationship based on these criteria.

In conclusion, this study showed significant change in outcome of pediatric IF and revealed IRP-introduction and omega-3 lipid emulsions to be the strongest contributors to the decreased disease-specific mortality. This is the first study to use time series analysis for IRP-evaluation and quantifies the change in clinical outcomes with IRP-maturation. This novel methodology allowed multivariable analysis of change over time adjusting for time lag of cause and effect, and quantification of independent effects of treatment options. This methodology is a valuable tool and should be used more frequently in clinical research.
5 Health-Related Quality of Life in Neonates and Infants: A Conceptual Framework

This chapter is formatted for publication in *Value in Health*.

5.1 Preamble

This study was initiated after being unable to conduct a cost-effectiveness analysis of multidisciplinary intestinal rehabilitation secondary to unavailability of valid methodology to quantify neonatal and infant HRQOL and utilities. A systematic literature review on HRQOL measurement in neonates and infants has been performed, which characterized the significant lack of HRQOL and utility methodology for neonatal and infant population in more detail, and provided the rationale for this study.\(^{572}\).

5.2 Cover Page

HEALTH-RELATED QUALITY OF LIFE IN NEONATES AND INFANTS: A CONCEPTUAL FRAMEWORK

Short title: Neonatal and Infant Quality of Life

Carol Oliveira\(^{1,2,3,4}\), Nicole T. de Silva\(^{1,2}\), Wendy J. Ungar\(^{3,4}\), Ahmed M. Bayoumi\(^{4,5}\), Yaron Avitzur\(^{1,6}\), Jeffrey S. Hoch\(^{4,7}\), Julia Maxwell\(^{1}\), Paul W. Wales\(^{1,2,3,4}\)
5.2.1 Affiliations

1Group for Improvement of Intestinal Function and Treatment Program (GIFT), The Hospital for Sick Children, Toronto, Canada

2Division of General and Thoracic Surgery, The Hospital for Sick Children, Toronto, Canada

3Child Health Evaluative Sciences (CHES) Program, The Hospital for Sick Children, Toronto, Canada

4Institute of Health Policy, Management and Evaluation (IHPME), University of Toronto, Toronto, Canada

5Centre for Research on Inner City Health, Li Ka Shing Knowledge Institute and Division of General Internal Medicine, St. Michael’s Hospital, Toronto, Canada

6Division of Gastroenterology, Hepatology and Nutrition, The Hospital for Sick Children, Toronto, Canada

7Centre for Excellence in Economic Analysis Research (CLEAR), St. Michael’s Hospital, Toronto, Canada

5.2.2 Corresponding author

Paul W. Wales, BSc, MD, MSc, FRCSC, FACS
Neonatal and Pediatric Surgeon
Director, The Group for Improvement of Intestinal Function and Treatment (GIFT)
Division of General and Thoracic Surgery
Associate Scientist, Child Health Evaluative Sciences, Research Institute
The Hospital for Sick Children
Associate Professor, Department of Surgery, University of Toronto
Room 1526, 555 University Ave
Toronto, ON M5G 1X8, Canada
Tel: +1 416 813 7654 x201490
Fax: +1 416 813 7477
E-mail address: paul.wales@sickkids.ca
5.2.3   Key Words

Infants, newborn; quality of life; qualitative research; concept formation

5.2.4   Funding Source

The Swiss National Science Foundation, The Walter and Gertrud Siegenthaler Foundation and The University of Zurich, Switzerland; and The Hospital For Sick Children, Toronto, Canada, have financially supported the first author with research scholarships during conduction of this research.

5.2.5   Financial Disclosure

All authors have no financial relationship relevant to this article to disclose.

5.2.6   Key Points

What’s known?

• The meaning of health-related quality of life in infants and neonates is largely undefined, leading to a lack of valid methodology for quality of life assessment in this population.

• With recent improvement of clinical outcomes and reduced mortality of many neonatal conditions, health-related quality of life measurement becomes imperative.

What does this paper add?

• Our study generated a conceptual framework of infant and neonatal quality of life based on focus groups and interviews with caregivers of chronically ill children and health care professionals.

• Infant and neonatal health-related quality of life is a definable entity, but strongly interrelated with the caregiving environment and dependent on the developmental stage of the infant.
• Health-related quality of life measurement instruments for infants, mature and premature neonates accounting for fast-changing developmental stages during infancy are warranted.

**How does it affect health care-related decision-making?**

• Understanding the meaning and determinants of HRQOL during the neonatal period and infancy allows optimization of healthcare provision and patient wellbeing by appreciating the dynamics between the neonate/infant, family and healthcare provider.

• It may also help guide end-of-life decisions and fair resource distribution.

### 5.3 Abstract

**Objective:** With reduced mortality of neonatal conditions, health-related quality of life (HRQOL) measurement has become an important clinical outcome. The meaning of HRQOL for this age-group remains largely undefined. We aimed to generate a conceptual framework of neonatal and infant HRQOL.

**Study Design and Setting:** Using qualitative methods, a concept was developed based on in-depth analysis of verbatim records of two focus groups (6 caregivers, 6 healthcare providers) and five interviews with caregivers of chronically ill neonates/infants (n=2), and healthcare professionals of a pediatric tertiary healthcare center (n=3). Two analysts independently performed thematic analysis using an inductive and contextual approach.

**Results:** Neonatal and infant HRQOL is a definable entity, which is strongly influenced by caregivers and society. It may be gauged by the perceived degree of effort required to achieve expected normalcy in everyday life for the neonate/infant and its family. The importance of individual HRQOL-factors is developmental-stage dependent.

**Conclusion:** Neonatal and infant HRQOL is a multidimensional, multilayered and interconnected concept, where the child’s needs contribute most directly, and the caregiver’s and society’s ability to meet those needs characterize the interdependence between the child
and its caregiving environment. Developmental-stage specific HRQOL-instruments for premature and mature neonates, and infants are warranted.

### 5.4 Introduction

Quality of life (QOL) is a wide-ranging multidimensional concept that consists of a subjective perception of wellbeing including all aspects of life. Health-related QOL (HRQOL) encompasses all aspects of QOL that interact with health, which as been defined by the World Health Organization (WHO) as a state of complete physical, mental and social wellbeing and not merely the absence of disease. Health-related QOL has become an increasingly important outcome in medicine and multiple measurement tools have been developed for adults and children. Valid methodology for neonates and infants is extremely scarce, despite HRQOL having become an increasingly important clinical outcome, as mortality rates of neonatal and infant conditions, such as prematurity, intestinal failure (IF) or critical congenital heart defects, have decreased significantly. This deficiency has led to HRQOL-measurement in neonates and infants being performed infrequently or using instruments neither developed for, nor validated in this population. Another common strategy is to measure HRQOL of persistent neonatal and infant conditions at a time point when the child is older and for whom a validated instrument is available. Either approach is suboptimal, however, as measuring HRQOL with an instrument developed for an older population puts the validity of this instrument under question and assumes that the underlying concept of HRQOL is not age-dependent. Measuring HRQOL at a later age may be of clinical interest and methodologically correct, but fails to measure the actual HRQOL of neonates and infants.

It is crucial to understand the meaning and determinants of HRQOL during the neonatal period and infancy in order to optimize healthcare, improve patient wellbeing, and guide end-of-life decisions. Without valid neonatal and infant HRQOL (NIHRQOL) definition and measurement, clinical decisions cannot be based on the same standards as pertaining to older populations. Additionally, HRQOL has become an important factor in policy-making and resource distribution. Lack of valid methodology may leave the neonatal and infant population with a disadvantage, as resource allocation is hard to justify without high-quality evidence of health benefits.
For these reasons, it is imperative to develop valid methods to measure NIHRQOL. It is the objective of this study to generate a conceptual framework of NIHRQOL.

5.5 Materials & Methods

5.5.1 Study Design & Setting

A qualitative study using two focus groups and five interviews with caregivers (i.e. parents or parental figures) of chronically ill neonates and infants, and healthcare professionals (HCP) at a Canadian pediatric tertiary healthcare center was conducted in 2012. Neonates were defined as children aged 0-28 days, while infants were regarded as children of 1-12 months of age. The study was conducted in English, as all participants were fluent despite diverse ethnic and cultural backgrounds, without this being a preset inclusion criterion. Generalizability may be limited to a multiethnic society of a developed country. Ethics board approval was obtained from the healthcare center, as well as the University of Toronto, along with individual consents for participation and audio recording.

5.5.2 Study Sample

Eight families with children who were diagnosed with severe chronic illness during their first year of life, and 14 HCPs with professional experience with neonatal/infant chronic disease were invited to participate. The goal was to include participants with variable characteristics in terms of their profession, family situation, age, gender, culture, experience and outcome. Two focus groups with 6 caregivers and 6 HCPs, respectively, were conducted. Participants willing to participate, but unavailable on the focus group date, were offered an individual interview. Five interviews with 2 caregivers and 3 HCP were held. One family declined to participate. For feasibility reasons participants were recruited from the institutional multidisciplinary intestinal rehabilitation program, the Group for Improvement of Intestinal Function and Treatment. Patients of this program are diagnosed with intestinal failure early in their life, have excessive morbidity and a substantial risk for mortality over several years, representing a neonatal/infant population with a chronic disease associated with high morbidity and recent mortality improvement\textsuperscript{49,569,587}. This sample was selected based on Prospect Theory, which states that losses are perceived more severely than gains\textsuperscript{602}. In this
sense, we concluded that participants had less difficulty identifying NIHRQOL-factors, which at some point in their life were disrupted.

5.5.3 Data Generation

The focus groups were assembled to address the research question, “What are important factors that contribute to NIHRQOL?” All focus groups were moderated by an experienced social worker (JM) who was known to the majority of the participants and the research fellow (CO) guided the interviews following predefined questions without pre-specifying a generic HRQOL-concept (Appendix 5-A). All participants were encouraged to name factors relevant to NIHRQOL drawing from their experience with affected neonates/infants. Participant characteristics were recorded. After achieving saturation, factors pertinent to HRQOL of older children, which were derived from a systematic review on pediatric HRQOL-instruments, were offered for further discussion. Subsequently the participants were asked to create a diagram depicting their conceptual understanding of NIHRQOL. Most relevant factors were listed and each participant weighted (high, medium, low significance) them for neonates and infants.

The focus groups were of 2 hours duration with a 10-minute break, while the interviews were 45 minutes in length. Financial compensation for out-of-pocket expenses was offered. All discussions were audio recorded and transcribed by a transcriptionist.

5.5.4 Analysis

Participant characteristics were described using non-parametric summary statistics. Thematic analysis was used to qualitatively analyze the discussions. Two analysts (CO, NTdS) performed the analysis independently following the guidelines published by Braun and Clarke. The two analysts discussed each analysis step to achieve consensus. Persistent disagreement was resolved by the senior author (PWW). An inductive (bottom-up) approach was used to elucidate the meaning of NIHRQOL. The analysis was limited to content related to the detailed account of NIHRQOL. Latent themes and subthemes were identified using a constructionist approach. Themes were viewed as factors contributing to NIHQOL (e.g. basic
needs), while subthemes consisted of perceptions on organization, weighting, interactions, and dynamics of factors.

The following terminology was used (*Appendix 5-B: A*): the term ‘factor’ was used for any kind of entity having an impact on NIHRQOL, without giving it a certain weight or relationship. Factors were organized using a circular diagram with hierarchical order: levels → domains → attributes → items. The term ‘level’ was chosen to describe the closeness of a factor to the neonate’s/infant’s HRQOL. Each level was divided into ‘domains’, which consist of ‘attributes’ that describe smaller categories of factors inside domain. Finally, each attribute embodies individual ‘items’. Items are not reported in detail, as it was not the objective of this study to produce a complete list of items relevant to NIHRQOL, but to give an overview over the conceptual understanding. Direct quotes from focus groups or interviews are marked with closed quotation marks, or referenced as \(Q_x\) and listed in *Appendix 5-C*.

### 5.6 Results

#### 5.6.1 Participants

The interest in NIHRQOL was high. Seven of eight families contacted were willing to participate and nine of fourteen HCPs were able to attend focus groups or interviews. Reasons for not participating were psychological overload for one family and conflicting schedules for HCPs. The refusing family did not differ significantly from the participating families in terms of infant severity of illness or health outcome.

Participant characteristics are listed in *Table 5-1*. 
Table 5-1: Participant Characteristics

<table>
<thead>
<tr>
<th>Group</th>
<th>Size</th>
<th>Age (IQR)</th>
<th>Gender</th>
<th>Children (IQR)</th>
<th>Personal experience*</th>
<th>Clinical outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCP – FG</td>
<td>6</td>
<td>44.5 years (35–52)</td>
<td>5 female, 1 male</td>
<td>0.5 (0-2)</td>
<td>2/6</td>
<td>NA</td>
</tr>
<tr>
<td>HCP – I</td>
<td>3</td>
<td>48 years (47-55)</td>
<td>1 female, 2 male</td>
<td>3 (2-3)</td>
<td>1/3</td>
<td>NA</td>
</tr>
<tr>
<td>CAR – FG</td>
<td>6</td>
<td>36 years (33-38)</td>
<td>5 female, 1 male</td>
<td>1.5 (1-2)</td>
<td>6/6</td>
<td>Diverse #</td>
</tr>
<tr>
<td>CAR – I</td>
<td>2</td>
<td>30 years (24-36)</td>
<td>2 female</td>
<td>1.5 (1-2)</td>
<td>2/2</td>
<td>PN-dependent, outpatient</td>
</tr>
</tbody>
</table>

Legend: HCP = healthcare professionals (2 surgeons, 1 neonatologist, 1 gastroenterologist, 1 intestinal rehabilitation nurses, 1 dietician, 1 occupational therapist/physical therapist, 1 child-life specialist); CAR = caregivers (3 stay-home-mothers, 2 students, 1 teacher, 1 sales associate, 1 educational assistant); FG = Focus Group; I = interview; PN = parenteral nutrition; * = personal experience with severe neonatal or infant illness in family; # = 1 PN-dependent and outpatient, 2 fully cured and released from care, 1 admitted, 1 transplant recipient

Healthcare professionals were older than caregivers and the male-to-female ratio was nearly 1:3 for the entire cohort. Except for three HCPs, all participants had 1-3 children. While all caregivers had personal experience with severe neonatal/infant illness, HCPs had extensive professional experience (4-25 years) and personal experience in three of nine cases. All HCPs had worked closely with the multidisciplinary intestinal rehabilitation program. All caregivers were parents of children with intestinal failure, defined as less than 25% of expected bowel length or more than 42 days of parenteral nutrition dependence, during their first year of life. At the time point of the discussions, all children were >1 year old, had diverse clinical outcomes and three had siblings.

Factors derived from a list that the participants generated at the end of the focus group/interview containing most relevant NIHRQOL-factors are depicted in Table 5-2.
Table 5-2: Most Important Factors to Participants

<table>
<thead>
<tr>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulation, entertainment, therapists, play, being touched, experiences</td>
</tr>
<tr>
<td>Social interaction, interaction with family/peers, relationships, visitors</td>
</tr>
<tr>
<td>Opportunity for development, development of new skills, learning</td>
</tr>
<tr>
<td>Pain, physical discomfort, symptoms</td>
</tr>
<tr>
<td>Oral stimulation, freedom from hunger, feeding, basic needs</td>
</tr>
<tr>
<td>Going out, seeing things, world exploration, unrestricted movement, freedom</td>
</tr>
<tr>
<td>Contact with caregiver, presence of caregiver, bonding, building trust</td>
</tr>
<tr>
<td>Environment, noise, light, circadian rhythm</td>
</tr>
<tr>
<td>Predictability, consistency, routine</td>
</tr>
<tr>
<td>Having comfort zone, feeling/being safe and secure</td>
</tr>
<tr>
<td>Emotions, psychological functioning</td>
</tr>
<tr>
<td>Family functioning, parental stress, parental interaction</td>
</tr>
<tr>
<td>Being cared for, being loved</td>
</tr>
<tr>
<td>Rest, sleep</td>
</tr>
<tr>
<td>Cognition</td>
</tr>
</tbody>
</table>

Legend: The list represents a summary of the factors listed by the participants at the end of the focus groups and interviews.

There was an understanding that the importance of individual factors was age-dependent, i.e. different for neonates and infants. No evident difference among the distinctive groups in terms of listed factors or their weighting became apparent. During the discussion, caregivers highlighted the importance of factors related to caregivers, families and society more than HCPs. Many caregivers chose to describe experienced situations, while HCPs were more theoretical in their exploration. The focus group and interviews revealed a great need of caregivers to make themselves being comprehensively understood by the HCPs. Therefore we anticipate caregivers of chronically and severely ill neonates and infants to be willing to complete lengthy questionnaires capturing all aspects of NIHRQOL.
5.6.2 Themes

The conceptual framework of NIHRQOL was derived from the focus group/interview transcripts using thematic analysis. Themes were identified and organized in levels, domains, attributes and items.

There was a basic understanding that NIHRQOL consisted of factors affecting the child directly and there were more distant factors influencing those primary factors. To respect this perception of a hierarchical order of interconnected NIHRQOL-factors, a circular diagram with four levels was generated (Figure 5-1, Appendix 5-D): Level 1 – Neonate/Infant Basic Needs and Extended Basic Needs, Level 2 – Neonate/Infant Non-Basic Needs, Level 3 – Caregivers & Family, Level 4 – Society & Community.
Fourteen domains (14 fields in circles) of NIHRQOL were categorized into four levels (four circles). Level 1 (blue circles) represents the child’s Basic Needs and Extended Basic Needs such as Food & Feeding or Sleep, without which life would be endangered. Level 2 (purple circle) consists of four domains standing for the child’s Non-Basic Needs. Level 3 and 4 (red and orange circle) include three and one domain of factors related to the child’s Caregivers & Family and Society & Community that impact its HRQOL.
5.6.2.1 Neonate/Infant Basic Needs & Extended Basic Needs

At the center of the circle, Level 1 represents the child’s Basic Needs, which consist of 6 domains: Safety, Hygiene, Sleep, Sensory Experiences, Physical Abilities, and Food & Feeding. This level incorporates needs that if not met, lead to the child’s death or are potentially life-threatening. To reflect this subtle difference between these two types of basic needs, Level 1 is divided by an interrupted line and the two sub-levels named Neonate/Infant Basic Needs and Extended Basic Needs (Figure 5-1, Appendix 5-D). Each domain has attributes, which may have items belonging to both sub-levels (e.g. under Food & Feeding, Feeding may include Being Fed pertaining to Neonate/Infant Basic Needs or Breastfeeding relating to Neonate/Infant Extended Basic Needs) (Figure 5-2, Appendix 5-D).
Figure 5-2: Attributes of Infant and Neonatal HRQOL Concept
For graphical clarity reasons, attributes NIHRQOL are illustrated using a hierarchical tree diagram. Thirty-five attributes of NIHRQOL were identified and organized into domains and levels. Level 1 – Neonate/Infant Basic Needs and Neonate/Infant Extended Basic Needs (blue boxes) consist of 6 domains and 11 attributes. Level 2 – Neonate/Infant Non-Basic Needs (purple boxes) include 4 domains and 12 attributes. Level 3 – Caregivers & Family (red boxes) comprise 3 domains and 10 attributes. Level 4 – Society & Community (orange boxes) contains 1 domain and 2 attributes.
Safety was perceived as a crucial NIHRQOL-factor. “Being safe”, “shelter” and “being cared for” were categorized under Safety and summarized in three attributes: Physical Safety (absence of physical harm/injuries, provision of medical care)\textsuperscript{02}, Psychosocial Safety (absence of emotional/verbal abuse, provision of emotional needs)\textsuperscript{03} and Environmental Safety (e.g. absence of harmful environment)\textsuperscript{04}.

Hygiene and Sleep are two domains of Neonate/Infant Basic Needs and Extended Basic Needs with only one attribute each. Statements of “being clean” and “being dry” were categorized under Hygiene, while Sleep included quantitative and qualitative aspects\textsuperscript{05}.

Sensory experiences were perceived as fundamental and grouped into Stimuli-Input and Stimuli-Output. Stimuli-Input represents the ability to sense stimuli with the human senses, and the quantity, quality and frequency of stimuli (e.g. “pain”, symptoms\textsuperscript{06}, disruption of physical integrity\textsuperscript{07}, physical and psychological “comfort”\textsuperscript{08}, “sounds”\textsuperscript{09} or “smell”\textsuperscript{10}). Stimuli-Output embodies the ability to respond to stimuli. Interaction of these two attributes is reflected by the stimuli-response mechanism, where an infant first has to be able to sense a stimulus (e.g. hearing ability), the stimulus has to be of a certain quantity and quality (e.g. loud noise) and he/she has to be able to respond to the stimulus (e.g. crying).

Closely related to Stimulus-Output is the domain of Physical Abilities, which was divided into Voluntary Abilities and Involuntary Abilities. Involuntary Abilities entail any physical function, which is beyond conscious control such as “breathing”, “digesting” or thermoregulation. Voluntary Abilities on the other hand include “moving” the body or “reaching for” something\textsuperscript{11}. Stimuli-Output and Physical Abilities are differentiated based on the context of the activity, in the sense that Stimuli-Output is defined as a response to a stimulus as part of the stimulus-response mechanism, while Physical Abilities refers to spontaneous physical activities.

Two aspects to Food & Feeding were identified: Nutrition representing the quality and quantity of food\textsuperscript{12}, and Feeding embodying a process, including the means of feeding and the performing person\textsuperscript{13}.
5.6.2.2 Neonate/Infant Non-Basic Needs

Level 2 exemplifies needs important to NIHRQOL without having fatal consequences if not fulfilled. Elucidated factors were grouped into 4 domains and 12 attributes: *Psychosocial Experiences, Development, Environment,* and *Higher-Order Function (Figure 5-1, Figure 5-2, Appendix 5-D).*

*Psychosocial Experiences* consists of 4 attributes: *Presence of Caregivers and other People, Interactions with Caregivers and other People, Emotional States* and *Ability to Communicate.* It is crucial for neonates and infants to have caregivers and other people present in order to allow psychosocial experiences\(^{Q14}\). *Interactions With Caregivers And Other People* includes opportunity and ability to interact with human beings, and the quantity, quality and frequency of such interaction (e.g. “bonding”, recognition of achievements)\(^{Q15}\). Positive and negative *Emotional States* involve any emotion from “being happy” to “being sad”\(^{Q16}\). The fourth attribute of *Psychosocial Experiences*, the *Ability To Communicate*, signifies the ability to send out cues, to have “someone to read” and “respond” to them and in return, to be able to read someone else’s cues and have the “ability to respond”\(^{Q17}\). This attribute is distinguished from *Sensory Experiences* in the sense that latter reflects a crude mechanism of stimuli perception and response, while *Ability To Communicate* includes the mechanism of exchanging information.

Any mental function that is above basic functioning such as *Cognition, Awareness* and *Volition* is incorporated in *Higher-Order Function.* The participants regarded *Higher-Order Function* as essential determinant of NIHRQOL. *Cognition* is seen as ability to interpret stimuli and make associations beyond simple conditioning including the ability to “learn”, accumulate experiences and “recognize” people, things and places\(^{Q18}\). *Awareness* embodies self-awareness and awareness of others including the ability to compare\(^{Q19}\), while *Volition* stands for knowing what you want and the “satisfaction from getting what you wanted”\(^{Q20}\).

Similarly to *Higher-Order Function, Development* seems to have a fundamental role in NIHRQOL. Three aspects of development were identified: *Physical Development, Cognitive Development,* and *Psychosocial Development.* The opportunity to learn including play and availability of age-appropriate stimuli promoting physical, cognitive and psychosocial
Development are crucial features\textsuperscript{Q21}. Cognitive Development also includes development of coping mechanisms and the ability to self-soothe\textsuperscript{Q22}. Development and Stimuli-Input or Ability to Communicate is differentiated, as it refers to stimuli promoting development rather than stimuli as part of stimuli-response mechanism or communication. While the ability to learn is categorized under the attribute Cognition, the opportunity to learn and availability of appropriate stimuli to promote development is part of Development.

Environment influences NIHRQOL in two ways. While Physical Environment portrays the nature of the physical environment\textsuperscript{Q23}, Psychosocial Environment refers to the emotional, psychological and social environment\textsuperscript{Q24}. Both attributes are closely related to Sensory Experiences, where environmental stimuli may impact the stimuli-response mechanism.

5.6.2.3 Caregivers & Family

Level 3 represents caregiver and family factors that affect NIHRQOL. They were grouped into 3 domains and include 10 different attributes: Caregiver & Family Safety, Caregiver & Family Psychosocial Function and Caregiver & Family Development (Figure 5-1, Figure 5-2, Appendix 5-D). All three domains refer to the ability of caregivers and family to meet the neonate’s and infant’s Basic Needs, Extended Basic Needs and Non-Basic Needs.

The Caregiver & Family Safety domain includes four attributes representing all aspects of safety such as Psychosocial Safety, Physical Safety, Environmental Safety and Financial Safety. All four attributes represent the caregivers’ and family’s state of “safety”. Psychosocial Safety includes absence of emotional and verbal abuse, and meeting the family’s psychosocial and emotional needs\textsuperscript{Q25}. Physical Safety revolves around the need for protection from physical harm and provision of medical care\textsuperscript{Q26}, while Environmental Safety exemplifies protection from environmental harm to the family\textsuperscript{Q27}. The ability to assure financial resources for the entire family is subsumed under Financial Safety\textsuperscript{Q28}.

Caregiver & Family Psychosocial Function consists of two attributes: Society & Caregiver Expectations & Values and Chronicity. Society and caregivers have expectations and values, which significantly impact NIHRQOL\textsuperscript{Q29}. The caregivers’ and family’s ability to compare to other neonatal or infant experiences with other children, to other families’ experiences or to the social norm, may influence how caregivers and family interact with the neonate/infant.
thereby influencing his/her HRQOL. Furthermore, the ability and opportunity to fulfill social roles and to belong to social groups is an important part of Caregiver & Family Psychosocial Function. Being a caregiver to all children of the family, the children’s role of being a child within the family and not overtaking adult responsibilities, the adults’ role in a relationship, in a profession, in society, and in the extended family may influence how caregivers and family function and interact with the neonate/infant. Disruption of these roles may impact the ability of the caregiver to meet the neonate’s/infant’s needs thereby affecting his/her HRQOL. Chronicity of any disrupting factor such as illness has its own effect on NIHRQOL. Although these aspects may not be comprehensible to the neonate/infant, the knowledge that there is no cure or there are repeated disruptions influences caregivers’ and families’ psychosocial function and thus the neonate’s/infant’s HRQOL.

Caregiver & Family Development was understood to play an important role in NIHRQOL consisting of four attributes: Caregiver’s Socio-Economic Status, Caregiver’s Maturity, Coping Mechanisms and the Ability & Opportunity To Learn Caregiver Skills. These attributes represent the caregivers’ and families’ ability to develop and extend their capacity to cope with events that disrupt their lives. Caregiver’s Socio-Economical Status includes the caregiver’s educational, intellectual and financial levels, which are linked to the caregiver’s ability to provide safety, to understand and influence the neonate’s/infant’s situation and the availability of a strong social network. Closely related is the amount of resilience, and ability to develop and the quality of Coping Mechanisms that caregivers and families bring with them. Their ability to manage stress and handle unpredictability and uncertainty directly impacts their capability to meet the neonate’s/infant’s needs thereby influencing NIHRQOL. Another facet of Caregiver & Family Development is Caregiver’s Psychological Maturity. It represents the caregiver’s psychological developmental state, amount of life experience and ability of self-reflection. Ability & Opportunity To Learn New Caregiver Skills is additional characteristic that influences the caregivers’ capability to understand and meet the neonate’s/infant’s needs.

5.6.2.4 Society & Community

Level 4 refers to aspects of society and community that influence NIHRQOL (Figure 5-1). These factors were categorized into a domain named Societal Structure & Mores and two
attributes: *Societal Values & Structure* and *Availability & Access to Support* (*Figure 5-2, Appendix 5-D*).

*Societal Values & Structure* determines how society cares and provides for neonates and infants and their caregivers. Society’s values regulate whether resources will be made available to fulfill the neonate’s/infant’s needs directly (e.g. low-cost health care) or indirectly (e.g. leave of absence for caregivers). The society’s attitude towards its dependent population and its tendency to judge and discriminate impacts NIHRQOL directly, but also indirectly, by imposing those values onto the caregiver and family\(^{Q37}\). The amount of stigmatization and effort to integrate all members of a society and community will further influence how the neonate’s/infant’s needs are met\(^{Q38}\). These phenomena relate to the other attribute of this domain, *Availability & Access to Support*, which represents the physical, psychological and financial support society and the community provides to neonates and infants, and their caregivers\(^{Q39}\). The type, availability and functionality of social networks including extended family, friends and self-help groups influence NIHRQOL by supporting caregivers in their ability to meet the neonate’s/infant’s needs\(^{Q40}\). If however caregivers are prevented from meeting those needs, then society’s and community’s willingness and ability to act as replacement caregiver will determine whether the neonate’s/infant’s needs are being fulfilled, thereby impacting significantly on their HRQOL\(^{Q41}\).

### 5.6.3 Subthemes

#### 5.6.3.1 Weighting, Quality & Quantity of Factors

*Neonate/Infant Basic Needs* were perceived to be fundamental for NIHRQOL, and the importance of *Neonate/Infant Non-Basic Needs* increases with the age of the child. For example, for a 1-week and a 10-month old child *Basic Needs* must be met for survival, while *Non-Basic Needs* such as *Psychosocial Experiences* are more relevant to older infants. In this sense, HRQOL-change in a neonate is simpler but also more endangering as it affects the *Basic Needs* primarily. Health-related QOL alteration of an older infant is more complex, but less substantial as it affects *Basic Needs* and *Non-Basic Needs*. For example, a neonate may have a small HRQOL reduction by being admitted to the hospital where all *Basic Needs* are being met. An older infant may have more severely reduced HRQOL in the same health-
state, as their Non-Basic Needs are more relevant to them and more severely affected by the hospital stay.

Although we regard all identified HRQOL-factors as relevant to children between 0 and 365 days of age, their importance is age-dependent. In addition to the weighting, the meaning of individual factors seems to be age-dependent as well. For instance, the presence of certain stimuli may be more or less important depending on the infant’s age, and the quality and quantity of these stimuli may or may not be beneficial. A premature neonate may have decreased HRQOL with the same stimulus that increases HRQOL of a term neonate. Ideally, weighting, quality and quantity of stimuli should be judged based on the developmental stage instead of the linear age to account for the fact that normal development might be disrupted. This is underlined by the perception that Development is fundamental to NIHRQOL. Higher-Order Function is another domain that is crucial, as with impaired Awareness, Volition and Cognition weighting of factors becomes significantly different. It may be extremely difficult for a proxy to judge HRQOL of a child with Higher-Order Function and Development at a neonatal or infantile stage, as their own cognitive function and development is typically at an adult level and there may be a significant interdependence between the child and the proxy, particularly if latter is one of the caregivers.

5.6.3.2 Interconnectivity

Caregivers struggled to discriminate NIHRQOL from their own HRQOL, while health care providers made this differentiation more easily. When asked directly, the majority of participants agreed that NIHRQOL was an individual entity, which was closely related and strongly influenced by caregivers and family. Also society and community at large are interconnected with NIHRQOL, either directly through societal structure, values and mores, which define how neonatal and infant needs shall be met, or indirectly by influencing caregiver and family factors. Dependence is more pronounced the younger the child is. In return, caregivers and family, as well as, society and community are influenced by the wellbeing of neonates and infants, but in a less significant way, particularly on the Society & Community level (Appendix 5-B: B). This reverse dependence becomes stronger the older the child becomes as the child becomes a more active contributor to the family and society.
5.6.3.3 Striving for Normalcy

Caregivers in particular emphasized the importance of their ability to function and meet their neonate’s/infant’s needs, and it became evident that they were constantly striving for normalcy in every aspect of their child’s life with the goal to achieve ‘normal’ NIHRQOL \(^{Q46}\). Based on this understanding HRQOL may be regarded as a surrogate for the amount of effort required by the caregiver and society to counteract HRQOL disrupting factors and achieve expected normalcy in every aspect of the child’s and its family’s life (e.g. everyday activities, physical and mental health, social roles). The effectiveness of these efforts and the amount of resilience amongst the child, its family and society may represent the quantity of achieved HRQOL despite disrupting factors. This also illustrates how caregivers’ and society’s expectations and values influence the neonate’s/infant’s HRQOL. For instance, if an infant is not meeting developmental milestones, his/her HRQOL may be reduced because the caregiver will be worried and maybe disappointed. Another example is if a neonate is always sick and interrupting normal daily activities of the family, there may be frustration and discontentment. One HCP put it in these words, “It may be that we are biased by normal and i.e. if you have a kid who [...] is born with no arms, they probably can be pretty happy not having arms because they figure out other ways of getting what they want.” This further illustrates how relevant Higher-Order Function and Development is to NIHRQOL. If the child is not aware of the fact that he/she is ‘not normal’ or disrupting normalcy by his/her health-state, or if he/she cannot compare to others, then the only way to feel the consequences is by reactions and the ability of their caregivers, family and society to meet the neonate’s/infant’s needs.

The attainment of normalcy is typically orchestrated by primary caregivers thereby assuring that the neonate’s/infant’s needs are being met. If the abnormal condition or the caregiver’s ability to create normalcy is insufficient, then substitutes are found to support this process. For example, if the child is severely sick, support from nurses and doctors are necessary to meet the Basic Needs. Other examples are, if there are insufficient financial recourses to assure financial safety then the government as representative of society’s values and mores will provide financial support, or if caregivers are forced to be at the hospital because of a sick infant, the social network may provide support such as babysit siblings or doing grocery
shopping. All examples illustrate how increased effort by the caregivers, family and society is required to offset the disruptive situation endangering normalcy and increase the probability that the neonate’s/infant’s needs are being met, and to ensure their HRQOL is as good as possible.

5.7 Discussion

The advancement of the medical care in neonatology has made HRQOL-quantification increasingly warranted. Only one HRQOL-instrument has been developed and undergone initial validation for children aged 1-12 months and none have been developed for term and preterm neonates\(^{572,605}\). Varni et al. based their questionnaire on the WHO HRQOL-concept, on instruments developed for older children and on expert opinion, assuming that the conceptual framework did not change significantly throughout early childhood. We aimed to address the lack of evidence around this assumption and propose a conceptual framework of NIHRQOL, identifying all relevant factors and their relationship and highlight particularities of HRQOL-measurement in this age-group.

The proposed NIHRQOL-concept does not aim to be final, but rather represent a seminal idea, which may be used in the development of age-specific, generic or disease-specific HRQOL-measurement tools. *Ability to Trust, Ability to Love, Predictability and Continuity* were difficult to incorporate. It remained unclear whether a neonate/infant has the cognitive ability to love and trust, or if it is part of the fundamental dependence from the caregiver and therefore an inert ability that may be disrupted by traumatic experiences. Both mechanisms may be integrated in the proposed NIHRQOL-concept, either under *Higher-Order Function* or *Psychosocial Experiences*. The results support the understanding that *Predictability and Continuity* in personal relationships and environment are important to older children and adults, their role in NIHRQOL remains uncertain. The ability to adapt to new health-states and thereby improve HRQOL is documented in adults, but it is unclear if neonates and infants can adapt at all, or if they adapt so fast and completely that the adaptation process becomes non-relevant to their HRQOL\(^{606}\). In the proposed NIHRQOL-concept, adaptation plays a role at the *Caregiver & Family* level, where their ability to adapt to a chronically ill child impacts their ability to meet the neonate’s/infant’s needs and therefore his/her HRQOL.
The high complexity of the proposed NIHRQOL-concept is partially caused by the absolute dependence on caregivers and fast-changing developmental stages. Based on our results we are inclined to believe that although individual factors and their relationships remain nearly unchanged between the age of 0 to 12 months, the importance of individual factors changes significantly depending on the child’s age and probably more importantly its developmental stage. It is out of the scope of this study to determine the ideal age-groups or developmental stages for which individual instruments should be developed, but we suggest using as narrow categories as possible to increase validity. Ideally, an HRQOL-instrument would use a mathematical logarithm to weigh the factors along the age or developmental continuum. In addition to rapid developmental changes and absolute dependence on caregivers, NIHRQOL-assessment is based on proxies. Although intuitively the best proxy is perceived as the primary caregiver, the absolute dependence may make it impossible for him/her to judge the neonate’s/infant’s HRQOL independently and thereby as true proxy. On the other hand, if interdependence with the family is part of NIHRQOL, then it may have to be assessed as a whole as proposed by other authors. Our participants however perceived NIHRQOL to be measureable on its own, but being influenced by the family’s wellbeing and most importantly determined by their ability to meet the neonate’s/infant’s needs. This perception is reflected in the presented NIHRQOL-concept, which organizes caregiver and family factors around the more direct factors of NIHRQOL. Due to the strong interconnectivity of caregivers and neonates/infants, it may be beneficial to use multiple proxies to reliably quantify NIHRQOL. Another particular challenge of proxy-measurement in neonates and infants is that proxies do not remember their own infancy, and may therefore be unable to “slip into the shoes” of the neonate/infant as easily. In contrast to older children or adults, neonates and infants have never had the opportunity to state their preferences or show their personality, which may or may not differ from those of their caregivers. The latter is particularly relevant for preference-based utility measurement, which in a neonate/infant may not be feasible for this reason. Despite the limitations of proxy-reporting, we regard NIHRQOL-measurement as crucial for clinical decision-making and outcome assessment particularly for chronically and severely ill neonates and infants, and advocate for development of developmental-stage appropriate methodology and HRQOL-instruments.
5.8 Conclusion

Neonatal and infant HRQOL is a multidimensional, multilayered and interconnected concept, where the child’s basic and non-basic needs represent HRQOL most directly, and the caregivers’, family’s and society’s ability to meet those needs characterize the interdependence between the child and the caregiving environment. The effort exerted by the caregiving environment to achieve normalcy in every aspect of the child’s life may be regarded as a proxy measure for HRQOL when adjusted for resilience among the child, its family and society. Neonatal and infant HRQOL-instruments accounting for fast-changing developmental stages by weighting HRQOL-factors accordingly are warranted.
Chapter 6

6 Concluding Remarks

6.1 Review of Thesis Rationale

Management and outcomes of pediatric IF have changed dramatically over the last decade\textsuperscript{44,46,49,165}. Multidisciplinary intestinal rehabilitation programs (IRPs) have been established and new medical and surgical treatment options have been introduced successfully\textsuperscript{40,41,44,46,49,58,101,156,165,190,253,286,335,336,381-384,386,388,609}. Multidisciplinary intestinal rehabilitation programs promote coordinated provision of care, leading to improved continuity, and earlier referral for transplant assessment if required\textsuperscript{49}. Strategies to reduce omega-6 lipid exposure in PN by, replacing it with solutions rich in omega-3 lipids led to a change of paradigm, as it became evident that the most significant complication of chronic PN, IFALD, could be reversed and mortality from IFALD reduced with adjusted lipid protocols\textsuperscript{58,190,336}. For patients who plateaued in their gut adaptation process, a new surgical approach was developed, that permitted lengthening and tapering short, dilated bowel, to enhance intestinal adaptive capacity\textsuperscript{40,41,101,156,253,286}. Most recently, ethanol locks have become available for pediatric IF-patients with frequent sepsis related to their central venous catheter. As sepsis represents a major complication of IF and contributes significantly to disease-specific mortality, reducing its incidence has been hypothesized to impact significantly on pediatric IF-outcomes as well. While these changes have been observed clinically, high-quality research to objectively evaluate this evolution has remained a challenge in pediatric IF due to multiple methodological difficulties associated with the disease (Chapter 2). Understanding these challenges is fundamental, however, as it drives the pursuit of improved or alternative methodology to ultimately allow creation of high-quality research evidence.

The objective of this thesis was to quantify the change in outcome and determine the independent impact of individual treatment options. Initially, there was an intention to assess outcome into relation to quality of life and cost in a comprehensive cost-utility analysis. As
valid age-specific HRQOL or utility measurement tools are unavailable, we needed to consider an alternative approach. Clinical effectiveness of three different treatment eras was first quantified using univariate and time-to-event analysis. Limitations of this analysis, as well as, the perception that change happened continuously over time made interrupted time series analysis a valuable methodological expansion that allowed adjustment of co-variables despite the small sample size. Once mortality of pediatric IF was found to be significantly decreased, HRQOL-measurement became more important. Lack of age-specific HRQOL-measurement tools prevented such research, however, and made more fundamental work necessary. Unavailability of an age-specific HRQOL-measurement tool applicable to the neonatal and infant IF-population led to the development of a conceptual framework of neonatal and infant HRQOL as a first step to conceptualize the meaning of HRQOL in this population. Development of such a conceptual framework helped understanding of pediatric IF-care in a comprehensive way. Such an understanding is invaluable as it recognizes the entire functional system of the neonate/infant, family and healthcare provider, and its dynamics. It also has the potential to facilitate development of a formal, age-specific HRQOL-measurement tool, applicable to a neonatal and infant population suffering from a chronic and severe medical condition such as pediatric IF.

The quantification of the change in clinical outcomes of pediatric IF, and the effect of IRP-maturation, as well as systematically introduced treatment options have not been performed prior to this PhD-thesis. With a significant lack of age-specific HRQOL-methodology for the pediatric IF-population, such valid quantification has remained impossible. Categorical comparison and survival analysis of different management eras, multivariate analysis of new treatments and their effect on outcomes over time, and development of HRQOL-methodology applicable to the pediatric IF-population was indicated at the time point of this PhD-thesis.

6.2 Review of Thesis Objective

The objective of this PhD-thesis was to evaluate the recent change in management and outcome of pediatric IF. This objective was achieved through three related projects. First, the clinically perceived change in pediatric IF outcomes was quantified comparing three distinct eras. Change in mortality was quantified using univariate time-to-event analysis for disease-
specific mortality, assessing overall case-fatality rates and evaluating mortality causes, while change in morbidity was assessed by comparing rates of CVC-complication, CRBSI and IFALD-development.

Second, the independent effect of systematically introduced treatment options was determined. This was achieved by performing a multivariable, interrupted time series analysis using ARIMA models. Clinically, most meaningful treatment options were assessed, consisting of an IRP in place, and availability of omega-3 lipids, STEP and ethanol locks for patients requiring these treatment options. Adjustments were made for patient characteristics, as well as, for severity of illness. This analysis allowed assessment of dynamic changes of treatment and outcomes over time.

Finally, a conceptual framework of neonatal and infant HRQOL was developed, as age-specific HRQOL-measurement of IF-patients for a comprehensive cost-utility analysis was not feasible due to unavailability of well-developed methodology. Qualitative analysis of data derived from focus groups and interviews with healthcare professionals and parents experienced with severely and chronically ill neonates and infants, enabled understanding of the meaning of HRQOL to a neonatal and infant population suffering from a chronic and severe disease. Factors contributing to neonatal and infant HRQOL, and their relationships and interactions were identified and arranged to reflect the conceptual understanding of neonatal and infant HRQOL. This understanding is fundamental for future neonatal and infant HRQOL measurement tool development.

6.3 Synthesis of Thesis Findings

This PhD-thesis reviewed medical aspects of pediatric IF in Chapter 1. In Chapter 2, methodological challenges related to pediatric IF-research were identified and solutions proposed. Methodological challenges were mainly found to be related to the surgical and rare nature of the disease. The rapidly evolving clinical care of pediatric IF and clinical urgency of the disorder, as well as, vulnerability of the population and sensitivity of relevant data makes high-quality pediatric IF-research additionally problematic. Heterogeneous terminology and inconsistent outcome selection and definition, as well as absence of age-specific methodology further intensify challenges encountered when aiming to conduct high-
quality pediatric IF-research. Solutions were proposed and attempts made in this PhD-thesis to promote high-quality pediatric IF-research despite methodological challenges.

Changes in outcome of pediatric IF were quantified in Chapter 3 by comparing three distinct eras of pediatric IF management. The improved clinical outcome was mainly quantifiable in terms of ICU admission, complications from central venous catheters, complications from chronic PN and disease-specific mortality. Introduction of the IRP led to an initial increase of ICU admission time from 86 days (IQR 59-145) to 91 days (IQR 72-126), while maturation of the program reduced this time duration to 59 days (IQR 17-102; *p*=0.002)). Sepsis from central venous catheter declined steadily from 22 events/1000 catheter-days (11-35) to 12 (2-25; *p*=0.014). Most importantly introduction and maturation of the IRP led to a significant decrease of disease-specific mortality from 26% Pre-IRP to 3.5% post-IRP (*p*=0.004). These findings illustrate how clinical outcomes improved as the IRP evolved. As a concurrent control group was not possible, this comparison of three eras does not permit conclusions on the cause-effect relationship of individual factors. Also, valid multivariable analysis was not feasible in this study due to the small sample size, although different methods such as variable transformation and propensity scores were explored. The available sample size was maximized, however, as the prospective database used in the study, is one of the largest and with the longest follow-up reported in literature. As the eras represent a compound exposure, that includes measurable treatment interventions such as alternative lipid emulsions and more intangible factors that occur when there is an experienced team managing these patients on a dedicated unit with specific resources and treatment guidelines, it remained unclear what the effect of individual treatment options were, as they were introduced and administered dynamically during one era. In our study, we were able to show that the introduction of our IRP initially led mainly to a significant increase in transplantation and STEPs, while with maturation of the program, surgical interventions became a less important therapeutic option. With time, disease-specific mortality decreased significantly as did CVC-complication and sepsis rates.

Due to the sample size, multivariable analysis was not feasible in the first thesis study (Chapter 3). To address this limitation, and to quantify the effect of individual treatment options, a multivariable time series analysis using ARIMA models was performed (Chapter 4). This study confirmed that disease-specific mortality decreased continuously over time (-
0.02 deaths/quarter; 95%-CI -0.03, -0.01). Introduction of IRP and omega-3 lipids independently decreased disease-specific mortality by 0.6 deaths/quarter each (95%-CI -1.23, -0.02; 95%-CI -0.77, -0.45;). In contrast, serial transverse enteroplasty and ethanol locks did not have a significant impact on mortality caused by intestinal failure-associated liver disease or sepsis. Parenteral nutrition weaning, transplant assessments and transplantations remained unchanged, while catheter sepsis and complication rates decreased significantly by 0.3 episodes per 1000 catheter-days each (95%-CI -0.43, -0.2; p<0.001; 95%-CI -0.45, -0.24; p<0.001). This study is the first to clarify the independent roles of treatment options in the clinical change of pediatric IF. It is also the first study to apply time series analysis to evaluate the evolution of an intestinal rehabilitation program in a longitudinal manner. This statistical approach was perceived as particularly valuable for assessment of a rare, complex disease undergoing dramatic and rapid change to contribute to establish causal relationships between exposures and outcomes. By approaching data analysis from a different perspective, such as examination of data longitudinally, the number of observations was increased allowing for multivariable analysis to be performed. Also, by introducing a time factor into the analysis, temporality between exposure and affect was accounted for. This is the first study to use time series analysis in the setting of pediatric IF-research. This methodology is a valuable tool and should be used more frequently in clinical research.

The inability to measure HRQOL in our neonatal and infant cohort led to the third PhD-thesis study that aimed to explain the meaning of neonatal and infant HRQOL. Chapter 5 described factors contributing to neonatal and infants HRQOL and organized them in a circular diagram. The proposed conceptual framework of neonatal and infant HRQOL was derived from focus groups and interviews with parents and healthcare professionals, who have personal or professional experience with severely and chronically ill neonates and infants. Neonatal and infant HRQOL is perceived as a multidimensional, multilayered and interconnected concept, where the child’s needs contribute most directly, and the caregiver’s and society’s ability to meet those needs characterize the interdependence between the child and its caregiving environment. This study is the first to propose a formal conceptual framework of HRQOL in neonates and infants. It suggests a fundamental understanding of neonatal and infant HRQOL, and identifies factors contributing to a neonate’s and infant’s HRQOL, and defines relationships and interactions between these factors. This basic concept
does not allow for HRQOL measurement per se, but may facilitate development of HRQOL-measurement tools in the future. It may also support understanding outcomes of pediatric IF in a comprehensive way and help elucidate functional mechanisms in families, such as coping mechanisms or noncompliance with treatments. Developmental-stage specific HRQOL-instruments for premature and mature neonates, and infants are urgently required to evaluate clinical outcomes, guide clinical-decision making and justify policy-making.

6.4 Implications

The three studies included in this PhD-thesis impact clinical decision-making, policy-making, methodology development, and affected patients and families.

6.4.1 Implications for Clinicians

Chapter 3 showed that IRPs are associated with decreased mortality in pediatric IF, while Chapter 4 revealed that the IRP itself as well as the omega-3 lipids were the main factors contributing to this change in outcome. These results support the clinically perceived benefit of IRPs that have been founded in several pediatric tertiary healthcare centers across North America. The results also showed that STEP and ethanol locks do not significantly impact disease-specific mortality, while omega-3 lipid prevented such mortality as effectively as the IRP itself. Whether these results represent a simple association between treatment options and improved outcomes, or a true causal relationship is debatable. Based on the methods used, results from the first study realistically depict observed associations between the IRP presence and experience, and the clinical outcomes. Results from the second study may, however, hold additional information about the relationship between the tested treatment options and measured outcomes. When following the Bradford Hill criteria to assess causal inference, the results may be regarded as reflecting causality for the primary outcome of interest (disease-specific mortality): first, the effect size was large and achieved statistical significance despite the small sample size; second, a decrease in mortality from liver disease and/or sepsis was observed by several different clinical and research groups; third, the study cohort was specific despite not being in an experimental setting, meaning that the cohort only included clearly defined participants with one medical condition managed all at the same medical center; fourth, there is a time-relationship between the exposure and the effect in the
study design as well as the statistical analysis applied; fifth, a biological gradient between exposure and effect was not observed due to the nature of the tested exposures, however, a gradient response was observed for the maturation of IRP, that is, the more mature the IRP, the more pronounced the decrease in disease-specific mortality; sixth, the cause-effect relationship between the tested treatment options and disease-specific mortality is plausible; seventh, the findings are believed to be coherent in different settings, although formal laboratory results are not available for any of the four tested treatment options due to limited applicability (e.g. the effect of STEP availability on disease-specific mortality is not assessable in a laboratory setting); eighth, availability of experimental evidence is limited due to methodological difficulties as outlined in Chapter 2; and finally, similar factors result in similar results, such as for example, novel lipid strategies using omega-3 lipids and such minimizing lipid dose exposure in PN result in improved liver function and ultimately in decreased disease-specific mortality. Although the application of Bradford Hill criteria are debatable, because they do not account for confounders and bias, assessing relationships between exposures and outcomes using these criteria may still help estimate the likelihood of a causal relationship when separately discussing the risk for bias and confounders. The associations found between the four treatment options and primary outcome of interest using time series analysis in Chapter 4 may likely represent a causal relationship based on these criteria. On the other hand, the two studies reported in Chapter 3 and Chapter 4 are vulnerable to bias introduced through the observational study design. As identified and discussed in Chapter 2, these study designs are particularly at risk for selection bias, sampling bias, recall bias, confounding bias, information bias and measurement bias. Although all feasible efforts were made to minimize these systematic errors, when generalizing results from the studies reported in Chapter 3 and Chapter 4, these potential weaknesses should be appreciated and weighed against the paucity of higher-level evidence and clinical urgency in pediatric IF. Despite the obvious limitations of these two studies, the produced knowledge is valuable to clinicians, as it may guide clinical decision-making around what type of treatment to offer to pediatric patients with IF. Depending on the patient’s situation, omega-3 lipids may be indicated to prevent advancement of IFALD, while recommendations for STEP and ethanol locks may be given more selectively knowing the limited impact on disease-specific mortality of these treatments. These other treatments may be found to have a larger impact on intestinal autonomy given longer follow up. On the other
hand, it should not be forgotten that these results fail to include other important aspects of outcome, such as morbidity and HRQOL. For example, ethanol locks may improve morbidity from septic episodes in a patient with frequent CRBSI. In addition, some results may have remained statistically non-significant for several reasons including a small a number of patients, who received the treatment, insufficient follow-up, and more complex anatomy in more recent patients, as discussed in the respective chapters (Chapter 3, Chapter 4).

6.4.2 Implications for Policy-Makers

Although no formal cost-effectiveness analysis could be performed, the findings of Chapter 3 and Chapter 4 are expected to provide valid information to policy-makers in developed countries with a similar healthcare setting. When deciding how to distribute resources, the results may be valuable as they show that omega-3 lipids are as effective in preventing disease-specific mortality as a formal IRP. Omega-3 lipids may be less expensive and more readily implementable than a formal IRP that involves numerous administrative and personnel changes. The added benefit of IRPs that is not captured in disease-specific mortality, the effect on morbidity, HRQOL and neurocognitive function, may however only become quantifiable in a formal cost-utility analysis. It is therefore recommended to not make overly strong conclusions on treatment modalities and generally offer them the setting of an IRP. Less specialized healthcare centers may benefit from these results as well, as they may direct clinicians and policy-makers to focus resources and energy on developing an IRP, and introduction of omega-3 lipids when optimizing management of pediatric IF-patients. A formal cost-effectiveness analysis for IRP and new treatment options is necessary, to put improved survival in context with cost of care and morbidity. The goal is not to create severely morbid patients at a high cost with poor long-term outcome and miserable HRQOL. To prevent this situation and optimize cost-effectiveness, formal economic evaluation using valid methodology is required. Although this necessity was acknowledged and attempts were made to conduct a cost-effectiveness analysis of IRPs, insufficient availability of age-specific methods to measure HRQOL or utilities in neonates and infants made it impossible. For the same reason, morbidity could only be quantified in terms of clinical outcomes, such as central venous catheter sepsis, IFALD or length of stay at the ICU, in the two outcome studies.
6.4.3 Implications for Researchers

A systematic literature review was performed by this author, but not reported in this PhD-thesis, and revealed that HRQOL has been measured previously in neonates and infants by using instruments developed for older children or has been evaluated by measuring HRQOL once the children were older. Although these approaches may provide some information, they do not represent the highest quality standard and may threat internal validity. For example, measuring HRQOL in neonates and infants with an instrument developed for older children may not capture all domains of HRQOL that are relevant to this specific age-group, may inaccurately measure certain features and represent a biased HRQOL estimate. Rather than employ sub-optimal methods to produce an urgently needed cost-effectiveness analysis, a qualitative study was designed instead to organize the fundamental understanding of neonatal and infant HRQOL in a conceptual framework (Chapter 5). The proposed conceptual framework of neonatal and infant HRQOL will allow development of age-specific HRQOL-instruments, and support clinicians and researchers in their understanding of neonatal and infant HRQOL. It identifies the diverse aspects of HRQOL that affect neonates and infants, and illustrates how far-reaching and complex neonatal and infant HRQOL is. Furthermore, awareness of limited high-quality methodology in this area may have been increased in clinicians and researchers as a result of this study and thereby prevent completion of further sub-optimal HRQOL-measurement studies in neonates and infants.

Finally, the detailed review of methodological challenges encountered in pediatric IF-research resulting from the disorder’s rarity, surgical and rapidly evolving nature, as well as, involvement of a vulnerable population and need for sensitive data as described in Chapter 2, may be valuable to other researchers of similar diseases when identifying difficulties, finding solutions and using a reference for such challenges. The proposed methods to address these limitations and application of novel methodology may guide researchers, reviewers and editors in the conduction and publication of such research. Similarly, the successfully applied time series analysis in a pediatric IF-study as described in Chapter 4, and a conceptual framework for neonatal and infant HRQOL as proposed in Chapter 5 may work as a methodological guide and reference for future studies.
6.4.4 Implications for Patients & Families

The thesis results are meaningful to affected patients and families, mainly through before-mentioned guidance of clinicians, policy-makers and researchers. As new medical knowledge is created in Chapter 3 and Chapter 4, clinical decision-making is facilitated with high-quality evidence and health care provision to patients and families is optimized. Based on the results of Chapter 3 and Chapter 4, showing that IRPs are associated with decreased mortality in pediatric IF and that the IRP itself as well as the omega-3 lipids were the main factors contributing to this change in outcome, clinicians as well as affected patients and families may arrange for consultation of healthcare centers with an IRP in place and omega-3 lipids available. The results of Chapter 3 and Chapter 4 may also support affected patients and families when trying to advocate for availability of an IRP and other pediatric IF-treatment options at their healthcare center and justify for coverage by health insurances of costs related to these treatments.

The results of Chapter 5 impact affected patients and families as the developed conceptual framework helps understanding pediatric IF-care in a comprehensive way accounting for the entire functional system of the neonate/infant, family and healthcare provider, and its dynamics. These results will also facilitate future HRQOL-measurement tool development, which will further benefit affected patients and families as it will allow quantification of their HRQOL and its incorporation into medical decision-making.

Also Chapter 1 and Chapter 2 may be significant to affected patients and families, as it summarizes the current medical knowledge on pediatric IF and explains methodological difficulties associated with its research. This is primarily achieved by providing free, condensed information to the general public, clinicians, policy-makers and researchers, thereby facilitating understanding of pediatric IF and associated methodological difficulties. In addition, Chapter 2 suggests solutions for the methodological challenges related to pediatric IF-research, thus providing support to researchers and increasing the quality level of evidence on this medical condition, which ultimately benefits the affected patients and families.

In summary, the thesis results impact the affected patients and families by making condensed medical knowledge available, summarizing methodological difficulties and suggesting
solutions, optimizing medical care by increasing high-quality evidence, and outlining the meaning of the patient’s HRQOL in the context of the entire family and care-providing environment.

### 6.5 Future Research Directions

Certain aspects of pediatric IF are under-researched and numerous research questions are still unanswered. Reasons for these gaps in knowledge have been discussed thoroughly in *Chapter 2*. With improved mortality, aspects related to morbidity and HRQOL have become progressively important outcomes and increasingly relevant to patients and families. Studies quantifying and characterizing morbidity including HRQOL in pediatric IF are therefore urgently warranted. Methodological studies developing valid age-specific neonatal and infant HRQOL-measurement tools are, however, fundamental to enable such outcome studies. Methodology around neonatal and infant HRQOL is therefore likely the most pressing and broad future research topic derived from this PhD-thesis. A significant lack of measurement tools for neonatal and infant HRQOL impedes evaluation of morbidity of all medical conditions affecting neonates and infants. High-stakes clinical decisions such as end-of-life decisions require measurement of current and expected HRQOL, which with the current state of knowledge are not measurable. Based on the proposed conceptual framework, an HRQOL-instrument for premature and mature neonates, and infants may be developed, that will be most applicable to pediatric IF as well as any other chronic, severe disease affecting neonates and infants. The proposed conceptual framework identified factors that contribute to HRQOL in neonates and infants, and revealed their relationship to each other. In future studies, an HRQOL-instrument for premature and mature neonates, and infants may be developed by reducing the number of these factors, weighting and summarizing them into a score using qualitative (e.g. focus groups for different age groups) and quantitative methods (e.g. exploratory factor analysis). Subsequent studies are necessary to then evaluate these age-specific neonatal and infant HRQOL-instruments in external cohorts.

Similar to the results from the two outcome studies of this PhD-thesis (*Chapter 3* and *Chapter 4*), several other research groups have assessed short-term outcomes in pediatric IF. However, long-term outcomes are still widely unknown. For example, it remains unclear what the long-term effect of chronic malabsorption during early childhood on the
neurocognitive development in children with IF is. Or closely related, whether fat minimization as a strategy to prevent end-stage IFALD affects neurocognitive development in children with IF. Also, long-term outcome questions around enteral function, bone health and skeletal growth, liver function, secondary malignancy development, and recurrent need for surgical intervention for bowel obstructions/perforations/strictures are completely unknown at this current time point. Ultimately, the question whether the life expectancy is normal for children who completely adapted their bowel function after neonatal IF needs to be answered, and, if not, by how many years is it reduced and what are the limiting factors. Also, long-term psychological functioning and socio-economical aspects will need to be evaluated, such as for example what their expected cognitive function will be as an adult.

As pediatric IF is a rapidly evolving disease with new treatment options being introduced continuously, coordinated and up-to-date outcome studies are necessary to evaluate treatment effectiveness and safety. Methods to improve efficiency and coordination of clinical research include inception of expert consortia and centralized national clinical trials offices. Such efforts have been made in pediatric IF-research in North America, as for example the Pediatric IF Working Group under the umbrella of the American Society of Parenteral and Enteral Nutrition was created 2 years ago. The group has met to discuss pediatric IF-specific research and clinical care topics, e.g. planning of new clinical guidelines, determination of relevant clinical outcomes, development of supportive patient tools such as apps, identification of funding difficulties. The American Society of Parenteral and Enteral Nutrition has been active in publication of various clinical guidelines for the management of patients on chronic PN support. Another way to coordinate clinical research and more importantly pool data is an international disease registry. While there is a registry in place for patients who underwent STEP, discussions are currently happening between experts for creating a disease registry for pediatric IF. Such a registry would be beneficial, as it would allow data pooling, resource maximization, research coordination and generalizability improvement. On the other hand, the cost and workload of registry maintenance should not be underestimated. Presently, the International Intestinal Transplant Registry captures data on essentially all bowel transplants that are performed worldwide. As patients who receive intestinal transplantation are those suffering intestinal failure, perhaps it could be possible to add an IF platform to the existing infrastructure. Funding may be challenging for such a
disease registry, but financial resources may be offset from professional associations or contributing institutions.

The results of the two outcome studies (Chapter 3 and Chapter 4) may also stimulate future studies. The conclusions should be validated by other IRPs, as the results are only generalizable to a highly specialized, tertiary health-care setting in Canada or any other comparable country. In the presented studies, all measured clinical outcomes are of relatively short nature due to the recent change in disease-specific mortality. With increased survival, long-term and functional outcomes (e.g. oral tolerance), as well as, developmental delay, quality of life aspects and economic parameters become more important. There are also numerous clinical research questions that remain to be answered. Examples include clear medical indications for treatment options and ideal time point of treatment initiation (e.g. STEP, omega-3 lipids), role of transplantation in pediatric IF in the era of omega-3 lipids, efficacy and safety of new treatment modalities such as trophic factors (e.g. GLP-2, GH) and standardized management of associated medical conditions (e.g. bacterial overgrowth, dysmotility). This list is not all-inclusive and only aims to illustrate the extent of possible future clinical research directions. Decision-analysis based research and economic evaluations in addition may provide evidence to guide administrative decisions, resource allocation and policy-making. Cost-effectiveness of different pediatric IF-managements would be one study to conduct as soon as valid, age-specific utility and HRQOL-measurement becomes available.

As there are numerous methodological challenges in pediatric IF and many unanswered questions, it would be helpful if a consortium of expert pediatric IF centers created a working document to outline research priorities for the field. An attempt to derive consensus definitions for the key outcomes in pediatric IF would be an excellent beginning. A manuscript that systematically identifies methodical challenges pertinent to pediatric IF-research and suggests solutions, could act as an important reference for researchers, reviewers and editors. In addition to the approaches proposed in such an article, development of novel methods or import of approaches from other academic disciplines could be proposed to address these methodological issues appropriately.
There are instances when clinical research is not feasible in human subjects for ethical and methodological reasons. Pre-clinical studies employing well-designed animal models have a role to answer questions and also be able to dive deeper into mechanism of outcome. Animal models represent a valuable alternative and provide translatable evidence. Animal models of pediatric IF (e.g. rodent and porcine models of SBS) have proven to be useful. Population heterogeneity and numerous clinical confounders can be controlled using animal models. The ability to harvest tissues permits determination of disease mechanisms that is not possible in humans. Translational animal studies do not replace human trials, but they serve an important purpose to advance the field.

One specific method to address difficulties that arise from small samples would be to develop a disease-specific, pediatric IF severity of illness index. A disease-specific, pediatric IF severity of illness index would be based on variables that are routinely available and measured, it would be well calibrated, have a high level of discrimination, would be applicable to all pediatric IF-populations, useable in different countries and able to predict not only mortality, but also functional status and HRQOL. A pediatric IF severity score would serve several purposes. First, it would help clinically with disease prognostication and aid in counseling families. Second, it could be used as a research tool to stratify patients for analysis. If widely adopted, outcomes of patients at different institutions could be adjusted by disease severity to make population comparisons more homogeneous. Third, such a severity score could be used for patient recruitment into clinical trials to ensure homogeneity of study cohorts. After creation of a disease severity score with internal validation, it would be necessary to complete external validation from an independent data set.

6.6 Conclusion

Clinical management and research of pediatric IF is challenging and complex. Multidisciplinary intestinal rehabilitation programs and omega-3 lipids have led to a rapid and dramatic change of pediatric IF-outcomes. Health-related quality of life, cost-effectiveness and long-term outcomes of neonates and infants with IF remain undetermined. Creative solutions for clinical management and research of such populations are required. The three PhD-thesis projects illustrate potential solutions to challenges related to research of
a rare, surgical, rapidly evolving disease involving a vulnerable, neonatal and infant population.
References


94. Nightingale JM, Lennard-Jones JE, Gertner DJ, Wood SR, Bartram CI. Colonic preservation reduces need for parenteral therapy, increases incidence of renal stones,


134. Szkudlarek J, Jeppesen PB, Mortensen PB. Effect of high dose growth hormone with glutamine and no change in diet on intestinal absorption in short bowel patients: a


198. Carricato MN. *Quality of life in pediatric patients with intestinal failure on home parenteral nutrition [Thesis (M Sc )], University of Toronto (Canada), 2010.; 2010*


403. StataCorp. 2009. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP.


406. Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *Lancet*. 2014;384(9938):189-205.


419. Broviac Jw Fau - Cole JJ, Cole Jj Fau - Scribner BH, Scribner BH. A silicone rubber atrial catheter for prolonged parenteral alimentation. 1973(0039-6087 (Print)).

420. Hickman Ro Fau - Buckner CD, Buckner Cd Fau - Clift RA, Clift Ra Fau - Sanders JE, Sanders Je Fau - Stewart P, Stewart P Fau - Thomas ED, Thomas ED. A modified right atrial catheter for access to the venous system in marrow transplant recipients. 1979(0039-6087 (Print)).


441. Oliveira C, de Silva NT, Wales PW. Enteral Omega-3 Lipids To Control Intestinal Haemorrhage After Serial Transverse Enteroplasty: A N-Of-1 Trial. Annual meeting of the Canadian Association of Paediatric Surgery (CAPS); 2014; Montreal, QC, Canada.


487. Gordon JS, Akers L, Severson HH, Danaher BG, Boles SM. Successful participant recruitment strategies for an online smokeless tobacco cessation program. Nicotine &


501. Kavanaugh K, Ayres L. "Not as bad as it could have been": assessing and mitigating harm during research interviews on sensitive topics. Research in nursing & health. 1998;21(1):91-97.


Appendices

Appendix 4-A: Linear Trend Of Disease-Specific Mortality

Disease-specific mortality decreased from 1 death/quarter to 0 deaths/quarter in average.
Appendix 4-B: Linear Trend Of Parenteral Nutrition Weaning & Advanced Liver Disease

A/B) The linear trend of complete weaning from parenteral nutrition and advanced liver disease incidence indicate no significant change over the 15 year time period.
Appendix 4-C: Linear Trend Of Central Venous Catheter & Transplantation Related Outcomes

A/B) Central venous catheter complications and CRBSI decreased by 0.3 episodes/1000 catheter-days per quarter.

C/D) The linear change of transplant assessments and transplantations was not significant.
Appendix 5-A: Guideline Used By The Moderator

1. What are important factors that contribute to the quality of life of an infant?
2. After saturation, what is the role of the following factors in HRQOL:
   a. Physical
   b. Emotional
   c. Family
   d. Social
   e. Cognitive
   f. Spiritual
   g. Psychological functioning
   h. Environment
3. Do you think there are factors that play a role in an infant’s quality of life that we did not discuss so far?
4. How are these factors organized in relation to each other?
5. How do they influence each other (diagram)?
6. How do rate the significance of these factors for infants?
7. Are there any differences for neonates in terms of the diagram or the importance of individual factors?
Appendix 5-B: Terminology & Interconnectivity
A: The factors impacting on NIHRQOL identified by the focus groups/interviews were organized using a circular diagram with hierarchical order: levels (circles) → domains → attributes (small white circles) → items.

B: One subtheme of the NIHRQOL concept is the interconnectivity of factors with each other. Basic Needs and Extended Basic Needs interact with Non-Basic Needs and factors related to Caregivers & Family and Society & Community (continuous lines). Although Society & Community is also influenced by individual neonatal and family well-being, this effect is much less significant (dashed line).
Appendix 5-C: Direct Quotes Of Participants

Q1. “basic things”, “basic needs”, “basic physical comforts”
Q2. “I didn’t want […] [something] that could potentially harm her.”, “It needs to be clean. We don’t want them to get sick.”, “The most important things were her TPN [total parenteral nutrition].”, “A sense of security by knowing who is going to take care of you.”
Q3. “[…] basic things of holding him, taking care of him”, “I didn’t want […] [something] that could potentially harm her.”, “caring”, “Making sure […] they are feeling good.”
Q4. “shelter”, “warmth”, “age appropriate toys”, “safe and secure environment”
Q5. “being able to rest”, “uninterrupted sleep”, “sleep duration”
Q6. “cramps”, “vomiting”
Q7. “getting poked”, “One of the big things […] [is when] he has no wires connecting him to anything.”
Q8. “something good”, “comfort”, “satisfaction”,
Q9. “mom’s voice”, “music”, “noise”
Q10. “breast milk smell”
Q11. “roll over”, “sit up”
Q12. “something to eat”, “warm food available”, “hunger”, “thirst”
Q13. “breast”, “bottle”, “tube”, “[…] they allowed me to start feeding her.”
Q14. “Having a mother or father present with you”
Q15. “socializing”, “encouragement”, “being loved”, “[…] positive vibes i.e. good for you, you’re so pretty.”
Q16. “frustrated”, “unhappy”, “feeling safe”, “being anxious”, “She really enjoyed […]”
Q17. “your family’s responsiveness to your cues”, “read the child’s cues and understand their needs”, “How is it affecting them? […] They are not comfortable. They are not happy”, “readability”
Q18. “… the ability to experience and the issue of cognition. Can you experience something”, “I think of anticipation. They have to know it is coming. […] And
that is part of cognition.”, “If you don’t have appropriate cognition, you can’t develop relationships and you can’t experience your world the way it is, or you can’t think about it or anticipate it or do anything.”

Q19. “to be aware of others”, “to compare experiences”

Q20. “[…] continuously wanting a bottle. It was like a never-ending thing. The bottle was like her happiness.”

Q21. “play is […] how they learn”, “age-appropriate toys”, “It took awhile to learn […] playing with other kids […] until she realized that other kids were around and she was able to play with other people.”, “Development is sort of ongoing learning.”

Q22. “… becoming calm. The issue of self-regulation”, “whatever calming or self-regulatory mechanism”

Q23. “unpleasantly noisy environment“, “Colorful things to look at, sparkly lights, etc. Having toys […]”, “I’m also thinking about sounds of music.”

Q24. “… parents […] stressed out […] they sense it„, “Growing up in a chaotic environment may place the foundation of their perception of the world.”, “A baby is able to feel that stress of the environment.”, “If the parent is feeling good and comfortable and relaxed, the child is going to feed off of that.”

Q25. “It is not just the patient living in that house. There are 14-year olds, there are 12-year olds and there are grandparents.”, “They should be at home playing with their brother and sister and just being a normal child.”, “It was more just being around a family environment with all the commotion and everything about that. I don’t necessarily know if it is for them, but it helped us.”, “We are like “I’m fine, I’m fine, I’m fine, we’re fine, we’re fine, we’re fine” but we’re not fine.”, “There are also our other two children who also suffer from it.”, “I had another little girl at home so I could only spend so much time here. Trying to manage the time away from one versus the time away from the other.”

Q26. “To be at home and be able to be with your child at home in your own comfort and all that kind of thing.”,

Q27. “family environment“, “home”
Q28. “You need to be able to support this child”, “I think resources contribute greatly to everyone’s individual situation”, “Now I’m not even having an income, but I have to be with my child”, “It is not just the patient living in that house.”

Q29. “They always say babies should be able to do certain things by certain ages […] how far behind my child was.”, “I just have to be conscious of […] what are my expectations in here.”, “I think it [the understanding that the child is behind with the development] is more for the parents.”

Q30. “I compare […] There are so many things that I didn’t do with him that I did with her.”, “They are telling me there child has a cough and they are crying. I’m like ‘are you kidding me’.”

Q31. “That is what mothers do […] This is part of the mother’s role. The baby’s role is to respond.”, “There is no time to work and to worry about your child at the hospital.”, “If they don’t have their basic needs, they are not going to develop. It is the mom and dad’s role.”

Q32. “My 14-year old says ‘Just cancel. It is not like we are going to be able to go because my sister is going to get sick. Just cancel it.’”, “Grade 1 - we spent half of it again in hospital.”, “If I wasn’t so paranoid, I would have let her play.”

Q33. “Being involved in the care […] understanding what is going on.”, “[…] I can’t talk to my girlfriends because they don’t know what I’m going through.”, “I understand that the doctor can’t be here 24/7.”, “[…] kind of understanding helped me.”, “It was kind of daunting to bring [him] home. [What helped] was just having resources available to assist.”, “I would come in and I would literally have charts of my child’s weight i.e. his input, his output, whatever.”

Q34. “Stress of the parent. That’s going to influence that [NIHRQOL]. The main point is how the caregivers then deal with the baby”, “I get very angry.”

Q35. “I know from my experience I find it normal to be here [in the hospital]”, “I asked myself that everyday ‘Why me? Why him? Why us?’", “You do ask yourself, but it is like ‘it’s destiny.’ That is going to happen.”

Q36. “The teaching and education […] understanding helped me.”, “[…] understanding what is going on and then being able to help.”

Q37. “Here my daughter is normal […] just like all the other kids”, “Just being a normal child.”, “[…] we are biased by normal.”, “You feel normal in an abnormal
world.”, “[…] as healthcare providers, we are dealing with an abnormal situation, we are dealing with a child in a hospital.”

Q38. “I brought him to […] Ontario Early Year Centers since he was a little baby. […] They get to play”

Q39. “Having resources […] to make sure he doesn’t fall too far behind.”, “She [the home nurse] still comes to the house and helps out. Even just scheduling of certain things.”, “[…] resources, we had a lot of help at the beginning.”

Q40. “It felt so good to speak to someone that was going through the same problems.”

Q41. “It was basically me and him that were the only ones here, and the nurses of course.”, “If not a parent be here, who else is going to be here?”

Q42. “Some of the weighting would change, but I don’t think the structure or the components would.”

Q43. “Stimulation […] that’s appropriate to their level of functioning.”

Q44. “The fact that they might not be cognitively aware of that [their HRQOL] plays a role in quality of life itself.”

Q45. “I think that they are separate […] a baby is one person with a different developmental level and ability to understand their needs and the parents are their own people. They have to be considered together, but separately.”, “I was even thinking about things for myself that may not necessarily have directly impacted him, but also helped me care for him.”

Q46. “Once she got out of the hospital she was more able to continue on with her life and do normal things.”, “I’m offering them some sort of normalcy.”, “Just try to make it as normal as possible.”, “Everything normal.”, “Something much more fun. Something that feels normal.”, “I just wanted to be able to feed my baby in a normal way.”, “They can’t interact the way a normal baby would. They are not comfortable. They are not happy.”, “The idea of the quality of life and what is normal to you is not normal for us anymore.”
Appendix 5-D: Levels, Domains & Attributes of Infant and Neonatal HRQOL Concept

Thirty-five attributes of NIHRQOL were identified.
Level 1: Neonate/Infant Basic Needs and Neonate/Infant Extended Basic Needs (central blue circles) consist of 6 domains (partial circle) and 11 attributes (small white circles): A) three attributes contribute to the domain Safety: Physical Safety (SA-A1), Psychosocial Safety (SA-A2), Environmental Safety (SA-A3); B) one attribute contributes to the domain Hygiene: Hygiene (HY-A1); C) one attribute contributes to the domain Sleep: Sleep (SL-A1); D) three attributes contribute to the domain Sensory Experiences: Sensory Input (SE-A1), Sensory Output (SE-A2); E) two attributes contribute to the domain Physical Abilities: Involuntary Abilities (PA-A1), Voluntary Abilities (PA-A2); E) two attributes contribute to the domain Food & Feeding: Feeding (FF-A1), Nutrition (FF-A2).

Level 2: Neonate/infant Non-Basic Needs (purple circle) include 4 domains (quarter-circular bands) and 12 attributes (small white circles): G) four attributes contribute to the domain Psychosocial Experiences: Presence (PE-A1), Interactions (PE-A2), Emotional States (PE-A3), Communication (PE-A4); H) three attributes contribute to the domain Higher-Order Function: Cognition (HOF-A1), Awareness (HOF-A2), Volition (HOF-A3); I) three attributes contribute to the domain Development: Psychosocial Development (DE-A2), Cognitive Development (DE-A2), Psychosocial Development (DE-A3); J) two attributes contribute to the domain Environment: Physical Environment (EN-A1), Psychosocial Environment (EN-A2).


Level 4: Society & Community (orange circle) consists of 1 domain (circular bands) and 2 attributes (small white circles): N) two attributes contribute to the domain Society & Community: Values and Structure (SC-A1), Availability & Access to Support (SC-A2).