Post-Thrombotic Syndrome in Pediatrics: Development of a Measurement Index

by

Maria Laura Avila

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy in Clinical Epidemiology

Institute of Health Policy, Management and Evaluation

University of Toronto

© Copyright by Maria Laura Avila, 2017
Abstract

Objectives: Post-thrombotic syndrome (PTS) is the most common long-term complication of upper and lower extremity (UE/LE) deep vein thrombosis (DVT). The objective of this thesis was to develop an index for diagnosis and severity rating of pediatric PTS.

Methods: Separate sub-indexes were developed for UE and LE. First, a preliminary study was conducted to determine normal values of limb fluid content in children, using bioimpedance spectroscopy (BIS, Study 1); this information was used for item piloting in Study 2. Next, the signs and symptoms (items) to be included in each sub-index were obtained from a survey answered by an international panel of pediatric thrombosis experts and interviews to pediatric patients with PTS; the items were then defined and piloted in children with UE or LE-DVT to test their diagnostic performance and correlation with PTS diagnosis and severity, assessed by a pediatric thrombosis expert (Study 2). Last, the final items of the sub-indexes were voted by the expert panel using a Delphi process; a scoring system was then developed taking into account item importance to health care providers and to patients/parents, using multiple-criteria decision analysis (MCDA); the meaning of the scores was explored by comparing scores with PTS...
diagnosis and with parental satisfaction/dissatisfaction with the clinical condition of their child (Study 3).

**Results:** Normal BIS values were obtained from 223 healthy children. Twenty-three experts and 16 patients with PTS identified 34 items, 32 of which fitted our measurement framework and were tested in children with UE-DVT (n=70) or LE-DVT (n=70). Aided by the obtained information on the diagnostic performance of each item, the pediatric thrombosis experts voted to retain 10 and 7 items for the LE and UE sub-indexes, respectively (81% response rate). Item weights were developed using MCDA (122 participants, 82% response rate). Both sub-indexes showed good discriminative ability to differentiate between patients with and without PTS. Higher scores were associated with parental dissatisfaction with the clinical condition of their child.

**Conclusion:** We present an index for diagnosis and severity rating of pediatric PTS, developed using expert consensus. The tool provides a meaningful score and has proven diagnostic properties.
Acknowledgments

Since “the process of time is a web of effects and causes” (*Borges, In Praise of Shadow*), it would be impossible to accurately acknowledge all the people who knowingly and unknowingly led me here and helped me develop this work.

I am especially thankful to Dr. Brian Feldman, my supervisor, who constantly challenged me to think further, and to Dr. Victor Blanchette, the most remarkable legacy builder I have ever met.

I am deeply grateful to the multiple builders of this thesis, some of whose names appear on the following pages. Dr. Jennifer Stinson, Dr. Alex Kiss, Dr. Leonardo Brandao, and Dr. Leigh Ward are among them, and I thank them for so generously sharing their tips of the trade.

I would also like to acknowledge Madeline Montoya, who stood by me through the despairs and glories of this work.

My gratitude also extends to the funders of the study, the Thrombosis Team, and to the many anonymous study participants who selflessly dedicated their time to develop an idea.
7. Blessed is he who does not insist on being right, for no one is or everyone is.

*Jorge Luis Borges*

Fragments of an Apocryphal Gospel, In Praise of Darkness, 1969
# Table of Contents

Acknowledgments ........................................................................................................ iv  
Table of Contents ........................................................................................................ vi  
List of Tables ................................................................................................................ viii  
List of Figures ................................................................................................................. ix  
List of Appendices ......................................................................................................... x  
Thesis Outline ............................................................................................................... xii  

Chapter 1 Purpose and Rationale of the Study ................................................................. 1  
1.1. Chapter Overview ........................................................................................................ 1  
1.2. Purpose of the Thesis ................................................................................................... 1  
1.3. Venous Thromboembolism ......................................................................................... 2  
  1.3.1. Definition and Epidemiology in Adults ................................................................. 2  
  1.3.2. Etiology .................................................................................................................. 4  
  1.3.3. Thromboembolic Events in Children ..................................................................... 4  
1.4. Post-thrombotic Syndrome ......................................................................................... 4  
  1.4.1. Definition ................................................................................................................. 4  
  1.4.2. Frequency in Adults and Children .................................................................... 5  
  1.4.3. Normal Hemodynamics in the Venous Circulation ............................................. 6  
  1.4.4. Pathophysiology .................................................................................................... 6  
  1.4.5. Diagnosis and Severity Rating of PTS ................................................................. 8  
  1.4.6. Measurement of Signs in the Pediatric PTS Assessment Instruments ............... 10  
  1.4.7. Measurement of Symptoms in the Pediatric PTS Assessment Instruments ........ 12  
  1.4.8. Measurement Properties of the Pediatric PTS Assessment Instruments .......... 14  
1.5. Why is Diagnosis and Severity Rating of PTS Important? ...................................... 19  
1.6. Objectives of this Thesis .......................................................................................... 22  

Chapter 2 Theoretical Background ............................................................................... 23  
2.1. Chapter Overview ...................................................................................................... 23  
2.2. Abstract ..................................................................................................................... 25  
2.3. Background .............................................................................................................. 26  
2.4. Methods ................................................................................................................... 28  
2.5. Results ..................................................................................................................... 29  
2.6. Discussion ............................................................................................................... 41  
2.7. Conclusion ............................................................................................................... 44  
2.8. Footnotes ............................................................................................................... 45  

Chapter 3 Study 1 ........................................................................................................... 47  
3.1. Chapter Overview ...................................................................................................... 47  
3.2. Abstract ................................................................................................................... 49  
3.3. Introduction .............................................................................................................. 49  
3.4. Methods ................................................................................................................. 50  
3.5. Results .................................................................................................................... 51  
3.6. Discussion .............................................................................................................. 55  
3.7. Footnotes .............................................................................................................. 58  

Chapter 4 Study 2a ........................................................................................................ 62  

Chapter 4 Study 2b ........................................................................................................ 63
## List of Tables

**Table 3.1.** General Characteristics of the Cohort (Study 1)

**Table 3.2.** Extracellular Resistance (R0/R0) Ratios for Arms and Legs by Age Group

**Table 3.3.** Extracellular to Intracellular Resistance (Ri/R0) Ratios by Body Segment

**Table 3.4.** Difference in Circumference (in cm) for Arms and Ankles

**Table 4.1.** Items Proposed by Patients and Expert Physicians

**Table 4.2.** General Characteristics of the Patients

**Table 4.3.** Item-PTS Correlations, Area Under the Curve, Sensitivity, and Specificity for the UE-DVT Stratum (Items Observed Only)

**Table 4.4.** Item-PTS Correlations, Area Under the Curve, Sensitivity, and Specificity for Each Item, LE-DVT Stratum

**Table 5.1.** Correlation Coefficients for Collateral Circulation and Other Variables for the Upper and Lower Extremity Groups

**Table 6.1.** Item Endorsement in the First and Second Round of the Delphi Survey

**Table 6.2.** Endorsement for Limb Measurement in the First and Second Round of the Delphi Survey

**Table 6.3.** Item Weights for the Upper Extremity Sub-Index

**Table 6.4.** Item Weights for the Lower Extremity Sub-Index

**Table 6.5.** Performance of Sub-index Overall Scores, and of the MVS and MJI Items
List of Figures

Figure 2.1. Direction of Causality in Reflective Models

Figure 2.2. Direction of Causality in Formative Models

Figure 2.3. Multiple Indicators and Multiple Causes (MIMIC) Models

Figure 6.1. Example of the Pairwise Ranking Question Used in the Study (from 1000minds.com)

Figure 6.2. Receiver Operating Characteristic Curve for Upper Extremity Sub-index Scores

Figure 6.3. Receiver Operating Characteristic Curve for Lower Extremity Sub-Index Scores
List of Appendices

Appendix 1.1. Villalta Scale

Appendix 1.2. Clinical, Etiological, Anatomical, and Pathophysiological Classification System

Appendix 1.3. Manco-Johnson Instrument

Appendix 1.4. Modified Villalta Scale

Appendix 2.1. Search Strategies and Terms

Appendix 2.2. Study Flow Diagram Showing Search Results, Screening, and Inclusions/Exclusions

Appendix 3.1. Research Ethics Board Approvals

Appendix 3.2. Consents and Assent Forms, Study 1

Appendix 4.1. Definition and Measurement Procedures for the Signs and Symptoms of Post-thrombotic Syndrome

Appendix 4.2. Item Generation Survey (REDCap)

Appendix 4.3. Consents and Assent Forms, Study 2a and Study 2b

Appendix 5.1. Vein Prominence Scale

Appendix 5.2. Simplified Scoring System

Appendix 6.1. Item Reduction Survey, First Round: Screenshot of the Survey, as Seen by the Respondent

Appendix 6.2. Item Reduction Survey, First Round: Detailed Instrument
Appendix 6.3. Item Reduction Survey, Second Round: Screenshot of the Survey, as Seen by the Respondent

Appendix 6.4. Item Reduction Survey, Second Round: Detailed Instrument

Appendix 7.1. Evolution of Test Validity Theory

Appendix 7.2. Representation of Post-thrombotic Syndrome
Thesis Outline

The rationale of the study and the theoretical background underlying the development of the index are introduced in Chapters 1 and 2.

Chapters 4 and 6 report the different phases of tool development, whereas Chapters 3 and 5 present supportive studies that were relevant for the development of the tool.

The final chapter, Chapter 7, summarizes the salient features of the work as a whole, its limitations, and discusses potential future directions.
Chapter 1

Purpose and Rationale of the Study

1.1. Chapter Overview

This chapter discusses the overall purpose of the thesis, and provides a background review of venous thrombosis and one of its most frequent complications: post-thrombotic syndrome (PTS). The evolution of knowledge on the pathophysiology of these conditions is summarized.

The background review describes the basic aspects of PTS and shows the complex heterogeneity that underlies PTS. This is the necessary backdrop to understand concepts relevant to the measurement of a condition that is solely defined on the basis of signs and symptoms, as well as the importance of assessing PTS. The latter is addressed in the closing sections of this chapter, after familiarizing the reader with PTS.

Adult data are compared to pediatric data throughout the chapter. The contrast provides a reference to understand PTS as a growing problem in the pediatric population.

As holds true for many medical conditions that are more frequent among adults, concepts in thrombosis and PTS have been extrapolated and adapted from adults to children. Pediatric tools for the diagnosis and severity rating of PTS are no exception: they were adapted from adult instruments over a decade ago. However, some of the tradeoffs made in the adaptation process have led to specific problems, which are carefully examined in the chapter and serve as the starting point for the development of the new tool.

Advancements in the field of measurement provide more solid theoretical grounds that can be applied to improve earlier attempts to measure pediatric PTS. That is the overall goal of this thesis.

1.2. Purpose of the Thesis

The series of studies presented herein describes the development of an index for the diagnosis and for the rating of severity of PTS affecting the upper extremity (UE) or lower extremity (LE)
of pediatric patients. In agreement with current international standards, PTS is defined as the signs and symptoms that follow a deep vein thrombosis (DVT). The purpose of the index is to discriminate between pediatric patients with and without PTS, and to evaluate changes in the severity of pediatric PTS over time. Given the existing definition of PTS, we focused on obtaining and measuring signs and symptoms of the syndrome and, therefore, the construct PTS was considered to belong to the Body Functions and Structures component of the International Classification of Functioning, Disability and Health, Children and Youth Version (conceptual framework). We followed a formative measurement approach for the development of this tool. These concepts are further explained in this chapter, as well as Chapter 2 and 4.

1.3. Venous Thromboembolism

1.3.1. Definition and Epidemiology in Adults

The term thrombosis refers to the semi-solidification of blood within a blood vessel in vivo [1]. Typically, venous thromboembolism (VTE) includes DVT, which arises in the deep veins of the extremities, and pulmonary embolism (PE).

VTE is considered a major contributor to global disease burden [2]. In the European Union, the number of VTE-related deaths has been estimated to be more than two-fold the deaths due to acquired immunodeficiency syndrome, breast and prostate cancer, and motor vehicle accidents, combined [3]. VTE is a leading cause of preventable deaths in hospitalized patients and of maternal mortality in the United States [4].

The annual incidence of DVT in the general population is largely affected by age, exponentially increasing from 2-3 cases/10,000 in the 30-49 years age range and 5/10,000 between the ages of 50-59 years, to 20/10,000 in the 70-79 age bracket [5]. It is remarkable, however, that the first undisputed case of DVT affected the lower limb of a young adult, Raoul. The case was reported in 1271 AC by Guillaume de Saint-Pathus [6].

1.3.2. Etiology

The first steps towards the elucidation of the potential etiology of VTE are credited to Richard
Wiseman, who is believed to have postulated in the late 1600s that the phenomenon was secondary to “coagulation of the serum or grumomness of the blood” [1, 7, 8]. From the early 1700s to mid-1800s, the dominant interpretation was that a thrombus was the manifestation of phlebitis, an inflammatory process associated with the presence of pus (white material) [9].

Notably, in the 18th century, the terms inflammation and pus were associated, but did not have the connotation of “infection”, an interpretation that belongs to the post-Pasteur era [1]. Malone and Agutter caution the reader about shifts in the meaning of the terms throughout time: “The words ‘inflammation’, ‘phlebitis’ and ‘pus’ have denoted different entities and carried markedly different connotations at various times in history” [1].

The greatest contribution to the field of thrombosis came by the hand of Rudolph Virchow in the 1850s. He moved away from the theory of inflammation, as understood at the time, towards the “coagulation of blood”, substituting the word “thrombosis” for “phlebitis”. He also gave the word thrombosis its more contemporary meaning.

The current consensus considers VTE a multifactorial disease [10], and classifies its multiple risk factors under three main components: blood stasis, damage to the vein wall, and hypercoagulability. Although these components of the causes of venous thrombosis are classically referred to as the “Virchow’s triad”, it is now acknowledged that the concept was misconstrued from Virchow’s reports on the consequences of PE [1, 8, 11, 12].

Subsequent scientific advances led investigators to weigh the specific contribution of each of the interrelated triad components to the development of VTE. Whereas initially vein wall damage was thought to be less important, the re-discovered interrelation between thrombosis and inflammation – in its modern sense [1] - gave a new meaning to the role of the vessel wall in the etiology of VTE.

New findings indicate that inflammation or local damage can modify the endothelial surface from non-thrombogenic to pro-thrombotic [13]. Inflammation has a prominent role in the development of PTS, as will be explained later.
1.3.3. Thromboembolic Events in Children

The first reports of VTE in children appeared in the literature in the late 1800s [14]. By the mid-1900s, case reports and small case series on pediatric PE documented some events of DVT [15, 16]. In 1975, Jones et al identified only 36 cases of pediatric VTE in hospital discharge codes in Scotland over a 4-year period [17].

Around the same time as the aforementioned pediatric study was published, detection of DVT in adult patients was steadily increasing on account of Dr. Eugene Strandness’ adaptation of Doppler ultrasound to evaluate vascular diseases. In contrast, a similar increase in the frequency of detection of VTE in children was only observed in the last few decades [18]. Whereas the Canadian registry reported an incidence of pediatric PE and DVT of 5.3 cases/10,000 hospital admissions over a 30-month period from July 1990 to December 1992, a study looking into hospital discharge codes in tertiary care centers in the United States reported that the diagnosis of VTE in children increased by 70%, from 34 events/10,000 hospital admissions in 2001, to 58/10,000 in 2007 [19]. This trend can be explained by the better diagnostic techniques now available for the detection of VTE in pediatrics, the improved survival of children with chronic diseases, and the widespread use of central venous lines (CVL), which is the main risk factor for VTE development in pediatrics.

As technology advances, the incidence of VTE in children is expected to continue to increase. In turn, potential complications of VTE, including the development of PE, recurrent thrombosis, PTS, and death [20], are likely to become more prevalent.

1.4. Post-thrombotic Syndrome

1.4.1. Definition

PTS is the most common chronic complication of DVT in both adults and children [21]. A formal definition was only recently proposed: the European Venous Forum described PTS as “chronic venous symptoms and/or signs secondary to deep vein thrombosis and its sequelae”. The International Society of Thrombosis and Haemostasis (ISTH) endorses this definition [22, 23]. In keeping with this description, it could be thought that young Raoul, the first reported case
of VTE mentioned above, might also have been the first case of PTS. Indeed, some time after the miraculous cure of his initial episode, “his foot throbbed a little” [24].

It is of note that the definition of PTS described above is limited in that it does not endorse any specific signs or symptoms. Nonetheless, it allows delineating the conceptual framework of the study, as mentioned above (i.e., Body Functions and Structures component of the International Classification of Functioning, Disability and Health, Children and Youth Version).

The most commonly acknowledged signs and symptoms of PTS are described later in this chapter (Section 1.4.5).

1.4.2. Frequency in Adults and Children

According to estimates, PTS affects 15 to 50% of adults diagnosed with LE DVT [25], and is clinically significant in 10-15% of cases. The mean frequency of UE PTS in adults is 15% (range: 7-46%) [26].

The variability of PTS frequency reported in adults is due to the different criteria used to define PTS. In fact, one study found that the frequency of LE PTS in 124 adult patients examined by a single rater using four classification systems (Widmer criteria, Venous Clinical Severity Score, Villalta Scale, and Brandjes criteria) ranged between 31% and 66% [27].

UE PTS has been less studied and reported in adults, in view of the much lower frequency of UE DVT. Whereas approximately 95% of DVT in adults occur in the LE, only 5% involve the UE [28]. In contrast, 30 to 50% of cases of DVT in children involve the UE, which reflects the role of CVL as a major risk factor for DVT in the pediatric population [29-31].

A systematic review of the frequency of pediatric PTS following LE or UE DVT, reported an overall frequency of 26% [95% Confidence Interval (CI): 23, 28], and a frequency of 17% (95% CI 14, 20) when only considering prospective studies [32]. This systematic review analyzed the frequency of PTS following DVT regardless of which extremity was affected (UE/LE). It was not until more recently that pediatric studies focused on PTS following DVT circumscribed to a single venous territory (i.e., either the UE or the LE). They showed a 49% frequency of PTS in the UE [33] and a 47% frequency in the LE [34].
1.4.3. Normal Hemodynamics in the Venous Circulation

Venous return from the LE in the standing position is accomplished by the action of venous valves, the peripheral muscle pumps (foot, calf, and thigh), and a small effect of the gradient in dynamic pressure generated by the heart. These mechanisms are able to work against gravity and offset hydrostatic pressure returning the blood from the periphery to the heart [35], preventing reflux and fluid accumulation [36]. The peripheral calf pump in conjunction with normal venous valves constitute the most efficient mechanism for blood return, and their combined action decreases venous pressure in the leg during walking [37].

In contrast, venous return in the UE is mainly regulated by the dynamic pressure gradient generated by cardiac pumping [38], and the peripheral muscle pump does not play a significant role [39]. In addition, there are fewer valves in the UE than in the LE, and they are less relevant to the physiology of the UE than the valves in the LE [39].

These physiologic differences have a major impact on the pathophysiology of PTS and, consequently, on the development of the signs and symptoms we attempt to measure, and will be further explained later on this chapter.

1.4.4. Pathophysiology

a. Macroscopic level. LE PTS is a form of secondary chronic venous insufficiency (CVI), characterized by persistent flow obstruction and valvular reflux [39].

The relevance of the combined role of both these components was only understood after a series of seminal studies conducted in the early 1990s by Strandness et al [40-43], the same investigators who revolutionized the field of vascular imaging by introducing the use of Doppler ultrasound, as mentioned earlier. The work of these researchers advanced the then existing notion that centered the pathophysiology of PTS on valvular reflux alone [39]. As with the development of knowledge on VTE pathophysiology, subsequent studies weighed the role of residual obstruction differently, suggesting it may even have a more prominent role than valvular reflux [44].

The combination of obstruction and valvular reflux can result in sustained elevation of venous
pressure (venous hypertension) and changes in shear stress affecting the surface of endothelial cells [45], which in turn lead to a diverse array of clinical manifestations seen in LE PTS, such as edema, skin inflammation, skin induration, skin pigmentation, and venous ulcers.

The pathophysiology of UE PTS is far less understood. Valvular function does not appear to have a prominent role in PTS affecting this territory, which may explain why certain clinical features of LE PTS are not seen in UE PTS, as discussed later [46].

b. Microscopic level. The pathophysiology of PTS is tightly dependent on the relation between the orchestrated inflammatory response observed in the context of DVT and thrombus resolution [47]. During the acute DVT phase, the thrombus reduces blood flow and stretches the endothelium, inducing hypoxia, inflammation, and migration of inflammatory cells to the peri-vasculature [48]. Neutrophil and subsequent monocyte infiltration following acute DVT promotes thrombus resolution and organization [10]. However, both inflammation and recanalization can also result in vein wall injury and valvular reflux by mechanisms not entirely understood [49-51], probably related to the release of growth factors, proteinases, and cytokines [52].

Venous hypertension, the hallmark of all forms of CVI, develops in the chronic DVT phases as a consequence of inadequate vessel recanalization, impaired fibrinolysis, and vein wall stiffness, which result in the ongoing venous flow obstruction described in the previous section [48]. Animal studies have shown that sustained elevated venous pressure results in inflammation and valve remodeling, with loss of function [37].

Venous hypertension can cause endothelial glycocalyx shedding with endothelial cell activation. Activated endothelial cells expose adhesion molecules to leukocytes, release cytokines, and express pro-coagulant, pro-migration, pro-apoptotic, and pro-growth genes [45]. Attracted leukocytes migrate and infiltrate the sub-endothelial extra cellular matrix, where they further amplify the inflammatory response by releasing cytokines [53]. This cascade of events triggers chronic inflammatory injury, alteration in capillary perfusion, interstitial accumulation of degradative enzymes, immune host reaction, and alteration of wound healing and/or of tissue remodeling. These events are thought to lead to the myriad of signs and symptoms characteristic of PTS [48, 49, 52].
1.4.5. Diagnosis and Severity Rating of PTS

Diagnosis refers to the act of establishing the nature of a problem, based on the presence of compatible clinical features. Through the process of diagnosis, we offer patients a label to identify their disease [54]. The concept of severity refers to the degree of the disease, to the extent of organ involvement, and to the risk of an unwanted outcome [55, 56].

There are no laboratory, imaging or functional tests to establish the diagnosis of PTS, which is solely defined by the presence of signs and symptoms [57, 58] combined with objectively confirmed DVT (i.e., by Doppler ultrasound, contrast venography, computerized tomographic venography, or magnetic resonance venography), irrespective of the presence of vein abnormalities detected by invasive or non-invasive tests. As summarized by Henke and Comerota, “imaging is neither part of the definition of PTS, nor does it correlate with severity”[52]. Physical findings are characteristic, and diagnosis is established by visual inspection of the limb [35].

Non-invasive diagnostic imaging such as plethysmography and Doppler ultrasound can reveal venous abnormalities, including valvular reflux, vein obstruction, and calf pump dysfunction [51, 59]. However, although documentation of reflux and obstruction may be a predictor of severe PTS in adult patients, these abnormalities may not present in patients with PTS, and may be encountered in patients without PTS [51, 60]. In consequence, experts stress that PTS should not be diagnosed in the absence of clinical findings, regardless of the results of imaging studies [57, 58].

Several clinical indexes and instruments have been developed to evaluate PTS in adult patients. The Villalta Scale (VS, Appendix 1.1) and the Clinical, Etiological, Anatomical and Pathophysiological classification (CEAP, Appendix 1.2) are the most widely used instruments for evaluating LE PTS [27]. The VS has been modified for the assessment of UE PTS in this population [61].

The structure and measurement purpose of the VS and CEAP are explored in more detail below, since these tools represent the foundations on which pediatric instruments were developed.
a. Measuring PTS in adults. The ISTH recommends the VS as the standard instrument to diagnose and establish the severity of LE PTS in adults. The discriminative and evaluative purposes of the VS are supported by its measurement properties (reliability, validity, and responsiveness) [22]. It assesses five subjective symptoms (heaviness, pain, cramps, pruritus, and paresthesia), and seven objective signs (pretibial edema, induration of the skin, hyperpigmentation, new venous ectasia, redness, pain during calf compression, and ulceration of the skin). Each item, except for ulcers, is scored using a 0 to 3 scale (none to severe). The presence of ulcers is equivalent to a score of 15. A score ≥5 is diagnostic of PTS [22]. A final score of 5-9 represents mild PTS, 10-14, moderate PTS, and >14, severe PTS [22].

As defined by the authors, the CEAP is a descriptive classification system for venous diseases in general [62]. It was developed to provide a basis for uniformity in reporting CVI, rather than a discriminative or evaluative PTS-specific instrument. Its basic form classifies patients according to clinical classes (seven classes), etiologic factors (four classes), anatomic distribution (four classes), and pathophysiological dysfunction (four classes) [62, 63]. Clinical classes are as follows: C0 (No visible or palpable signs of venous disease), C1 (Telangiectasies or reticular veins), C2 (Varicose veins; distinguished from reticular veins by a diameter of 3 mm or more), C3 (Edema), C4 (Changes in skin and subcutaneous tissue secondary to chronic venous disease, C4a pigmentation or eczema, C4b lipodermatosclerosis or atrophie blanche), C5 (Healed venous ulcer), C6 (Active venous ulcer). Each clinical class can be recorded as symptomatic or asymptomatic, although symptoms are not specifically defined.

b. Measuring PTS in children. The Modified Villalta Scale (MVS) and the Manco-Johnson Instrument (MJI) are the current reference instruments proposed by the ISTH for the “definition and outcome measurement” of LE PTS in children [64]. Their use for the “definition and outcome assessment” of UE PTS has been acknowledged, but not explicitly endorsed by the ISTH [65].

The MVS [29] (Appendix 1.3), the most frequently used instrument in pediatric studies [66], is a modification of the VS for adults, intended for the diagnosis and establishment of severity of pediatric PTS. Adaptation of the VS to the pediatric population included the change of the 4-point scoring system utilized to indicate the severity of signs and symptoms (i.e., none to severe)
to a dichotomous score (absent/present) for most items, and the addition of three signs: head edema, increase in limb circumference, and venous collaterals. The five symptoms of PTS measured by the VS (pain, cramps, heaviness, paresthesia, and pruritus), were simplified to: 1) pain or abnormal use, and 2) swelling.

The MJI [67] (Appendix 1.4), adapted from the CEAP classification, is intended for diagnosis of pediatric PTS, and combines clinical findings (edema, collateral circulation, skin changes, and skin ulcer) with the Wong-Baker FACES Pain Rating Scale for the evaluation of pain. The pediatric pain assessment uses parental report for children younger than 7 years and self-report for older patients. Pain is evaluated at three different levels: rest, daily activities, and aerobic activities.

The signs and symptoms (items) of the MVS and MJI and, when available, the measurement approach to assess the items are detailed in the next sections. Items are grouped according to the different aspects of PTS. The knowledge on the pathophysiology of each item is summarized, given that the pathophysiology underpins our subsequent measurement approach. UE and LE are contrasted when applicable.

1.4.6. Measurement of Signs in the Pediatric PTS Assessment Instruments

a. Segmental edema. “Increase in limb circumference”, “pitting edema”, and “swelling” are the items used to determine the presence of limb edema in the MVS and MJI.

Limb edema is the clinical sign common to all instruments developed to measure PTS in both children and adults [29, 67-70], and is one of the most common findings reported in children with UE [33] and LE PTS [34].

The pathophysiology of edema, which has been described in the context of LE PTS, is linked to the inflammatory response associated with VTE. This response leads to destruction of vein valves, venous hypertension, and increased transcapillary filtration. Inflammation opens gaps between endothelial cells, further enhancing the filtration of fluid and extravasation of plasma proteins [71]. Edema develops when the transcapillary filtration exceeds the lymphatic drainage, the “safety valve” that prevents interstitial edema [72]. Deteriorated lymphatic function has been described in patients with severe CVI [72].
The MVS measures “swelling” as a symptom, though the measurement procedure is not explained. In addition, the MVS does not define the technique to measure limb circumference. In contrast, the MJI provides a training video showing how limb measurements should be performed. The technique to determine pitting edema is not described in either tool.

Although the pediatric tools measure limb circumference, measurement of limb volume has been the standard method to determine the presence of edema in adults in other diseases characterized by the presence of localized fluid accumulation, such as lymphedema [73]. There are three main methods to measure limb volume: geometric volume assessment, water displacement, and perometry. These methods only provide a measurement of the total volume of the extremity, which is then inferred to reflect edema. There is, however, a fourth method, bioimpedance spectroscopy (BIS), which allows obtaining information on the fluid content of an extremity [73]. However, BIS has not been used for the measurement of pediatric extremity edema.

b. Observable veins. “Collateral vessels” and “varicosities” are the items included in the MVS for the assessment of observable veins. The MJI includes “dilated collateral circulation of extremity only”.

Whereas the formation of collateral vessels is regarded as a physiologic event triggered in the presence of DVT to bypass the obstruction [39], varicose veins [62] are considered pathologic in nature, with structural or functional abnormalities that result in reflux, incompetent valves, and dilatation of vein walls [74, 75].

Neither the MVS nor the MJI provide formal definitions or measurement techniques for these items. In the MVS, varicosities are scored as absent, moderate or severe, and the distinction among categories depends on the interpretation of the user.

The measurement of observable veins is addressed in Chapter 5, Study 2b.

c. Skin manifestations. The MVS includes the following skin signs: “change in skin color”, “pigmentation”, and “ulceration”. The MJI evaluates “skin changes ascribed to venous disease (i.e., pigmentation, venous eczema)” and “skin changes as [before] with ulceration or superior vena cava syndrome”. The exact meaning of the items is not explained in either scale. In addition, the MVS does not define the difference between the items “change in skin color” and
“pigmentation”. The spectrum of the skin changes that are measured in adult patients with CVI will be use as a guide to illustrate the differences between these items, and will help describe the pathophysiology of skin manifestations.

Skin findings in adult patients are classified as skin pigmentation, skin inflammation, skin induration, and venous ulcers [62, 76, 77]. Although venous hypertension is central to the pathophysiology of CVI and PTS, the relationship between venous hypertension and skin manifestations is not fully understood.

Skin pigmentation is attributed to the deposits of hemosiderin that result from extravasation and lysis of red cells in the interstitium [62, 78], and/or to melanin deposits that are seen in chronic irritation, which could be mediated by inflammatory cytokines [79], and could be influenced by individual differences in the tendency of melanocytes to react to inflammation with hyperpigmentation [79, 80].

Skin induration is secondary to dermal fibrosis, which is in turn thought to be related to the presence of leukocytes in the dermal tissue, or to the presence of pro-fibrotic agents released in the presence of venous hypertension [39]. Venous ulcers are also attributed to the presence of trapped leukocytes in combination with altered wound healing and tissue remodeling [10, 39]. Among the enzymes that are released as a response to venous hypertension, matrix metalloproteinases (MMP) appear to have a more prominent role in the genesis of skin induration and leg ulcers [39, 53]. MMP are proteinases that can degrade and remodel the extracellular matrix [49]. MMP polymorphisms may convey a higher risk of developing some of the clinical manifestations in adult patients [81].

Skin inflammation includes erythema and venous eczema. Patients with eczematosous changes are thought to have an added autoimmune component to their CVI, whereas patients with skin induration show cutaneous changes that are more consistent with chronic inflammation and impaired tissue remodeling [39].

1.4.7. Measurement of Symptoms in the Pediatric PTS Assessment Instruments

In contrast with the VS, which assesses five symptoms [68], the MJI only evaluates one symptom, namely “pain”, and the MVS assesses two symptoms, “pain or abnormal use” and
“swelling”. Whereas the MJI uses a scale (Wong-Baker FACES Pain Rating Scale) to measure pain, the MVS does not specify a structured question or tool to evaluate either of the items.

Pediatric thrombosis experts have pointed out the limited number of symptoms that are assessed in both instruments [64]. This is largely attributed to the difficulties that evaluating symptoms in children of different age ranges entails.

Pain in venous diseases is multi-faceted. Proposed mechanisms once again are centered on local inflammatory mediators, which activate nociceptors in the microcirculation [82]. Interestingly, studies in adults have found a lack of correlation between symptoms (including pain, edema, heaviness, cramps, and paresthesia) and clinical signs, valvular reflux, or systemic markers of inflammation. The pathophysiology of the remaining symptoms of PTS (i.e., heaviness, cramps, pruritus and paresthesia, according to the VS) is unknown [83].

**Key point I:**

The complex underlying mechanisms that result in PTS *signs* have not been fully elucidated. Far less is known about PTS *symptoms*.

In addition, the current state of knowledge on the pathophysiology of signs and symptoms of PTS is almost entirely specific to the LE. In comparison, our understanding of UE PTS is extremely limited.

The wide range and diversity of possible mechanisms leading to a given sign or symptom of PTS, and the differences between UE and LE make it difficult to consider PTS a unique and homogeneous disorder.

In the following sections, the structure and measurement properties of both pediatric tools, the MVS and MJI, are examined. This analysis will disclose the limitations and weaknesses of these instruments, thus providing the rationale for the development of the new index.
Measurement Properties of the Pediatric PTS Assessment Instruments

The quality of measurement instruments is crucial because it translates not only into quality clinical practice but also into quality public health practices and medical research [84]. Quality in health care is defined as the degree to which health care increases the likelihood of a favorable outcome, and is consistent with current knowledge [85].

It is known that there are limitations to the MVS and MJ, and the ISTH advocates for further investigation in order “to identify [their] advantages and disadvantages” [64]. Analysis of the MVS and MJ reveals the following potential shortcomings:

a. Item selection. The MVS uses several items to measure edema, a key attribute of PTS. These items (“swelling”, “increase in limb circumference”, and “pitting edema”) overlap to some extent. A fourth item, “head edema”, shares the same pathophysiological principle (i.e., fluid accumulation), but refers to a different territory of the body. However, most of the remaining items of the MVS (e.g., pain, varicosities, skin ulcers) capture unique information that refers to different attributes or dimensions of PTS. The redundancy of closely related items that measure the same dimension of a construct may lead to results that are overly weighted in certain dimensions, potentially overcalling the severity of the disease. This is particularly relevant to tools in which a final overall score is obtained by simple summation of item scores.

Whereas the goal of psychometricians is to create a homogeneous unidimensional scale, items included in a clinical index might not be expected to be homogeneous, given that they define different attributes of a complex clinical phenomenon, rather than different facets of the same attribute [86]. According to Feinstein, if certain variables are so closely related to each other that one can be substituted for the other, one of these variables can be eliminated because it does not make a distinctive contribution to a clinical index [87]. These ideas, which underlie the clinimetric approach, are remarkably similar to that of formative models used in social research [88, 89]. These concepts are discussed in greater detail in Chapter 2, and are at the core of the present thesis.

b. Operationalization and reliability. Operationalization refers to the specification of how the items or variables of an instrument are defined and measured [90]. Reliability is traditionally
defined as the correlation between scores on two equivalent forms of a test, and indicates the fraction of variance in an observed score that can be attributed to the true score of the underlying construct [91, 92]. In a more general sense, it refers to the consistency of scores obtained in different replications or instances of the testing [91].

According to a pediatric study, the inter-rater reliability of the MVS and MJI was moderate (kappa of 0.60), when applied by non-content experts [93]. When considering individual items, the greatest inter-rater discrepancy was observed in the items requiring limb measurement in young patients. A second study reported a higher kappa (0.88) between non-expert raters when using the MJI (overall score). In the study, the authors attributed the discrepancy to the assessment of collateral circulation [94]. The MVS does not provide a description of the meaning of the items or a standardized procedure to measure them (i.e., item operationalization or operational definitions), which might in part explain the discrepancies. Although the MJI provides more detail on some items, the assessment of collateral circulation is not defined.

The lack of clear operational definitions results in different interpretations of the meaning and measurement technique of each item. In fact, it has been described that judgmental rather than factual decisions may be the prime source of variability in the use of a clinical index, which in turn affects the reliability of the instrument [87].

The importance of defining terms resonates with the words of Malone and Agutter when discussing the shifts of meaning of VTE-related terms across time (Section 2.2). The authors stated: “Rigorous use of words is a sine qua non of scientific discourse. What a word denotes depends on theoretical presumptions underlying its use, so the same word might have different meanings – connotations as well as denotation – in different contexts.”[1].

c. Item sub-scores and overall scores. The items of the MVS are scored as 0 (absent) or 1 (present), except for ulcers (absent=0, present >8) and edema of the head and varicosities (absent=0, moderate=1, severe=2). Assessment of head edema and varicose veins as mild is not included as an option. Moreover, whereas present items (except ulcers) are scored 1 point regardless of intensity, head edema and varicose veins are scored only if moderate or severe. Thus, the weighting of these items is not consistent. Ulcers are weighted differently, as with the original VS.
In terms of severity rating, a final score of 1-3 in the MVS indicates mild PTS, 4-8 indicates moderate PTS, and 9 or higher, severe PTS. In terms of diagnosis, the MVS establishes the presence of PTS based on a score of 1 or higher, which implies that the presence of any sign or symptom, regardless of intensity and frequency (except for ulcers and some severities of varicose veins and head edema), indicates PTS. In comparison, the original VS requires at least one severe clinical finding and one moderate clinical finding to classify a patient as having mild PTS (score of 5), except for the presence of ulcers.

As regards the MJI, signs are scored 1 if present. Although pain is scored on a 0-5 point scale, only 1 point is given if pain of any severity is present, following the interpretation provided by the ISTH Pediatric/Neonatal Thrombosis and Hemostasis Subcommittee [32, 64]. Pain is assessed three times: “pain with aerobic exercise only”, “pain with activities of daily living”, and “pain at rest”. The pain score corresponding to each of these situations equally contributes to the final score, and none is assigned a higher weight (e.g., “pain at rest” and “pain with aerobic activities” are equal). Nonetheless, weights are irrelevant in the MJI, given the way the overall score is read. A final MJI score ≥ 1 results in diagnosis of “any PTS present”, and when both symptoms and signs are ≥ 1, the MJI indicates “physically and functionally significant PTS”. This means that pain of any intensity occurring at any level (either at rest, with exercise, or with activities of daily living), plus any single clinical sign is indicative of “physically and functionally significant PTS”. Lastly, although the MJI inquires whether the pain interferes with activity (yes/no), the answer is not considered in the final score.

The equal weight assigned to most clinical findings in both tools has been addressed in the literature. Pediatric thrombosis experts have recognized that in clinical practice they might intuitively place more emphasis on certain clinical features of PTS, such as pain intensity [95], to determine PTS severity. This differential item weighting is not accomplished in either of these tools.

Since essentially any sign or symptom of PTS results in PTS diagnosis, the interpretation of the overall score obtained in both pediatric tools leads to over detection of PTS as compared to the VS. Raffini et al [66] compared the MVS, MJI, and VS in 44 children, and found that although the proportion of children diagnosed with PTS was identical with both the pediatric tools (66%),
the VS only diagnosed 11% of participants as having PTS. The authors of the study argued that the pediatric tools may be “too sensitive” for the diagnosis of PTS. In terms of PTS severity, according to the MVS 52% of patients had mild PTS, as compared to 9% of patients when using the VS. The authors also suggested considering a third category, “post-thrombotic changes”, to indicate patients with mild clinical features, but likely not severe enough to be diagnosed as true PTS.

d. Validation and implications of scores. An international consensus in the terminology used in health-related patient reported outcomes defined validity “the degree to which a health related patient reported outcome instrument measures the construct(s) it purports to measure” [96]. The original VS was validated in terms of the correlations and associations of the overall score with the score of a questionnaire assessing the degree of interference with daily life. The cut-off scores representing none, mild to moderate, and severe PTS when applying the VS were selected by estimating values that resulted in interference with daily life, measured using an ordinal scale (none to severe); the scores of the VS were shown to have an excellent discriminatory power to differentiate among the different levels of interference [68]. In addition, a number of studies found poorer self-reported health-related quality of life (HRQoL), assessed using generic and disease-specific instruments, in patients with PTS according to the VS; HRQoL scores decreased with increasing severity of disease [97].

In contrast, there is limited research on the validation of the MJI and MVS. The validation process of the MJI consisted of describing the absence of its items in healthy children, and the infrequent occurrence of its items in children with a history of CVL but no VTE [98, 99]. In regard to the MVS, studies have relied on demonstrating its strong correlation and almost perfect agreement with the MJI [66, 93].

In terms of the implications of scores, Luceri et al reported a 64.5% prevalence of PTS among 62 children evaluated using the MJI. According to the authors, the high prevalence of PTS they found was worrisome as it appeared to question the validity of the MJI when applied by non-hematologists [94].

The implications of the MVS scores were analyzed by Brandão et al [93] in a cross-sectional evaluation of 101 children. The authors demonstrated that the MVS and MJI had a high negative
predictive value (NPV, 89% and 94%, respectively) but low positive predictive value (PPV, 53% and 45%, respectively) for diagnosing PTS, as compared with expert opinion. In addition, the results showed no difference in HRQoL scores measured with a generic instrument, when comparing patients with and without PTS. Importantly, all patients were found to have mild PTS, which could explain the results. In a different study, Kumar et al used a self-reported version of the MVS to assess PTS severity and a generic instrument to measure HRQoL, and found similar HRQoL scores in children with mild PTS and without PTS, and significantly lower HRQoL scores in patients with moderate-severe PTS [100].

In summary, as compared to the VS, which uses a score ≥5 for PTS diagnosis, both pediatric tools use a threshold score of only 1 point. Furthermore, whereas mild PTS corresponds with a score of 5-9 in the VS, it corresponds to a score of 1-3 on the MVS. The use of lower cut-off scores in the MVS and MJI as compared with the VS appears to affect the implications of both these pediatric tools.

e. Interpretability. The markedly higher frequency of LE DVT in adults as compared to UE DVT explains why the available instruments were developed specifically to assess LE PTS. However, both the MVS and MJI were adapted to evaluate not only LE but also UE PTS in children. The adaptation process may have affected interpretability of the pediatric instruments.

Due to the physiological and pathophysiological differences in hydrostatic pressure and in venous hypertension between UE and LE explained in the opening sections of this chapter, the severity and frequency of clinical manifestations of UE and LE PTS are expected to be different. Clinical features classically described during the natural course of LE PTS (at least in adults), such as varicose veins, hyperpigmentation, or skin ulcers, would be unusual in UE PTS. As a result, most children are expected to have low scores when using the MVS to assess severity of UE PTS. These low scores can either indicate that UE PTS has indeed fewer clinical manifestations than LE PTS, or that the MVS does not adequately reflect the severity of UE PTS. The latter scenario is compatible with the presence of floor effect. Floor effect has been defined as a restriction of the range of scores at the lower end of the tool [101], and constitutes a limitation to the measuring tool, decreasing its ability to differentiate scores at the bottom of the instrument. In the context of pediatric PTS, such an effect could hamper the ability of the MVS
to rate the severity of UE PTS.

**Key point II:**

The careful selection of the items that should be included in a tool, the development of operational definitions of the items as well as of an adequate scoring system for the items are essential to the construction an instrument.

As seen throughout the last sections of this chapter, these processes have a profound impact on the measurement properties of a tool.

**1.5. Why is Diagnosis and Severity Rating of PTS Important?**

Once PTS is established, particularly if complicated by skin ulcers, it can lead to significant disability and poor HRQoL in adult patients, even poorer than that of patients with other forms of CVI [102]. It has been reported that patients with severe PTS have worse generic physical HRQoL than patients with chronic lung disease, osteoarthritis, and angina [103].

In addition, PTS sustains a particularly high direct medical cost as well as an indirect cost driven by loss of productivity. In 1997, a Swedish study estimated that the average cost of treating DVT complications was 75% of the estimated cost of treating the index DVT [104]. The total health care cost has been estimated to be 32%-50% higher in patients with PTS than in patients with DVT who do not develop PTS [105, 106]. Leg ulcers have been associated with time lost from work and job loss, particularly in young adults [107].

The negative repercussion of PTS is largely due to the limited options to treat this syndrome. PTS treatment is challenging and frustrating as it can only target the improvement of symptoms, rather than restoring the normal physiology of the venous system. In fact, there is no effective treatment for PTS to date. For this reason, PTS prevention is considered paramount and is the focus of research efforts, especially among adult patients. Two modalities have been used in this context:

**a.** Until recently, graduated elastic compression garments were indicated for the prevention of
PTS in adults at risk of developing this complication (i.e., patients who sustained DVT) [28]. Supporting their use, a pooled analysis of randomized controlled trials (RCT) conducted between the 1990s and the early 2000s [69, 108-111], which included a total of 628 patients (320 assigned to the intervention group, 308 to the control group), concluded that compression garments reduced the incidence of PTS [112] by 50%. Even though the studies differed at several levels (type and pressure of the garment, time to starting the intervention, length of follow up, blinding, and comparator group – placebo vs. no intervention), they all showed a lower absolute frequency of PTS in patients assigned to the compression garment arm. For more than 10 years these studies provided the rationale for the recommendation of prescribing compression garments to prevent PTS as an evidence-based therapeutic option [28, 48]. However, the efficacy of garments for PTS prevention in adults at risk has since been challenged. The SOX trial [113], a more recent and larger RCT that enrolled 806 patients (410 assigned to the intervention and 396 to placebo garments), showed compression garments to have no effect. Although poor patient compliance has surfaced as one of the potential underlying factors that could explain these conflicting results [36, 114, 115], the findings of the SOX trial shifted the opinion of experts, as reflected in the change in the recommendations regarding the indication of compression garments with preventive purposes [116]. No pediatric data on the use of garments to prevent PTS are available.

b. Aggressive treatment of the index thrombosis using catheter directed or pharmaco-mechanical thrombolysis is now being intensively investigated in adult patients, and it appears to have a role in the prevention of PTS in selected cases in this population [13, 103]. The results of the Acute Venous Thrombosis: Thrombus Removal With Adjunctive Catheter-Directed Thrombolysis (ATTRACT) trial and the CAtheter Versus Anticoagulation Alone for Acute Primary (Ilio)Femoral DVT (DUTCH CAVA) trial are anticipated to address this pressing issue [48, 117].

The role of thrombolysis for the prevention of PTS in children is unclear. Published experience consists of a few small case-series and cohort studies that investigated the safety and efficacy in terms of short-term outcomes (thrombosis resolution). A study by Goldenberg et al showed a lower frequency of PTS among nine patients who received this preventive treatment, as compared to 13 children receiving anticoagulation alone [99]. In contrast, review of the data on
500 children seen at our institution showed no benefit from the use of this therapeutic modality in the prevention of PTS in either UE [33] or LE [34] DVT, although confounding by indication cannot be ruled out. Nevertheless, thrombus resolution was found to be a predictor of PTS in our patients, which is not surprising, given the link between thrombus resolution and inflammation and the bearing of both on the pathophysiology of PTS. This finding suggests that, if thrombolysis leads to complete resolution of VTE, it could prevent PTS.

However, there are some risks associated with thrombolysis. One study reported minor and major bleeding in 6% and 3% of children who underwent directed thrombolysis [118]. In our local experience, mechanical/pharmaco-mechanical and catheter-directed thrombolysis for treatment of UE and LE DVT was associated with minor bleeding in 17% of cases and with major bleeding in 7% of children. Major bleeding is defined by the presence of fatal bleeding, clinically overt bleeding with a hemoglobin decrease ≥ 20 g/L, bleeding on specific sites, or requiring surgical intervention [119].

Despite the paucity of data on safety and efficacy, thrombolysis is still used to treat limb DVT in pediatric patients in order to prevent PTS. In a survey conducted among members of the American Society of Pediatric Hematology/Oncology (ASPHO) in 2008, 81% of 285 respondents answered they had seen at least one child who had undergone thrombolysis (systemic or directed) for treatment of thrombotic events at their institution in the previous year, and 5% of respondents reported that their institution had treated more than 15 cases in the previous year [120].

The growing incidence of DVT in children will result in an increase in the number of PTS cases in the near future. The more prolonged life expectancy of children as compared to adults anticipates that they will endure long-term sequelae of PTS for a longer period of time [121]. Lastly, it remains to be explored whether PTS in children limits and interferes with normal development. All these considerations might explain, at least in part, the increasing trend to use upfront aggressive therapies for the prevention of PTS in pediatrics, despite the limited evidence to support their use. Hence, the expected impact of PTS in children highlights the need of pursuing further research on this topic. In this regard, a group of investigators is currently planning a RCT to assess the efficacy of thrombolysis in preventing PTS in children [122].
However, to successfully conduct further research in this area, better tools for the diagnosis and evaluation of severity are required.

1.6. Objectives of this Thesis

In view of the aforementioned limitations to the pediatric instruments currently in use, especially the limited item operationalization and item and overall scoring, a new and rigorously developed clinical index is urgently needed. Our goal was to develop a tool that allowed diagnosis of PTS in the UE and in the LE, as well as evaluation of PTS severity, following the key points highlighted above. Such a tool would be instrumental to future research on the natural history, prevention, and treatment of this syndrome.

The specific objectives of this thesis were to define which items should be included in the new index (item generation), to develop a definition and measurement technique for each item (item operationalization), to evaluate the diagnostic properties of each item and the correlation of items with PTS severity (item piloting), to select the items that should be included in the final UE and LE sub-indexes (item reduction), to develop a scoring system for the tool (item weighting), and to test the meaning of the scores (Chapters 4 and 6).

To achieve these objectives, additional groundwork was required:

We explored and compared different measurement models and reviewed the options proposed in the literature for summarizing instrument scores, which is a necessary step in the design of an instrument (Chapter 2).

In view of the central role of edema in PTS, we also sought to explore the use of segmental BIS to measure normal fluid content in the extremities of healthy children, in order to determine a pediatric protocol and establish normal ranges in pediatric patients (Chapter 3).

Lastly, we assessed the effect of simplifying the measurement of one of the items, collateral circulation, by testing the inter-rater reliability of the simplified measurement (Chapter 5).

The studies are presented in the order they were executed throughout the development of this thesis.
Chapter 2
Theoretical Background

2.1. Chapter Overview

This chapter focuses on the theoretical concepts that are relevant to the development of the new tool.

The first part of the chapter describes key aspects that underlie the different measurement models (in other words, the connection between what needs to be measured and how it is measured), comparing the traditional psychometric approach to less traditional models.

The second part of the chapter addresses how an overall score can be obtained in a manner that is consistent with the measurement model. Exploring the theory that supports the different scoring options is pertinent, since planning the design of a tool also involves selecting a suitable scoring system.

The chapter is an original publication reproduced verbatim. The footnotes indicated with roman numerals are comments relevant to the thesis and are shown at the end of the chapter (Section 2.8).
A Critical Review of Scoring Options for Clinical Measurement Tools

(As published in BMC Research Notes 2015; 8: 612. Open access)

Authors: Maria Laura Avila\textsuperscript{1}; Jennifer Stinson\textsuperscript{1}; Alex Kiss\textsuperscript{2}; Leonardo R. Brandão\textsuperscript{1}; Elizabeth Uleryk\textsuperscript{1}; Brian M. Feldman\textsuperscript{1}

\textsuperscript{1} The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

\textsuperscript{2} Department of Research Design and Biostatistics, Institute for Clinical Evaluative Sciences, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Correspondence: Brian M. Feldman, Division of Rheumatology, Hospital for Sick Children, 555 University Ave., Toronto, Ontario, Canada, M5G-1X8. E-mail: brian.feldman@sickkids.ca. Tel: 1-416-813-5828; Fax: 1-416-813-4989
2.2. Abstract

**Background:** The aim of this paper is two-fold: 1) to describe the fundamental differences between formative and reflective measurement models, and 2) to review the options proposed in the literature to obtain overall instrument summary scores, with a particular focus on formative models.

**Methods:** An extensive literature search was conducted using the following databases: MEDLINE, EMBASE, PsycINFO, CINAHL and ABI/INFORM, using “formative” and “reflective” as text words; relevant articles’ reference lists were hand searched.

**Results:** Reflective models are most frequently scored by means of simple summation, which is consistent with the theory underlying these models. However, our review suggests that formative models might be better summarized using weighted combinations of indicators, since each indicator captures unique features of the underlying construct. For this purpose, indicator weights have been obtained using choice-based, statistical, researcher-based, and combined approaches.

**Conclusion:** Whereas simple summation is a theoretically justified scoring system for reflective measurement models, formative measures likely benefit from the use of weighted scores that preserve the contribution of each of the aspects of the construct.
### 2.3. Background

From a holistic perspective [123], measurement has been described as an empirical process of “assigning numbers to objects or events according to a rule” [124] as well as an intellectual activity of “giving meaning to the theoretical variables”.

A measurement model describes the relationship between a *construct* and its *indicators*[125]. A construct can be defined as an abstract phenomenon of interest, and indicators as the observable elements used to assess this construct [125, 126]. For example, *melancholia* is a construct, and “depressed mood”, “tiredness”, and “sleep disturbance” are some of the indicators used to assess melancholia [127].

*Psychometrics*, or the study of the theories and techniques concerned with the measurement of mental manifestations and phenomena [127, 128], has influenced the design of the measurement tools used in social and health sciences for more than a century [129]. However, it has been stated that “the foundations of psychometric theory are full of theoretical tensions and fissures that mostly go unnoticed in the daily activity of test construction and use” [130].

One of these fissures, which has received increasing attention for the past three decades, is the meaning of indicators in a measurement model.

In general, instruments developed under psychometric theory (typically for the measurement of mental characteristics[131]) aim to capture the entirety of an underlying construct [132, 133], for example *melancholia*. A battery of homogeneous and positively intercorrelated indicators are thus selected because they all reflect the construct being measured [88] – for example, “tiredness” and depressed mood” may be items of a melancholia scale. As defined by Fayers, homogeneity refers to the fact that the indicators are expected to equally tap into the same construct [88, 134]. However, the assumption that the indicators used in a measurement tool are homogeneous and positively intercorrelated does not hold true in some cases [135]. For example, the construct *life stress* can be measured by indicators such as “job loss”, “divorce”, and “death in the family” [129]. In contrast to the indicators used to assess melancholia, each of these indicators can be seen as a more distinct and unique aspect of the construct.
It was in the social sciences that indicators that were not necessarily homogeneous and positively intercorrelated were first formally used in their measurement tools (in view of the specific characteristics and different nature of the constructs studied in this field). These indicators were termed “cause/causal” indicators, as opposed to “effect” indicators, which prevail in the psychometric tradition [136, 137]. Indeed, in the 1960s, Curtis et al noted that the traditional psychometric approach was not fully appropriate to measure aspects of research in sociology – in which there were valid but unrelated or even inversely correlated indicators of the same construct [137]. The differences between the types of indicators were further explored in the field of sociology by Hubert Blalock Jr., who was the first to describe the distinction between cause (formative) and effect (reflective) indicators [136, 138]. Similarly, in the field of marketing, cause/causal and effect indicators [126] were adopted and referred to as “formative” and “reflective”, respectively. More recently, the terminology of formative and reflective indicators was introduced into the health sciences in the 2000s by the work of Fayers and Hand [88] for the measurement of Quality of Life (QoL).

Whereas reflective models represent the classical concept of measurement used in psychometrics [139, 140], formative measurement models were proposed as an alternative to measure constructs for which the application of a traditional reflective measurement approach would have violated its theoretical foundation. Formative models apply to constructs that are represented by different facets (domains or dimensions) [133], so that constructs in formative models are not unidimensional, but rather result from the combination of heterogeneous indicators [129, 141].

Understanding the difference between reflective and formative measurement models is highly relevant during the development of a measurement tool. The choice of the scoring method is an important step in the development of an instrument and should be consistent with the choice of a measurement model. The scores of a tool are in fact an essential component of the validity of the instrument. Messick defined validity as a property not of the test, but of the meaning, interpretation, and implications of the test scores [142]. Therefore, decisions regarding the choice of a scoring system are deeply attached to the nature of the construct, and have implications for the validity of any instrument. Researchers developing a measurement tool should be aware of the different perspectives regarding measurement models and their impact on scoring systems, in order to decide which approach better corresponds to his or her objective.
The objective of this paper is to offer a brief summary of the fundamentals of formative and reflective measurement models, and to review the different approaches used to obtain summary scores that have been proposed in the literature.

This review is particularly intended for the clinical researcher and practitioner since it focuses on the less traditional formative models \[^{iii}\], which may be of more value in the clinical setting.

2.4. Methods

An extensive literature search was conducted with the assistance of an experienced research librarian to identify technical papers or manuscripts that have described and/or discussed the issue of formative and reflective models. The search strategies and terms are shown in Appendix 2.1.

The searches were run using 1) the OvidSP search platform using the following databases: MEDLINE, EMBASE, and PsycINFO; 2) the EBSCOHost search platform using the following database: CINAHL and 3) the ProQuest search platform using the following database: ABI/INFORM to include articles indexed as of February 25, 2013. The references of identified articles were screened for additional studies.

All articles discussing conceptual issues related to scoring methods in formative and reflective models were included in this narrative review in order to address the second objective.

It is worth noting that although the literature was searched in a systematic manner and all the papers matching the inclusion criteria were retrieved, the theoretical and abstract nature of the subject of the present study did not allow following some of the usual steps involved in a systematic review. For example, the PRISMA checklist and tools for assessing risk of bias were developed to assess health-related interventions or outcomes, and cannot be used in the setting of our study. For this reason, the term “systematic” was avoided when describing the methodology followed herein.

Ethics approval was not required for this study. All the data collected are presented in the manuscript.
2.5. Results

Part I. Theoretical foundations (fundaments) of reflective and formative measurement models

The distinction between formative and reflective models is not only of theoretical nature; it has implications in the design and validation of measurement instruments [143].

The **reflective measurement model** stems from classical test theory (CTT), and is the basis for factor analysis [144]. According to CTT, the observed score (O), or test score obtained from a measurement instrument, comprises two parts: the true underlying score (T), which represents the hypothetical unobservable value that a subject has for a construct, and random error (E), which is the part of the observed score that can be attributed to measurement error [145]:

\[ O = T + E \]  \hspace{1cm} (1)

Consistent with CTT, the observable indicators \( y_i \) in reflective models are considered to be a manifestation of a hypothetical construct (or latent variable) \( \eta \).

\[ y_i = \lambda_i \eta + \varepsilon_i \]  \hspace{1cm} (2)

where \( \lambda \) represents a coefficient capturing the effect of the construct \( \eta \) on an indicator \( y_i \), and \( \varepsilon_i \) represents the measurement error for \( y_i \) [125]. Thus, according to this regression model, the observable indicator \( y_i \) is a function of the latent variable \( \eta \) and of measurement error \( \varepsilon_i \). Variation in the scores of the indicators is assumed to be a function of a true score plus measurement error at the indicator level [146, 147] (equation 1 and 2).

Coltman et al distinguish reflective and formative models based on theoretical and empirical features [148]. Following are the characteristics of reflective models:

- **Nature of the construct**: The underlying latent construct is thought to exist separately from its measures [133]. This concept is akin to “philosophical realism”, and it will be further examined in the discussion section.

- **Direction of causality**: The direction of causality flows from the construct to the indicators (Figure 2.1). A critical aspect of these models is that an *underlying construct*...
influences its indicators [149], and changes in the underlying construct are reflected by simultaneous changes in all the indicators.

Figure 2.1: Direction of Causality in Reflective Models

Legend: $\eta$ represent the construct; $y_i$, the observable indicators; $\lambda$, the coefficients linking the construct $\eta$ to the indicators $y_i$; $\varepsilon$, the error term associated with $y_i$.

- **Characteristics of indicators and indicator intercorrelation**: Because the underlying latent variable or construct influences the indicators, the indicators are intercorrelated. Thus, covariance among indicators reflects variation in the latent variable. Moreover, it is expected that all the indicators will have a high positive correlation and high internal consistency [148-151].

- **Measurement error**: Reflective models include an error term that, as shown in equation (2), is associated to each indicator. Edwards defines this term as “uniqueness” of the indicator, which combines measurement error and indicator specificity [133].

- **Indicator relationship with construct antecedents and consequences**: The meaning of a construct depends not only on its relationship with its indicators, but also on its relationship with other constructs to which it is connected through a complex network of
interlocking laws, known as a nomological network [123]. These laws can link constructs to other constructs (e.g., the construct of self-esteem to the construct of emotional stability), constructs to observed measurement (the construct of self-esteem to the measurement of positive attitude towards self), or observed measurement to observed measurement (the measurement of positive attitude towards self to the measurement of being satisfied with self) [152]. The nomological network helps define a theory, where the meaning of a construct is dependent on its antecedents, or causes, and on its consequences, implications or results. Because the indicators of a reflective model are assumed to be interchangeable, the theoretical implication is that they have a similar relationship with the antecedents and consequences [141].

**Formative models** abandon the idea of a single latent variable causing all the indicators, assuming, essentially, the opposite – that in certain cases the indicators jointly determine the meaning of the construct. Therefore, this model has *indicators x causing the underlying construct η* [129, 153]:

\[ \eta = \gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_i x_i + \zeta \]  

(3)

where \( \gamma \) represents the effect of the indicator \( x_i \) on the underlying construct \( \eta \). \( \zeta \) is a disturbance term that represents all the remaining causes of the construct that are not explained by the indicators [125]. As opposed to equation (2), the construct \( \eta \) is the dependent variable, which is explained by its indicators \( x_i \).

Based on the criteria delineated by Coltman et al [148], the characteristics of formative models are as follows:

- **Nature of the construct**: The construct being measured is defined (formed) according to the indicators the researchers select to measure it.
- **Direction of causality**: The relationship flows from the indicators to the construct, as shown in Figure 2.2.
Figure 2.2: Direction of Causality in Formative Models

Legend: \( \eta \) represent the construct; \( y_i \), the observable indicators; \( \gamma_i \), the coefficients indicating the contribution of \( x_i \) to the construct \( \eta \); \( \zeta \), the disturbance term; \( r \), the correlations between \( x_i \).

- **Characteristics of indicators and indicator intercorrelation:** It is a change in the indicators that determines a change in the value of the underlying construct [141]. However, a change in one indicator is not necessarily accompanied by a change in all indicators. A typical example of this model is socio-economic status (SES) [154], which can be defined as a combination of occupation, education, residence, and income: If one of the indicators changes, SES changes, but if SES changes, not all indicators will necessarily change. There are no specific expectations about the correlations between/among formative indicators: they may display positive, negative, or zero correlation. Positive correlations may exist only because the indicators are capturing the same concept. Determination of internal consistency is therefore not appropriate, and indicators are not interchangeable as each captures a specific aspect of the construct. Therefore, elimination of one indicator carries the risk of changing or affecting the meaning of the construct [125, 135, 141, 148, 150].
• **Measurement error:** Formative models do not incorporate measurement error, but they specify a disturbance term at the construct level which, as noted above, represents all the aspects or determinants of the construct that have not been specified [155].

• **Indicator relationship with construct antecedents and consequences:** Because of their potential heterogeneity or diversity, indicators of formative models do not necessarily have the same relationship with construct antecedents and consequences [148]. Each indicator of formative models conveys unique and distinct information. Importantly, this difference between measurement models as regards the relations of the construct with antecedents and consequences should affect the approach to obtain the overall summary of an instrument, as will be discussed in the next section.

**Part II. Scoring methods in reflective and formative models:**

Results of the search: 1104 citations were retrieved (Appendix 2.1). References were saved in an EndNote X6 library, which was used to identify 357 duplicates. The remaining 747 unique references were reviewed against the inclusion criteria; 136 were retrieved in full for assessment. Finally, 23 unique references offered methodological perspectives on the approach to obtain summary scores in formative and reflective models, and constitute the core of this review (Appendix 2.2).

Synthesis of results: In reflective models, the underlying construct determines the score of each indicator [156], whereas in formative models, the indicators are the determinants of the underlying construct. This difference in the relationship between indicators and construct influences the methods used to obtain an overall score, and applies to instruments that consist of more than one indicator (i.e., multi-indicator or multi-item instruments) [157]. As most available scaling guidelines and textbooks refer to the development of reflective models, we will pay special attention to the methods pertaining to formative models. The scoring concepts that apply to reflective models are explained briefly to better understand the theory behind score generation in formative models.

**Reflective models:**
According to measurement theory, in reflective measurement models the underlying construct contributes to each indicator, and each indicator is an estimate of the construct. As such, reflective models are most frequently scored by means of simple summation [88, 135, 157]. Summation is one of the most commonly used techniques in social sciences, and its invention is attributed to Rensis Likert [145]. The theoretical foundation for summation comes from CTT. As can be seen in equation (1), the observed score in CTT is considered to be a function of the true score plus random error, which has a normal distribution with a mean of 0. Hence, with the summation of several indicators, error will tend to average to 0 [126, 145]. Thus, summation of the reflective indicators is considered a sensible method of estimation [143]. In this process, individual indicators are given a score, and the scores are then added up.

Scores of instruments with multiple subscales that use different metrics in each one of the subscales can be transformed (standardized). Hence, standardized subscales and subscales that have the same metrics can also be added up, which implies equal contribution (or weighting) of each subscale.

Indicator weighting is employed to gauge the contribution of each of the indicators of an instrument to the overall score. In order to implement weights, indicator scores are multiplied by a factor and then added up; factors can be either chosen by the researcher (“theoretical” or “judgment derived weights”) or obtained from the beta coefficients in a regression analysis, or from factor loadings in factor analysis (“empirical weights”) [135, 157, 158]. Despite its logical appeal, the use of weights in reflective models has been reported to have little impact on results [88, 135, 143]. This holds true particularly for scales with highly intercorrelated and/or a large number of indicators [135, 157]. The low impact of weighting is not unexpected since, according the underlying theory, indicators should be highly intercorrelated and interchangeably important [158].

Instruments developed using Structural Equation Modeling (SEM) techniques [159], and even those based on modern psychometric methods such as item-response theory (IRT), also use aggregate sum scores. Even though IRT models allow more complex scoring approaches, it remains unclear whether these approaches yield superior results, and summation remains a simple viable method [88]. It is important to note that what CTT and the more modern
psychometric methods, including IRT, have in common is that their analyses nearly always assume the use of reflective indicators [138].

Summation is straightforward in scales based on reflective models that capture a unidimensional construct. In these cases, all the items in the scale relate to a single construct and a variation of the global scale score is easily understood to reflect a variation in the underlying construct. Some researchers also advocate for the use of global summed scores in complex multidimensional instruments composed of multiple subscales, particularly when the subscales are highly intercorrelated or when there are concerns about the performance or reliability of a subscale. In such cases, researchers may prefer reporting a total score, since it is based on more indicators [160, 161]. In the context of reflective models, multidimensional instruments are instruments that measure “higher-level” constructs using reflective indicators at all levels. The concepts pertaining to construct structure (i.e., first or second order constructs) [162] are not addressed here, as they are beyond the scope of this work.

Some experts consider that multidimensionality does not necessarily justify the scoring and reporting of subscales, because subscales may not always provide accurate, unique, and reliable information about the corresponding subdimension v [160]. In contrast, other experts highlight the interpretational ambiguities that summed scores can create vi [163] and, therefore, the issues regarding scoring in formative models discussed below may also apply to the scoring of multidimensional scales composed of reflective indicators.

**Formative models:**

There is no consensus about the approach to summarizing formative instruments. Some researchers consider that formative indicators can be dealt with using simple summation to obtain an overall rating [144, 148, 154, 164]; adding up each indicator in an overall score (simple summation) or obtaining an average score dividing the total score by the number of indicators has been proposed in order to facilitate the use of these instruments in applied research [165].

However, a major concern is that whereas aggregation of indicators achieves the objective of model parsimony, the distinct and unique information each indicator provides can be lost [148]. It is the opinion of some experts that when formative indicators are involved, neither simple
summation nor weighted sums are easy to justify, because each indicator refers to a different aspect of the construct [88], and some indicators may be more important than others [157].

In addition to the loss of information, the use of average scores can potentially result in a cancellation effect. Cancellation occurs when there is a high score in one indicator and low scores in the remaining indicators, leading to a lower overall score [88, 166] and obscuring the contribution of indicators that may be of particular relevance. Summation lumps together respondents that have the same overall score, independently of their pattern of indicators [167]. This issue should be considered if discriminating subgroups of patients or respondents is relevant to the objective of the measurement instrument [167].

Howell et al [144] have further elaborated on the issue of loss of information when adding up formative uncorrelated indicators. The researchers explained that the number of possible combinations of the scores of every indicator in an index (e.g., $5^3 = 125$ in the case of an index consisting of 3 indicators, each one measured using a 5-point ordinal scoring system) means loss of information, as there are fewer possible overall results when the individual indicator scores are summed (15 in this case). Moreover, each of the possible 125 combinations may be unique, yet this uniqueness is lost by only considering 15 possible values. When the indicators of a model are highly correlated, the number of observed configurations will be substantially smaller because most configurations will be rather homogenous. This is not necessarily the case for formative indicators, and more possible configurations can therefore be expected [144].

Simple summation implies equal weighting $^{vii}$. Indices that contain relatively more indicators for one particular aspect of the construct in a formative measure are implicitly weighting that aspect differently [166]. The weights of formative indicators convey information about their relative contribution to the construct [168].

Following, are different weighting techniques reported in the literature in the context of formative measures:

1. **Choice-based approach**: It has been suggested in the literature that preferences derived from individuals or groups may be particularly important for weighting combinations in formative models [143]. Preference-based methods such as utility analysis and discrete choice
experiments, and the Schedule for the evaluation of individual QoL have been reported as weighting techniques for formative models viii.

Preference-based methods are based on the judgment of the value that is placed on a particular outcome (e.g., a particular pattern of indicator responses).

The terms preference, values, and utility are linked to these methods, and though sometimes used interchangeably, according to some, they represent different concepts [169]. “Preference” is a more general term that describes the “desirability of a set of outcomes” [170]. According to Drummond et al, “Values” refers to the preferences elicited under conditions of certainty and are evaluated with methods such as rating scales (RS) and time-trade-off (TTO). “Utility” refers to the preferences elicited under conditions of uncertainty and is measured using methods like standard gamble (SG) [169, 170].

RS, TTO, SG are the most commonly used methods to measure preferences. The basic form of RS uses simple scales asking respondents to rate a given health condition (e.g., from 0 to 10). SG and TTO involve choice, exploring the willingness of an individual to take a risk in order to gain a benefit [171]. The SG technique requires the individual to hypothetically choose between a certainty (e.g., continuing life in the current health state) and a gamble (which has a probability of resulting in perfect health or death). As for TTO, the aim of the choice task is to elicit the amount of time a participant is willing to sacrifice in order to avoid a worse condition (e.g., a worse health state). A number of authors have addressed these techniques in detail [172-174].

There is a long-standing debate on which method should be used, in view of theoretical concerns regarding the inconsistency of results and the difficulty of some of the tasks. These considerations highlight the complexity of the human judgment process [171, 173]. Furthermore, it is not yet clear whose preferences should be elicited (e.g., for health scales, whether it should be patients/actual users or the general population) [173, 174].

Regardless of the method the researcher uses to elicit preferences, choice-based techniques are considered to be particularly important for obtaining weights in formative models [143].

According to the results of the present review, two techniques have been used in the context of formative models:

a. Utility analysis has mainly been used in QoL assessment as an alternative to the psychometric approach. According to Lenert et al, utility in this context reflects the willingness of an individual to take risks in order to gain a benefit, and is used as a numeric measure to
address significance in a systematic manner, using the judgment of an individual [175]. Multi-attribute utility theory, which has been used in formative models [176], is an extension of the traditional utility theory, and allows quantifying the utility derived from each attribute and combining utilities in a summary measure [177]. For example, this approach was applied to the Health Utilities Index Mark 2 [176].

b. Discrete choice experiment is a preference-based method that derives from behavioral theory, and has been applied in the context of QoL [178]. The premise is that a construct can be described by its attributes (i.e., relevant factors that affect the decisions of an individual [179]), and the value assigned by individuals to those attributes can be used to elicit the value of the construct [178, 180]. This method can be used to estimate the relative importance or the weights of attributes by using a judgmental task based on paired comparisons [178]. Respondents are requested to choose between paired hypothetical scenarios that compare, for example, attributes related to cancer treatment (e.g., improvement in survival and urinary function). Each paired comparison combines different levels of the investigated attributes (e.g., improvement in survival 4, 8, or 12 years and urinary function unimpaired, somewhat, or severely impaired). Choices are then analyzed using regression methods [180].

c. Schedule for the evaluation of individual Quality of Life (SEIQoL): This method is a quantitative technique that has been used to elicit preferences in health care [174]. It stems from the idea that people define and evaluate the aspects of their lives in different ways, and therefore, they estimate the relative importance of each aspect differently. In short, SEIQoL consists of having respondents nominate the five areas of life that they consider most important, and rate their satisfaction/functioning in each of these areas. Finally, the relative importance, or weight, of each area is determined using the SEIQoL-direct weighting technique – respondents fill in a pie chart in which the weight of each aspect is equivalent to the proportion of each sector of the pie; weights are read on the chart circumference [178, 181].

2- **Statistical approach:** Structural Equation Modeling refers to an expanding family of statistical methods that provide a quantitative test for a theoretical model specified by the researcher. It depicts how a set of indicators relate to a construct and how constructs relate to each other using information about their variances or covariances [182]. The hypothetical relationships that the researcher conceptualizes when specifying a model can be expressed as parameters. (To estimate these parameters, a basic principle states that the number
of unknown parameters cannot be larger than the number of pieces of information provided by the variance-covariance matrix. This concept is known as model identification. The problem is that the basic formative model \textit{per se} is not identified. To achieve identification, it has been suggested that at least two reflective indicators must be added as consequences of the formative construct [162, 183, 184]. When two reflective indicators are added directly to the construct, a multiple indicator multiple cause (MIMIC) model is obtained [133]. Thus, MIMIC models are special cases of SEM proposed to operationalize formative indicators that classically involve reflective indicators \( x_i \), directly or indirectly caused by the underlying construct \( n \) (Figure 2.3), as well as formative indicators \( y_i \). For example, the formative indicators \textit{task performance}, \textit{job dedication}, and \textit{interpersonal facilitation} can be considered different facets of the construct \textit{job performance}, whereas reflective indicators may include indicators such as “\textit{overall, this employee performs the job well}” or “\textit{this employee fulfills job requirements}” [133, 162]. In MIMIC models, the construct is summarized as the sum of the regression coefficients or betas of its formative indicators (i.e., weighted sum) [88].

\textbf{Figure 2.3: Multiple Indicators and Multiple Causes (MIMIC) Models}

![Diagram showing MIMIC model](image)

\textbf{Legend:} \( \eta \) represent the construct; \( y_i \), the observable reflective indicators; \( \lambda \), the coefficients linking the construct \( \eta \) to the reflective indicators \( y_i \); \( \varepsilon \), the error term for \( y_i \); \( x_i \), the observable formative indicators; \( \Upsilon_i \), the coefficients indicating the contribution of \( x_i \) to the construct \( \eta \); \( \zeta \), the disturbance term; \( r \), the correlations between \( x_i \).
However, the adequacy of adding reflective measures to a model in order to achieve identification, independently of the conceptual relevance and impact of these measures in the construct, has been subject of high controversy for the past years. The central problem is that the meaning of the construct in MIMIC models is now a function of both $x_i$ and $y_i$. According to Bagozzi, the construct functions figuratively, linking the information contained in $x_i$ to that contained in $y_i$. This makes the model valuable for the prediction of $y_i$ by $x_i$, but hinders the possibility of interpreting the construct in a meaningful way [185]. Moreover, the choice of reflective indicators $x_i$ can have a profound effect on the construct, because choosing a different set of reflective indicators can substantially alter the empirical meaning of the construct. This issue could create further problems in construct interpretation (i.e., interpretational confounding) [133, 144, 186], which in turn affects the comparability of measurements between/among studies (i.e., generalizability) [186]. All these issues have led experts to challenge the suitability of current approaches to deal with formative models in the context of SEM [133, 187], and to propose alternative models to solve these issues [133, 188]. Hence, MIMIC models should be used with caution in the estimation of formative constructs.

3- **Researcher-determined approach:** This category includes arbitrary, literature-driven (theory), or consensus-based weights. The use of these approaches seems to be supported by the opinion of experts, according to whom data analysis is neither needed nor appropriate to decide how to combine indicators in certain models, and the importance of indicators must be defined not by the data but by the objectives of the researchers developing the instrument [88].

In fact, an approach proposed in the literature to deal with the problem of parameter estimation in SEM is to predetermine the contribution of the indicator to the construct ($\gamma$ in equation (3)) [144]. Experts have suggested that weights could be determined *a priori*, according to the theoretical contribution of the indicators to the construct [144, 189].

According to Cadogan et al, if there is no theory suggesting the contrary, formative indicators should have equal weightings [156]. For example, in the earlier version of the Human Development Index, which combined three areas (longevity, educational attainment, and standard of living), researchers intentionally gave equal weights to each one of the aspects [190]. All these recommendations are in keeping with the underlying theory, as Lee states: “a formative variable is simply a researcher-defined composite of subdimensions, and testing these models is
unnecessary” [188].

4- Mixed approaches:

a. Impact or relevance: Indicators related to symptoms may have particular implications due to individual differences in disease expression and the impact that each symptom can have [158]. Hirsch et al evaluated the impact or relevance of symptoms using logistic regression analysis [191]. In brief, respondents to a disease screening survey underwent a physical exam and had a battery of disease-related tests. Clinical experts, blinded to the responses to the survey, rated each patient’s probability of having the disease by assessing their test results. The experts’ responses were combined using Bayesian methods. Individuals with 50% or higher probability of disease were considered disease positive, whereas the remaining patients were considered controls. A logistic regression model was then used to obtain weights that reflected the importance of each question to predict the outcome, allowing calculation of weighted scores [191].

b. Another approach that incorporated the importance of a domain to the measurement of QoL was proposed by Hsieh. Conceptualizing QoL from a formative perspective, he proposed a variation of simple multiplicative weights for patient-reported outcomes that included both importance and satisfaction scores [192]. However, there is evidence that this strategy may not be superior to unweighted schemes [193], in keeping with the idea that the responses of an individual to indicators measuring satisfaction already include an implicit estimation of the importance of the indicator to the subject [194].

2.6. Discussion

More than a century ago, the pioneer work of Charles Spearman on correlation methods in the study of intelligence established the foundations of CTT and factor analysis [127, 195].

Classical test theory, at the heart of traditional psychometrics, is the foundation of reflective measurement models. It focuses on the observed scores, which are considered to reflect true scores plus random error [196]. Item-response theory is a family of contemporary psychometric methods that seek to explain or predict the performance of an item or indicator as a function of an underlying latent variable or construct [197]. Despite the differences between CTT and IRT, they share some principles – the observable measures (i.e., indicators) are a function of an
underlying construct, variation in the latter precedes variation in the former [144], and all the measures of an instrument share “one and only one” underlying construct [92]. Homogeneity of indicators is a desired property, and statistical methods are used to evaluate this property [88]. The reflective measurement model is based on these principles [198].

Decades ago, however, researchers in sociology recognized that not all constructs can be measured with positively intercorrelated indicators, thus laying the foundations for formative models. These models were later extrapolated to other social sciences, and the theoretical and empirical aspects of formative and reflective measurement models continued to develop [155].

The evolution of concepts explained above shows that the problems and concerns regarding the adequacy of the traditional measurement approach are common to a number of research fields. However, although the theory underlying formative measurement models has reached clinical research, it is not widely known ix. Indeed, the formative approach is seldom used in applied medical research despite the fact that many measurements in this field can be conceptualized as composite indexes.

An important task for the clinical researcher developing a measurement instrument pertains to the choice of a measurement model. This choice is dependent on the ontology (this is, the nature of being or existence) of the underlying construct [148].

From an ontological point of view, if the construct is assumed to exist independent of measurement, it corresponds to the school of philosophical realism, which states that reality is independent of our conceptual schemes or perceptions. On the other hand, if the construct is considered a construction of the human mind and does not necessarily exist independent of measurement, it corresponds to philosophical constructivism [139], in which “the truth is what we create to better negotiate the world of our experience” [199]. Whereas in the reflective model, ascribed to realism, a construct determines its indicators, in the formative model, which is closer to constructivism, constructs are understood to be a summary of the indicators [139].

For example, the construct anxiety is measured as a real entity using correlated questions in the 10-item Anxiety Symptom Scale (i.e., reflective measurement model), and the construct gender inequality is measured using the Gender Inequality Index, a researcher-created tool composed of
heterogeneous indicators such as reproductive health, empowerment, and labor market participation (i.e., formative measurement model). A formative measurement is therefore seen as a theoretical entity that is not real beyond what is defined by the indicators, and that does not exist independent of its measurement [133].

Since the goal of a measurement instrument is to provide a score by combining the values of its indicators, the considerations surrounding the nature of indicators are critical to the result of a measurement tool. In general, it can be said that, whereas reflective measures can be handled by simple summation, formative measures benefit from the use of weighted scores that preserve the contribution of each of the aspects of the construct. We have reviewed different approaches to obtain weights as a means to preserve the relevance of each indicator.

Each of the techniques described here has advantages and disadvantages, and the choice of a weighting method should rest on contextual factors.

There are limitations to our study that must be pointed out. We limited our search to the terms “formative” and “reflective”, since the inclusion of the terms “causal/cause/effect”, which are commonly used the English language, resulted in the retrieval of a great quantity of irrelevant publications. However, the references of the retrieved articles were hand-searched in order to find related and relevant literature.

The present study attempts to disseminate measurement concepts introduced in health research by the work of investigators such as Feinstein, de Vet, Fayers, and Hand [87, 143, 157], while also offering essential concepts in measurement that would allow the healthcare practitioner to better appraise and understand the measurement tools that are used in everyday clinical assessment.

In an era when medicine is centered on the measurement of clinical outcomes [200], with the assessment of patient satisfaction, quality of care, and efficient use of resources providing the evidence that drives modern health care systems [201], the present work was deemed timely and relevant.
2.7. Conclusion

In conclusion, it is important for the clinical researcher to be familiar with the differences between reflective and formative measurement models, including the different approaches to obtaining a summary score. Summary scores are an integral part of the validity of a measurement tool. Whereas simple summation is a theoretically sound scoring method in reflective models, formative models likely benefit from a weighting scheme that preserves the contribution of each aspect of the construct.

List of abbreviations

CTT: classical test theory

SES: socio-economic status

SEM: structural equation modeling

IRT: item-response theory

QoL: quality of life

SEIQoL: schedule for the evaluation of individual quality of life

MIMIC: multiple indicator multiple cause

SEER: surveillance, epidemiology, and end results

Competing Interests

The authors declare that they have no competing interests.

Authors’ Contributions

MLA designed the study, collaborated with the literature search, reviewed the literature, and wrote the manuscript. JS, AK, LRB critically reviewed the manuscript. EU conducted the literature search. BMF designed the study and wrote the manuscript.
Acknowledgments

Dr. Avila was supported by a Baxter Bioscience Fellowship in Pediatric Hemostasis and Thrombosis at the Hospital for Sick Children, Toronto. We thank Dr. Dorcas Beaton for her critical review of the manuscript and Ms. Elizabeth Uleryk for her help with the literature search.

2.8. Footnotes

• i: Reference [123].
• ii: Objective here refers to the concept to be measured and the measurement purpose.
• iii: Sometimes referred to in plural in the literature, e.g., “In formative models, items do not necessarily correlate with each other” [157]; “For instance, although the present validity concept can be applied directly to reflective latent variable models used in psychological measurement, it seems that formative models (Bollen & Lennox, 1991; Edwards & Bagozzi, 2000) do not allow for such application” [140].
• iv: It must be pointed out that high item-item correlation (usually, higher than 0.9) could indicate that items are almost identical and that one of them can be deleted, for example to reduce task burden [157].
• v: “Subscale” refers to a collection of items that is part of a larger instrument; “subdimension” refers to one of the facets of a multidimensional construct.
• vi: To cite the author: “Use of multidimensional scales as if they are unidimensional (i.e., summing or averaging item composites) may result in interpretational ambiguities of the relationships among constructs in a test of theory. That is, if a construct is multidimensional but all item scores are summed/averaged across dimensions into a single composite score and correlated with a criterion variable, such a correlation is ambiguous.”
• vii: For simple summation to imply equal weighting, each indicator should have the same response option.
• viii: The instrument SEIQoL has been described as a method, procedure, or technique by some authors. For example: “The [SEIQoL] is a well-established method of assessing QOL which incorporates the value system of the individual respondent” [181]; “The [SEIQoL]: a Direct Weighting procedure for Quality of Life Domains (SEIQoL-DW)” [202]; “A variation of the [conjoint analysis] rating technique, which shares the same theoretical foundations, is
the schedule for the evaluation of individual quality of life (SEIQoL)”[174].

- ix: An earlier version of this work cited the conceptual correspondence between formative models and clinimetrics, according to the following paragraph: “In the 1980s, Feinstein coined the term clinimetrics to refer to the study of clinical phenomena [203]. Interestingly, the differences between formative and reflective models have been equated to those found between the clinimetric (formative) and psychometric (reflective) assessment approaches [88].”

- x: The choice of one weighting technique over another is difficult to justify, and the researcher should consider his/her goal as well as other factors such as the feasibility of applying these techniques.
Chapter 3
Study 1

3.1. Chapter Overview

This chapter presents some of the groundwork conducted prior to tool development.

The overall aim of the study shown in this chapter was to obtain information on the use of segmental bioimpedance spectroscopy (BIS) in children.

The study allowed establishing ranges of fluid content in the extremities of healthy children, while also developing and perfecting a pediatric protocol for the use of BIS in children.

This work was the first step towards the subsequent investigation of the role of segmental BIS in the assessment of limb edema in pediatric PTS (shown in Chapter 4).

As in the previous chapter, footnotes added to the original publication are shown in Section 3.7.
Normal Values for Segmental Bioimpedance Spectroscopy in Pediatric Patients


Authors: Maria Laura Avila¹; Leigh C. Ward²; Brian M. Feldman¹,³; Madeline I. Montoya¹; Jennifer Stinson⁴,⁵; Alex Kiss⁶; Leonardo R. Brandão¹

¹ Department of Paediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

² School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, Brisbane, Australia

³ Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

⁴ Child Health Evaluative Sciences, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

⁵ Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Ontario, Canada

⁶ Department of Research Design and Biostatistics, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Correspondence: Leonardo R. Brandão, The Hospital for Sick Children, Toronto, Ontario, Canada, M5G-1X8. E-mail: leonardo.brandao@sickkids.ca. Tel: 1-416-813-7758; Fax: 1-416-813-5327
3.2. Abstract

Introduction: Localized limb edema is a clinically relevant sign in diseases such as post-thrombotic syndrome and lymphedema. Quantitative evaluation of localized edema in children is mainly done by measuring the absolute difference in limb circumference, which includes fat and fat-free mass. Bioimpedance spectroscopy (BIS) provides information on the fluid volume of a body segment. Our objective was to determine normal ranges for segmental (arm and leg) BIS measurements in healthy children. Additionally, we determined the normal ranges for the difference in arm and ankle circumference and explored the influence of handedness and the correlation between techniques.

Methods: Healthy children aged 1-18 years were recruited. The ratio of extracellular fluid content between contralateral limbs (estimated as the inter-arm and inter-leg extracellular impedance ratio), and the ratio of extracellular to intracellular fluid content for each limb (estimated as the intracellular to extracellular impedance ratio) were determined with a bioimpedance spectrometer. Arm and ankle circumference was determined with a Gulick II tape.

Results: We recruited 223 healthy children (48 infants, 54 preschoolers, 66 school-aged children, and 55 teenagers). Normal values for arm and leg BIS measurements, and for the difference in arm and ankle circumference were estimated for each age category. No influence of handedness was found. We found a statistically significant correlation between extracellular impedance ratio and circumference difference for arms among teenagers.

Conclusion: We determined normal BIS ranges for arms and legs and for the difference in circumference between arms and between ankles in children. There was no statistically significant correlation between extracellular impedance ratio and difference in circumference, except in the case of arms in adolescents. This may indicate that limb circumference measures quantities other than fluid, challenging the adequacy of this technique to determine the presence of localized edema in most age groups.
3.3. Introduction

Post-thrombotic syndrome (PTS) is a potential long-term complication of deep vein thrombosis (DVT) estimated to affect 1 in 3 adults [59] and 1 in 4 children [32] who have sustained limb DVT.

Since localized edema is a typical clinical feature of PTS, all clinical tools and criteria used to assess PTS in both adult and pediatric patients include the measurement of edema [29, 67-70, 77, 204].

Detection of localized edema is important for the diagnosis and follow-up not only of PTS, but also of other conditions that feature excessive limb fluid content, such as other forms of chronic venous disease and lymphedema[205]. Moreover, early diagnosis of segmental edema has been shown to allow timely therapeutic interventions and better results in patients with lymphedema[206].

Currently, the assessment tools most commonly used in the clinical setting for objective evaluation of localized edema are based on the measurement of either limb circumference or limb volume [207]. The measurement of limb circumference involves determining the difference between the circumferences of both limbs at one or two anatomical points, using a measuring tape. Limb circumference determination has been used predominantly to assess edema in children at risk of PTS [29, 98, 99, 208]. Limb volume can be assessed directly or indirectly using a number of techniques, the four most common being the use of a measuring tape and mathematical formulae, water displacement [205, 209], perometry, and bioimpedance. Although these methods have been shown to be accurate and precise for the assessment of limb volume in adult patients affected by lymphedema [210], less is known about their accuracy in children.

Each method has advantages and disadvantages. Limb volume determination with measuring tape is inexpensive and highly acceptable to patients and clinicians, but it can be tedious and time-consuming to perform, and identification of bony landmarks can be difficult in some patients [211]. Water displacement is a low cost technique, but can be time-consuming, sensitive to movement artefacts, and may be associated with infection control problems [212, 213]. Perometry is a rapid, accurate, and precise method, but the equipment is very expensive, and
may not be suitable for young patients since they have to adopt a stable position for the measurement[210].

Importantly, the major limitation of these three methods is that they provide a measurement of total limb volume, and an increase in volume is then inferred to reflect limb edema. However, the total volume of a limb is composed of not only fluid content, but also fat and muscle. Bioimpedance spectroscopy (BIS) is able to provide information on the amount of extracellular, intracellular, and total fluid content of a body segment, therefore avoiding the inaccuracy introduced by changes in fat and muscle tissue in a swollen limb [210]. An additional advantage of BIS is that several determinations can be taken in a few seconds, which is particularly important in pediatric patients, who may find it difficult to stay still.

The main objective of our study was to determine the reference ranges for BIS measurements for upper and lower limbs (i.e., arms and legs) in a population of healthy children. The secondary outcomes comprised 1) determination of reference ranges for the difference in circumference between both arms and both ankles in these patients, 2) investigation of the influence of handedness on BIS measurements, and on the difference in arm and ankle circumference, and 3) determination of the correlation between BIS measurements and the difference in circumference.

3.4. Methods

Healthy children aged 1-18 years were recruited in this cross-sectional study that was conducted at the Hospital for Sick Children (pilot), and at the Ontario Science Centre, Toronto, ON.

Four age-categories were included: toddlers (1-2 year-olds), pre-schoolers (3-4 year-olds), school-aged children (5-11 year-olds), and adolescents (12-18 year-olds) iii.

Ethics statement: The University of Toronto and the Hospital for Sick Children Research Ethics Boards approved the study. Written informed consent was obtained from parents/legal guardians or participants, as appropriate iv.

The following variables were collected: age, sex, height, weight, handedness, BIS measurements, and anatomical measurements (arm and ankle circumference, as described below).
1. Height was measured using a portable stadiometer to the nearest 0.1 cm; weight was measured with a digital scale to the nearest 0.1 kg. Handedness was determined as per parental or participant’s report. Body mass index percentile or weight-for-length percentile was estimated in participants older or younger than 2 years, accordingly. Body mass index percentiles were estimated using a SAS macro available through the Centers for Disease Control and Prevention website [214].

2. Biompedance spectroscopy measurements: the principle underlying bioimpedance is that the opposition of biological tissues to the flow of imperceptible electrical current passing through is inversely proportional to the fluid volume of the tissue [215]. Biompedance spectroscopy involves the application of a range of frequencies (usually from around 3 kHz to 1 MHz) and the measurement of the impedance or resistance to the passage of this electric current. Whereas the impedance at low frequencies, ideally at zero frequency (R0), reflects the extracellular fluid content alone, the impedance at high frequencies, ideally at infinite frequency (Rinf), reflects total fluid content. Both R0 and Rinf are used to derive Ri, the impedance of the intracellular fluid compartment. Segmental impedance values are determined from impedance measurements made of the body segments according to the principle of equipotentials [216]. Importantly, the measurement of extracellular fluid content of an arm or leg (reflected by the R0 value) is normalized by comparison to the value of the contralateral limb, and expressed as a ratio [217]. The normalization of the R0 values allows accounting for the wide individual biological variation in impedance values, which depends on factors such as length of the arms or legs, exercise, and diet [218]. Hence, it follows that the R0/R0 ratio is only useful in patients with unilateral arm or leg edema [219-221]. However, when patients have bilateral arm or leg edema, the extracellular fluid content can be normalized against the intracellular fluid of the same limb, and because resistance is inversely related to volume, the extracellular to intracellular fluid volume is calculated as the inverse ratio of their resistances (extracellular/intracellular fluid content = intracellular/extracellular impedance ratio = Ri/R0 ratio) [217, 218].

Segmental BIS (arm: wrist to axilla; leg: ankle to groin) was measured following standard protocols, in accordance with the principle of equipotentials [216, 222]. At the time of the testing, the subjects were asked to remove all jewellery (except rings and earrings) and shoes; skin was prepared with an alcohol wipe and allowed to dry prior to placing the surface electrode. Participants aged 3 years and older were positioned lying supine on a non-conductive
examination table, arms at sides with palms facing down, and legs slightly abducted. Toddlers were allowed to sit up during measurements to encourage cooperation. Lightly adhesive silver/silver chloride EKG-style resting electrodes were placed as follows: 1) voltage sensing electrodes were placed on the dorsum of the wrists adjacent to the ulnar styloid process, and anterior to the ankle joints between the malleoli; 2) current drive electrodes were placed distal to the voltage sensing electrodes, on the dorsum of the hands and feet. Impedance measurements were performed with an SFB7 bioimpedance spectrometer (ImpediMed Ltd, Brisbane, Australia). Device calibration was checked regularly, following the manufacturer’s instructions.

3. Circumference measurements were determined with a Gulick II anthropometric tape, a no-stretch retractable tape with a tensioning device that allows application of constant tension, thus minimizing measurement error that can be ascribed to differences in the applied tension [223].

   a. Arm circumference was measured at a point equidistant to the acromium process and the olecranon process. This technique was chosen based on a previous pilot study conducted at our institution, which showed better intra and inter-rater reliability using this technique (unpublished data). Measurements were recorded in cm, rounded to the nearest first decimal place. The absolute difference in cm between right and left arm and between dominant and non-dominant arm were estimated.

   b. Given the exposed setting of the study, and for privacy reasons, we decided to measure the ankles of the participants following the “Figure-of-8” technique [223] instead of measuring mid-calf and mid-thigh circumferences. For this measurement, the participant was positioned lying supine on a table, and (whenever possible) keeping the ankle in neutral dorsiflexion. The Gulick II anthropometric tape was placed midway between the tendon of the tibialis anterior and the lateral malleolus, pulled medially toward the tuberosity of the navicular, then pulled laterally across the longitudinal arches of the foot toward the base of the 5th metatarsal. The tape was then pulled around the ankle joint, encircling an imaginary line below the medial malleolus, across the Achilles tendon and below the lateral malleolus, ending where the measurement started. A video showing how this technique is performed is available [224]. The results were recorded in cm, rounded to the nearest first decimal place. The absolute difference in cm between right and left ankle and between dominant and non-dominant ankle were estimated.
It is of note that arm and ankle circumference measurements of the healthy subjects who were recruited at the Hospital for Sick Children to participate in the pilot study were not taken.

A trained researcher performed all the anatomical measurements and a second trained researcher undertook all BIS determinations. Both researchers were blinded to each other’s results.

**Data Analysis**

Categorical data were summarized as percentages and ratios. Means and standard deviations (SD) or medians and interquartile ranges (IQR) were used to report continuous data depending on data distribution.

The BIS data were processed using the manufacturer’s software (Bioimp version 5.4.0.3, Copyright© 2012 ImpediMed) to obtain extracellular (R0) and intracellular (Ri) resistance values for each limb. The R0/R0 and Ri/R0 ratios for arms and legs were estimated from these values.

The coefficient of variation for replicate R0 and Ri measurements was estimated. The percentile distributions of the R0/R0 ratios for arms and legs and of the difference in circumference for arms and ankles were determined in relation to age group.

The impact of handedness on R0/R0 ratios was investigated within each group category by comparing the right limb R0 to left limb R0 ratio vs. the dominant hand side R0 to non-dominant hand side R0 ratio using paired t-tests. A similar analysis was performed to compare the impact of dominance on the difference in circumferences (right – left circumference and dominant – non-dominant hand side). The potential effect of sex on R0/R0 ratios was explored using t-tests.

The effect of age and hand side dominance on Ri/R0 ratios of each limb was analyzed with two-way analysis of variance (ANOVA).

Lastly, the relationship between arm R0/R0 ratios and arm circumference difference, and between leg R0/R0 ratios and ankle circumference difference was estimated in each age group using the Pearson product-moment correlation coefficient.

Statistical significance was set at $\alpha=0.05$. 
3.5. Results

Two hundred and twenty three healthy children were recruited during the study period. Two percent of the population (n=6) was recruited at the Hospital for Sick Children, and the remaining participants (n=217) were recruited at the Ontario Science Centre, Toronto, Ontario. Age distribution was as follows: 48 infants, 54 preschoolers, 66 school-aged children, and 55 teenagers. The general characteristics of the cohort are presented in Table 3.1.

Table 3.1. General Characteristics of the Cohort

<table>
<thead>
<tr>
<th>Male:female ratio</th>
<th>Toddlers</th>
<th>Preschoolers</th>
<th>School age</th>
<th>Teenagers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1:1</td>
<td>1.1:1</td>
<td>1.3:1</td>
<td>0.8:1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weight-for-length percentile for &lt;2 year-olds, median (IQR)</th>
<th>Toddlers n=20</th>
<th>64 (76)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI percentile for ≥2 year-olds, median (IQR)</td>
<td>Toddlers n=28</td>
<td>62 (33)</td>
</tr>
<tr>
<td></td>
<td>Preschoolers</td>
<td>53 (50)</td>
</tr>
<tr>
<td></td>
<td>School age</td>
<td>44 (47)</td>
</tr>
<tr>
<td></td>
<td>Teenagers</td>
<td>61 (52)</td>
</tr>
</tbody>
</table>

| Handedness, right/left (% left handed)                       | Preschoolers   | 46/4 (8) |
|                                                             | School age     | 53/13 (27) |
|                                                             | Teenagers      | 48/7 (13) |

Legend: IQR refers to interquartile range; BMI, to body mass index

The coefficient of variation of R0 ranged between 0.33% and 0.46%, whereas that of Ri ranged between 1.46% and 4.00%. The distribution of R0/R0 ratios for arms and legs per age category is displayed in Table 3.2.
Table 3.2. Extracellular Resistance (R0/R0) Ratios for Arms and Legs by Age Group

<table>
<thead>
<tr>
<th>Age group</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>P3</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
<th>P97</th>
<th>Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toddlers</td>
<td>48</td>
<td>1.01</td>
<td>0.047</td>
<td>0.93</td>
<td>0.96</td>
<td>0.98</td>
<td>1.00</td>
<td>1.03</td>
<td>1.08</td>
<td>1.12</td>
<td>Right/Left Arm</td>
</tr>
<tr>
<td>Preschoolers</td>
<td>54</td>
<td>1.01</td>
<td>0.037</td>
<td>0.94</td>
<td>0.95</td>
<td>0.98</td>
<td>1.00</td>
<td>1.03</td>
<td>1.06</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>School age</td>
<td>66</td>
<td>1.01</td>
<td>0.033</td>
<td>0.95</td>
<td>0.96</td>
<td>0.98</td>
<td>1.01</td>
<td>1.03</td>
<td>1.04</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>Teenagers</td>
<td>55</td>
<td>0.99</td>
<td>0.044</td>
<td>0.92</td>
<td>0.94</td>
<td>0.95</td>
<td>0.98</td>
<td>1.01</td>
<td>1.04</td>
<td>1.09</td>
<td></td>
</tr>
</tbody>
</table>

| Toddlers      | 48   | 1.01 | 0.048 | 0.93 | 0.96 | 0.98 | 1.00 | 1.03 | 1.08 | 1.12 |      |
| Preschoolers  | 54   | 1.01 | 0.036 | 0.94 | 0.96 | 0.99 | 1.01 | 1.03 | 1.03 | 1.07 | Dominant/Non-dominant Arm |
| School age    | 66   | 1.01 | 0.033 | 0.95 | 0.97 | 0.98 | 1.01 | 1.03 | 1.04 | 1.08 |                  |
| Teenagers     | 55   | 0.99 | 0.045 | 0.92 | 0.94 | 0.96 | 0.99 | 1.02 | 1.05 | 1.09 |                  |

| Toddlers      | 48   | 1.00 | 0.053 | 0.91 | 0.94 | 0.97 | 0.99 | 1.03 | 1.09 | 1.11 | Right/Left Leg    |
| Preschoolers  | 54   | 1.00 | 0.034 | 0.93 | 0.96 | 0.98 | 1.00 | 1.01 | 1.03 | 1.07 |                  |
| School age    | 66   | 0.99 | 0.025 | 0.95 | 0.96 | 0.97 | 0.99 | 1.00 | 1.02 | 1.05 |                  |
| Teenagers     | 55   | 1.00 | 0.032 | 0.95 | 0.96 | 0.97 | 0.99 | 1.02 | 1.03 | 1.07 |                  |

| Toddlers      | 48   | 1.00 | 0.053 | 0.91 | 0.95 | 0.97 | 0.99 | 1.04 | 1.09 | 1.11 |      |
| Preschoolers  | 54   | 1.00 | 0.034 | 0.93 | 0.96 | 0.98 | 1.00 | 1.01 | 1.03 | 1.07 | Dominant/Non-dominant Leg |
| School age    | 66   | 0.99 | 0.026 | 0.95 | 0.96 | 0.97 | 0.99 | 1.01 | 1.02 | 1.04 |                  |
| Teenagers     | 55   | 0.99 | 0.031 | 0.93 | 0.95 | 0.97 | 0.99 | 1.02 | 1.03 | 1.05 |                  |

**Legend:** SD refers to standard deviation, P3 to the 3rd percentile, P10 to the 10th percentile, P25 to the 25th percentile, P50 to the 50th percentile, P75 to the 75th percentile, P90 to the 90th percentile, and P97 to the 97th percentile.

Regarding the influence of handedness on R0/R0 ratios, we found little variation in the distribution of values when comparing the right/left limb vs. dominant/non-dominant hand side. Indeed, paired t-tests showed no statistically significant difference between these two methods to estimate arm and leg R0/R0 ratios (toddlers t=1.0, 47 degrees of freedom [df], p=0.32; preschoolers t=-1.8, 53 df, p=0.08; school aged children t=-0.9, 65 df, p=0.38; teenagers t=-1.0, 54 df, p=0.30 for arms; toddlers t=1.0, 47 df, p=0.32; preschoolers t=-1.2, 53 df, p=0.23; school aged children t=-0.8, 65 df, p=0.43; teenagers t=0.42, 54 df, p=0.68 for legs).

Of note, we found no statistically significant difference in arm or leg R0/R0 ratios between males and females (arms t=1.41, 221 df, p=0.14; legs t=0.32, 218 df, p=0.75).

Table 3.3 shows the distribution of extracellular to intracellular fluid content, expressed as Ri/R0 ratio, for each body segment in each age category. There was a statistically significant difference...
in the Ri/R0 ratios among age groups, with no significant effect of handedness on any of the limbs (Table 3.3).

<table>
<thead>
<tr>
<th>Age group</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>P3</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
<th>P97</th>
<th>Segment</th>
<th>Two-Way ANOVA*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toddlers</td>
<td>48</td>
<td>3.33</td>
<td>0.50</td>
<td>2.17</td>
<td>2.67</td>
<td>3.06</td>
<td>3.39</td>
<td>3.66</td>
<td>3.85</td>
<td>4.21</td>
<td>Right Arm</td>
<td>Age group: F (3,217) = 32.7, MSE=5.1 p &lt;0.001; Handedness: F (2,217) =2.7, MSE=0.4 p= 0.07</td>
</tr>
<tr>
<td>Preschoolers</td>
<td>54</td>
<td>3.26</td>
<td>0.35</td>
<td>2.47</td>
<td>2.79</td>
<td>3.06</td>
<td>3.31</td>
<td>3.52</td>
<td>3.41</td>
<td>3.89</td>
<td>Left Arm</td>
<td>Age group: F (3,217) =27.0, MSE=4.4 p &lt;0.001; Handedness: F (2,217) =1.88, MSE=0.31 p= 0.15</td>
</tr>
<tr>
<td>School age</td>
<td>66</td>
<td>2.94</td>
<td>0.37</td>
<td>2.17</td>
<td>1.49</td>
<td>2.71</td>
<td>2.88</td>
<td>3.12</td>
<td>3.38</td>
<td>4.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenagers</td>
<td>55</td>
<td>2.66</td>
<td>0.38</td>
<td>1.98</td>
<td>2.17</td>
<td>2.38</td>
<td>2.66</td>
<td>2.99</td>
<td>3.16</td>
<td>3.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toddlers</td>
<td>48</td>
<td>3.43</td>
<td>0.48</td>
<td>2.68</td>
<td>2.79</td>
<td>3.05</td>
<td>3.44</td>
<td>3.79</td>
<td>3.96</td>
<td>4.27</td>
<td>Right Leg</td>
<td>Age group: F (3,217) =56.8, MSE=16.6 p &lt;0.001; Handedness: F (2,217) =0.24, MSE=0.07 p= 0.82</td>
</tr>
<tr>
<td>Preschoolers</td>
<td>54</td>
<td>3.39</td>
<td>0.38</td>
<td>2.59</td>
<td>2.93</td>
<td>3.15</td>
<td>3.37</td>
<td>3.63</td>
<td>3.78</td>
<td>4.06</td>
<td>Left Leg</td>
<td>Age group: F (3,217) =39.7, MSE=12.3 p &lt;0.001; Handedness: F (2,217) =0.16, MSE=0.08 p= 0.82</td>
</tr>
<tr>
<td>School age</td>
<td>66</td>
<td>3.05</td>
<td>0.35</td>
<td>2.39</td>
<td>2.60</td>
<td>2.83</td>
<td>3.03</td>
<td>3.21</td>
<td>3.48</td>
<td>3.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenagers</td>
<td>55</td>
<td>2.82</td>
<td>0.43</td>
<td>2.05</td>
<td>2.31</td>
<td>2.54</td>
<td>2.76</td>
<td>3.04</td>
<td>3.43</td>
<td>3.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toddlers</td>
<td>48</td>
<td>3.93</td>
<td>0.66</td>
<td>3.00</td>
<td>3.19</td>
<td>3.56</td>
<td>3.83</td>
<td>4.27</td>
<td>4.59</td>
<td>5.13</td>
<td>Right Leg</td>
<td>Age group: F (3,217) =39.7, MSE=12.3 p &lt;0.001; Handedness: F (2,217) =0.16, MSE=0.08 p= 0.82</td>
</tr>
<tr>
<td>Preschoolers</td>
<td>54</td>
<td>3.35</td>
<td>0.45</td>
<td>2.46</td>
<td>2.75</td>
<td>3.16</td>
<td>3.38</td>
<td>3.61</td>
<td>3.89</td>
<td>4.27</td>
<td>Left Leg</td>
<td>Age group: F (3,217) =56.8, MSE=16.6 p &lt;0.001; Handedness: F (2,217) =0.24, MSE=0.07 p= 0.82</td>
</tr>
<tr>
<td>School age</td>
<td>66</td>
<td>2.84</td>
<td>0.52</td>
<td>1.91</td>
<td>2.40</td>
<td>2.69</td>
<td>2.89</td>
<td>3.14</td>
<td>3.37</td>
<td>3.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenagers</td>
<td>55</td>
<td>2.67</td>
<td>0.54</td>
<td>1.95</td>
<td>2.13</td>
<td>2.36</td>
<td>2.73</td>
<td>3.01</td>
<td>3.30</td>
<td>3.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toddlers</td>
<td>48</td>
<td>4.11</td>
<td>0.80</td>
<td>2.44</td>
<td>3.28</td>
<td>3.60</td>
<td>4.18</td>
<td>4.65</td>
<td>4.95</td>
<td>5.33</td>
<td>Right Leg</td>
<td>Age group: F (3,217) =39.7, MSE=12.3 p &lt;0.001; Handedness: F (2,217) =0.16, MSE=0.08 p= 0.82</td>
</tr>
<tr>
<td>Preschoolers</td>
<td>54</td>
<td>3.63</td>
<td>0.63</td>
<td>2.57</td>
<td>2.95</td>
<td>3.25</td>
<td>3.56</td>
<td>3.85</td>
<td>4.25</td>
<td>5.40</td>
<td>Left Leg</td>
<td>Age group: F (3,217) =39.7, MSE=12.3 p &lt;0.001; Handedness: F (2,217) =0.16, MSE=0.08 p= 0.82</td>
</tr>
<tr>
<td>School age</td>
<td>66</td>
<td>3.06</td>
<td>0.54</td>
<td>1.98</td>
<td>2.52</td>
<td>2.81</td>
<td>3.12</td>
<td>3.32</td>
<td>3.60</td>
<td>4.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenagers</td>
<td>55</td>
<td>2.87</td>
<td>0.63</td>
<td>2.08</td>
<td>2.21</td>
<td>2.47</td>
<td>2.89</td>
<td>3.28</td>
<td>3.18</td>
<td>4.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend: SD refers to standard deviation, MSE, to mean square error, P3 to the 3rd percentile, P10 to the 10th percentile, P25 to the 25th percentile, P50 to the 50th percentile, P75 to the 75th percentile, P90 to the 90th percentile, and P97 to the 97th percentile.

*Two-way ANOVA analyzing the effect of two factors, age group and hand side dominance on Ri:R0 ratios

The distribution of the difference in circumference for arms and for ankles is shown in Table 3.4. As observed with R0/R0 ratios, there was little impact of hand side dominance on the distribution of values of circumference difference: paired t-tests showed no statistically significant difference when comparing right minus left limb vs. dominant minus non-dominant hand side (toddlers t=-1.2, 47 df, p=0.25; preschoolers t=-0.7, 53 df, p=0.51; school aged children t=-1.4, 64 df, p=0.18; teenagers t=0.7, 50 df, p=0.45 for arms; toddlers t=-1.0, 47 df, p=0.32; preschoolers t=-1.7, 53 df, p=0.10; school aged children t=1.25, 64 df, p=0.22; teenagers t=1.16, 50 df, p=0.25 for legs).
Table 3.4. Difference in Circumference (in cm) for Arms and Ankles

<table>
<thead>
<tr>
<th>Age group</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>P3</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
<th>P97</th>
<th>Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toddlers</td>
<td>48</td>
<td>-0.08</td>
<td>0.48</td>
<td>-1.00</td>
<td>-0.70</td>
<td>-0.45</td>
<td>-0.05</td>
<td>0.20</td>
<td>0.50</td>
<td>1.00</td>
<td>Right – Left Arm</td>
</tr>
<tr>
<td>Preschoolers</td>
<td>54</td>
<td>-0.08</td>
<td>0.52</td>
<td>-1.00</td>
<td>-0.70</td>
<td>-0.40</td>
<td>-0.10</td>
<td>0.40</td>
<td>0.60</td>
<td>1.00</td>
<td>Dominant – Non- dominant Arm</td>
</tr>
<tr>
<td>School age</td>
<td>64</td>
<td>-0.08</td>
<td>0.46</td>
<td>-0.90</td>
<td>-0.70</td>
<td>-0.30</td>
<td>0.00</td>
<td>0.20</td>
<td>0.50</td>
<td>1.00</td>
<td>Right – Left Ankle</td>
</tr>
<tr>
<td>Teenagers</td>
<td>51</td>
<td>0.03</td>
<td>0.77</td>
<td>-1.40</td>
<td>-0.80</td>
<td>-0.50</td>
<td>0.00</td>
<td>0.60</td>
<td>1.00</td>
<td>1.60</td>
<td>Right – Left Ankle</td>
</tr>
<tr>
<td>Toddlers</td>
<td>48</td>
<td>-0.11</td>
<td>0.70</td>
<td>-1.80</td>
<td>-1.00</td>
<td>-0.50</td>
<td>0.00</td>
<td>0.25</td>
<td>0.70</td>
<td>1.00</td>
<td>Right – Left Ankle</td>
</tr>
<tr>
<td>Preschoolers</td>
<td>54</td>
<td>-0.06</td>
<td>0.45</td>
<td>-1.00</td>
<td>-0.50</td>
<td>-0.30</td>
<td>0.00</td>
<td>0.20</td>
<td>0.50</td>
<td>1.00</td>
<td>Right – Left Ankle</td>
</tr>
<tr>
<td>School age</td>
<td>64</td>
<td>0.06</td>
<td>0.48</td>
<td>-1.00</td>
<td>-0.50</td>
<td>-0.20</td>
<td>0.00</td>
<td>0.40</td>
<td>0.60</td>
<td>1.00</td>
<td>Right – Left Ankle</td>
</tr>
<tr>
<td>Teenagers</td>
<td>51</td>
<td>-0.01</td>
<td>0.67</td>
<td>-1.00</td>
<td>-0.80</td>
<td>-0.50</td>
<td>0.00</td>
<td>0.50</td>
<td>0.65</td>
<td>1.50</td>
<td>Right – Left Ankle</td>
</tr>
</tbody>
</table>

Legend: SD refers to standard deviation, P3 to the 3rd percentile, P10 to the 10th percentile, P25 to the 25th percentile, P50 to the 50th percentile, P75 to the 75th percentile, P90 to the 90th percentile, and P97 to the 97th percentile.

The correlation between R0/R0 ratios and circumference difference was statistically significant in the case of arm measurements in teenagers (r=−0.44; 95% confidence interval [CI] -0.64 to -0.18; p=0.001). It must be pointed out that the negative direction of the correlation is due to the fact that resistance is inversely related to volume, as explained in the Methods section.

Conversely, correlations between arm R0/R0 ratio and difference in circumference in the remaining age groups were not statistically significant (toddlers r=−0.26; 95% CI -0.51 to 0.02; p=0.07; preschoolers r=−0.15; 95% CI -0.40 to 0.12; p=0.28, school aged children r=0.01; 95% CI -0.23 to 0.26; p=0.92). Similarly, we found no significant correlation between leg R0/R0 ratios and difference in ankle measurement in any age group (toddlers r=0.15; 95% CI -0.14 to 0.42; p=0.30; preschoolers r=−0.08; 95% CI -0.34 to 0.19; p=0.57, school aged children r=0.00; 95% CI -0.24 to 0.25; p=0.99, teenagers r=0.06; 95% CI -0.22 to 0.33; p=0.67).

3.6. Discussion

We determined the normal R0/R0 and Ri/R0 ratios for arms and legs in pediatric patients, as well
as the normal difference in circumference between arms and between ankles.

The R0/R0 ratios for arms that we observed in the pediatric population were similar to those reported in adult patients by Ward et al [220]. In both studies, the mean R0/R0 ratios for arms were slightly closer to one (unity) than those found by Ridner et al [221] and Cornish et al [225] in smaller study samples, perhaps reflecting that the effect of limb dominance has yet to assert itself in the young. Leg BIS R0/R0 ratios in the present cohort were also similar to those reported in healthy adults, especially in the group of teenagers, whose values were slightly below one [219].

Our findings showed that younger patients had a greater extracellular fluid content, as reflected by higher Ri/R0 ratios. This is in keeping with the physiological differences in extracellular and intracellular compartments described in pediatric patients [226]. In addition, the body water content of the fat-free mass has been reported to decline throughout childhood, reaching adult values around the age of 20 years [227]. This concept is consistent with the decline in ratios we found over the four age groups. Interestingly, adult patients have also been described to show extracellular fluid compartment expansion (i.e., higher Ri/R0 ratios) with age [218]. According to the literature, the difference in Ri/R0 ratios in adults is more marked in the legs, suggesting that, unlike the pediatric population, the increase in the amount of extracellular fluid in the elderly may be associated with the age-related increase in the frequency of vascular disorders [218].

Measurement of *limb circumference* has been the standard method to determine the presence of edema not only in pediatric PTS, but also in other diseases characterized by the presence of localized fluid accumulation, such as lymphedema. As regards the difference in circumference, Goldenberg et al [98, 99] described a cohort of 78 healthy children aged 1 to 21 years divided in three age categories (1-5 year-olds n=30; 6-12 year-olds n=28; 13-21 year-olds n=20), and observed that the normal upper limit for the absolute difference between right and left arm circumference (measured at mid-upper arm and mid-forearm) and between right and left leg circumference (measured at mid-thigh and mid-calf) was one cm, irrespective of the age group. Cut-off points were determined as [median + (1.5 x interquartile range)]. In our study, the 97th percentile for difference in arm and ankle measurements in children aged 12-18 years was
slightly higher (1.6 cm for arms and 1.4 cm for ankles). The higher upper limits for difference in circumference in our study were also observed when estimating the cut-off point in keeping with Goldenberg et al. The difference could be partly explained by the larger sample size in our study.

We observed no correlation between arm circumference difference and arm R0/R0 measurements, or between ankle circumference difference and leg R0/R0 ratio in 1 to 11 year-olds. A possible explanation to this finding might be that these methods do not measure the same elements of body composition [215]. Whereas R0/R0 ratios reflect extracellular fluid content, the difference in circumference includes all tissues (i.e., fat and fat-free mass). Therefore, a difference detected in the total circumference between arms or ankles is not necessarily due to a difference in fluid content, suggesting that reliance on circumference measurements in the assessment of segmental edema may be misleading in most age groups.

Of note, there is a well-recognized and consistently reported bilateral asymmetry (directional asymmetry) in the human body that favors the right side of right-handers and is more pronounced in the arms [228-230]. Directional asymmetry is tied to limb preference and is attributed to differential biomechanical stress during bone growth, due to muscular development [228, 231-236]. As a result, it is not surprising that directional asymmetry has been reported to gradually develop during childhood, and to be more marked in adults than in children [229, 230, 237]. Indeed, a greater right-sided asymmetry in triceps skinfold [231], bone-free lean tissue, bone mineral density [238], and in limb bone dimensions [228] has been reported in adolescent and adults, particularly among males. The effect of sex hormone stimulation during puberty drives lipogenesis, muscle tissue hypertrophy, and bone growth [239], thus accentuating the arm asymmetry among teenagers. Importantly, changes in the relative amounts of adipose and lean tissue may affect the extracellular and intracellular fluid volumes as measured by BIS, since increases in adipose tissue can contribute to expansion of the extracellular water compartment; this could, in part, explain the significant correlation between arm circumference difference and extracellular (R0) resistance ratio among adolescents shown by our results.

The present study showed no conclusive influence of hand side dominance on any of the measurements, which could be explained by the small number of left-handed participants in the study population, the less pronounced directional asymmetry in children in general, as well as the
less pronounced effect of directional asymmetry that has been described in left-handed, as opposed to right-handed children and adults [229].

Our results need to be interpreted considering possible limitations to our methods. Though foot dominance was not recorded, it is known that there is a high correlation between hand and foot dominance, and only 5% of people show combined or cross-lateral dominance [240, 241]. Secondly, the population studied here is not representative of obese patients, who have been reported to have a larger extracellular/intracellular water ratio [242]. Hence, further studies focusing on the impact of obesity on Ri/R0 ratios are warranted. Thirdly, unlike the older children, infants would often only cooperate if they were allowed to sit up during BIS measurements and the protocol was modified accordingly. It must be pointed out however, that even though it is not known whether this position has an impact on BIS results, it is unlikely that it significantly affects hydrostatic pressure, and any effect should be similar in both extremities, therefore not affecting the inter-limb ratios. Lastly, whereas we did not assess biological maturity using specific methods, we grouped the children according to calendar age in an attempt to control this factor, which could potentially influence the measured variables.

To conclude, we determined normal segmental BIS ranges for arms and legs, and the normal ranges for the difference in circumference between arms and ankles in pediatric patients. We only found a significant correlation between arm R0/R0 ratios and difference in arm circumference in teenagers, suggesting that BIS and measuring tape measure different elements of body composition.

Acknowledgments

We thank the Science Engagement Department, Andrea Mus, and the hosts of the Ontario Science Centre for their contribution to the conduction of the study. We thank Francesca Kim for her support with patient recruitment and assessment. Lastly, we thank Karen Chapman and Christy Greene for the guidance provided during the planning of the study.
3.7. Footnotes

• i: The correlation was established between the two techniques used in the study, namely circumference difference and BIS.
• ii: The term “quantities” here is used in the sense of “amount”.
• iii: 18 year olds were included in the adolescent category, in line with the World Health Organization definition of adolescence, which includes up to 19 year olds in this age range.
• iv: Research Ethics Board approvals and consents can be found in Appendices 3.1 and 3.2.
• v: The six participants recruited at The Hospital for Sick Children did not have their arm and ankle circumference measured because these measures were included as part of the study at a later date.
• vi: Age refers to age-group (i.e., categorical variable)
• vii: The sample size for the study was estimated at a minimum of 30 children per age group, in order to approximate a z-distribution (t distribution with 29 degrees of freedom), to allow parametric tests to be run.
Chapter 4

Study 2a

4.1. Chapter Overview

Chapter 4 presents the first steps of tool development, and encompasses item generation, item operationalization, and item piloting.

During item generation, we used different sources (health care providers and patients diagnosed with PTS during childhood) to identify the signs and symptoms that define PTS and that will allow monitoring over time.

The original publication reproduced in this chapter shows only a brief description of how items were operationalized (defined and measured), in compliance with the regulations of the journal where it was published. Therefore, a full report of item operationalization is shown in Appendix 4.1.

In the item piloting step, the items identified during item generation were measured in children with previous diagnosis of deep vein thrombosis (DVT). Since DVT is a necessary condition for having PTS, these children were deemed at risk of having PTS, and 20 to 25% of them were expected to have PTS. During the study, pediatric thrombosis experts assessed whether the enrolled patients had or did not have PTS (diagnosis), and they also assessed PTS severity.

The results of each item and of PTS assessment were then compared, and the data were used in the next steps of tool development, presented in Chapter 6.

As in the previous chapter, additional comments to the original publication are added as footnotes, shown at the end of the chapter (Section 4.6).
Pediatric Post-Thrombotic Syndrome in Children: Toward the Development of a New Diagnostic and Evaluative Measurement Tool

(As published in Thrombosis Research 2016, 144: 184-191, reproduced with permission)

Authors: Avila, M.L. ¹, Brandão, L.R. ¹, Williams, S. ¹, Ward, L. C. ², Montoya, M. I. ¹, Stinson, J. ³,⁴,⁵, Kiss, A. ³,⁶, Lara-Corrales, I. ¹, Feldman B. M. ¹,³,⁴

¹Department of Pediatrics, The Hospital for Sick Children, Toronto, ON, Canada

²School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, Brisbane, Australia

³Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

⁴Child Health Evaluative Sciences, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

⁵Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Ontario, Canada

⁶Department of Research Design and Biostatistics, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Correspondence: Leonardo R. Brandão, Division of Hematology/Oncology, The Hospital for Sick Children, Toronto, Ontario, Canada, M5G-1X8. E-mail: leonardo.brandao@sickkids.ca. Tel: 1-416-813-7758; Fax: 1-416-813-5327
4.2. Abstract

**Objective:** Our goal was to conduct the generation and piloting of items for a discriminative and evaluative tool for pediatric post-thrombotic syndrome.

**Methods:** We followed a formative model for the development of the tool, focusing on the signs/symptoms (items) that define post-thrombotic syndrome. For item generation, pediatric thrombosis experts and subjects diagnosed with extremity post-thrombotic syndrome during childhood nominated items. In the piloting phase, items were cross-sectionally measured in children with limb deep vein thrombosis to examine the performance of the items.

**Results:** Twenty-three experts and 16 subjects listed 34 items, which were then measured in 140 subjects with previous diagnosis of limb deep vein thrombosis (70 upper extremity and 70 lower extremity). The items with strongest correlation with post-thrombotic syndrome severity and largest area under the curve were pain (in older children), paresthesia, and swollen limb for the upper extremity group, and pain (in older children), tired limb, heaviness, tightness and paresthesia for the lower extremity group.

**Conclusion:** The diagnostic properties of the items and their correlations with post-thrombotic syndrome severity varied according to the assessed venous territory. The information gathered in this study will help experts decide which item should be considered for inclusion in the new tool.
4.3. Introduction

Venous thromboembolic events are increasingly being reported in children [18]. This is in part due to the increased use of invasive diagnostic and therapeutic procedures, and the improvement of the detection methods for thrombotic events in modern medicine. As a consequence, venous thromboembolic events have become “the new epidemic” of hospitalized children [243], affecting approximately 1 in 250 hospitalized patients [19].

Post-thrombotic syndrome (PTS) is the most frequent long-term complication of deep venous thrombosis (DVT) affecting the upper and lower extremities in children and adults [20, 33, 34]. PTS is estimated to occur in 15 to 50% of adults affected by lower extremity (LE) deep vein thrombosis [25], and 7 to 46% of adult patients after sustaining upper extremity (UE) deep vein thrombosis [26]. A systematic review of PTS in the pediatric population reported an overall frequency of 26% [95% Confidence Interval (CI): 23-28] and a frequency of 17% [95% CI: 14-20] among prospective studies [32], including both UE and LE venous territories.

It is well-established that PTS can lead to significant disability and poor quality of life in adult patients, particularly when complicated by skin ulcers [102]. The negative repercussion of PTS is largely due to its difficult, costly, and frustrating treatment [244, 245].

The significance of PTS requires specific considerations in the pediatric population, since it is expected that the medical consequences of this syndrome will be disproportionately higher in younger patients [246], who are anticipated to endure the sequelae of PTS for a longer period of time [121]. Nonetheless, the proper diagnosis and monitoring of PTS remains a difficult challenge in this young population.

A systematic review identified two outcome-assessment tools that are commonly used in pediatric patients: the Modified Villalta Scale and the Manco-Johnson Instrument vi [32]. Both were derived from tools used in adult patients and are the current reference instruments proposed by the International Society of Thrombosis and Hemostasis (ISTH) for the assessment of PTS in children vii [64].
The Modified Villalta Scale derives from the Villalta Scale. The Villalta Scale [68, 247] was developed for the assessment of LE PTS in adults, since 96 to 99% of DVT occurring in adult patients affects the venous territory of the LE viii, whereas only 1 to 4% affects the UE [26, 248]. In striking contrast, 30 to 50% of cases of DVT in children affect the UE, thus highlighting the need for a pediatric tool that can be used to assess both UE and LE. For this reason, the original Villalta Scale was modified for use in pediatric patients by incorporating items (head edema, increase in limb circumference, and collateral vessels), replacing the four-point scoring system used in the adult scale to indicate the severity of signs and symptoms with a dichotomous score (absent/present) for most items, and reducing the five symptoms measured by the Villalta Scale (pain, cramps, heaviness, paresthesia, and pruritus) to two symptoms: pain or abnormal use and swelling, to simplify its applicability in children of all ages [29].

The Manco-Johnson Instrument [67], adapted from a classification tool for chronic venous diseases in adults, the Clinical-Etiologic-Anatomic-Pathophysiologic (CEAP) classification ix [62], combines clinical signs (edema, collateral circulation, skin changes, and skin ulcer) and one symptom (pain). Pain is assessed with the Wong-Baker FACES Pain scale [249] and is evaluated at three different levels: rest, daily activities, and aerobic exercise [64, 67].

Both the Modified Villalta Scale and Manco-Johnson Instrument have been acknowledged to have limitations, which are mainly related to their measurement properties and the information they provide. The Pediatric PTS Task Force within the Pediatric and Neonatal Hemostasis Subcommittee of the ISTH advocated for further investigation in order “to identify [their] advantages and disadvantages” [64].

One of the main shortcomings of these tools is the lack of operationalization of the items (i.e., signs and symptoms), since the tools do not clearly define all the items and do not always specify how the items should be measured.

Importantly, physicians caring for adult patients with chronic venous disorders, including PTS, have already recognized that the lack of consistent definitions of the terms used in venous diseases negatively impacts the comparison of clinical outcomes across studies and the exchange of medical information [23], and can affect the reliability of the tools. To address this issue, an international and interdisciplinary group of experts in venous disorders developed several
consensus documents, including the CEAP, revised CEAP, and VEIN-TERM consensus, which provided a common language for the reporting of chronic venous diseases [23, 62].

In order to address the aforementioned limitations to the instruments currently used for the evaluation of pediatric PTS, we developed a new PTS assessment tool for this population. The present study includes the steps of item generation and piloting. The latter stage served to investigate the diagnostic properties of the items obtained and their association (correlation) with PTS severity.

4.4. Material and Methods

Underlying theory: The construct PTS was defined according to the literature as “chronic venous symptoms and/or signs secondary to DVT and its sequelae” [22, 23]. Therefore, PTS was considered to be defined precisely by the items (signs and symptoms) that are chosen to measure it [148]. This conceptualization of the nature of the construct is one of the features of the formative model, and is consistent with the theory underlying the measurement of clinical phenomena [88]. Moreover, under the formative model, the items that define PTS are not expected to be interchangeable, but rather each item represents a different facet or aspect of PTS.

Conceptual framework: Given our focus on signs and symptoms, the construct PTS was mapped to the Body Functions and Structures component of the International Classification of Functioning, Disability and Health, Children and Youth Version [250]. Signs were defined as “visible manifestations of venous disorders”, and symptoms as “complaints related to venous disease”, as defined in the VEIN-TERM consensus [23].

Measurement purpose: The instrument was intended to be used for discriminative and evaluative purposes (i.e., to assess and monitor pediatric PTS) [251].

4.4.1. First Phase: Item Generation and Operationalization

Source of the items: items were obtained from two different sources.

a. Expert survey: Physicians involved in the care of children with PTS for at least five years were identified using the ISTH Pediatric and Neonatal Hemostasis Subcommittee
directory, in combination with recent publications in pediatric PTS. We aimed to reach experts from different continents and centers, since a panel of members with different perspectives on PTS would render a high proportion of high quality responses, over and above that of a homogeneous group [252]. Experts were asked to complete a survey by listing the signs and symptoms they considered should be measured to define and monitor UE and LE-PTS in children. The survey was designed using REDCap (Research Electronic Data Capture) software. To avoid introducing bias in the study by imposing our views, we used open ended questions [253]. In view of the potential use of different terms by different respondents for items that were essentially the same and for subsequent item operationalization, the experts were asked to provide a brief description/meaning of the items and/or proposed measurement approach. The design and implementation of the survey followed the theory of social exchange to motivate participation [254]. Reminder e-mails and/or phone calls were used to enhance response rates.

b. Patient interview: We conducted semi-structured interviews with school-aged children and teenagers who had sustained UE or LE-DVT before the age of 18 years, and who were subsequently diagnosed with mild to severe PTS by a pediatric thrombosis clinician. Patients <10 years of age were interviewed using a dyad-approach, in which the parent or proxy and the patient are interviewed at the same time. This approach has been reported to improve the assessment of health status in children, as compared to parent/proxy report alone [255]. In fact, we did not seek parent/proxy report alone, as the correlation between parent and child response has been shown to be low to moderate [255, 256]. Young adults no longer followed-up at our institution were interviewed over the phone. Patients with cognitive impairment or not fluent in English were not enrolled in the study.

Sample size: Given the well-defined area of knowledge and the limited number of experts in the field, a sample size of at least 14 experts providing complete responses to the survey was considered adequate in the context of the study [257]. Patients were enrolled until reaching saturation (i.e., no new items were obtained), which usually entails between 5 and 25 patients, based on general guidelines for the conduction of qualitative studies using the phenomenological approach [258].
Analysis: Clinical judgment [87] was followed to identify and group items that measure the same aspect of PTS xiii.

4.4.2. Second Phase: Item Piloting

Under the formative model, each item (sign or symptom of PTS) captures a different aspect of the construct PTS. Hence, the challenge is to select the items that are most representative of the construct. To address this issue, we piloted the items obtained in the previous phase in a cohort of children at risk of PTS xiv, and assessed the diagnostic performance of the signs and symptoms of PTS and their relationship with PTS severity. The resulting data will be xv evaluated by experts in the following phases of the study to select the items that, in their opinion, best represent the construct PTS.

For item piloting, we conducted a cross-sectional single-center study involving 1 to 21 year-old children with a history of objectively confirmed DVT, at least six months before the study procedures. Subjects with bilateral DVT, or with arterial thrombosis, muscular atrophy, active local inflammatory processes, or vascular malformations affecting the limb with the index DVT or the contralateral limb, were excluded to avoid confounding signs or symptoms xvi. Patients with pacemakers were not enrolled since bioimpedance spectroscopy, one of the techniques used to measure localized swelling (see below), is contraindicated in these patients.

The signs and symptoms of PTS collected in phase 1 were measured as follows:

a. Symptoms of PTS were measured using a self- or proxy-reported questionnaire developed with all the items obtained in phase 1. The self-reported version was used for children aged 10 and older. The structure of the questionnaire followed that of the Edinburgh Vein Study (Prof. Gerry Fowkes, with permission). Three different tools were used for the measurement of pain, according to the age of the patient: 1) Visual analogue scale using proxy/parental report for children 1-3 years. Marks were then translated into a 0-10 final score, reflective of the worst pain episode in the previous 4 weeks; 2) Faces Pain Scale-Revised [259] for patients aged ≥4 years. The choice was translated into a 0-10 final score, reflecting pain intensity over the previous 4 weeks; 3) Adolescent Pediatric Pain Tool for patients aged ≥10 years. This tool includes a measure of pain intensity
(translated into a 0-10 score), pain descriptors and distribution [260]. Only data on pain intensity was used in the present study. The proxy-reported questionnaire initially only included symptoms that we thought the parent/proxy would be able to observe. In addition, throughout the study, patients and proxies were inquired about other symptoms they thought should be included in the questionnaire.

b. Skin signs of PTS were measured using the items of the Venous Clinical Severity Score [76], namely skin pigmentation, skin inflammation, skin induration, and ulcers. Additional skin signs were: skin discoloration (defined as reticulate erythema [261], erythema hyperemicum [262], and livedo reticularis [263]), trophic changes (defined as atrophy and lipoatrophy [261]), easy bruising [264], recurrent superficial infections/poor healing, and altered perfusion [265]. Skin perfusion was measured with a DermaTemp Infrared Surface Skin Scanner, a non-invasive high precision hand-held infrared thermographic scanner designed to detect the subtle variations in skin temperature due to changes in the underlying perfusion. Temperature was measured at three standardized anatomic points; values were averaged to calculate the affected to unaffected limb ratio. Lastly, photographic documentation of skin findings and/or a second assessment by a Pediatric Dermatologist (ILC) were also carried out.

c. Observable veins were measured using the Venous Clinical Severity Score, specifically with the item varicose veins, which also includes corona phlebectatica [266], and the Vein Prominence Scale [208]. Additional signs related to observable veins were telangiectasias [62] and reticular veins [62].

d. Limb edema was measured using different methods: 1) limb circumference was measured using a tension-controlled measuring tape (Gulick II, Country Technology Inc., Wisconsin). In children affected by UE DVT, the circumference of the mid-arm of the right and left UE was measured at a distance corresponding to the mid-point between the acromion and olecranon process. In children affected by LE DVT, the circumference of the mid-calf was measured at the mid-point between the tibial tuberosity and the medial malleolus. The mid-thigh circumference was measured at the mid-point between the tibial tuberosity and the anterior superior iliac spine. The absolute difference (in cm) between
the affected and unaffected arm or leg, and the ratio of the affected to unaffected arm or leg were estimated. 2) **Limb volume** was measured by assuming a truncated cone shape of the limb [209]. The volume of four or five truncated cones was added to estimate the volume of the UE and the LE, respectively, based on circumferential measures taken at five anatomic landmarks. Circumferences were measured with a Gulick II tape; the distance between these points (length of the segment) was measured using digital calipers. The value of the limb volume obtained was normalized by comparing the affected to the unaffected limb, and expressed as a ratio. 3) **Ankle size** was measured using the Figure-of-8 technique [223]. The difference in cm of the ankle of the affected and unaffected extremities, and the ratio of the affected to unaffected limb were determined. 4) **Limb fluid content** was measured using a bioimpedance spectroscopy SFB7 device (ImpediMed, Inc. California) [215, 218, 267], following a pediatric protocol reported previously [268].

Additional items related to edema were head edema, pretibial edema, and pitting edema [269]. Pretibial and pitting edema were measured by using a 4-point clinical scale [270] and limb durometry [271], estimated with a hand held digital force gauge. This device measures the resistance to tissue compression. Determinations were taken at one specific anatomic location in the extremities; values were averaged to calculate the affected to unaffected limb skin resistance.

e. The functional impact of PTS xvii was assessed as impairments, activity limitation, and participation restrictions [250] according to the framework of the International Classification of Functioning, Disability and Health, Children and Youth Version. Impairments were evaluated using self- or proxy-reported endurance and strength [272], as well as testing limb strength with a compact dynamometer (Commander™ muscle tester, JTECH Medical, Utah) in children who were able to comply with the instructions, following a standardized protocol. We determined the affected to unaffected limb strength ratio. Activity limitations and participation restrictions were evaluated with the Pediatric Outcomes Data Collection Instrument (PODCI) [273].
The severity of PTS was determined by a pediatric thrombosis expert (SW, LB) blind to the results of the measurement procedures mentioned above. PTS was classified as none, mild, moderate and severe, according to the assessors’ expert clinical assessment. Given the limitations of the current pediatric tools, neither the Modified Villalta Scale nor the Manco-Johnson Instrument were used in this study.

Signs and symptoms were measured by study investigators (LA, MM) blind to the PTS severity of the subject xviii. Skin manifestations were also assessed by a pediatric dermatologist (ILC).

Data were entered into a REDCap database. Double-data entry was carried out to ensure accuracy.

Sample size: Given the normal difference in hydrostatic pressure between UE and LE, we anticipated a different frequency and/or severity of clinical manifestations of UE and LE-PTS. Therefore, the diagnostic performance of every item was studied separately in UE and LE. A sample size of 70 subjects per stratum was estimated to produce a two-sided 95% confidence interval (CI) with acceptable widths that equal 0.47, 0.43, 0.36, 0.25 and 0.09 when the sample correlations (r) are 0.1, 0.3, 0.5, 0.7 and 0.9 respectively.

Statistical analysis: In order to assess the strength of association between each item and PTS severity, we used Spearman correlation coefficients xix. Additionally, we determined the sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve ([AUC]) of each item in order to establish their ability to discriminate subjects with and without PTS (i.e., mild, moderate or severe PTS vs. no PTS). The cutoff criterion used for the reported sensitivity and specificity was that corresponding to the Youden index (i.e., the optimal threshold point from the ROC curve). Significance level was set at 0.05.

Statistical analyses were performed using SAS software, version 9.2 (SAS Institute Inc. Cary, NC, USA) and MedCalc for Windows (MedCalc Software, Ostend, Belgium).

The study was approved by the Research Ethics Boards of The Hospital for Sick Children and of the University of Toronto. Informed consent and assent, when applicable, were obtained. The instruction letter included in the survey sent to the experts informed them that completion of the survey implied consent xx.
4.5. Results

4.5.1. First Phase: Item Generation and Operationalization

Twenty-five pediatric thrombosis experts were identified and invited to participate in the study; one expert declined participation. The overall complete response rate was 92% (i.e., 23 of the 25 experts completed the survey). The survey was conducted between July-November 2013. Respondents worked at North-American institutions (n=11), European institutions (n=6), Australian institutions (n=3), Latin-American institutions (n=2), and Asian institutions (n=1).

Sixteen patients (median age at the time of study enrolment 13.5 years, 25\textsuperscript{th}-75\textsuperscript{th} percentile 8.75-21 years) were interviewed between April-September 2013; only three additional items were obtained after interview number nine, and only one after interview number 12. Six of these patients (37%) had mild PTS; 7 (44%) had moderate PTS and 3 (19%) had severe PTS at the time of study participation or during previous follow-up. Ten patients (63%) had LE PTS.

The group of experts listed an average of 7 signs and symptoms each (166 in total) whereas patients listed an average of 3 items each (54 items in total). We analyzed all the listed signs and symptoms to identify similar responses. As a result, we obtained a total of 34 items, which were subsequently grouped according to PTS domains \textsuperscript{xxi}. The number of times each item was proposed (i.e., frequency) per domain is shown in Table 4.1.
Table 4.1. Items Proposed by Patients and Expert Physicians

<table>
<thead>
<tr>
<th>Domain</th>
<th>Item</th>
<th>Patients (n=16)</th>
<th>Experts (n=23)</th>
<th>Number of definitions</th>
<th>Suggested measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
<td>Pain</td>
<td>7 20</td>
<td>2 Scale/history</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heaviness</td>
<td>6 9</td>
<td>2 Scale/history</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paresthesia</td>
<td>8 3</td>
<td>2 History</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Edema (symptom)</td>
<td>7 2</td>
<td>- History</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cramps</td>
<td>5 3</td>
<td>1 Scale/history</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tightness</td>
<td>4 1</td>
<td>1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pruritus</td>
<td>1 5</td>
<td>1 Scale/history</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Change in skin color (redness/purple/blotchy skin)</td>
<td>6 -</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fatigue (tired limb)</td>
<td>2 1</td>
<td>1 Scale/history</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hot/warmer limb</td>
<td>1 -</td>
<td>- -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin signs</td>
<td>Venous ulcer</td>
<td>- 15</td>
<td>5 Exam/scale/visual aid/photography</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skin discoloration</td>
<td>- 11</td>
<td>4 Exam/scale/photography</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skin hyperpigmentation</td>
<td>- 9</td>
<td>5 Scale/visual aid/exam</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Venous eczema/stasis dermatitis</td>
<td>- 6</td>
<td>3 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trophic changes</td>
<td>- 3</td>
<td>2 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skin induration</td>
<td>- 2</td>
<td>1 Scale/visual aid</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Altered perfusion</td>
<td>1 1</td>
<td>1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Easy bruising</td>
<td>1 1</td>
<td>1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recurrent infections/ poor healing</td>
<td>- 1</td>
<td>1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observable veins</td>
<td>Observable veins/collateral circulation</td>
<td>4 18</td>
<td>9 Exam/scale/visual aid/imaging</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Varicose veins</td>
<td>- 5</td>
<td>2 Exam</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Telangiectasias</td>
<td>- 3</td>
<td>3 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reticular veins</td>
<td>- 1</td>
<td>1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corona phlebectatica</td>
<td>- 1</td>
<td>1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edema</td>
<td>Edema (sign)</td>
<td>- 23</td>
<td>11 Circumference/volume difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pitting edema</td>
<td>- 5</td>
<td>1 Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Head edema/svc syndrome</td>
<td>- 3</td>
<td>2 Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pretibial edema</td>
<td>- 1</td>
<td>1 -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact</td>
<td>Limitation in activities</td>
<td>- 4</td>
<td>- -</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional impairment</td>
<td>1 3</td>
<td>1 History</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Loss of venous access</td>
<td>- 2</td>
<td>- Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quality of life</td>
<td>- 1</td>
<td>- Tool</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Pain on palpation</td>
<td>- 2</td>
<td>- Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Venous claudication</td>
<td>- 1</td>
<td>- -</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experts proposed definitions and/or measurement techniques for some items. For example, the Wong-Baker Faces Pain scale, the Faces Pain Scale, a numerical scale, a Visual analogue scale, or, more generally, the use of “standard or validated scales” were suggested for the measurement of pain. Measurement of edema was the only item mentioned by all the experts who answered the survey. Experts suggested measurement of limb circumference and of limb volume. Some experts considered that >1 centimeter (cm) difference between limbs should be considered abnormal, whereas others used ≥1 cm, or >1 cm for children <5 years of age and >2 cm for older children. In addition, the use of a variety of anatomical landmarks was proposed.

Table 4.1 also shows the number of definitions and suggested measurement approaches for each item, which were used to develop a set of definitions and measurement techniques (i.e., item operationalization) for the following phase of the study.

In the majority of cases experts suggested the use of a scale but did not specify which. Therefore, we searched in the literature and contacted experts in the field to find a suitable tool for every item. The diagnostic properties of the items thus defined and characterized were subsequently investigated in the second part of the study. The complete list of definitions and measurement techniques generated is available from the authors upon request xix.

4.5.2. Second Phase: Item Piloting

Thirty-two of the 34 items collected were investigated in this part of the study (see discussion for details on the remaining two items), which ran between March 2014 and May 2015. The characteristics of the enrolled patients are shown in Table 4.2.
Table 4.2. General Characteristics of the Patients

<table>
<thead>
<tr>
<th>Variable</th>
<th>Upper extremity DVT, N=70</th>
<th>Lower extremity DVT, N=70</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at the time of study participation, median (p25-p75), in years</td>
<td>6.8 (3-11)</td>
<td>8.3 (4-12)</td>
</tr>
<tr>
<td>Time from DVT diagnosis to PTS exam, median (p25-p75), in months</td>
<td>42 (22-77)</td>
<td>60.8 (32.4-103.1)</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>43 (61%)</td>
<td>43 (61%)</td>
</tr>
<tr>
<td>Underlying condition, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac disease</td>
<td>25 (36%)</td>
<td>35 (50%)</td>
</tr>
<tr>
<td>Congenital/genetic conditions a</td>
<td>10 (14%)</td>
<td>10 (14%)</td>
</tr>
<tr>
<td>Prematurity</td>
<td>8 (12%)</td>
<td>4 (6%)</td>
</tr>
<tr>
<td>Cancer</td>
<td>10 (14%)</td>
<td>1 (1.5%)</td>
</tr>
<tr>
<td>Healthy</td>
<td>3 (4%)</td>
<td>7 (10%)</td>
</tr>
<tr>
<td>Infection</td>
<td>4 (6%)</td>
<td>6 (8.5%)</td>
</tr>
<tr>
<td>Other</td>
<td>10 (14%)</td>
<td>7 (10%)</td>
</tr>
<tr>
<td>Limb dominance, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>48 (68.5%)</td>
<td>47 (67%)</td>
</tr>
<tr>
<td>Left</td>
<td>11 (16%)</td>
<td>7 (10%)</td>
</tr>
<tr>
<td>Cross dominance</td>
<td>1 (1.5%)</td>
<td>0</td>
</tr>
<tr>
<td>Unknown</td>
<td>10 (14%)</td>
<td>16 (23%)</td>
</tr>
<tr>
<td>Right vs. left DVT</td>
<td>50 (71%) vs. 20 (29%)</td>
<td>46 (66%) vs. 24 (34%)</td>
</tr>
<tr>
<td>Line-related DVT, n (%)</td>
<td>64/70 (91%)</td>
<td>59/70 (84%)</td>
</tr>
<tr>
<td>PTS severity, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>45 (64.3%)</td>
<td>34 (48%)</td>
</tr>
<tr>
<td>Mild</td>
<td>17 (24.3%)</td>
<td>25 (36%)</td>
</tr>
<tr>
<td>Moderate</td>
<td>7 (10%)</td>
<td>11 (16%)</td>
</tr>
<tr>
<td>Severe</td>
<td>1 (1.4%)</td>
<td>0</td>
</tr>
</tbody>
</table>

Legend: DVT, deep vein thrombosis; p25-p75, 25th and 75th percentile; PTS, post-thrombotic syndrome

Throughout the testing of the proxy-version of the symptoms questionnaire, some parents suggested including symptoms that were originally only part of the self-reported version. Hence, the proxy-reported questionnaire was expanded to include optional questions for those non-observable symptoms.

Items such as pitting edema, pain on palpation, venous claudication, and skin signs (aside from collaterals in UE/LE and varicose veins in LE) were not observed in this cohort of patients.

The items with strongest correlation with PTS xxii were pain intensity in older children, as per the Faces Pain Scale-Revised and the Adolescent Pediatric Pain Tool, paresthesia and swollen limb
for the UE stratum, and pain intensity, as per the Adolescent Pediatric Pain Tool, tired limb, heaviness, paresthesia, tightness and endurance for the LE stratum (Tables 4.3 and 4.4). The items with largest AUCs \(^{xxiii}\) were pain in older children, as per the Faces Pain Scale-Revised and Adolescent Pediatric Pain Tool, tired limb, paresthesia, and observable veins. In the LE stratum, the best performing items were pain (measured by the Adolescent Pediatric Pain Tool), heaviness, tightness, thigh circumference difference and ratio, and Pediatric Outcomes Data Collection Instrument (Tables 4.3 and 4.4).

Table 4.3. Item-PTS Correlations, Area Under the Curve, Sensitivity, and Specificity for the UE-DVT Stratum (Items Observed Only)

<table>
<thead>
<tr>
<th>Item UE (n)</th>
<th>r, (95% CI)</th>
<th>p-value</th>
<th>AUC (95% CI)</th>
<th>p-value</th>
<th>Best cut-off</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain APPT (27)</td>
<td>0.57 (0.24 – 0.78)</td>
<td>0.001</td>
<td>0.74 (0.53 – 0.89)</td>
<td>0.02</td>
<td>&gt;=2</td>
<td>60 (26 – 88)</td>
<td>94 (70 – 99)</td>
</tr>
<tr>
<td>Pain FPS-R (41)</td>
<td>0.61 (0.37 – 0.77)</td>
<td>&lt;0.001</td>
<td>0.72 (0.56 – 0.85)</td>
<td>0.002</td>
<td>&gt;=1 (^a)</td>
<td>47 (21 – 73)</td>
<td>96 (80 – 100)</td>
</tr>
<tr>
<td>Pain VAS (29)</td>
<td>0.36 (-0.01 – 0.64)</td>
<td>0.05</td>
<td>0.55 (0.36 – 0.73)</td>
<td>0.32</td>
<td>&gt;=1 (^a)</td>
<td>10 (0.3 – 45)</td>
<td>100 (82 – 100)</td>
</tr>
<tr>
<td>Tiredness (40)</td>
<td>0.47 (0.18 – 0.58)</td>
<td>0.002</td>
<td>0.69 (0.53 – 0.83)</td>
<td>0.01</td>
<td>&gt;=1 (^b)</td>
<td>47 (21 – 73)</td>
<td>88 (69 – 98)</td>
</tr>
<tr>
<td>Heaviness (27)</td>
<td>0.47 (0.10 – 0.71)</td>
<td>0.01</td>
<td>0.67 (0.47 – 0.84)</td>
<td>0.05</td>
<td>&gt;=1 (^b)</td>
<td>40 (12 – 74)</td>
<td>94 (71 – 100)</td>
</tr>
<tr>
<td>Paresthesia (40)</td>
<td>0.55 (0.28 – 0.73)</td>
<td>&lt;0.001</td>
<td>0.70 (0.54 – 0.83)</td>
<td>0.002</td>
<td>&gt;=1 (^b)</td>
<td>40 (16 – 68)</td>
<td>100 (86 – 100)</td>
</tr>
<tr>
<td>Cramps (27)</td>
<td>0.33 (-0.06 – 0.63)</td>
<td>0.09</td>
<td>0.57 (0.37 – 0.76)</td>
<td>0.33</td>
<td>&gt;=1 (^b)</td>
<td>20 (2.5 – 56)</td>
<td>94 (71 – 100)</td>
</tr>
<tr>
<td>Redness (70)</td>
<td>0.41 (0.19 – 0.59)</td>
<td>&lt;0.001</td>
<td>0.61 (0.48 – 0.72)</td>
<td>0.02</td>
<td>&gt;=1 (^b)</td>
<td>24 (9.4 – 45)</td>
<td>98 (88 – 100)</td>
</tr>
<tr>
<td>Swelling (70)</td>
<td>0.51 (0.31 – 0.66)</td>
<td>&lt;0.001</td>
<td>0.62 (0.50 – 0.73)</td>
<td>0.006</td>
<td>&gt;=1 (^b)</td>
<td>24 (9.4 – 45)</td>
<td>100 (92 – 100)</td>
</tr>
<tr>
<td>Hot/Warm (70)</td>
<td>0.40 (0.18 – 0.58)</td>
<td>&lt;0.001</td>
<td>0.59 (0.47 – 0.71)</td>
<td>0.03</td>
<td>&gt;=1 (^b)</td>
<td>20 (6.8 – 41)</td>
<td>98 (82 – 100)</td>
</tr>
<tr>
<td>Pruritus (70)</td>
<td>0.15 (-0.09 – 0.37)</td>
<td>0.23</td>
<td>0.53 (0.41 – 0.65)</td>
<td>0.34</td>
<td>&gt;=1 (^b)</td>
<td>8 (1 – 26)</td>
<td>98 (88 – 100)</td>
</tr>
<tr>
<td>Tightness (27)</td>
<td>0.38 (-0.01 – 0.66)</td>
<td>0.05</td>
<td>0.55 (0.35 – 0.74)</td>
<td>0.32</td>
<td>&gt;=1 (^b)</td>
<td>10 (0.3 – 45)</td>
<td>100 (81 – 100)</td>
</tr>
<tr>
<td>Delayed healing (27)</td>
<td>0.24 (-0.16 – 0.57)</td>
<td>0.22</td>
<td>0.55 (0.35 – 0.74)</td>
<td>0.32</td>
<td>&gt;=1 (^a)</td>
<td>10 (0.3 – 45)</td>
<td>100 (81 – 100)</td>
</tr>
<tr>
<td>Skin temperature (69)</td>
<td>-0.03 (-0.26 – 0.21)</td>
<td>0.81</td>
<td>0.51 (0.39 – 0.63)</td>
<td>0.90</td>
<td>&lt;=0.99</td>
<td>8 (1 – 26)</td>
<td>75 (60 – 87)</td>
</tr>
<tr>
<td>Durometry (11)</td>
<td>-0.11 (-0.66 – 0.53)</td>
<td>0.74</td>
<td>0.53 (0.23 – 0.82)</td>
<td>0.87</td>
<td>&lt;=1.06</td>
<td>67 (22 – 96)</td>
<td>60 (15 – 95)</td>
</tr>
<tr>
<td>Collaterals (70)</td>
<td>0.39 (0.17 – 0.57)</td>
<td>&lt;0.001</td>
<td>0.71 (0.59 – 0.81)</td>
<td>&lt;0.001</td>
<td>&gt;=1 (^a)</td>
<td>64 (43 – 82)</td>
<td>71 (56 – 84)</td>
</tr>
<tr>
<td>Circumference ratio (70)</td>
<td>0.23 (-0.01 – 0.44)</td>
<td>0.06</td>
<td>0.62 (0.50 – 0.73)</td>
<td>0.10</td>
<td>&gt;=1.01</td>
<td>76 (55 – 91)</td>
<td>56 (40 – 70)</td>
</tr>
<tr>
<td>Circumference difference (70)</td>
<td>0.26 (0.02 – 0.46)</td>
<td>0.03</td>
<td>0.63 (0.51 – 0.74)</td>
<td>0.07</td>
<td>&gt;=0.4</td>
<td>64 (43 – 82)</td>
<td>62 (47 – 76)</td>
</tr>
<tr>
<td>Volume ratio (70)</td>
<td>0.27 (0.03 – 0.47)</td>
<td>0.03</td>
<td>0.65 (0.53 – 0.76)</td>
<td>0.03</td>
<td>&gt;=1.01</td>
<td>88 (69 – 98)</td>
<td>44 (30 – 60)</td>
</tr>
<tr>
<td>BIS ratio (70)</td>
<td>0.21 (-0.02 – 0.46)</td>
<td>0.08</td>
<td>0.62 (0.50 – 0.74)</td>
<td>0.08</td>
<td>&gt;=1.02</td>
<td>72 (51 – 88)</td>
<td>60 (44 – 74)</td>
</tr>
<tr>
<td>Strength (70)</td>
<td>-0.09 (-0.31 – 0.15)</td>
<td>0.48</td>
<td>0.51 (0.39 – 0.63)</td>
<td>0.84</td>
<td>&lt;=5 (^c)</td>
<td>8 (1 – 26)</td>
<td>96 (85 – 100)</td>
</tr>
<tr>
<td>Endurance (70)</td>
<td>-0.47 (-0.63 – -0.26)</td>
<td>&lt;0.001</td>
<td>0.66 (0.54 – 0.77)</td>
<td>0.002</td>
<td>&lt;=6 (^d)</td>
<td>36 (18 – 58)</td>
<td>96 (85 – 100)</td>
</tr>
<tr>
<td>Strength tested (21)</td>
<td>-0.02 (-0.45 – 0.42)</td>
<td>0.94</td>
<td>0.56 (0.33 – 0.77)</td>
<td>0.69</td>
<td>&lt;=1.06</td>
<td>60 (26 – 88)</td>
<td>64 (31 – 99)</td>
</tr>
<tr>
<td>Global PODCI (63)</td>
<td>-0.24 (-0.46 – 0.01)</td>
<td>0.06</td>
<td>0.63 (0.50 – 0.74)</td>
<td>0.09</td>
<td>&lt;=84</td>
<td>42 (22 – 63)</td>
<td>87 (73 – 96)</td>
</tr>
</tbody>
</table>
Legend: UE-DVT, upper extremity deep vein thrombosis; $r$, Spearman correlation coefficient; CI, confidence interval; AUC, area under the curve; APPT, adolescent pediatric pain tool; FPS-R, Faces Pain Scale, Revised; VAS, visual analogue scale; BIS, bioimpedance spectroscopy ratio; PODCI, pediatric outcome data collection instrument

\* Symptom/sign present, any intensity or severity; \*\* Present, any frequency; \*\*\* Strength slightly below the other arm or worse; \*d Endurance just below that of the other arm or worse

Table 4.4. Item-PTS Correlations, Area Under the Curve, Sensitivity, and Specificity for Each Item, LE-DVT Stratum

<table>
<thead>
<tr>
<th>Item LE (n)</th>
<th>$r_s$ (95% CI)</th>
<th>p-value</th>
<th>AUC (95% CI)</th>
<th>p-value</th>
<th>Best cut-off</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain APPT (25)</td>
<td>0.78 (0.54 – 0.89)</td>
<td>&lt;0.001</td>
<td>0.83 (0.63 – 0.95)</td>
<td>&lt;0.001</td>
<td>≥1</td>
<td>67 (41 – 87)</td>
<td>100 (59 – 100)</td>
</tr>
<tr>
<td>Pain FPS-R (43)</td>
<td>0.46 (0.18 – 0.67)</td>
<td>0.002</td>
<td>0.67 (0.51 – 0.81)</td>
<td>0.01</td>
<td>1</td>
<td>50 (31 – 69)</td>
<td>80 (52 – 96)</td>
</tr>
<tr>
<td>Tiredness (44)</td>
<td>0.57 (0.32 – 0.74)</td>
<td>&lt;0.001</td>
<td>0.66 (0.51 – 0.80)</td>
<td>0.009</td>
<td>≥1</td>
<td>41 (24 – 61)</td>
<td>93 (68 – 100)</td>
</tr>
<tr>
<td>Heaviness (26)</td>
<td>0.75 (0.50 – 0.88)</td>
<td>&lt;0.001</td>
<td>0.76 (0.56 – 0.91)</td>
<td>&lt;0.001</td>
<td>≥1</td>
<td>53 (29 – 76)</td>
<td>100 (59 – 100)</td>
</tr>
<tr>
<td>Paresthesia (42)</td>
<td>0.55 (0.19 – 0.67)</td>
<td>0.002</td>
<td>0.67 (0.51 – 0.80)</td>
<td>&lt;0.001</td>
<td>≥1</td>
<td>33 (17 – 54)</td>
<td>100 (78 – 100)</td>
</tr>
<tr>
<td>Cramps (26)</td>
<td>0.32 (-0.08 – 0.63)</td>
<td>0.11</td>
<td>0.61 (0.40 – 0.79)</td>
<td>0.28</td>
<td>≥2</td>
<td>26 (9.1 – 51)</td>
<td>100 (59 – 100)</td>
</tr>
<tr>
<td>Redness (70)</td>
<td>0.40 (0.18 – 0.58)</td>
<td>&lt;0.001</td>
<td>0.60 (0.47 – 0.71)</td>
<td>0.004</td>
<td>≥1</td>
<td>19 (8.2 – 33)</td>
<td>100 (90 – 100)</td>
</tr>
<tr>
<td>Swelling (70)</td>
<td>0.50 (0.30 – 0.66)</td>
<td>&lt;0.001</td>
<td>0.68 (0.56 – 0.79)</td>
<td>&lt;0.001</td>
<td>≥1</td>
<td>42 (26 – 59)</td>
<td>94 (80 – 99)</td>
</tr>
<tr>
<td>Warmer (70)</td>
<td>0.48 (0.28 – 0.64)</td>
<td>&lt;0.001</td>
<td>0.60 (0.47 – 0.71)</td>
<td>0.004</td>
<td>≥1</td>
<td>19 (8.2 – 33)</td>
<td>100 (90 – 100)</td>
</tr>
<tr>
<td>Pruritus (70)</td>
<td>0.47 (-0.01 – 0.43)</td>
<td>0.06</td>
<td>0.57 (0.45 – 0.69)</td>
<td>0.05</td>
<td>≥1</td>
<td>17 (6.4 – 33)</td>
<td>97 (85 – 100)</td>
</tr>
<tr>
<td>Tightness (26)</td>
<td>0.61 (0.28 – 0.80)</td>
<td>&lt;0.001</td>
<td>0.71 (0.50 – 0.87)</td>
<td>&lt;0.001</td>
<td>≥1</td>
<td>42 (20 – 67)</td>
<td>100 (59 – 100)</td>
</tr>
<tr>
<td>Delayed healing (30)</td>
<td>0.16 (-0.22 – 0.49)</td>
<td>0.41</td>
<td>0.52 (0.33 – 0.71)</td>
<td>0.71</td>
<td>≥1</td>
<td>12 (1.2 – 36.4)</td>
<td>92 (64 – 100)</td>
</tr>
<tr>
<td>Skin temperature (70)</td>
<td>0.06 (-0.18 – 0.29)</td>
<td>0.55</td>
<td>0.53 (0.40 – 0.65)</td>
<td>0.66</td>
<td>&gt;1.00</td>
<td>56 (38 – 72)</td>
<td>65 (47 – 80)</td>
</tr>
<tr>
<td>Durometry (11)</td>
<td>0.20 (-0.46 – 0.71)</td>
<td>0.22</td>
<td>0.53 (0.23 – 0.82)</td>
<td>0.97</td>
<td>&gt;0.98</td>
<td>67 (22 – 96)</td>
<td>60 (15 – 95)</td>
</tr>
<tr>
<td>Varicose veins (70)</td>
<td>0.12 (-0.12 – 0.34)</td>
<td>0.32</td>
<td>0.53 (0.10 – 0.65)</td>
<td>0.32</td>
<td>≥1</td>
<td>8 (1.8 – 23)</td>
<td>97 (85 – 100)</td>
</tr>
<tr>
<td>Collaterals (70)</td>
<td>0.09 (-0.14 – 0.32)</td>
<td>0.44</td>
<td>0.51 (0.39 – 0.64)</td>
<td>0.81</td>
<td>&gt;2</td>
<td>14 (4.7 – 30)</td>
<td>88 (73 – 97)</td>
</tr>
<tr>
<td>Thigh circ. ratio (70)</td>
<td>0.33 (0.10 – 0.53)</td>
<td>0.005</td>
<td>0.71 (0.59 – 0.81)</td>
<td>&lt;0.001</td>
<td>&gt;1.02</td>
<td>93 (67 – 94)</td>
<td>53 (35 – 70)</td>
</tr>
<tr>
<td>Thigh circ. difference (70)</td>
<td>0.43 (0.21 – 0.60)</td>
<td>&lt;0.001</td>
<td>0.74 (0.62 – 0.84)</td>
<td>&lt;0.001</td>
<td>&gt;1.2</td>
<td>64 (46 – 79)</td>
<td>74 (56 – 87)</td>
</tr>
<tr>
<td>Calf circ. ratio (70)</td>
<td>0.11 (-0.11 – 0.35)</td>
<td>0.29</td>
<td>0.63 (0.51 – 0.74)</td>
<td>0.06</td>
<td>&gt;0.99</td>
<td>94 (81 – 99)</td>
<td>35 (20 – 54)</td>
</tr>
<tr>
<td>Calf circ. difference (70)</td>
<td>0.22 (-0.02 – 0.43)</td>
<td>0.07</td>
<td>0.68 (0.56 – 0.79)</td>
<td>0.005</td>
<td>&gt;0.4</td>
<td>100 (90 – 100)</td>
<td>29 (15 – 48)</td>
</tr>
<tr>
<td>Ankle circ. ratio (69)</td>
<td>-0.10 (-0.33 – 0.14)</td>
<td>0.40</td>
<td>0.55 (0.43 – 0.67)</td>
<td>0.48</td>
<td>&lt;0.99</td>
<td>28 (14 – 45)</td>
<td>88 (72 – 97)</td>
</tr>
<tr>
<td>Ankle circ. difference (69)</td>
<td>-0.08 (-0.31 – 0.16)</td>
<td>0.53</td>
<td>0.54 (0.41 – 0.66)</td>
<td>0.59</td>
<td>&lt;0.5</td>
<td>28 (14 – 45)</td>
<td>88 (72 – 97)</td>
</tr>
<tr>
<td>Leg volume ratio (70)</td>
<td>0.17 (-0.07 – 0.39)</td>
<td>0.16</td>
<td>0.61 (0.47 – 0.71)</td>
<td>0.182</td>
<td>&gt;1.02</td>
<td>83 (67 – 94)</td>
<td>44 (27 – 62)</td>
</tr>
<tr>
<td>Leg BIS ratio (70)</td>
<td>0.34 (0.12 – 0.53)</td>
<td>0.003</td>
<td>0.66 (0.49 – 0.72)</td>
<td>0.02</td>
<td>&gt;1.02</td>
<td>86 (71 – 95)</td>
<td>44 (27 – 62)</td>
</tr>
<tr>
<td>Strength (70)</td>
<td>-0.30 (-0.49 – -0.06)</td>
<td>0.01</td>
<td>0.61 (0.48 – 0.72)</td>
<td>0.02</td>
<td>≤6</td>
<td>25 (12 – 42)</td>
<td>94 (80 – 99)</td>
</tr>
</tbody>
</table>
The items with largest AUCs were pain in older children, as per the FPS-R and APPT, tired limb, paresthesia, and observable veins. In the LE stratum, the best performing items were pain (measured by the APPT), heaviness, tightness, thigh circumference difference and ratio, and PODCI (Tables 4.3 and 4.4).

Overall, items had higher specificity than sensitivity.

4.6. Discussion

In this report we present the first steps of the development of a new tool for the discrimination and evaluation of pediatric PTS affecting the UE and LE.

We considered that the theory underlying the formative model was appropriate to study our construct. As a consequence, we first elicited the signs and symptoms of PTS from a group of experts and affected patients.

The items most commonly endorsed by both patients and experts were pain and edema, which are items included in the tools developed to measure PTS in adults [68-70] and children, and are the only items of the Ginsberg criteria for the diagnosis of PTS in adult patients [70].

Experts also proposed several additional symptoms, including for example heaviness and pruritus, which contrast with the fact that the Modified Villalta Scale and Manco-Johnson Instrument only assess one or two symptoms. The limited number of symptoms that the Modified Villalta Scale and Manco-Johnson Instrument assess in comparison with the Villalta Scale, which has been acknowledged [64], roots on the difficulty of measuring symptoms in young patients. Still, when tailored to the age of the patient, the measurement of symptoms
provided valuable information. In fact, the best performing items in the item-piloting phase of the study were pain (as measured in older children) and other symptoms, such as paresthesia for the UE and heaviness and tightness for the LE.

Patients in the LE stratum reported a higher frequency of symptoms than the patients on the UE stratum, and diagnostic properties of the symptoms differed between UE and LE. Overall, LE symptoms showed better diagnostic properties than UE symptoms. These facts suggest the need for evaluating UE and LE differently, using tailored tools. Indeed, the severity and frequency of clinical manifestations of UE and LE PTS are expected to differ. Moreover, clinical signs that usually present later during the natural course of LE PTS (e.g., varicose veins, hyperpigmentation, skin ulcers) may not be seen in UE PTS. A potential unwanted effect of using the same tool for assessment of both the upper and lower venous territory is that most children will likely have low scores when assessing UE PTS, leading to floor effects [157]. If patients with UE PTS are less symptomatic and cluster on the lower end of a tool, then the tool will have a low discriminatory power, which will affect its sensitivity and responsiveness [274]. The responsiveness of the tool is relevant to monitor patients over time, and to assess the efficacy of interventions.

Several items, including skin manifestations (pigmentation, inflammation, induration, ulcers), pain on palpation, pitting edema, and venous claudication were not observed in this study, which is probably due to the relatively milder nature of PTS manifestations in children, as compared to adults. In addition, five of the nine signs of the Modified Villalta Scale and two of the four signs of the Manco-Johnson Instrument were not seen in our patients, suggesting their infrequent occurrence and potential problem with floor effects when applying these tools. These findings suggest that the new pediatric PTS tool currently being developed may be simpler than the current tools, at least in terms of clinical signs.

In general, individual items were more specific than sensitive. This is likely due to the fact that PTS is a syndrome composed of signs and symptoms, and the combination of items is expected to be more useful for the determining the presence or absence of PTS than any sign or symptom by itself [275]. With the probable exception of limb swelling, no single sign or symptom is a common denominator of PTS. In fact, limb swelling was the only item proposed by all the
experts in the first part of the study, and was also the most sensitive in both strata. Symptoms, on the other hand, are harder to identify as younger patients may not recognize or express them.

Our results should be interpreted in the light of possible study limitations. For example, the reliance on clinical expertise to define PTS instead of the existing tools may be debatable. However, this was necessary to overcome the problems with the existing instruments. Since PTS is a syndrome essentially diagnosed on clinical grounds [245], and no test has been able to replace clinical assessment in establishing the presence or absence of PTS, the evaluation of the patients enrolled in our study only by experienced clinicians was considered the best approach for determination of PTS severity. This approach was also consistent with the theory behind the formative model. In addition, patients were assessed by one of the clinical experts, as assessment of the entire group of patients by both experts was not feasible in the context of the study.

It must be pointed out that quality of life and loss of venous access were not included among the measured items. The International Classification of Functioning, Disability and Health, Children and Youth Version acknowledges that the concepts of quality of life and disease/disability are related. However, whereas quality of life pertains to how patients “feel” about their health, disease/disability refers to “objective and exteriorized signs of the individual” [250]. In consequence, quality of life was considered to be outside the framework of our study. Nonetheless, we assessed two other more closely related components of functioning, namely activity and participation, using the Pediatric Outcomes Data Collection Instrument [273]. The measurement of loss of venous access was deemed not directly assessable in the clinical setting and was not included in the present work.

In conclusion, these two steps identified the signs and symptoms that experts and patients consider that should be used to assess and monitor pediatric PTS. Each sign and symptom differed in its strength of association with PTS severity, as well as its diagnostic properties, when assessing the UE vs. the LE, or younger vs. older patients (as in the case of pain assessment). This information is not only relevant to the following steps of the tool development, but also a thought provoking analysis of the measurement theory and practice of pediatric PTS. The following steps comprise the selection of the items to be included in the tool and the
development of a weighting scheme for the items that allows an overall interpretation of scores consistent with the measurement approach followed herein.

Acknowledgements

We thank Karen Chapman, Christy Greene, Dodge Baena, Sophia Song, Batool Abidi, Jennifer Vincelli, Joy Qu, Aronela Benea, Bassant Gobin, Marilyn Savedra, Prof. Gerry Fowkes, and all the pediatric experts for their contribution to this project. The present study was funded through a Health Research Grant from the Physicians’ Services Incorporated Foundation.

During the conduct of this research Dr. Laura Avila was a research fellow in the Pediatric Thrombosis and Hemostasis Program at The Hospital for Sick Children. Salary support was provided, in part, by an Education grant from Pfizer, Canada.

Author contribution


Conflicts of interest

The authors report no conflicts of interest.

4.7. Footnotes

- i: The tool is intended for the diagnosis and for the severity rating of pediatric PTS.
- ii: PTS is defined as the “chronic venous symptoms and/or signs secondary to DVT and its sequelae” [22, 23].
- iii: The children included in the study were patients with past DVT, who could therefore have PTS at the time of the study. A history of DVT is an essential inclusion criterion, since PTS, by definition, requires a prior DVT. Approximately 20 to 25% of patients were expected to have a diagnosis of PTS. Two experts assessed the enrolled children at
the time of study participation to determine whether they had PTS (diagnosis) and its severity (none to severe).
• iv: We analyzed the relationship between items and PTS diagnosis (yes/no) estimating sensitivity, specificity and ROC curves, as well as the correlation of each items and PTS severity.
• v: PTS severity was determined by two experts (as explained in iii).
• vi: Both the Modified Villalta Scale and the Manco Johnson Instrument are used for making a diagnosis of PTS and establishing PTS severity.
• vii: The use of the term “assessment” is in line with the terms used in the cited publications of the ISTH (i.e., “definition” and “outcome assessment” of pediatric PTS) [64, 65].
• viii: Since PTS by definition follows a DVT, the VS was developed for the measurement of PTS affecting the LE, the most common location of DVT in adults – as indicated by the quoted 96-99% frequency. Importantly, although not specifically indicated in the original publication of the VS, this tool is now accepted for the diagnosis and rating of severity of PTS in adults.
• ix: The CEAP is defined by its developers as a descriptive “classification of chronic venous disorders” [23]. According to its domains, it classifies types of venous disorders on the basis of signs, anatomic distribution, etiology, and pathophysiology.
• x: The “VEIN-TERM consensus” was a transatlantic interdisciplinary consensus developed during the first and second VEIN-TERM meetings.
• xi: The survey is shown in Appendix 4.2.
• xii: The young adults enrolled in the study were diagnosed with PTS during childhood.
• xiii: Items were grouped in domains according to the clinical judgment of one of the researchers (LA).
• xiv: Patients were at risk of having PTS at the time of the study, as explained in point iii.
• xv: The results were evaluated in the following study (item reduction step), and are shown in Chapter 6.
• xvi: In keeping with the inclusion criteria, all enrolled patients had a history of DVT, which is a necessary condition for PTS development, and were therefore at risk of having PTS at the time enrollment. Participants enrolled in the study were assessed for the
presence or absence of signs and symptoms of PTS in the limb affected by their past DVT.

- xvii: Even though the panel of experts was asked to focus on PTS signs and symptoms, which correspond to the Body Functions and Structures component of the ICF-CY, they nominated items that were outside our conceptual framework. As explained in the discussion section of the present paper, some of these items, such as quality of life, were not assessed. Nonetheless, we did measure items more closely related to the Body Functions and Structures component, namely Activity and Participation, in order to address areas that are of concern to experts.

- xviii: All the items were measured by LA. MM provided support for item measurement.

- xix: The full report of item operationalization is shown in Appendix 4.1.

- xx: Research Ethics Board approval and informed consents are included as Appendices 4.3 and 4.4.

- xxi: PTS domains, shown in Table 4.1, were determined using clinical judgment, as mentioned in the Methods section.

- xxii: Correlation with PTS severity, as in iii.

- xxiii: The classification variable for ROC curve analysis was PTS yes/no (i.e., mild, moderate or severe PTS vs. no PTS, as explained in the Methods section).

- xxiv: Best performing items refer to the items with largest AUC. Nevertheless, it must be pointed out that this is a general statement of the results and tables. The aim of the testing was to provide data to an international panel of experts in the following phase of the study so they could choose the items for the sub-indexes based on their own opinion as well as the data (discussed in Chapter 6).
Chapter 5
Study 2b

5.1. Chapter Overview

Chapter 5 is a brief report (letter to the editor) that separately evaluated collateral circulation, one of the signs of PTS.

The Vein Prominence Scale was used in Study 2a to determine the presence and severity of collateral circulation. However, copyright restrictions precluded the re-use of the Vein Prominence Scale (Appendix 5.1) in the new index.

In view of this restriction, Study 2b was designed to develop a simplified version of the Vein Prominence Scale. Inter-rater reliability of the simplified scale and the Vein Prominence Scale was tested.

The simplified scale was incorporated in the new index, as reported in Chapter 6.

Additional comments to the original published paper are added as footnotes, and shown in Section 5.2.
Collateral Circulation In Pediatric Post-Thrombotic Syndrome

(Letter to the Editor, as published in Thrombosis Research 2016, 144: 210-212, reproduced with permission)

Authors: Avila, M.L. 1, Feldman B. M. 1,2,3, Williams, S. 1, Baena, D. 4, Montoya, M. I. 1, Stinson, J. 2,3,5, Kiss, A. 3,6, Lara-Corrales, I. 1, Brandão, L.R. 1

1 Department of Pediatrics, The Hospital for Sick Children, Toronto, ON, Canada

2 Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

3 Child Health Evaluative Sciences, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

4 Creative Services Studio, The Hospital for Sick Children, Toronto, ON, Canada

5 Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Ontario, Canada

6 Department of Research Design and Biostatistics, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Correspondence: Leonardo R. Brandão, Division of Pediatrics, The Hospital for Sick Children, Toronto, Ontario, Canada, M5G-1X8. E-mail: leonardo.brandao@sickkids.ca. Tel: 1-416-813-5401; Fax: 1-416-813-5327.
Dear Editor,

Post-thrombotic syndrome (PTS) is a chronic complication of upper and lower extremity (UE/LE) deep vein thrombosis (DVT) that can affect both children and adults. Its diagnosis is essentially clinical, and is based on the presence or absence of specific signs and symptoms. The Modified Villalta Scale (MVS), and the Manco-Johnson Instrument (MJI) are the current tools used to evaluate LE-PTS in children, as recommended by the International Society of Thrombosis and Haemostasis (ISTH) [64]. These two tools are also used for the evaluation of UE-PTS [65].

Both the MVS and MJI include an item that assesses collateral circulation, termed “venous collaterals on skin” and “dilated superficial collateral veins” [64], respectively. The Vein Prominence Scale [208] is a scoring system more recently proposed for the measurement of vein prominence in UE in children.

Although visible collateral circulation is regarded as an indicator of obstruction, lack of collaterals does not rule out the presence of blood flow obstruction [10]. A study comparing ultrasound findings and collateral circulation, as detected on physical examination, in children with a history of previous central venous lines showed low sensitivity but high specificity of pathologically increased superficial veins for the detection of occlusive thrombosis or absent venous segments [276].

Besides adding to the severity of PTS by increasing the score of the MVS and MJI by one point, the implications of observing collateral circulation on the skin of a child in the setting of PTS are unknown. The present work aimed to investigate the role of collateral circulation, as assessed by physical examination, in children who sustained UE or LE-DVT. To this end, we investigated the association between the results of the Vein Prominence Scale and other PTS signs and symptoms, and between the Vein Prominence Scale and functional impact of this syndrome. In addition, we investigated the inter-rater reliability of the Vein Prominence Scale, and of a simplified scoring system in UE-DVT.

Our study was part of a cross-sectional study investigating the diagnostic properties of 32 signs and symptoms of UE and LE-PTS in 140 pediatric patients aged 1 to 19 years, diagnosed with
unilateral UE (n=70) or LE-DVT (n=70), and who were therefore at risk of PTS. In the present study, we focused on the variables explained below.

Collateral circulation was assessed using the Vein Prominence Scale [208] with the patient lying (whenever feasible), and both sides of the body uncovered to facilitate comparison with the unaffected contralateral extremity. The comparison with the contralateral extremity allowed the effect of skin translucency when evaluating visible veins to be taken into account.

Symptoms of PTS included heaviness, tired limb, tightness, cramps, pruritus, hot or warm limb, paresthesia, skin redness/purple or blotchy skin, swollen limb, and pain, in the affected extremity only. These symptoms were measured using a structured patient-reported or proxy-reported questionnaire for children 10 years of age and older or parents/care givers of younger children, respectively. Symptoms, except for pain, were measured on the basis of their frequency, following the model of the Edinburgh vein study questionnaire (courtesy of Prof. Gerry Fowkes, with permission). Pain intensity was measured using the Faces Pain Scale – Revised (FPS-R) [259].

Signs of PTS (excluding collateral circulation) comprised of thigh or arm circumference difference, skin changes and ulcers. Limb circumference was measured with a Gulick II anthropometric measuring tape. Thigh circumference was measured at the mid-distance between the anterior superior iliac spine and the tibial tuberosity. Arm circumference was measured at the mid-distance between the acromion and the olecranon process. The difference in centimeter between the affected and unaffected extremity was then estimated. Skin changes and ulcers were defined according to the Venous Severity Clinical Score and the Revised Clinical Grade, Etiology, Anatomy and Pathophysiology Classification [62, 76].

The functional impact of PTS was determined using the Pediatric Outcomes Data Collection Instrument (PODCI) [273], as well as a modified version of self-reported endurance and strength [272]. The latter were included in the self- and proxy-reported versions of the symptoms questionnaire. For these instruments, higher scores signified higher functioning.

We also investigated the inter-rater reliability of the Vein Prominence Scale and of a 3-point simplified scoring system in a subset of 24 patients who had sustained unilateral or bilateral UE-
DVT and consented to have their photograph taken. The simplified scoring system was as follows: 0 = No observable veins; 1 = Comparable to levels one and two of the Vein Prominence Scale (i.e., visible veins up to 10 cm long); 2 = Comparable to levels three and four of the Vein Prominence Scale (i.e., visible veins at least 10 cm long, or palpable, or highly distinctive on the skin).

On the same day as the clinical assessment, photographs of their UE and neck area were taken by a professional medical photographer (DB) in a photographic studio, following a protocol standardizing lighting, distance, and patient posture. Three investigators (SW, ILC, LB) independently evaluated the 24 photographs, and scored the collateral circulation of all the cases using the Vein Prominence Scale first, and then using the simplified scoring system.

We used Spearman correlation coefficients to assess the strength of association between collateral circulation and individual signs and symptoms of PTS and between collateral circulation functional and impact of PTS. The analysis was stratified for UE and LE-DVT, as we considered that the clinical manifestations of PTS are likely different in these venous territories. Inter-rater reliability among the three raters was measured using the Krippendorff’s alpha coefficient [277].

Statistical analyses were performed using SAS software, version 9.2 (SAS Institute Inc. Cary, NC, USA), and R (R Foundation for Statistical Computing, Vienna, Austria).

The study was approved by the Research Ethics Boards at The Hospital for Sick Children and the University of Toronto. Informed consent and assent were obtained, as appropriate.

The median age of the 70 patients in the UE-DVT group at the time of study participation was 7 years (25th-75th percentile 3-11 years). Patients were assessed at a median of 3.5 years after DVT (25th-75th percentile 1.8-6.4 years); 91% of DVT were line-related. Median age in the LE-DVT group was 8.3 years (25th-75th percentile 4-12 years). Patients were assessed at a median of 5.1 years after DVT (25th-75th percentile 2.7-8.6 years); 84% of DVT were line-related. Male to female ratio was 1.6 on both groups.

Classically, the development of collateral circulation following DVT has been described to influence the hemodynamics of the venous system, therefore impacting the severity of PTS [10].
However, collateral circulation as determined by clinical assessment showed no correlation, or only weak correlation, with other signs and symptoms of PTS in children affected by UE or LE-DVT (Table 5.1). This finding can be explained by the fact that clinical examination is not sufficient to assess the extent of collateral pathways that develop after a thrombotic event. In fact, several different pathways can develop after DVT, and are affected by the level of obstruction. Enhanced imaging techniques such as magnetic resonance angiography and computed tomographic angiography provide a better assessment of the extent of collateral pathways than clinical examination alone.

Table 5.1: Correlation Coefficients for Collateral Circulation and Other Variables for the Upper and Lower Extremity Groups

<table>
<thead>
<tr>
<th>Variable (n)*</th>
<th>Spearman Rho (95% confidence interval)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upper extremity group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heaviness (27)</td>
<td>-0.16 (-0.51 – 0.24)</td>
<td>0.43</td>
</tr>
<tr>
<td>Tired limb (40)</td>
<td>-0.04 (-0.35 – 0.28)</td>
<td>0.81</td>
</tr>
<tr>
<td>Tightness (27)</td>
<td>-0.14 (-0.49 – 0.26)</td>
<td>0.50</td>
</tr>
<tr>
<td>Cramps (27)</td>
<td>0.26 (-0.13 – 0.58)</td>
<td>0.19</td>
</tr>
<tr>
<td>Pruritus (70)</td>
<td>0.02 (-0.22 – 0.25)</td>
<td>0.89</td>
</tr>
<tr>
<td>Hot/warm limb (70)</td>
<td>0.01 (-0.23 – 0.24)</td>
<td>0.96</td>
</tr>
<tr>
<td>Paresthesia (40)</td>
<td>0.20 (-0.12 – 0.48)</td>
<td>0.21</td>
</tr>
<tr>
<td>Skin redness (70)</td>
<td>0.36 (0.14 – 0.55)</td>
<td>0.002</td>
</tr>
<tr>
<td>Swollen limb (70)</td>
<td>0.26 (0.02 – 0.46)</td>
<td>0.03</td>
</tr>
<tr>
<td>Faces Pain Scale – revised (41)</td>
<td>-0.08 (-0.38 – 0.24)</td>
<td>0.63</td>
</tr>
<tr>
<td>Arm circumference difference</td>
<td>0.21 (-0.03 – 0.42)</td>
<td>0.09</td>
</tr>
<tr>
<td>Endurance (70)</td>
<td>-0.07 (-0.30 – 0.17)</td>
<td>0.57</td>
</tr>
<tr>
<td>Strength (70)</td>
<td>-0.10 (-0.32 – 0.14)</td>
<td>0.42</td>
</tr>
<tr>
<td>PODCI (63)</td>
<td>0.01 (-0.24 – 0.26)</td>
<td>0.95</td>
</tr>
<tr>
<td><strong>Lower extremity group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heaviness (26)</td>
<td>0.44 (0.06 – 0.70)</td>
<td>0.02</td>
</tr>
<tr>
<td>Tired limb (44)</td>
<td>0.17 (-0.14 – 0.44)</td>
<td>0.28</td>
</tr>
<tr>
<td>Tightness (26)</td>
<td>0.29 (-0.12 – 0.60)</td>
<td>0.16</td>
</tr>
<tr>
<td>Cramps (26)</td>
<td>0.00 (-0.39 – 0.39)</td>
<td>0.99</td>
</tr>
<tr>
<td>Itchiness (70)</td>
<td>0.16 (-0.08 – 0.38)</td>
<td>0.19</td>
</tr>
<tr>
<td>Hot/warm limb (70)</td>
<td>0.15 (-0.09 – 0.37)</td>
<td>0.22</td>
</tr>
<tr>
<td>Tingling (42)</td>
<td>0.01 (-0.30 – 0.31)</td>
<td>0.97</td>
</tr>
<tr>
<td>Redness (70)</td>
<td>0.28 (0.05 – 0.49)</td>
<td>0.02</td>
</tr>
<tr>
<td>Swollen limb (70)</td>
<td>0.24 (0.01 – 0.45)</td>
<td>0.04</td>
</tr>
<tr>
<td>Faces Pain Scale – revised (43)</td>
<td>-0.11 (-0.40 – 0.20)</td>
<td>0.48</td>
</tr>
<tr>
<td>Thigh circumference difference (70)</td>
<td>0.14 (-0.10 – 0.36)</td>
<td>0.26</td>
</tr>
<tr>
<td>Endurance (70)</td>
<td>-0.23 (-0.44 – 0.01)</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Despite our findings, collateral circulation can still be regarded as a relevant item in the measurement of PTS. First, observable collateral circulation fits the definition of PTS [62]. Second, collateral circulation can be considered a unique aspect of the syndrome, depending on the approach that researchers follow for the measurement of PTS and how this construct is conceptualized within the chosen measurement model. If PTS is conceptualized as a formative construct, its indicators are not necessarily intercorrelated, and each indicator captures a unique aspect of the construct [125, 148, 154, 159]. For example, patients may be concerned about the aesthetic appearance of prominent veins, and therefore, it could be considered relevant to measure this item, regardless of the correlation with other clinical findings. Further work is required to understand the relative importance of collateral circulation in the assessment of pediatric PTS.

Inter-rater reliability of the Vein Prominence Scale evaluated by three raters in a subset of 24 children with UE-DVT was 0.78 (95% confidence interval (CI) 0.59-0.87). Inter-rater reliability of the simplified collateral circulation scale was 0.81 (95% CI 0.62-0.93). Hence, agreement among users of both scoring systems ranged from substantial to almost perfect, as per Landis and Koch [278].

Collateral veins are more easily recognized in patients with fair skin [208]. In our study, we systematically compared the skin of the affected side to the contralateral extremity (if unaffected), or a different territory (in the case of bilateral DVT), in order to evaluate the impact of skin translucency, and we included patients of different ethnicities. Nevertheless, the clinical evaluation of collateral circulation in some ethnic groups remains challenging.

| Strength (70) | -0.26 (-0.47 - -0.03) | 0.03 |
| PODCI (61)   | 0.00 (-0.25 – 0.25)   | 0.97 |

**Legend:** PTS refers to post-thrombotic syndrome; PODCI, to pediatric outcome data collection instrument

*Note that sample sizes differ according to the whether the patient’s age was within the appropriate range to provide an answer for the variable (i.e., sample size reflects the number of patients who were able to provide a response, the remaining patients were too young)
There are limitations to our study. The evaluation of inter-rater reliability was performed on photographs; this is an artificial scenario as opposed to a real-life direct assessment in the clinical setting. In the latter, relevant factors like the position of the patient can vary. Still, the substantial agreement found herein advocates for the consistency of results, and the interchangeability of raters when a standardized approach is followed for the measurement of clinical findings that rely on subjective interpretation.

In sum, we found that there are no or only weak correlations between collateral circulation, as measured by physical examination, and PTS signs and symptoms, or functional impact of PTS. Nonetheless, this does not necessarily mean that collateral circulation is not relevant in PTS assessment, since its relevance depends on how PTS is conceptualized. The simplified scoring system for the measurement of collateral circulation used here appears to be reliable among raters.

Acknowledgements

We thank Sophia Song, Batool Abidi, Jennifer Vincelli, Joy Qu, Tori Zimmerman, Robert Teteruck, and Prof. Gerry Fowkes for their contribution to this study. The study was funded by a Health Research Grant from the Physicians’ Services Incorporated Foundation.

5.2. Footnotes

• i: The children included in the study were patients with past DVT, who could therefore have PTS at the time of the study.
• ii: Reference [279].
• iii: These patients could or could not have had PTS. Like the patients included in the first part of the study, they were assessed to determine diagnosis and severity of PTS at the time of study participation. Approximately 20-25% of them were expected to have PTS.
• iv: Spearman correlation coefficients were used as the data did not follow a normal distribution.
• v: Item operationalization is shown in Appendix 4.1.
Chapter 6
Study 3

6.1. Chapter Overview

This chapter describes the final steps of tool development. It includes item reduction, item weighting, and a preliminary analysis of the implications of the scores.

We used a modified Delphi technique for item reduction and Multiple Criteria Decision Analysis for item weighting. A more detailed description of how weights are derived from the latter technique is offered in the closing chapter (Chapter 7).

Of note, the index was designed with the ultimate purpose of discriminating between pediatric patients with and without PTS and evaluating changes in the severity of PTS over time. The performance of the tool on the latter aspect will be evaluated in future research.

Additional comments to the original publication are added as footnotes, and shown in Section 6.7.
Development of CAPTSure™, a New Index for the Assessment of Pediatric Post-Thrombotic Syndrome

(As published in Journal of Thrombosis and Haemostasis 2016, 14 (12): 2371-2385, reproduced with permission)

Authors: Avila, M.L. ¹, Brandão, L.R. ¹, Williams, S. ¹, Montoya, M. I. ¹, Stinson, J. ²,³,⁴, Kiss, A. ²,⁵, Feldman B. M. ¹,²,³

¹Department of Pediatrics, The Hospital for Sick Children, Toronto, ON, Canada

²Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

³Child Health Evaluative Sciences, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

⁴Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Ontario, Canada

⁵Department of Research Design and Biostatistics, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Correspondence: Maria Laura Avila, Division of Haematology/oncology, The Hospital for Sick Children, 555 University Avenue, Toronto, ON, Canada, M5G 1X8.

E-mail: pts.research@sickkids.ca
Essentials

• We developed a discriminative and evaluative index for pediatric postthrombotic syndrome (PTS)
• A Delphi-survey was used for item reduction and multi-criteria decision analysis for weighting
• The new index assess limb PTS based on the relative severity of each sign and symptom
• Higher scores related to higher odds of parental dissatisfaction with their child’s condition
6.2. Summary

**Background:** Post-thrombotic syndrome (PTS) is a complication of deep vein thrombosis defined by the presence of characteristic signs and symptoms. We developed a discriminative and evaluative index for the assessment of upper extremity (UE) and lower extremity (LE) pediatric PTS.

**Methods:** The items to be included in the index were voted for by 26 pediatric thrombosis experts invited to participate in a Delphi survey. Subsequent item weighting was based on item importance elicited using multi-criteria decision analysis (MCDA); 122 health care providers and patients/parents were invited to participate in item weighting. The implications of the overall scores were explored by comparison with PTS diagnosis (independently assessed by two clinical experts) and parental satisfaction/dissatisfaction with their child’s current condition.

**Results:** Items voted for inclusion by at least 70% of the Delphi survey respondents (81% response rate) were pain, paresthesia, swelling, heaviness, endurance, collateral circulation and arm circumference difference for the UE, and pain, paresthesia, swelling, heaviness, tightness, tired limb, redness/purple or blotchy skin, endurance, ulcers, and thigh/calf circumference difference for the LE. Items were then weighted by the use of MCDA (82% response rate). The index had an excellent discrimination for patients with/without PTS. For every 10-point increase in index scores (with higher scores being indicative of worse PTS), the odds of parental dissatisfaction increased by 75% and 92% in the UE and LE, respectively.

**Conclusion:** We report the development of the CAPTSure™ (index for the Clinical Assessment of Post-thrombotic Syndrome in children), which reflects collective judgment of the severity of pediatric PTS. We also provide information on the meaning of the scores.
6.3. Background

The frequency of diagnosis of venous thromboembolism (VTE), which classically includes deep vein thrombosis (DVT) and pulmonary embolism, is rising in pediatric patients. A study reported a 70% increase in VTE diagnosis in tertiary care institutions in the USA between 2001 and 2007 [19]. This number is expected to increase even further, accompanying advances in medical and surgical interventions.

Post-thrombotic syndrome (PTS) is a common chronic complication that can follow DVT affecting the upper extremities (UEs) or lower extremities (LEs) of adult and pediatric patients [20, 33, 34]. This syndrome has a negative effect on the self-perceived health status of adults and children [100, 280], and results in economic burden on the health care system [50].

PTS in children may not only interfere with their normal development, but also lead to earlier manifestations of chronic venous disorders later in life. The morbidity associated with this long-term complication in pediatrics is likely to become a relevant health issue, especially in view of the observed rise in global life expectancy, which increased 6 years between 1990 and 2013, and is expected to increase another 10 years by 2030 [281].

The Modified Villalta Scale (MVS) and the Manco-Johnson Instrument (MJI) are recommended (or at least acknowledged) for the evaluation of LE and UE PTS [32, 65] in children.

The MVS and MJI were developed on the basis of clinical knowledge and experience [87], but they have some limitations, as recognized by members of the Pediatric/Neonatal Thrombosis and Hemostasis Subcommittee of the ISTH [64]. One limitation is their scoring system, as both tools equally weight most of their items. Although equal weighting of the items and simple summation to obtain a final score facilitates application of the tools in clinical practice, equal weighting implies that each sign/symptom contributes the same to the final score, regardless of its nature. In addition, simple summation results in the same overall score in patients with distinct manifestations [282]. This suggests that the modifications made to facilitate the application of tools in pediatrics could have compromised their ability to convey disease severity.
In order to address the limitations to the tools currently used for the assessment of pediatric PTS, we developed a new consensus-based evaluative and discriminative index, for the diagnosis and severity rating of PTS. The aim of the present study is to describe the item-reduction and item-weighting phases of the CAPTSure™ (index for the Clinical Assessment of Post-thrombotic Syndrome in children), and to investigate the meaning of the overall scores.

6.4. Materials and Methods

PTS, the construct of interest, was defined in accordance to the literature as “chronic venous symptoms and/or signs secondary to DVT and its sequelae” [22, 23]. We used a formative approach to measurement. Formative models, which are akin to the clinimetric approach, are most appropriate for the measurement of clinical phenomena [88].

The purpose of the CAPTSure™ was to discriminate between pediatric patients with and without PTS, and to evaluate changes in PTS over time.

Given the understanding of PTS as the signs and symptoms that follow a DVT, our measurement framework focused on the Body Functions and Structures component of the International Classification of Functioning, Disability and Health, Children and Youth Version (ICF-CY) [250].

The design of the CAPTSure™ took into account the differences between the upper and lower venous territories, and two sub-indexes were developed in parallel: one for the assessment of UE PTS, and one for the assessment of LE PTS.

A series of cross sectional studies was carried out, as follows.

6.4.1. Item Reduction

Signs and symptoms (items) of PTS were gathered in a previous study comprising item-generation and item-piloting [283]. The item-generation phase involved patients with a diagnosis of PTS during childhood and pediatric thrombosis experts. The item-piloting phase assessed the diagnostic performance of the signs/symptoms of PTS generated in the first phase, and their
relationship with PTS diagnosis and severity in a cross-sectional study involving 140 children with UE or LE DVT [283].

In the present study, the item-reduction phase involved 26 pediatric thrombosis experts, who were asked to select the items that best represented the construct PTS in children. Clinicians with ≥5 years of experience in the care of children with PTS were selected from the ISTH Pediatric/Neonatal Hemostasis Subcommittee directory and from authors of recent publications in pediatric PTS.

Item reduction was conducted with a modified Delphi technique, in the form of an electronic survey developed with Research Electronic Data Capture (REDCap) software [284] (Appendices 6.1 to 6.4). The survey listed different signs/symptoms of PTS, along with a brief summary of the measurement technique and information on the ability of each item to discriminate between subjects with and without PTS, gathered during item piloting [283]. The information included sensitivity and specificity according to the optimal cut-off point corresponding to the Youden index, the area under the receiver operating characteristic (ROC) curve (AUC), and ROC curves, and was presented separately for UEs and for LEs.

Two rounds were planned. In each round, experts were asked to vote to either “keep” or “remove” each item, therefore indicating the items that should and should not be considered in the sub-indexes, on the basis of their own clinical experience and the information provided in the survey. A third option, “not sure”, was only included in the first round (Appendices 6.1 and 6.2). Items receiving less than a pre-specified 70% endorsement for “keep” or for “remove” were circulated a second time (Appendices 6.3 and 6.4). During the second round, experts were provided with the percent endorsement achieved on the first round (partial results). Items that did not meet the 70% endorsement for “keep” on the second round were removed from the sub-indexes.

In parallel with asking which item should be kept or removed, the survey showed the experts an overall estimate of the time required to apply the CAPTSure™ according to the items they chose, in case experts wished to consider the time factor when choosing between/among different measurement techniques for the same or similar signs/symptoms.
Quasi-anonymity of participants [285] was maintained during the rounds.

6.4.2. Item Weighting

The scoring of the final items of the UE and LE sub-indexes was developed with the Potentially All Pairwise RanKings of all possible Alternatives (PAPRIKA) method for multi-criteria decision analysis.

In brief, the PAPRIKA method presents the respondent with a pair of alternatives for a given question. One alternative includes a higher ranked category for at least one item and a lower category for the other item as compared to the other alternative [286]. Only two items and two criteria are compared at a time (Figure 6.1). The PAPRIKA method minimizes the number of comparisons by identifying and eliminating pairs that are implicitly ranked by using the explicitly ranked pairs, via the principle of transitivity [286]. The weights obtained with this method represent the relative importance that each level of each item has to the respondent iii [287].

Figure 6.1. Example of the Pairwise Ranking Question Used in the Study (From 1000minds.com)

A total of 122 participants, including health care providers (HCPs) and patients/parents, were invited to participate in item weighting. Given that the aim of the CAPTSure™ was specifically
to assist HCPs in the assessment of UE and LE PTS, and in view of the relative complexity of the task, more HCPs than patients/parents were invited to participate (4.5:1 ratio).

The HCPs (n=100) who were invited to participate included: (i) clinicians (i.e., excluding basic laboratory researchers) identified by screening the ISTH Pediatric/Neonatal Hemostasis Subcommittee directory; (ii) speakers on pediatric thrombosis listed in the ISTH Subcommittee Report Minutes of meetings held between 1995 and 2015; and (iii) former trainees identified by contacting current and past pediatric thrombosis and hemostasis fellowship programs around the world.

The patients/parents (n=22) who showed interest in participating in further PTS studies during the item-piloting phase were included.

Two separate models were built with the 1000Minds software platform [286], one for UE PTS and one for LE PTS. In each model, respondents were asked to choose “which of these two patients’ signs/symptoms represent the WORSE case of [arm/leg] PTS” (or to indicate their indifference regarding the alternatives, Figure 6.1). The hypothetical cases were stated to be otherwise identical in all other respects.

To decrease the burden of the task, the number of items and item levels was condensed until a reasonable number of possible decisions were reached. In addition, HCPs were randomized to complete only one of the two models. By use of a true random number generator [288], the first half of the participants were assigned to the UE model, and the remainder were assigned to the LE model. Patients/parents were assigned to complete the model corresponding to the location of the DVT they/their child had sustained, to facilitate the task.

As a check of consistency, participants were asked to re-answer two questions. Participants with no identical answers to the two questions, participants who spent less than a median of 2 s on each response or participants who answered “they are equal” to all questions were removed from the analysis.
6.4.3. Meaning of the Overall Scores

In order to test the implications of the scores of the UE and LE sub-indexes, scores were calculated for 140 children with UE DVT (n=70) and LE DVT (n=70), who were assessed during the item-piloting phase [283]. The ability of each sub-index to discriminate between subjects with and without PTS was investigated by calculating the AUC and the sensitivity and specificity according to different cut-off points, including that corresponding to Youden index (optimal threshold point). The presence/absence of PTS was assessed by two pediatric thrombosis experts (L. B. and S. W.), independently of the measurement of individual items included in the sub-indexes vi. In view of the limitations to the current tools, the expert assessors did not apply either the MVS or the MJI in the study.

We also estimated the scores that would be obtained if only the items of the MVS and MJI were considered, but operationalized as established in the new tool presented here [283]. We investigated the ability of the items of the MVS and MJI to discriminate between subjects with and without PTS, as assessed by the pediatric thrombosis experts vi.

Finally, we used simple logistic regression to estimate how scores predicted parental satisfaction, as measured by a modified global question from the Pediatric Outcome Data Collection Instrument (PODCI) [273]. The question asked how parents would feel in the event that their child had to spend the rest of his/her life with his/her “blood clot” condition as it was at the time of assessment. Responses were scored on a five-point Likert item ranging from “very satisfied” to “very dissatisfied”, and were categorized as neutral/satisfied versus dissatisfied. We estimated the sensitivity and specificity of different cut-off scores of each sub-index to classify parents in the two categories.

Analysis was performed using SAS, version 9.2 (SAS Institute Inc. Cary, NC, USA) and MedCalc (MedCalc, Ostend, Belgium).

The study was approved by the Research Ethics Board at The Hospital for Sick Children vii. In the surveys, the invitation e-mail informed participants them that completion of the survey implied consent.
6.5. Results

6.5.1. Item Reduction

Twenty of the 26 experts invited to participate in the first round of the Delphi survey provided complete responses, with one participant opting out from participating (overall 81% response rate). Twenty of the 25 remaining experts (80%) provided complete responses in the second round.

Table 6.1 and 6.2 show the endorsement for the different items and limb measurement methods in each round of the modified Delphi survey.
### Table 6.1. Item Endorsement in the First and Second Round of the Delphi Survey

<table>
<thead>
<tr>
<th>Item</th>
<th>Upper extremity</th>
<th>Lower extremity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frist round</td>
<td>Second round</td>
</tr>
<tr>
<td></td>
<td>K %</td>
<td>R %</td>
</tr>
<tr>
<td>Heaviness</td>
<td>65</td>
<td>20</td>
</tr>
<tr>
<td>Tired limb</td>
<td>65</td>
<td>20</td>
</tr>
<tr>
<td>Swelling</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>Pruritus</td>
<td>25</td>
<td>65</td>
</tr>
<tr>
<td>Paresthesia</td>
<td>65</td>
<td>20</td>
</tr>
<tr>
<td>Redness/purple or blotchy skin</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>Warm/hot limb</td>
<td>35</td>
<td>50</td>
</tr>
<tr>
<td>Cramps</td>
<td>25</td>
<td>55</td>
</tr>
<tr>
<td>Tightness</td>
<td>15</td>
<td>60</td>
</tr>
<tr>
<td>Delayed healing</td>
<td>20</td>
<td>70</td>
</tr>
<tr>
<td>Reported endurance</td>
<td>75</td>
<td>15</td>
</tr>
<tr>
<td>Reported strength</td>
<td>15</td>
<td>70</td>
</tr>
<tr>
<td>Pain, visual analogue scale</td>
<td>30</td>
<td>55</td>
</tr>
<tr>
<td>Pain FPS-R</td>
<td>85</td>
<td>10</td>
</tr>
<tr>
<td>Pain APPT</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>Skin temperature (perfusion)</td>
<td>0</td>
<td>85</td>
</tr>
<tr>
<td>Collateral circulation</td>
<td>90</td>
<td>5</td>
</tr>
<tr>
<td>Varicose veins</td>
<td>20</td>
<td>70</td>
</tr>
<tr>
<td>Durometry</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Skin Pigmentation</td>
<td>30</td>
<td>70</td>
</tr>
<tr>
<td>Skin Induration</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>Skin inflammation</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Venous ulcers</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>Pitting edema</td>
<td>20</td>
<td>75</td>
</tr>
<tr>
<td>Pain on palpation</td>
<td>10</td>
<td>85</td>
</tr>
<tr>
<td>PODCI</td>
<td>65</td>
<td>25</td>
</tr>
<tr>
<td>Strength tested</td>
<td>20</td>
<td>75</td>
</tr>
</tbody>
</table>

**Legend:** K refers to keep; R, to remove; U, to unsure; FPS-R, to Faces Pain Scale – Revised; APPT, to Adolescent and Pediatric Pain Tool; PODCI, to Pediatric Outcome Data Collection Instrument.
### Table 6.2. Endorsement for Limb Measurement in the First and Second Round of the Modified Delphi Survey

<table>
<thead>
<tr>
<th>Limb segment and measurement method</th>
<th>First Round, Measurement Choice %</th>
<th>Second Round, Measurement Choice %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm circumference difference</td>
<td>45</td>
<td>90</td>
</tr>
<tr>
<td>Arm circumference ratio</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>Arm volume ratio</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Not sure</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>None</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Thigh, calf circumference difference</td>
<td>45</td>
<td>95</td>
</tr>
<tr>
<td>Thigh, calf circumference ratio</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>Thigh, calf, ankle circumference ratio</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Thigh, calf, ankle circumference difference</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Thigh circumference ratio</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Thigh circumference difference</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Calf circumference difference</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Not sure</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

#### 6.5.2. Item Weighting

In order to prepare the PAPRIKA models for the UE and LE sub-indexes, the final items were condensed as follows.

Pain intensity was measured using the Faces Pain Scale – Revised (FPS-R) [259], and was categorized as follows: 0 = no pain, 1.0-3.9 = mild pain, 4.0-6.9 = moderate pain, and 7.0-10.0 = severe pain [289]. These sets of cut-points approximate a flat histogram in a general population, with ~ 25% of children in each category (C. von Baeyer, personal communication).

Collateral circulation was measured with a simplified scoring system that classifies visible veins in three levels (none, mild, moderate-severe) [290].
As the PODCI provides information on the Activity and Participation component of the ICF-CY, it was considered to be a separate instrument to assess the impact of PTS, complementing the evaluation of signs/symptoms, and was not included in item weighting.

Finally, given that the number of items in the LE sub-index (n=10) would have resulted in an overly demanding cognitive exercise, increasing the number of possible decisions by 25-40% as compared with the UE model, the items heaviness, tightness, and tired leg were condensed, as they appear to convey close sensorial concepts. The item-item Spearman correlation coefficients estimated in a sample of 27 patients who answered the three questions in the item-piloting dataset were 0.73 (95% confidence interval (CI) 0.48-0.87) for tightness-heaviness, 0.63 (95% CI 0.32-0.81) for tired limb-heaviness, and 0.53 (95% CI 0.19-0.76) for tightness-tired limb. Tightness was chosen to represent these three concepts.

Hence, the final models contained seven and eight items for the UEs and LEs, respectively (Tables 6.3 and 6.4).
Table 6.3. Item Weights for the Upper Extremity Sub-Index

<table>
<thead>
<tr>
<th>Items</th>
<th>Item weight (sum to 100)</th>
<th>Level</th>
<th>Single item score</th>
<th>Rounded scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain intensity</td>
<td>24.3 (95% CI 21.0-27.6)</td>
<td>None</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mild</td>
<td>31.5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moderate</td>
<td>61.9</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severe</td>
<td>100.0</td>
<td>24</td>
</tr>
<tr>
<td>Endurance (compared with the other limb)</td>
<td>16.0 (95% CI 14.2-17.8)</td>
<td>The same or better</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slightly worse</td>
<td>55.9</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Worse</td>
<td>100.0</td>
<td>16</td>
</tr>
<tr>
<td>Swelling</td>
<td>13.0 (95% CI 11.1-15.1)</td>
<td>Never</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
<td>35.8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
<td>100.0</td>
<td>13</td>
</tr>
<tr>
<td>Heaviness</td>
<td>12.9 (95% CI 11.4-14.4)</td>
<td>Never</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
<td>42.7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
<td>100.0</td>
<td>13</td>
</tr>
<tr>
<td>Visible veins (collateral circulation)</td>
<td>12.6 (95% CI 10.4-14.8)</td>
<td>No</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mild</td>
<td>41.1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moderate/severe</td>
<td>100.0</td>
<td>13</td>
</tr>
<tr>
<td>Paresthesia (tingling)</td>
<td>12.1 (95% CI 10.1-14.1)</td>
<td>Never</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
<td>40.4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
<td>100.0</td>
<td>12</td>
</tr>
<tr>
<td>Arm circumference difference</td>
<td>9.1 (95% CI 7.1-11.1)</td>
<td>0-1 cm</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.1-2 cm</td>
<td>56.1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;2 cm</td>
<td>100.0</td>
<td>9</td>
</tr>
</tbody>
</table>

Legend: CI refers to confidence intervals
Table 6.4. Item Weights for the Lower Extremity Sub-Index

<table>
<thead>
<tr>
<th>Items</th>
<th>Item weight (sum to 100)</th>
<th>Level</th>
<th>Single item score</th>
<th>Rounded scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain intensity</td>
<td>25.1 (95% CI 22.7-27.5)</td>
<td>None</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mild</td>
<td>25.9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moderate</td>
<td>59.6</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Severe</td>
<td>100.0</td>
<td>25</td>
</tr>
<tr>
<td>Venous ulcers</td>
<td>20.4 (95% CI 18.1-22.7)</td>
<td>Absent</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>100.0</td>
<td>21</td>
</tr>
<tr>
<td>Endurance (compared with the other limb)</td>
<td>13.2 (95% CI 11.7-14.7)</td>
<td>The same or better</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slightly worse</td>
<td>49.7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Worse</td>
<td>100.0</td>
<td>13</td>
</tr>
<tr>
<td>Swelling</td>
<td>10.3 (95% CI 8.8-11.8)</td>
<td>Never</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
<td>36.9</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
<td>100.0</td>
<td>10</td>
</tr>
<tr>
<td>Paresthesia (tingling)</td>
<td>9.1 (95% CI 7.5-10.8)</td>
<td>Never</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
<td>39.1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
<td>100.0</td>
<td>9</td>
</tr>
<tr>
<td>Tightness (tightness/ heaviness/tired limb)</td>
<td>7.8 (95% CI 6.9-8.7)</td>
<td>Never</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
<td>38.4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
<td>100.0</td>
<td>8</td>
</tr>
<tr>
<td>Redness/ purple or blotchy skin</td>
<td>7.3 (95% CI 6.1-8.5)</td>
<td>Never</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
<td>39.9</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
<td>100.0</td>
<td>7</td>
</tr>
<tr>
<td>Thigh/calf circumference difference</td>
<td>6.8 (95% CI 5.4-8.3)</td>
<td>0-1 cm</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.1-2 cm</td>
<td>52.3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;2 cm</td>
<td>100.0</td>
<td>7</td>
</tr>
</tbody>
</table>

**Legend:** CI refers to confidence intervals

The overall response rate for both final models was 82%, comprising 7% participants who opted-out; 78% of HCPs and 73% of parents/patients allocated to the UE model provided complete responses. Corresponding percentages for the LE model were 80% (HCPs) and 55% (parents/patients).
Three of 47 responses in the UE model (6%), and one of 46 responses in the LE model (2%) were excluded from the analysis, owing to inconsistency.

The final items, item weights, and a simplified rounded scoring for the levels of each item are shown in Tables 6.3 and 6.4. The final scoring of CAPTSure™ is based on a 0-100-point scale, and is obtained by adding up the weights corresponding to the level of each item (shown as “rounded scoring” in Tables 6.3 and 6.4). Higher scores represent greater PTS severity (worse PTS). A score of 100 would correspond to a patient with all clinical features of PTS at their highest frequency/severity (worst possible PTS) at the time of assessment, whereas a score of 0 would correspond to a patient presenting none of the clinical features of PTS.

### 6.5.3. Meaning of the Overall Scores

The characteristics of the 140 children with UE and LE DVT were reported previously [283]. In brief, the median ages of the patients at the time of assessment were 7 years (25th-75th percentile, 3-11 years) for the UE DVT group, and 8 years (25th-75th percentile, 4-12 years) for the LE DVT group. Patients were evaluated at a median of 3.5 years (25th-75th percentile, 1.8-6.4 years) post-DVT in the UE group, and 5.1 years (25th-75th percentile, 2.7-8.6 years) in the LE group. Sixty-one percent of patients in both groups were male, and 91% and 84% of DVTs were line-related in the UE and LE groups, respectively.

Median CAPTSure™ scores in patients with UE DVT and LE DVT were 5 (range, 0-50) and 4 (range, 0-45) points, respectively.

The performance of each sub-index in discriminating between children with and without PTS, and between dissatisfied and neutral/satisfied parents, and the performance of the items of the MVS and MJI are shown in Table 6.5. Different cut-off scores, optimizing sensitivity, optimizing specificity, and the optimal threshold point, are presented. ROC curves are shown in Figures 6.2 and 6.3.
Table 6.5. Performance of Sub-index Overall Scores, and of the Modified Villalta Scale (MVS) and Manco-Johnson Instrument (MJI) Items

<table>
<thead>
<tr>
<th>Comparison</th>
<th>AUC</th>
<th>Cut-off score</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upper extremity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UE sub-index and PTS y/n</td>
<td>0.88 (0.78-0.94)</td>
<td>&gt;0</td>
<td>92.0 (74.0-99.0)</td>
<td>48.0 (33.7-64.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;8.9*</td>
<td>72.0 (50.6-87.9)</td>
<td>93.3 (81.7-98.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;14.1</td>
<td>48.0 (27.8-68.7)</td>
<td>97.8 (88.2-99.9)</td>
</tr>
<tr>
<td>UE sub-index and parental satisfaction</td>
<td>0.74 (0.63-0.85)</td>
<td>&gt;0</td>
<td>91.7 (61.5-99.8)</td>
<td>43.1 (29.3-57.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;13.1*</td>
<td>58.3 (27.7-84.8)</td>
<td>88.2 (76.1-95.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;17.7</td>
<td>33.3 (9.9-65.1)</td>
<td>90.2 (78.6-96.7)</td>
</tr>
<tr>
<td>MVS items and PTS y/n</td>
<td>0.78 (0.67-0.87)</td>
<td>&gt;0*</td>
<td>88.0 (68.8-97.5)</td>
<td>53.3 (37.9-68.3)</td>
</tr>
<tr>
<td>MJI items and PTS y/n</td>
<td>0.78 (0.67-0.87)</td>
<td>&gt;1*</td>
<td>60.0 (38.7-78.9)</td>
<td>88.9 (75.9-96.3)</td>
</tr>
<tr>
<td><strong>Lower extremity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LE sub-index and PTS y/n</td>
<td>0.83 (0.72-0.91)</td>
<td>&gt;0</td>
<td>91.7 (77.5-98.2)</td>
<td>50.0 (32.4-67.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;6.5*</td>
<td>69.4 (51.9-83.7)</td>
<td>82.4 (65.5-93.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;13.1</td>
<td>50.0 (32.9-67.1)</td>
<td>97.1 (84.7-99.9)</td>
</tr>
<tr>
<td>LE sub-index and parental satisfaction</td>
<td>0.81 (0.69-0.90)</td>
<td>&gt;0</td>
<td>92.9 (66.1-99.8)</td>
<td>32.7 (19.9-47.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;9.6*</td>
<td>78.6 (49.2-95.3)</td>
<td>79.6 (65.7-89.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;17.1</td>
<td>57.1 (28.9-82.3)</td>
<td>87.8 (75.2-95.4)</td>
</tr>
<tr>
<td>MVS items and PTS y/n</td>
<td>0.76 (0.64-0.85)</td>
<td>&gt;0*</td>
<td>91.7 (77.5-98.2)</td>
<td>44.1 (27.2-62.1)</td>
</tr>
<tr>
<td>MJI items and PTS y/n</td>
<td>0.75 (0.63-0.84)</td>
<td>&gt;2*</td>
<td>41.7 (25.5-59.2)</td>
<td>100.0 (89.7-100.0)</td>
</tr>
</tbody>
</table>

Legend: AUC refers to area under the curve; PTS, to post-thrombotic syndrome; UE, to upper extremity; LE, to lower extremity
* optimal threshold point
Figure 6.2. Receiver Operating Characteristic Curve for Upper Extremity Sub-Index Scores

Note: The white circle represents the optimal threshold point reported in Table 6.5

Figure 6.3. Receiver Operating Characteristic Curve for Lower Extremity Sub-Index Scores

Note: The white circle represents the optimal threshold point reported in Table 6.5
Simple logistic regression showed that, for every 10-point increase in CAPTSure™ scores, the odds of being dissatisfied with the current condition increased by 1.75 in the UE sub-index (odds ratio 1.06, 95% CI 1.01-1.11), and by 1.92 in the LE sub-index (odds ratio 1.07, 95% CI 1.03-1.11).

6.6. Discussion

We report the development of the CAPTSure™, a discriminative and evaluative index for pediatric PTS. CAPTSure™ consists of two sub-indexes: one for the assessment of UEs and one for the assessment of LEs. Each sub-index has a clinician-reported and a patient/parent-reported component, for evaluation of signs and symptoms, respectively.

The signs/symptoms included in each sub-index were selected according to expert opinion, combining clinical experience and the information provided in the survey. The results suggest that, in some cases, item selection was based on clinical experience. For example, ulcers were included despite the lack of information (the item was not observed during item piloting). Similarly, arm circumference difference and skin redness were chosen despite their relatively lower discriminative ability in the pilot study [283].

Both sub-indexes include more symptoms than the MVS and MJI. These tools only assess pain/swelling as symptoms, on account of the challenge that is entailed assessing subjective findings in the pediatric population [64]. However, the definition and measurement of most clinical signs in current pediatric tools heavily relies on the users’ judgment, and the degree of judgment involved in the evaluation of an item is used to distinguish between objective and subjective measures, with more objective measures involving no human judgment [291]. Importantly, our systematic approach for the measurement of symptoms with a standardized questionnaire, and the development of operational definitions for PTS signs, will probably help to decrease the variability inherent in outcome assessment involving raters, whether patient or HCPs [292]. This will be the subject of future research.

Another difference from the MVS and MJI is the exclusion of collateral circulation from the LE sub-index. It is of note that a previous study showed almost no correlation between visible collateral circulation and any other clinical manifestation of PTS [290]. Pain in the UE and LE sub-indexes is only measured in children aged ≥4 years. Our previous item-piloting study showed that pain was rarely described by parents of younger children, and
that the discriminative ability of this item (measured using a visual analogue scale) was poor. Moreover, a previous study has described the difficulty parents of younger children have in assessing pain in the context of PTS [94].

Following item-reduction, we developed a multicriteria model to elicit the relative importance (or weight) of each item in determining PTS severity \(^{xxiii}\), as judged by HCPs and patients/parents. The ultimate goal was to combine the individual weights in an overall score, in keeping with the theory underlying formative instruments, such as CAPTSure™. The weight of each item in a formative instrument conveys information on the relative and absolute contribution of the item to the construct [168]. Although many clinical indexes score items equally, it is likely that HCPs intuitively assign weights to different signs/symptoms when evaluating PTS severity [95].

An advantage to the method that we used to derive weights is that PAPRIKA allows comparing individual responses. For example, it is interesting to note that patients/parents and HCPs ranked pain as the most important item for both UEs and LEs, and circumference difference as the least important item for UEs. For LEs, patients/parents ranked ulcers as the least important. This could be explained by the infrequent occurrence of ulcers, and hence the fact that patients/parents were not familiarized with this clinical feature and its natural history. Nonetheless, since the number of patients/parents surveyed was relatively small, further work is warranted to obtain generalizable data regarding their preferences.

Both the MVS and MJI classify children with a final score \(\geq 1\) as having PTS. Hence, finding even one sign or symptom, regardless of severity, is indicative of PTS. This very sensitive interpretation of the overall score led HCPs to suggest addition of the category “post-thrombotic changes” to classify mild clinical features when using the MVS [66], or using an MJI score of \(\geq 2\) points (equivalent to physically and functionally significant PTS) to diagnose PTS [94]. These reports suggest a problem when interpreting the scores of the existing pediatric tools.

Current mainstream views on validity conceive it as the interpretations made on the basis of the score of the tool, rather than an intrinsic property of the tool or the score [293, 294]. Validation is the task of generating empirical evidence to support or refute the plausibility of the inferences made, based on the intended use of the tool [295]. Therefore, we evaluated how the overall scores of the UE and LE sub-indexes related to PTS diagnosis and to parental satisfaction. It is of note that this approach is a preliminary attempt to understand the implications of overall scores, as the scores were estimated on the data originally used to guide item reduction.
In addition, we estimated the scores that would have been obtained if the items of the MVS or MJI had been applied. The UE and LE sub-indexes had a larger AUC as compared to the AUC of the items of the MVS or MJI, further suggesting that experts use more elements when evaluating patients than those included in these two pediatric tools. It must be kept in mind that this is not a direct comparison to the MVS and MJI but to their items, as items were operationalized according to the standards of the new tool presented here. The cut-off score to be used for the new sub-indexes depends on the context in which the tool is applied. Given that aggressive therapies are currently used to prevent PTS, a score equal or higher than the cut-off point that optimizes specificity (~15 points, as shown in Table 6.5) could be applied in the setting of research in PTS. In contrast, a score that optimizes sensitivity (i.e., score > 0, Table 6.5), could be used by general practitioners or other HCPs without expertise in pediatric thrombosis for screening purposes, when deciding whether a patient should be referred to a specialist. The score corresponding to the optimal threshold point could be applied when a treatment decision must be made, as these scores are associated with the best trade-off between sensitivity and specificity.

There are some potential limitations of our study. As the number of decisions in PAPRIKA increases with the number of items and levels, levels were condensed and three items (tightness, heaviness, and tired limb) of the LE sub-index were rated under tightness. The higher ranking of heaviness in the UE sub-index might imply that the importance of heaviness would be underestimated in the LE sub-index if it were assigned the same weight as tightness. Nevertheless, although collapsing levels and items leads to a sub-optimal understanding of their individual value, eliminating the item for task simplicity would have resulted in entirely losing the information that the item provides. A future refinement of both sub-indexes might consist of eliminating unusual categories to allow the evaluation of more levels and items. In addition, we used the cut-off values of the MJI (i.e., >1 cm) to categorize the lower levels of limb circumference difference. In a previous study we showed that 1 cm corresponded to the 97th and 90th percentile for right-left arm circumference difference among healthy 1-11-year-olds and 12-18-year-olds, respectively. A 1-cm cut-off may result in slight item level overestimation among right-handed teenagers with right UE DVT, owing to directional asymmetry [268]. Finally, relying on clinical expertise to define PTS diagnosis as opposed to using existing tools may be debatable. It is of note that this approach was necessary to avoid the problems inherent in the
existing pediatric tools. Given that PTS is a syndrome defined and diagnosed solely on clinical grounds [245], and that no test has been able to replace clinical assessment in establishing the presence or absence of PTS, clinical examination remains the only way to ascertain a PTS diagnosis. Given this scenario, we considered that the evaluation of the patients enrolled in our study only by clinicians with > 10 years of experience in the area of pediatric thrombosis was the best approach for determination of PTS diagnosis and severity.

In summary, we report a new index developed following a formative approach. Distinctive features of CAPTSure™ are the separate measurement of UE and LE PTS, the construction of each sub-index using patient and clinician expert opinion, the inclusion of operational definitions for item measurement, and the scoring of the items according to the collective values of HCPs and patients/parents. These features are expected to better represent the construct PTS, improving the measurement properties of the tool. This is particularly important in the setting of research on invasive treatments to prevent PTS, as “new, improved, and better-understood tools for assessing the patient can move therapy development forward” [292].

Acknowledgements

We thank Franz Ombler, Paul Hansen, Dodge Baena, Sophia Song, Batool Abidi, Jennifer Vincelli, Joy Qu, Marilyn Savedra, Gerry Fowkes, Carl von Baeyer, and Janice Nicholson for their help. Laura Avila was recipient of a research fellowship in the Pediatric Thrombosis and Hemostasis Program supported in part by an education grant from Pfizer, Canada. The study was funded by a Health Research Grant from the Physicians’ Services Incorporated Foundation.

Author Contribution

Study design: LA, LB, AK, JS, BF. Surveys/models design: LA, MM. Statistical analysis: LA. Manuscript writing: LA. Critical review of the manuscript: LB, MM, AK, JS, BF.

6.7. Footnotes

- i: Future research will investigate tool performance in evaluation of PTS changes over time.
- ii: Photographs of the measurement devices were also included.
- iii: The theory underlying PAPRIKA is further explained in the Chapter 7.
• iv: The sampling strategy aimed at identifying HCP with interest in pediatric thrombosis worldwide, in order to sample from the targeted population of intended users.
• v: Each patient was assessed by one of the experts.
• vi: In a separate analysis (not included in the original paper), we estimated the correlation between the sub-index scores and PTS severity (none, mild, moderate, severe), as measured by the same experts (LB, SW), using the Spearman correlation coefficient. The results are shown in footnote xii.
• vii: Research Ethics Board approval is shown in Appendix 3.1.
• viii: The final items of each sub-index are shown in Tables 6.3 and 6.4.
• ix: The levels of the items shown in Tables 6.3 and 6.4 were based on the item scoring levels described in Appendix 4.1.
• x: i.e., participants with no identical answers to the same question, participants who spent less than a median of 2 seconds on each response, or who answered “they are equal” to all questions were removed from the analysis (as described in the Methods section).
• xi: Note that the lowest possible CAPTSure™ score is 0 and the highest possible score is 100, with 0 corresponding to a patient without clinical features of PTS, and 100 corresponding to the most severe PTS.
• xii: The coefficient correlation (described in point vii) was 0.67 (95% CI 0.52-0.78) for both sub-indexes.
• xiii: PTS severity here reflects the severity given by the combined weights of the items. When assessing a patient using CAPTSure™, the weights corresponding to the levels of PTS signs and symptoms shown in Tables 6.3 and 6.4 are added up to estimate an overall score that indicates the severity of PTS for that patient.
Chapter 7
Overview and Implications

7.1. Chapter Overview

This thesis reports the development of a new index that allows the diagnosis and determination of the severity of post-thrombotic syndrome (PTS) in pediatric patients, according to the relative importance or weight of clinical findings, as defined by collective opinion.

The aim of the present closing chapter is to summarize the relevance and highlights of this work, addressing special issues and challenges encountered during the development of the tool. The theoretical concepts underlying these issues are explained.

The final sections summarize the overall limitation to this thesis work and present the author’s view on future steps.

7.2. Limitations of the Work

The limitations to the different studies of this thesis have been highlighted in each chapter. This section aims to further discuss one of the main challenges faced during the design of CAPTSure™, namely the use of expert opinion to diagnose and rate PTS severity during item piloting.

As described in Chapter 3, the use of expert assessment can be considered suboptimal and a weakness of the tool. However, it must be pointed out that, to date, no test has been able to replace clinical assessment in establishing the presence or absence of PTS. Whereas many authors propose that expert opinion is a suitable reference standard when objective pathognomonic evidence is not available [296-298], its acceptability as such might be challenged. Not surprisingly, the most challenging aspect of assessing a diagnostic tool is often identifying a reference standard that will be widely accepted. Feinstein argued that although clinicians prefer hard data, even the meaning of hard data had to be established by experts at some point. Lastly, the existence of a true “gold standard” is debatable: even objective tests such as pathology or radiology depend on interpretation [299]. Given that clinical examination
remains the diagnostic tool for PTS, having only experienced clinicians assess the patients enrolled in the study was considered the best approach for the determination of PTS status.

More subtle advantages as well as challenges and limitations to the new index that were not discussed in the original publications of this thesis are discussed in the next sections. Additional theoretical concepts intended to help the reader put these issues in context are shown, following the same order as the steps of tool development, i.e., from item to overall score, to meaning.

7.3. Items and their Measurement

7.3.1. Measuring Objective Signs and Subjective Symptoms

Constructs that can be defined with sufficient specificity can be measured with great precision [300], and are said to be “operationally defined”. Examples of such constructs are time or weight.

Operational definitions specify how to concretely measure an item, so users can replicate the observations and obtain the same results [301]. Operational definitions also allow communication between researchers and subjects, and advance research by providing a transparent link between items and their measures, which can be questioned and perfected over time [301].

Since clinical signs can be seen as items that are more easily operationalized than symptoms, they are usually regarded as being “objective”. However, in many of the clinician reported instruments that evaluate a given aspect of health, the items used to assess “objective signs” do not include defined and precise questions/instructions, and the user must then follow his/her own judgment to measure and rate the sign [291].

The contrast between objective and subjective measures is complex. In psychophysics, which is concerned with the measurement of sensations, some researchers distinguish “objective” from “subjective” on the basis of the data collection method, whereas others consider that a judgment is objective if it can be classified as correct or incorrect according to an external benchmark [302].
When considering measurement instruments used for the assessment of health, McDowell establishes a methodological difference: “objective measures” are based on mechanical methods (laboratory tests) and involve minimal judgment, perhaps only to interpret results. “Subjective measures” involve human judgment in the collection, processing, and interpretation of information. This is in sharp contrast with the prevailing view of “objective signs” vs. “subjective symptoms” in clinical practice, as can be seen in the classification of the items of the Villalta Scale and the Modified Villalta Scale (Appendices 1.1 and 1.4).

According to the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), clinical assessments “depend on the patient or another person, for example, a trained medical professional, spouse, caregiver, or teacher, to integrate observations, transform them into a rating, and record (report) the result (measurement)” [292]. Clinical assessments can be influenced by patient volition and motivation during the assessment process as well as by the judgment of the rater, and can fluctuate beyond the variations in the condition of the patient. The judgment of the rater can also be influenced by prior experiences and biases. Akin to the definition of objective measures provided by McDowell, the term “biomarkers”, in the ISPOR view, refers to an assessment that is not or is only slightly influenced by judgment.

Although psychophysics has shown that people can accurately judge the relative magnitude of stimuli, it has also shown that there is always a tendency to over or underestimate them. Statements about personal health may not only be affected by perceptual under and overestimation, but also by bias given the more personal nature of the subject. Biases can root from the personality of the patient and from the way the questionnaire is perceived [291].

In sum, the measurement of signs and symptoms is challenging, and even “objective signs” such as those included in the measurement of PTS can be influenced by the judgment of the rater.

As briefly discussed in Chapter 6, the development of a consistent approach and operational definitions for the measurement of signs and symptoms of PTS in the new sub-indexes is expected to decrease the variability inherent to the degree of judgment their assessment involves. In addition, the broader evaluation of the patient/parent’s perspective of PTS by including more symptoms, as compared to the traditional tools, may prove valuable to assess the outcome.
Lastly, the more rigorous approach to item measurement is likely to improve the reliability of the tool, which remains to be evaluated.

### 7.3.2. Measuring Symptoms: Frequency vs. Intensity

Although frequency and intensity might be the most common aspects of symptoms assessed in clinical practice, there are many other dimensions to consider, including duration over time, fluctuation, character, coping strategies, and the level of distress they cause [303, 304].

While the number of categories, labeling, and anchors of rating scales have been investigated in depth, only more recently has the effect of measuring frequency vs. intensity of a symptom received more attention [305]. A study that applied Rasch analysis to investigate the difference between patient-rated frequency and intensity of the different items of the Functional Assessment of Chronic Illness Therapy tool showed a high correlation between frequency and intensity. Frequency scales appeared to perform better than intensity scales in differentiating people at the lower end of the construct (less symptomatic patients) [306].

Another study suggested that recalling frequency and duration of pain appeared to be more accurate than recalling intensity [307]. Frequency scales were also found to be slightly more stable across time than intensity scales in the measurement of items related to depression, probably because they are easier to use [305, 308].

In the new PTS upper and lower extremity sub-indexes, pain is measured and weighted in terms of intensity and the remaining symptoms are measured and weighted in terms of frequency (the latter according to the questionnaire of the Edinburgh Vein Study). The potential marginal advantage of measuring more aspects of a symptom, like assessing both frequency and intensity, should be weighted against patient burden.

### 7.3.3. Measuring Edema: The Role of Bioimpedance

Study 1 (Chapter 3) investigated normal values of fluid content in the upper and lower extremities of healthy children using bioimpedance spectroscopy (BIS), a non-invasive, simple, and fast to perform method to assess fluid content.
Whereas methods based on isotopic dilution are considered the reference for the assessment of body fluid distribution, BIS is an indirect method. The foundations of BIS date back to the late 1960s. This method gained popularity in the context of evaluation of body composition in the 1980s and has, since then, become widely used to assess the total water content of the body and body composition for fitness and weight loss purposes. Advances in technology during the past decade led to the development of multiple and single frequency devices that have allowed the measurement of body segments.

The theoretical advantage to using BIS in the field of PTS is that BIS measures fluid content only, overcoming the bias introduced by weight/fat and muscle changes in a swollen limb. Experts in adult clinical practice consider BIS to be the reference standard for measuring limb volume in disorders where fluid accumulation is a central feature, such as the case of PTS.

In consequence, the use of BIS in the measurement of limb edema in pediatric PTS seems theoretically sound. However, we found a low correlation between fluid content ratio, measured by BIS, and PTS severity in Study 2 (Chapter 4), perhaps suggesting that BIS is not a good method to assess fluid accumulation in pediatric PTS. The finding could be explained by the study design: PTS severity was established by a thrombosis expert who measured limb circumference with a regular measuring tape, and then used the information together with other clinical features to grade the severity of PTS according to the expert’s opinion. However, limb fluid content, determined by either BIS or circumference difference, is only one aspect of PTS, which is a composite construct. In Study 3 (Chapter 6), we saw that limb circumference difference was the item collectively ranked least important in determining PTS severity. It is likely that the thrombosis experts who determined PTS severity also ranked limb circumference least important in their assessment of PTS. Therefore, even if BIS did reflect fluid content more accurately than a measuring tape, its role would be obscured by the low value attributed to the item it measures (i.e., limb circumference difference).

Measurement of limb circumference with a measuring tape remained the item and technique of choice for the new sub-indexes, despite the theoretical limitations described in Study 1 (Chapter 3). The choice is likely due to the simplicity of performing circumferential measurements in children, as compared to BIS assessment, in addition to the lower cost.
The use of BIS in pediatric PTS remains to be further investigated. In lymphedema, for example, this technique has evolved to play a central role in early detection [73]. The same potential application is yet to be explored in pediatric PTS.

7.3.4. Measuring the Impact of Disease: Functioning and Function

The International Classification of Functioning, Disability, and Health (ICF) provides a framework and standardized language for the classification of health and health-related domains.

The Children and Youth Version (ICF-CY) [250] considers the following components: Body Structures (i.e., parts of the body, such as the upper or the lower extremity), Body Functions (i.e., physiological and psychological functions, like power or strength), Activity (i.e., execution of a task or action by an individual, for example running), and Participation (i.e., involvement in a life situation, such as playing baseball).

Since by definition PTS is a collection of signs and symptoms, it conceptually pertains to the impairment of Body Structures component of the ICF-CY [250]. However, experts and researchers often intuitively attempt to evaluate the disease from a more global perspective. For example, during item generation (Chapter 4, Study 2), despite being instructed to provide signs and symptoms of PTS, one-third of the experts who participated suggested a number of different measures that were neither signs nor symptoms (as shown in Table 4.1).

Many instruments in health assessment attempt to evaluate global functioning, sometimes defined as “how far the individual is able to function normally and to carry out typical daily activities” [291], and thus analyze the impact of the signs and symptoms of a disease. In the ICF, functioning refers to Body Structures and Functions, Activity, and Participation. In contrast, disability refers to Impairments, Activity Limitations and Participation Restrictions. Clinicians often refer to functional outcomes as the impact of the disease on the Activities and Participation components [309].

In view of the importance of defining terms clearly, as stressed in the opening chapter of this thesis, the new index was developed considering the construct PTS to correspond with the Body Structures component of the ICF-CY. In addition to the signs and symptoms elicited during item generation, items for the measurement of Body Functions in the form of strength (tested and
reported) and endurance were also assessed. The value of the Pediatric Outcome Data Collection Instrument (PODCI), a tool that measures Activities and Participation in children, was also explored. Nevertheless, consistent with the design of the new instrument, PODCI was considered a separate and complementary instrument that could be appended when measuring health outcomes in PTS.

Although the items of PODCI assess activities in both upper and lower extremities, it is likely that a child with impaired endurance but not strength, as shown by the findings of Study 2, may find that most of the upper limb activities listed in PODCI (such as lifting heavy books, pouring half a gallon of milk, putting on a coat, writing with a pencil) are simply very easy to perform, and may find that certain lower limb activities (such as walking more than a mile, riding a bicycle or tricycle, or climbing three flight of stairs) that require endurance are very difficult.

Therefore, although PODCI might be a reasonably good approximation to assess the impact of PTS on the lower extremities, it is not sensitive enough to measure the impact of PTS on the upper extremities.

7.4. The Overall Score

According to the ISPOR, when assessing outcomes in health care, the information can be obtained from different sources: from patients (patient-reported outcome assessment), trained investigators (clinician-reported outcome assessment), and untrained observers (observer-reported outcome assessment) [292]. In clinician- and observer-reported outcome assessment, external raters apply their judgment to rate the response of the patient and the relative weight of different observations [292].

The rater assigns a numerical score to each item, usually based on ordinal sub-scales, and sub-scales are then commonly summed up to present the information in the form of an overall score, which is interpreted in a linear manner: higher scores represent a more severe disease.

Disease severity can be measured for different purposes: to establish prognosis, to determine the impact of the disease on the patient, to determine a course of treatment, or to establish disease activity and treatment response [310].
Given the relevance and implications of determining disease severity, it could be argued that tools that apply equal item weighting and simple summation to obtain an overall score, which is then interpreted as a linear representation of disease severity, can be problematic. This specially applies to tools that use ordinal sub-scales to rate the items, as do most instruments, because of the mathematical properties of ordinal scales. As McDowell states “adding ordinal answer scales may lead to incorrect conclusions, and this is the main motivation for developing more accurate scale weights for answer categories” [291]. Forman further argues that adding ordinal numbers can lead to meaningless results [311].

In the case of the new PTS sub-indexes, individual items were weighted according to the opinion and preferences of health care providers and patients/parents. This perspective is in keeping with the underlying measurement theory, presented in Chapter 2, and results in an overall score that more accurately reflect PTS severity. To this end, we used the PAPRIKA method, an acronym for **Potentially All Pairwise RanKings of all possible Alternatives** (Study 3, Chapter 6), which belongs to the multiple criteria decision analysis (MCDA) family, and produces weights that represent the relative value or utility of each level of the items [287].

MCDA is a family of stated preference methods that uses choice modeling to elicit preferences and values, to determine the weight of items or criteria. Stated preference methods refer to the preferences that are determined according to the decisions the respondents make by ranking, rating, and choosing in a hypothetical scenario [287].

MCDA has been defined as “a methodology for appraising alternatives on individual, often conflicting criteria, and combining them into one overall appraisal” [312, 313]. It allows aggregating the opinion of individuals into a collective judgment, using a wide range of techniques. Judgments made using MCDA are referred to as priorities, preferences, importance, and values.

MCDA methods have slowly been adopted by health care research. A systematic review identified 66 publications in health care between 1981 and 2013, half of which were published between 2008 and 2013 [314].
In health research, the opinion of members of regulatory committees, boards of directors, general public and patients is sought to reach decisions related to four broad categories: benefit-risk assessment, health technology assessment, portfolio decision analysis, and commissioning decisions/priority setting frameworks [312]. The use of MCDA to establish disease severity in the present thesis falls outside these more common categories.

There are a number of approaches that MCDA methods can use to generate weights. A summary of the different weighting and scoring methods has recently been reported by the ISPOR Task Force [315]. The PAPRIKA method requires each respondent to answer a series of pairwise-ranking questions (i.e. choosing one option from two). Each question involves ranking two hypothetical alternatives defined on two criteria at-a-time and involving a trade-off.

Because the alternatives that the respondent is asked to choose between combine two criteria at-a-time, the PAPRIKA method can be thought of as a type of conjoint analysis. Conjoint, in the present context, means that the criteria are combined together. Thus, conjoint analysis is a method used in MCDA to estimate weights on criteria. Although conjoint analysis is sometimes used to refer to choice modelling and discrete choice experiments, in the present context "conjoint" means that the criteria are combined together. Moreover, PAPRIKA can be thought of as a type of adaptive conjoint analysis, because each time a respondent answers a pairwise-ranking question, a new question is chosen based on all the previous answers (i.e. the method adapts).

Whereas most other conjoint analysis methods determine the weights on the criteria using regression analysis, PAPRIKA uses quantitative methods based on linear programming (also known as linear optimization) [286].

Another difference between the PAPRIKA and most other methods is that PAPRIKA generates a set of weights for each individual participant in a conjoint survey, in contrast with other methods, which mostly produce aggregated data only. Individual-level data enables subsequent investigations into the extent to which participants’ responses have similar patterns of weights. Some of these differences have been noted in Chapter 6, Study 3. Nonetheless, it must be pointed out that the subgroup of patients/parents included in the item weighting study was too small to draw definitive conclusions.
The theory of conjoint analysis allows transforming ordinal preferences into interval and ratio scales [316]. Therefore, in addition to conveying the stated preferences of health care providers and patients/parents, the final score of CAPTSure™ has sound mathematical properties, which would allow inferring the higher precision of CAPTSure™ as a tool to measure PTS. In fact, “the precision of scoring methods depends on whether they display interval properties” [315]. Nevertheless, whether preference weighted scores result in improved decisions for patient management remains to be studied.

### 7.5. The Meaning of Scores: On Validity and Validation

#### 7.5.1. Current Views on Validity

In the late 1980s and 1990s, the concept of validity shifted from being represented by different types [317], aspects [318], or categories [319], according to the definitions provided in the first versions of the *Technical Recommendations and Standards*, to a more comprehensive and unified notion that relied on various sources of evidence to support the use of a test for a given purpose. The conceptual shift is reflected as such in the nomenclature of the modern *Standards* [91, 320].

Messick proposed the currently accepted view of validity, which is defined as “an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores or other modes of assessment” [321]. More simply, “score validation is empirical evaluation of the meaning and consequences of measurement” [142]. This concept emphasizes a more prominent role of the user in the validation process [294]. In Messick’s view, “validity is an evolving property and validation a continuing process” [142].

The idea that validity is not a property of a test or of a score, but an inferential interpretation of the scores in the context of the use of the test, has been stated and emphasized over the past 70 years [293, 294]. According to the current version of the *Standards* it is not correct to employ the unqualified phrase “the validity of the test” [91], unless the interpretation or the use of the tool is fixed [293]. Borsboom et al note that the current mainstream view of validity that theorists have is “strangely divorced” from the view the actual researchers have in mind [140].
The daunting task of validating the inferences of scores can be achieved by the pragmatic approach proposed by Kane in 2006 [293, 322, 323]. The approach is based on the development of an interpretative argument by defining the inferences to be made on the scores, establishing hypotheses or sources of evidence to prove or disprove these inferences, and collecting and analyzing the evidence [293]. “The best that can be done is to show that the interpretative argument is highly plausible, given all available evidence” [295].

One of the problems encountered frequently in research is that modest interpretative arguments postulated during the validation of tools are followed by more ambitious interpretations in practice [293], leading to a “begging the question” fallacy, according to Kane.

A summary of the evolution of measurement theory up until the 1980s can be found in Appendix 7.1.

7.5.2. Validity and Validation in Medicine

Although measurement in medicine and other sciences differs from psychology in its line of reasoning and methodology [140], the approach to validity and validation in medicine has been largely adapted from early psychology theory [291]. For example, an international consensus on taxonomy and terminology for health instruments (the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) taxonomy [96]) recently defined validity as the “degree to which a health-related patient reported outcome instrument measures the construct(s) it purports to measure”. Nevertheless, some authors offer a more modern view. In fact, McDowell states that “the core idea in validity concerns the meaning, or interpretation, of scores on a measurement” [291].

Over thirty years ago, Feinstein coined the term clinimetrics to refer to the science of measurement of clinical phenomena, using indexes and rating scales to describe signs, symptoms and other phenomena in medicine [87, 203]. In line with Feinstein, Fayers argued that psychometric instruments were different from and clinimetric instruments, and that their development “should follow different paths” [88]. He further stated that the difference between reflective and formative indicators (discussed in Chapter 2) accounted for the difference between
psychometrics and clinimetrics [88]. de Vet et al later discussed the role of formative and reflective models in measurement in medicine [84].

Not surprisingly, validity in formative models is a complex topic. Borsboom et al state that all measurement procedures have a common characteristic, namely “the either implicit or explicit assumption that there is an attribute out there that, somewhere in the long and complicated chain of events leading up to the measurement outcome, is playing a causal role in determining what values the measurements will take” [140]. Since the relation between formative indicators and their constructs is reversed, with the items pointing towards the construct rather than otherwise, the authors argue that validity does not apply to formative models. According to the authors, validity requires the existence of an attribute that, upon variation, causally produces variations in the outcome of a test or measurement procedure. In fact, they state that formative models “might be better conceptualized as models for indexing or summarizing the indicators, or as causal models that do not involve a measurement structure for which one can ask the question of validity” [140]. In contrast, Zumbo considers that the Borsboom’s reliance on causal (reflective) models results in a restricted view of measurement [293], highlighting the ongoing discussions in this area.

Fayers et al proposed the following aspects as relevant in the validation of instruments containing formative indicators: Clinical sensibility (equivalent to content validity), comprehensive coverage of items, emphasis upon items that patients rate as important, and emphasis upon items that patients experience frequently [274].

Our strategies for item generation and reduction resonate with those proposed by Fayers, while the preliminary investigation on the implications of the scores in PTS suggests the new tool can be interpreted in a more meaningful manner, as compared to current pediatric instruments.

Indeed, even though the MJI has gone through a generally accepted process of validation and the MVS highly correlates with the MJI, the implications of the scores of both tools are problematic in clinical practice, as discussed in Chapter 1 [66, 94]. These facts illustrate the relevance of the modern mainstream concept of validity and validation.
Kane states that the evidence required for validation depends on the interpretation and use of the tool. If the scores of a test have a direct and simple interpretation, little or no evidence would be required for validation [293]. In the case of the new pediatric PTS sub-indexes developed in the present thesis, their interpretation can indeed be considered simple: higher scores are considered to reflect worse PTS. Nevertheless, Chapter 6, Study 3, discusses the implications of the scores in terms of their association with independent PTS diagnosis and severity rating, and with parental satisfaction.

Understanding the meaning of scores is particularly important in the measurement of PTS, given the ongoing emphasis on its prevention with aggressive treatment of the index deep vein thrombosis, and the expected increase in the frequency of PTS in pediatrics.

7.6. Relevance of the Work

CAPTSure™ is a new PTS index for the diagnosis and severity rating of pediatric PTS affecting the upper and lower extremity.

The steps followed for tool development included PTS conceptualization (i.e., the process of specifying the meaning and the use of a term [324]) and item generation, item operationalization, item piloting, reduction, and scoring. The thesis also explored the implications of the scores. The formative measurement approach was followed to link the construct PTS and the items that define it.

The final product is a well-defined measurement tool that assesses the upper and lower venous territories separately, and considers the input of both clinicians and patients or proxies/parents. In addition, the final score is mathematically and methodologically sound, increasing precision in the assessment of patients.

All these features suggest that CAPTSure™ offers an improved approach to the measurement of pediatric PTS. This, in turn, expected to translate into better decisions in the clinical care setting and more appropriate research designs, ultimately resulting in improved health outcomes. In the opening chapter, it was mentioned that quality in health care is defined as the degree to which health care increases the likelihood of a favorable outcome, and is consistent with current
knowledge. CAPTSure™ is expected to enhance both the quality of research and of patient care, but this remains to be proven in further empirical research.

The new tool stands in contrast to the existing pediatric instruments. As discussed in Chapter 1, the Modified Villalta Scale and the Manco-Johnson Instrument have limited item operationalization and an oversimplified approach to scoring each item and obtaining an overall score. These limitations result in excessive reliance on the judgment of the users, and therefore on their expertise, and in difficulties in the interpretation of overall scores.

7.7. Future Directions

Each of the steps followed to construct CAPTSure™ is expected to contribute to the field of PTS, improving the diagnosis and assessment of severity of the syndrome. However, further testing of the new index is required, and future studies should aim to investigating different aspects of the new tool. Some examples are discussed below.

In formative models, analysis of test-retest reliability of the items and of inter-rater reliability are recommended [125, 325]. These aspects are essential to the temporal stability of the new index, and are akin to the “reliability/precision” aspect of reliability described by the Standards [91, 92]. Whether scores remain constant from one measurement to another is especially relevant in the case of items that assess symptoms, since in addition to the reliability of the measurement procedure, symptom scores can differ due to differences in subjects [92]. In contrast, the variability in the scores of items that measure signs will likely be associated with differences among clinical raters; therefore, inter-rater reliability needs to be explored.

Since validation (in its mainstream conceptualization) is a continuous process, additional steps towards deepening our understanding of the meaning and implications of scores in the new sub-indexes are warranted. CAPTSure™ should be tested in a larger spectrum of patients (e.g., including patients with arterial thrombosis).

Following the model proposed by Kane [326], an interpretative argument can be envisioned for CAPTSure™. This argument is presented below within a framework developed by Thurler et al
[327, 328] that is useful to represent normal and abnormal states (i.e., health and disease) of a biological system.

In Thurler’s framework, PTS is considered a dysfunctional state of the vascular system. This state is represented by the collection of values of each of the signs and symptoms of PTS at a given instant. Changes in the values of signs and symptoms reflect a new state for that patient. Within this framework, PTS is represented as a trajectory across different states, from health to disease, normal to abnormal, non-pathological to pathological (shown in Appendix 7.2). CAPTSure™ scores reflect different PTS states, with increasing CAPTSure™ scores, from lower to higher (Appendix 7.2) reflecting greater PTS severity.

In Chapter 6 we explored the cut-off score that would represent the transition of individuals from a non-diseased to a diseased state. A practical and relevant cut-off yet to be established is that used by clinicians to consider treating PTS. In this case, the interpretative argument would be as follows:

(1) The presence of one or more of the clinical signs or symptoms of PTS assessed in CAPTSure™ (i.e., score >0) implies that the patient is moving to the right in the PTS trajectory (Appendix 7.2), from a less severe to a more severe state.

(2) A cut-off score for treatment would indicate that children at or above that score are at a level of the PTS trajectory that would justify the use of a compression garment.

Since the scoring of CAPTSure™ weighs the opinion of health care providers about the items that represent severity, it would be expected that the treatment cut-off would be higher than the cut-off that separates diseased from non-diseased. Analyzing what clinicians actually do in clinic, and estimating the scores of patients who are prescribed a compression garment is a possible source of evidence for this argument.

Lastly, given that the assessment of other aspects of functioning is recommended by the ISTH in the context of PTS [65], comparison of scores with those of tools that assess activity limitation and participation restriction, as separate constructs, is of relevance. Nonetheless, if and how exactly PTS interferes with functioning remains poorly explored. In our work, we found low sensitivity of one of the existing tools (PODCI) to measure activity limitation and participation
restriction, especially in the upper extremity. Of note, our search of the literature did not identify a better-suited pediatric tool, since most do not particularly target endurance. Therefore, to fully explore the impact of PTS on functioning, an instrument able to detect the consequences of impaired endurance should be developed. Ideally, the items of such a tools should reflect the patient’s everyday life, i.e., what a “patient does or would want to do” [292], following an approach that considers the individual’s perspective on what is important.
References


214. CDC. A SAS Program for the CDC Growth Charts. 2011 Accessed April 6, 2105.
224. Eberley, C. Figure of Eight Method of Measuring Ankle Joint Swelling. Accessed April 6, 2015.


147


Appendices

Appendix 1.1. Villalta Scale

<table>
<thead>
<tr>
<th>Subjective symptoms</th>
<th>Objective signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heaviness</td>
<td>Pretibial oedema</td>
</tr>
<tr>
<td>Pain</td>
<td>Induration of the skin</td>
</tr>
<tr>
<td>Cramps</td>
<td>Hyperpigmentation</td>
</tr>
<tr>
<td>Pruritus</td>
<td>New venous ectasia</td>
</tr>
<tr>
<td>Paresthesia</td>
<td>Redness</td>
</tr>
<tr>
<td></td>
<td>Pain during calf compression</td>
</tr>
<tr>
<td></td>
<td>Ulceration of the skin</td>
</tr>
</tbody>
</table>

PTS is classified as mild if the score is 5–9, moderate if the score is 10–14 and severe if the score is ≥15 or a venous ulcer is present. Each sign or symptom is graded with a score between 0 and 3. The presence of ulcer is only noted.

Legend: PTS refers to post-thrombotic syndrome
Reproduced from Prandoni et al, with permission [329]

Appendix 1.2. Clinical, Etiological, Anatomical, and Pathophysiological Classification System

<table>
<thead>
<tr>
<th>Clinical signs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 0</td>
<td>No visible or palpable signs of venous disease</td>
</tr>
<tr>
<td>Class 1</td>
<td>Telangiectasia or reticular veins</td>
</tr>
<tr>
<td>Class 2</td>
<td>Varicose veins</td>
</tr>
<tr>
<td>Class 3</td>
<td>Oedema</td>
</tr>
<tr>
<td>Class 4</td>
<td>Skin changes ascribed to venous disease</td>
</tr>
<tr>
<td>Class 5</td>
<td>Skin changes as described above with healed ulceration</td>
</tr>
<tr>
<td>Class 6</td>
<td>Leg ulceration, skin changes as defined above</td>
</tr>
<tr>
<td>Aetiological</td>
<td>Congenital, primary, secondary</td>
</tr>
<tr>
<td>Anatomical</td>
<td>Superficial, deep, or perforator, alone or in combination</td>
</tr>
<tr>
<td>distribution</td>
<td>Pathophysiological dysfunction</td>
</tr>
<tr>
<td></td>
<td>Reflux or obstruction, alone or in combination</td>
</tr>
</tbody>
</table>

Legend: PTS refers to post-thrombotic syndrome
Reproduced from Prandoni et al, with permission [329]
Appendix 1.3. Manco-Johnson Instrument

<table>
<thead>
<tr>
<th>Patient ID: __________________________</th>
<th>Date of Birth: __________________________</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of Thrombus Diagnosis: __________________________</td>
<td>Date of Assessment: __________________________</td>
</tr>
<tr>
<td>Affected limb (circle): Arm □ Leg □</td>
<td>Affected side (circle): Right □ Left □</td>
</tr>
</tbody>
</table>

### PHYSICAL FINDINGS (Signs)

<table>
<thead>
<tr>
<th>Limb Circumference Measurements</th>
<th>Affected (circle: R L)</th>
<th>Unaffected (circle: R L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid-proximal limb</td>
<td>______ cm</td>
<td>______ cm</td>
</tr>
<tr>
<td>Mid-distal limb</td>
<td>______ cm</td>
<td>______ cm</td>
</tr>
</tbody>
</table>

### Basic CEAP: Mark an “X” where applicable/present.

<table>
<thead>
<tr>
<th>Physical Findings</th>
<th>Affected (circle: R L)</th>
<th>Unaffected (circle: R L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. No visible or palpable signs of venous disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Swelling, with or without pitting edema</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Dilated collateral circulation of extremity only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Skin changes ascribed to venous disease (i.e., pigmentation, venous eczema)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Skin changes as in 3 with ulceration or superior vena cava syndrome</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### FUNCTIONAL FINDINGS (Pain Symptoms)

**Wong-Baker Faces Pain Rating (Oucher) Scale:** Score 0-5 for each.

<table>
<thead>
<tr>
<th>Pain Outcome: Wong-Baker (Oucher) Scale</th>
<th>Affected (circle: R L)</th>
<th>Unaffected (circle: R L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>With aerobic exercise only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With activities of daily living</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At rest</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If pain is present (i.e., score 1-5): Does the pain interfere with activities? Yes □ No □

**Comments:**

### Wong-Baker Faces Pain Rating Scale

![Wong-Baker Faces Pain Rating Scale Image]

* Aerobic exercise only: implies that symptoms are present only when child engages in vigorous age-appropriate sport such as running, lap swimming, soccer, basketball or volleyball.

* Activities of daily living: implies that a child is symptomatic when engaging in ordinary age-appropriate activities in the home, school and community short of organized sports and vigorous aerobic activities. These symptoms limit and alter a child’s ordinary day-to-day activities such as walking at school, shopping with the family or participation in a birthday party.

* At rest: implies a constant presence of symptoms that is independent of activity. The child’s daily life is severely limited by symptoms.

Reproduced from Goldenberg et al, with permission [98]
Appendix 1.4. Modified Villalta Scale

<table>
<thead>
<tr>
<th>The PTS scoring sheet</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symptoms</strong></td>
<td><strong>Score</strong></td>
</tr>
<tr>
<td>Subjective symptoms</td>
<td></td>
</tr>
<tr>
<td>Pain or abnormal use</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Swelling</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Objective signs</td>
<td></td>
</tr>
<tr>
<td>Change in skin color</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Increase in limb circumference</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Pitting edema</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Collateral vessels</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Pigmentation</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Tenderness on palpation</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Edema of the head</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Varicosities</td>
<td>0 (absent)</td>
</tr>
<tr>
<td>Ulceration</td>
<td>0 (absent)</td>
</tr>
</tbody>
</table>

Reproduced from Kuhle et al, with permission [29]

**Interpretation:**

<table>
<thead>
<tr>
<th>Mild Post-Thrombotic syndrome</th>
<th>1-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate Post-Thrombotic syndrome</td>
<td>4-8</td>
</tr>
<tr>
<td>Severe</td>
<td>≥9</td>
</tr>
</tbody>
</table>

### Appendix 2.1. Search Strategies and Terms

<table>
<thead>
<tr>
<th>Database</th>
<th>Search Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>OvidSP MEDLINE (1946 to February 25, 2013)</td>
<td>((formative or reflective) adj2 (model* or measure* or indicator*)).ti,ab.</td>
</tr>
<tr>
<td>OvidSP EMBASE (1947 to 2013 Week 08)</td>
<td>((formative or reflective) adj2 (model* or measure* or indicator*)).ti,ab.</td>
</tr>
<tr>
<td>OvidSP PsycINFO &lt; 1967 to February Week 3 2013</td>
<td>((formative or reflective) adj2 (model* or measure* or indicator*)).ti,ab.</td>
</tr>
<tr>
<td>EBSCOHost CINAHL 1982 to February 25, 2013</td>
<td>(TX (formative OR reflective) N2 (model* or measure* or indicator*))</td>
</tr>
<tr>
<td>ProQuest ABI/INFOR 1960 to February 25, 2013</td>
<td>AB,TI (formative NEAR2 model*) OR AB,TI (formative NEAR2 measure*) OR AB,TI (formative NEAR indicator*) OR AB,TI (reflective NEAR2 model*) OR AB,TI (reflective NEAR2 measure*) OR AB,TI (reflective NEAR2 indicator*)</td>
</tr>
</tbody>
</table>

Reproduced from Avila et al, open access [282]

### Appendix 2.2. Study Flow Diagram Showing Search Results, Screening, and Inclusions/Exclusions

- **1104 citations initially retrieved**
- **747 unique citations after removing duplicates**
- **357 duplicates**
- **136 citations retrieved in full after potentially matching inclusion criteria**
- **23 included citations**
- **113 citations were excluded as they did not discuss issues on scoring methods**

Reproduced from Avila et al, open access [282]
Appendix 3.1. Research Ethics Board Approvals

Research Ethics Board (REB)

The Research Ethics Board for The Hospital for Sick Children is organized and operates according to the principles and practices outlined in the Tri-Council Policy Statement, the ICH Harmonized Tripartite Guidelines: Good Clinical Practice, and Division 5 and the Medical Devices Regulations of the Food and Drug Act as well as the Natural Health Products Regulations of Health Canada. This signed document is in lieu of the Health Canada Research Ethics Board Attestation Form.

Approval & Terms of Agreement

Investigators: Dr. Leonardo Brandao, B Feldman, J Stinson, A Kiss, M Avila

Study Title: Post-Thrombotic Syndrome in Children: Item Generated for the Development of an Evaluative and Discriminative Tool

REB File number: 1000036542 Level of Continuing Review: I B
Protocol Version Date: February 6 2013
Consent & Assent Form Version Date(s): Participant Consent - February 6 2013, Parent Consent - February 6 2013, Assent – January 6 2013
Investigator’s Brochure Version Date: N/A

I agree to carry out the proposed research involving human subjects in accordance with the above-noted guidelines and regulations (as applicable) and using only the REB-approved study protocol and consent/assent forms. I shall notify the division/department head and the REB prior to implementing any amendments to the protocol and consent/assent forms and of any deviations or any changes in study activity. I shall also notify the REB of any unexpected adverse events as per REB guidelines. As applicable, I certify that the research contracts and corresponding protocol are consistent and will inform the contract manager of any protocol amendments as required.

I agree that, in accordance with the Personal Health Information Protection Act of Ontario, I am responsible for adhering to all conditions and restrictions imposed by the REB governing the use, security, disclosure, return and disposal of the research subjects’ personal health information. I am also responsible for reporting immediately any privacy breaches to the REB Chair and to Janice Campbell, the Sick Kids privacy officer. I will ensure that the personal health information is used, only as necessary, to fulfill the specific research objectives and related research questions, described in this application and approved by the REB.

Signature of Principal Investigator

DATE

I approve of this research protocol, agree to share responsibility for its proper conduct, and will ensure that the REB is notified of concerns, as appropriate.

Signature of Division/Department Head

DATE

The REB of The Hospital for Sick Children has reviewed and approved the above-named research study.

Mr. Richard Sugarman, REB Chair
Vice-Chair
555 University Avenue, Toronto, Ontario, MSG 1X8
Tel: 416-813-6152 Fax: 416-813-5085 Email: richard.sugarman@sickkids.ca

DATE OF APPROVAL FEB 13 2013 EXPIRY DATE FEB 2014

153
Research Ethics Board (REB)

The Research Ethics Board for The Hospital for Sick Children is organized and operates according to the principles and practices outlined in the Tri-Council Policy Statement, the ICH Harmonized Tripartite Guidelines: Good Clinical Practice, and Division 3 and the Medical Devices Regulations of the Food and Drug Act as well as the Natural Health Products Regulations of Health Canada. This signed document is in lieu of the Health Canada Research Ethics Board Attestation Form.

Approval & Terms of Agreement

Investigators: Dr. Leonardo Brando, M.Aviña, B.Feldman, A.Kiiss, J.Stinson

Study Title: Fluid Accumulation in Pediatric Post-Thrombotic Syndrome

REB File number: 1000038967 Level of Continuing Review: II C
Protocol Version Date: April 2 2013
Investigator's Brochure Version Date: N/A

I agree to carry out the proposed research involving human subjects in accordance with the above-noted guidelines and regulations (as applicable) and using only the REB-approved study protocol and consent/assent forms. I shall notify the division/department head and the REB prior to implementing any amendments in the protocol and consent/assent forms and of any deviations or any changes in study activity. I shall also notify the REB of any unexpected adverse events as per REB guidelines. As applicable, I certify that the research contract and corresponding protocol are consistent and will inform the contract manager of any protocol amendments as required.

I agree that, in accordance with the Personal Health Information Protection Act of Ontario, I am responsible for adhering to all conditions and restrictions imposed by the REB governing the use, security, disclosure, return and disposal of the research subjects' personal health information. I am also responsible for protecting the research subjects' privacy and for reporting immediately any privacy breaches to the REB Chair and to Janice Campbell, the Sick Kids privacy officer. I will ensure that the personal health information is used, only as necessary, to fulfill the specific research objectives and related research questions described in this application and approved by the REB.

Signature of Principal Investigator: [Signature]
DATE: July 9, 2013

I approve of this research protocol, agree to share responsibility for its proper conduct, and will ensure that the REB is notified of concerns, as appropriate.

Signature of Division/Department Head: [Signature]
DATE: July 9, 2013

The REB of The Hospital for Sick Children has reviewed and approved the above-named research study.

Mr. Richard Sugarman, REB Chair 555 University Avenue, Toronto, Ontario, M5G 1X8 Tel: 416-813-6152 Fax: 416-813-5085 Email: richard.sugarman@sickkids.ca

DATE OF APPROVAL: July 12, 2013
EXPIRY DATE: July 12, 2014
PROTOCOL REFERENCE # 28717

March 7, 2013

Dr. Brian Feldman 
FACULTY OF DENTISTRY

Dr. Maria Laura Avila
FACULTY OF DENTISTRY

Dear Dr. Feldman and Dr. Maria Laura Avila,

Re: Administrative Approval of your research protocol entitled, "Post-thrombotic syndrome in children: Item generation for the development of an evaluative and discriminative tool"

We are writing to advise you that the Office of Research Ethics (ORE) has granted administrative approval to the above-named research protocol. The level of approval is based on the following role(s) of the University of Toronto (University), as you have identified with your submission and administered under the terms and conditions of the affiliation agreement between the University and the associated TAHSN hospital:

- Graduate Student research - hospital-based only
- Storage or analysis of De-identified Personal Information (data)

This approval does not substitute for ethics approval, which has been obtained from your hospital Research Ethics Board (REB). Please note that you do not need to submit Annual Renewals, Study Completion Reports or Amendments to the ORE unless the involvement of the University changes so that ethics review is required. Please contact the ORE to determine whether a particular change to the University's involvement requires ethics review.

Best wishes for the successful completion of your research.

Yours sincerely,

[Signature]

Daniel Gyewu
REB Manager
PROTOCOL REFERENCE # 29240
August 1, 2013
Dr. Brian Feldman
FACULTY OF DENTISTRY

Dear Dr. Feldman,

Re: Administrative Approval of your research protocol entitled, "Fluid accumulation in pediatric post-thrombotic syndrome"

We are writing to advise you that the Office of Research Ethics (ORE) has granted administrative approval to the above-named research protocol. The level of approval is based on the following role(s) of the University of Toronto (University), as you have identified with your submission and administered under the terms and conditions of the affiliation agreement between the University and the associated TAHBN hospital:

- Graduate Student research - hospital-based only
- Storage or analysis of De-identified Personal Information (data)

This approval does not substitute for ethics approval, which has been obtained from your hospital Research Ethics Board (REB). Please note that you do not need to submit Annual Renewals, Study Completion Reports or Amendments to the ORE unless the involvement of the University changes so that ethics review is required. Please contact the ORE to determine whether a particular change to the University’s involvement requires ethics review.

Best wishes for the successful completion of your research.

Yours sincerely,

Daniel Gyerwi
REB Manager

OFFICE OF RESEARCH ETHICS
McClelland Building, 12 Queen's Park Crescent West, 2nd Floor, Toronto, ON M5S 1A9 Canada
Tel. +1 416 946-3271 Fax. +1 416 946-5761 ethicsreview@utoronto.ca
http://www.research.utoronto.ca/for_researchers/research_reethylene
Appendix 3.2. Consents and Assent Forms, Study 1

Research Consent Form for Patients

Title of Research Project: Fluid accumulation in pediatric post-thrombotic syndrome

Investigator(s):
Principal Investigator:
Dr. Leonardo R. Brandao, MD, MSc
Tel: 416-813-7654 (ext. xxx)

Co-Investigators:
Dr. Brian Feldman, MD, MSc
Dr. Alexander Kiss, PhD
Dr. Jennifer Stinson, RN, PhD
Dr. Suzan Williams, MD, MSc
Dr. Irene Lara-Corrales, MD, MSc
Dr. Maria Laura Avila, MD*
*(Haematology Fellow, supervisor: Dr. Brian Feldman)

Research Assistant:
Ms. Madeline Montoya, BA (Hons.)
Tel: 416-813-7654 (ext. xxx)

Purpose of the Research:
The objective of this study is to learn about a potential complication that can present in children who have had a blood clot in a vein. “Blood clots” happen when the blood gets sticky and solid, blocking the vein. The complication we are studying is known as post-thrombotic syndrome. One out of 4 children who have had a blood clot in their legs or arms can have this complication.

One of the most common symptoms that tell us this may be happening is a swollen arm or leg. The swelling is caused by an accumulation of water. Doctors have been using a tape measure to know if a child’s arm or leg is swollen. However, we need to be sure that this is the best way to check for swelling.

To do so, we are using a body water machine called “bioimpedance device”, which has been used for decades to measure body fat. This time, we are using it to measure arm and leg swelling. In other words, to know how much water has accumulated.

We are asking you to take part in this study because you have never had a blood clot. The amount of water we find will tell us how much water there normally is in a child’s arm or leg.

Description of the Research:
We will measure your height and weight, and the circumference of your arms and ankles (the last two, using a measuring tape). After that, we will put stickers on your arms and legs to measure the amount of water, using the body water machine. This will take 10 minutes.
If you have any questions, you can reach the researchers by phone at any time.

Please indicate below if you give your permission to store the obtained information for potential use in future research:

Yes [ ] No [ ]

Potential Harms:
We know of no harm that taking part in this study could cause you. The body water machine does not cause any known harm. Like any other material, the stickers could cause a rash, but we know this is not common at all. Taking part in this study will cause no harm to pregnant mothers or their babies.

As in any study, there could be a breach of data confidentiality. We have taken high security measures to prevent this from happening.

Potential Discomforts or Inconvenience:
Potential inconvenience relates to the time commitment required to participate (20 minutes, including explaining the study and results).

Potential Benefits:

To individual subjects:
You may not benefit directly from participating in this study. However, participation will generate information that will help researchers and physicians learn more about this medical problem, which could potentially be of benefit to individuals, besides their contribution to society.

Because we are learning about bioimpedance, its results will not carry the risk of any unexpected or ‘incidental’ findings to be concerned about.

The results of the bioimpedance will be available to you at your request.

To society:
This study will generate information that will help researchers/physicians understand better ways to measure swelling in children. With the interpretation of the information collected in this study, we expect to generate new ideas that will be utilized in future studies.

Alternatives to participation:
You may choose not to participate in this study.

Confidentiality:
We will respect your privacy. No information about who you are will be given to anyone or be published without your permission, unless required by law. For example, the law could make us give information about you if you have an illness that could spread to others, if you or someone else talks about suicide (killing themselves), or if the court orders us to give them the study papers.

SickKids Clinical Research Monitors, employees of the funder or sponsor, or the regulator of the study may see your health record to check on the study. By signing this
consent form, you agree to let these people look at your records. We will put a copy of this research consent form in your patient health record and give you a copy as well.

The data produced from this study will be stored in a secure, locked location. Only members of the research team (and maybe those individuals described above) will have access to the data. This could include external research team members. Following completion of the research study the data will be kept as long as required then destroyed as required by SickKids policy. Published study results will not reveal your identity.

**Certificate of participation:**
A certificate of participation will be provided to you once you have completed the study.

**Participation:**
Your participation in this research study is voluntary. If you choose to take part in this study, you can leave the study at any time. The care you get at SickKids will not be affected in any way by whether you take part in this study or not.

Your signing this consent form does not interfere with your legal rights in any way. The staff of the study or the hospital is still responsible, legally and professionally, for what they do.

During this study we may create new tests, new medicines, or other things that may be worth some money. Although we may make money from these findings, we cannot give you any of this money now or in the future because you took part in this study.

**Sponsorship:**
The sponsor/funders of this research study are Dr. Brandao and The Hospital for Sick Children.

**Conflict of Interest:**
The Principal Investigator, Dr. Brandao, and the other research team members have no conflict of interest to declare.

**Consent (participants who can consent for themselves):**

“By signing this form, I agree that:
1) You have explained this study to me. You have answered all my questions.
2) You have explained the possible harms and benefits (if any) of this study.
3) I know what I could do instead of taking part in this study. I understand that I have the right not to take part in the study and the right to stop at any time. My decision about taking part in the study will not affect my health care at Sick Kids.
4) I am free now, and in the future, to ask questions about the study.
5) I have been told that my medical records will be kept private except as described to me.
6) I understand that no information about who I am will be given to anyone or be published without first asking my permission.
7) I have read and understood pages 1 to 4 of this consent form. I agree, or consent, to take part in this study.
Printed Name of Subject & Age

Subject’s signature & date

Printed Name of person who explained consent

Signature of Person who explained consent & date

Printed Witness’ name (if the subject/legal guardian does not read English)

Witness’ signature & date

If you have any questions about this study, please contact Laura Avila at 4168137654 (ext. xxx) or Madeline Montoya (ext. xxx). If you have questions about your rights as a subject in a study or injuries during a study, please call the Research Ethics Manager at ext. xxx.”
Research Consent Form for Parents

Title of Research Project: Fluid accumulation in pediatric post-thrombotic syndrome

Investigator(s):
Principal Investigator:
Dr. Leonardo R. Brandao, MD, MSc  Tel: 416-813-7654 (ext. xxx)

Co-Investigators:
Dr. Brian Feldman, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Alexander Kiss, PhD  Tel: 416-480-6100 (ext. xxx)
Dr. Jennifer Stinson, RN, PhD  Tel: 416-813-7654 (ext. xxx)
Dr. Suzan Williams, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Irene Lara-Correllas, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Maria Laura Avila, MD*  Tel: 416-813-7654 (ext. xxx)
*(Haematology Fellow, supervisor: Dr. Brian Feldman)

Research Assistant:
Ms. Madeline Montoya, BA (Hons.)  Tel: 416-813-7654 (ext. xxx)

Purpose of the Research:
The objective of this study is to learn about a potential complication that can present in children who have had a blood clot in a vein. “Blood clots” happen when the blood gets sticky and solid, blocking the vein. The complication we are studying is known as post-thrombotic syndrome. One out of 4 children who have had a blood clot in their legs or arms can have this complication.

One of the most common symptoms that tell us this may be happening is a swollen arm or leg. The swelling is caused by an accumulation of water. Doctors have been using a tape measure to know if a child’s arm or leg is swollen. However, we need to be sure that this is the best way to check for swelling.

To do so, we are using a body water machine called “bioimpedance device”, which has been used for decades to measure body fat. This time, we are using it to measure arm and leg swelling. In other words, to know how much water has accumulated.

We are asking you to let your child take part in this study because he/she has never had a blood clot. The amount of water we find will tell us how much water there normally is in a child’s arm or leg.

Description of the Research:
We will measure your child’s height and weight, as well as the circumference of his/her arm and ankle with a measuring tape. After that, we will put stickers on his/her arms and legs to measure the amount of water, using the body water machine. This will take 10 minutes.

If you have any questions, you can reach the researchers by phone at any time.

Please indicate below if you give your permission to store the obtained information for potential use in future research:
**Potential Harms:**
We know of no harm that taking part in this study could cause to your child. The body water machine does not cause any known harm. Like any other material, the stickers could cause a rash, but we know this is not common. Taking part in this study will cause no harm to pregnant mothers or their babies.

As in any study, there could be a breach of data confidentiality. We have taken high security measures to prevent this from happening.

**Potential Discomforts or Inconvenience:**
Potential inconvenience relates to the time commitment required to participate (20 minutes, including explaining the study and the results).

**Potential Benefits:**

**To individual subjects:**
Your child may not benefit directly from participating in this study. However, participation will generate information that will help researchers and physicians learn more about this medical problem, which could be of benefit to individuals, besides their contribution to society.

Because we are learning about bioimpedance, its results will not carry the risk of any unexpected or ‘incidental’ findings to be concerned about.

The results of the study will be available to you at your request.

**To society:**
This study will generate information that will help researchers/physicians understand better ways to measure swelling in children. With the interpretation of the information collected in this study, we expect to generate new ideas that will be utilized in future studies.

**Alternatives to participation:**
You may choose not to have your child participate in this study.

**Confidentiality:**
We will respect your child’s privacy. No information about who your child is will be given to anyone or be published without your permission, unless required by law. For example, the law could make us give information about your child if he/she have an illness that could spread to others, if he/she or someone else talks about suicide (killing themselves), or if the court orders us to give them the study papers.

SickKids Clinical Research Monitors, employees of the funder or sponsor, or the regulator of the study may see your health record to check on the study. By signing this consent form, you agree to let these people look at your child’s records. We will put a copy of this research consent form in your child’s patient health record and give you a copy as well.
The data produced from this study will be stored in a secure, locked location. Only members of the research team (and maybe those individuals described above) will have access to the data. This could include external research team members. Following completion of the research study the data will be kept as long as required then destroyed as required by SickKids policy. Published study results will not reveal your child’s identity.

**Certificate of participation:**
A certificate of participation will be provided to your child once he/she has completed the study.

**Participation:**
Your participation in this research study is voluntary. If you choose your child to take part in this study, he/she can leave the study at any time. The care your child gets at SickKids will not be affected in any way by whether he or she takes part in this study or not.

If your child is able to read a description of the study, we will provide a simplified form and allow him/her to ask any questions that he/she has before obtaining verbal consent.

Your signing this consent form does not interfere with your child’s legal rights in any way. The staff of the study or the hospital is still responsible, legally and professionally, for what they do.

During this study we may create new tests, new medicines, or other things that may be worth some money. Although we may make money from these findings, we cannot give you any of this money now or in the future because your child took part in this study.

**Sponsorship:**
The sponsor/funders of this research study are Dr. Brandao and The Hospital for Sick Children.

**Conflict of Interest:**
The Principal Investigator, Dr. Brandao, and the other research team members have no conflict of interest to declare.

**Consent (For parent/legal guardians or substitute decision makers consenting for their children):**

“By signing this form, I agree that:
1) You have explained this study to me. You have answered all my questions.
2) You have explained the possible harms and benefits (if any) of this study.
3) I know what I could do instead of having my child take part in this study. I understand that I have the right to refuse to let my child take part in the study. I also have the right to take my child out of the study at any time. My decision about my child taking part in the study will not affect my child’s health care at Sick Kids.
4) I am free now, and in the future, to ask questions about the study.
5) I have been told that my child’s medical records will be kept private except as described to me.
6) I understand that no information about my child will be given to anyone or be published without first asking my permission.
7) I have read and understood pages 1 to 4 of this consent form. I agree, or consent, to allow my child to take part in this study."

Printed Name of Parent/Legal Guardian

Parent/Legal Guardian’s signature & date

Printed Name of person who explained consent

Signature of Person who explained consent & date

Printed Witness’ name (if the parent/legal guardian does not read English)

Witness’ signature & date

If you have any questions about this study, please contact Laura Avila at 416-837-654 (ext. xxx) or Madeline Montoya (ext. xxx). If you have questions about your rights as a subject in a study or injuries during a study, please call the Research Ethics Manager at ext. xxx.
ASSENT FORM

• Title of Study: Fluid accumulation in pediatric post-thrombotic syndrome

• Investigator(s)

Principal Investigator:
Dr. Leonardo R. Brandao, MD, MSc  Tel: 416-813-7654 (ext. xxx)

Co-Investigators:
Dr. Brian Feldman, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Alexander Kiss, PhD  Tel: 416-480-6100 (ext. xxx)
Dr. Jennifer Stinson, RN, PhD  Tel: 416-813-7654 (ext. xxx)
Dr. Suzan Williams, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Irene Lara-Corrales, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Maria Laura Avila, MD*  Tel: 416-813-7654
*(Haematology Fellow, supervisor: Dr. Brian Feldman)

Research Assistant:
Ms. Madeline Montoya, BA (Hons.)  Tel: 416-813-7654 (ext. xxx)

• Why are we doing this study?

You are being asked to take part in this research study because we need your help to understand how water accumulates in the arms and legs of children. This information will help children with blood clots, which happens when blood gets sticky and solid inside your veins. Children with blood clots can develop complications with accumulation of water in their arms or legs.

This assent form describes the purpose of the study and what you will have to do if you want to take part. Please read this paper carefully and ask as many questions as you like. If you want to talk to the study doctor by yourself, please ask. If this form has any words you do not know, please ask the study doctor what they mean.

Taking part in this study is up to you and your parents. If you do not want to be in the study, it is OK for you to say no. Your study doctor will still take care of you and no one will be mad at you. You will not be punished if you do not want to be in the study.

• What will happen during the study?

A researcher will measure your height and weight, and the distance around the middle of your arms and your ankles with a measuring tape. After that, we will put stickers on your arms and on your legs and take the water measurement with a body water machine. This study does not cause any harm. This study does not require you to take any additional medication or have any pokes.

• Are there good things and bad things about the study?
The good thing about participating in this study is that you will help us improve the tests that we need to know if a child has post-thrombotic syndrome.

• Who will know about what I did in the study?
    If we feel your health is in danger, we may have to report your results to your doctor. You, the doctors who run the study, and the hospital group that checks on studies will know that you were in the study.

• Can I decide if I want to be in the study?
    Nobody will be angry or mad at you if you do not want to be in the study. You can say yes now and then change your mind later. We are talking to your parent/legal guardians about the study and you should talk to them about it too.

**Assent:**
The following section must be included at the end of the assent form:

"I was present when ___________________________ read this form and said that he or she agreed, or assented, to take part in this study”.

____________________________________  ____________________________________
Printed Name of person who obtained assent   Signature & Date
Appendix 4.1. Definition and Measurement Procedures for the Signs and Symptoms of Post-thrombotic Syndrome

The VEIN-TERM consensus [23] and the American Venous Forum International Ad Hoc Committee for Revision of the Clinical-Etiologic-Anatomic-Pathophysiologic (CEAP) Classification consensus statement [62], which generated recommendations for definition of terms related to venous chronic disorders in adults, were used as a guide to develop the definitions used in the present series of studies.

1. Symptoms

Symptoms of post-thrombotic syndrome (PTS) were measured using a questionnaire developed using the items provided by patients and experts in Study 2a. The structure of the questionnaire was based on that of the questionnaire developed by the investigators of the Edinburgh Vein study (Gerry Fowkes, with permission). A self-reported version for children aged 10 years and older and proxy-reported version for younger patients were developed. Data collected included the frequency of symptoms (5-point scale, never to every day) as well as intensity, anatomic location and descriptors of pain for older patients (see below). An expert on pain assessment (JS) was consulted for selecting of tools used to measure pain.

1.1. Pain:

1.1.1. Definition: “An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage” [330].

1.1.2. Measurement:

- Visual analogue scale (VAS) using proxy/parental report for children 1-3 years. Marks were then translated into a final score ranging from 0 to 10, with 10 representing the worst pain episode in the previous 4 weeks.

- Faces Pain Scale-Revised [259] for patients aged 4-19 years. The choice was translated into a final score ranging from 0 to 10, with 10 representing the worst pain over the previous 4 weeks.
• Adolescent Pediatric Pain Tool (APPT) for patients aged 10-19 years. This tool includes a measure of pain intensity (which was then translated into a **0 to 10 final score**, with 10 representing the worst pain over the previous 4 weeks), pain descriptors and distribution [260].

We started by asking about the worst episode of pain over the previous 4 weeks, to avoid recall bias. We also asked about frequency of pain when patients or parents reported the presence of this symptom, what worsened and what controlled the pain. Only pain intensity was used in the item properties section of the analysis.

**1.2. Heaviness:**

1.2.1. **Definition:** Sensation of increased weight in the affected arm/leg.

1.2.2. **Measurement:** 0 to 5 point scale (never to every day).

**1.3. Paresthesia:**

1.3.1. **Definition:** “Abnormal sensation, whether spontaneous or evoked. Not unpleasant, as opposed to dysesthesia” [330].

1.3.2. **Measurement:** 0 to 5 point scale (never to every day), and sensory neurological exam in patients referring paresthesia (normal vs. abnormal).

**1.4. Edema (patient-reported, swollen limb):**

1.4.1. **Definition:** Perceptible increase in the volume of fluid in the skin and subcutaneous tissue [62].

1.4.2. **Measurement:** 0 to 5 point scale (never to every day).

**1.5. Cramps:**

1.5.1. **Definition:** “Sudden, involuntary, painful muscle contractions, self-extinguishing within seconds or minutes, often accompanied by a palpable knotting of the muscle”[331]. According to the McGill pain questionnaire, cramps are also considered a quality of pain [332].

1.5.2. **Measurement:** 0 to 5 point scale (never to every day).

**1.6. Tightness:**
1.6.1. **Definition**: Perception of increased passive tension in the limb.

1.6.2. **Measurement**: 0 to 5 point scale (never to every day).

1.7. **Pruritus**:

1.7.1. **Definition**: “Unpleasant sensation that elicits the desire or reflex to scratch” [333].

1.7.2. **Measurement**: 0 to 5 point scale (never to every day).

1.8. **Fatigue (tired limb)**:

1.8.1. **Definition**: “Transient weariness during exertion, due to an imbalance between the metabolic demands of the working skeletal muscle and the availability of blood flow to deliver oxygen and remove products of muscle metabolism” [334].

1.8.2. **Measurement**: 0 to 5 point scale (never to every day).

1.9. **Hot/warm limb**:

1.9.1. **Definition**: Sensation of increased temperature in the affected arm/leg.

1.9.2. **Measurement**: 0 to 5 point scale (never to every day).

1.10. **Change in skin color**:

See discoloration below.

2. **Skin manifestations**

Skin findings were defined using the definitions of the items included in the revised Venous Clinical Severity Score [76] (VCSS) as well as the definitions provided in pediatric dermatology literature. A pediatric dermatologist (ILC) was involved in the development of the definitions in this domain. One of the researchers (LA) independently assessed the presence of all the items that required physical examination. In addition, photographic documentation was obtained and/or a second and independent physical examination was performed by the pediatric dermatologist. Pictures were taken by a professional medical photographer (DB) following a standard protocol.

2.1. **Skin hyperpigmentation**:

2.1.1. **Definition**: “Presumes venous origin. Brownish darkening of skin. Usually occurs in the ankle region, but may extend to the leg and foot. Does not include focal pigmentation over
varicose veins or pigmentation due to other chronic diseases” [76].

2.1.2. Measurement: Venous Clinical Severity Score (VCSS) using the item “skin pigmentation”, which is scored in a 0 to 3 point scale corresponding to the extension of the pigmentation, from perimalleolar to above the lower third of the calf. Skin pigmentation in the upper extremity was assessed by describing its location.

2.2. Skin discoloration:

2.2.1. Definition: Skin discoloration was defined according to the following terms: reticulate erythema, defined as a “flat network of intersecting red lines that enclose small patches of pale or normal-colored skin. The red lines vary in width from several millimeters to as much as 1 cm; these lines are not sharply demarcated and instead blend into the adjacent normal skin. This lack of sharp demarcation gives a mottled or blotchy appearance to the involved areas. The red color that makes up the lines is variable in hue but generally tends toward the dusky end of the spectrum and sometimes even appears violaceous. The lesions are not palpable, these are macules and patches, and there is no substance to them” [261]; erythema hyperemicum: “redness of almost any hue, which always disappears under pressure, no signs of induration, infiltration or elevation” [262]; and livedo reticularis, a form of reticulate erythema, was defined as cyanosis of the skin (marbled, net-like or lace-like) caused by stasis in the capillaries furthest from their arterial supply [263]. Some experts classify cutis marmorata as a physiologic type of livedo reticularis, whereas others classify them as separate entities; the difference between them noted by some is that the latter disappears with warming.

2.2.2. Measurement: Physical examination (present/absent) and an item asking about redness in the skin, or blotchy purple skin, measured in a 0 to 5 point scale (never to every day), part of the symptoms questionnaire.

2.3. Venous eczema/venous stasis dermatitis:

2.3.1. Definition: “Erythematous dermatitis, which may progress to blistering, weeping, or scaling eruption of skin of leg. Most often located near varicose veins, but may be located anywhere in the leg” [62]. Described under “inflammation” in the VCSS “more than just recent pigmentation (i.e., erythema, cellulitis, venous eczema, dermatitis)” [76]. It can be accompanied by fine whitish scale and xerosis (dryness) of the skin.
2.3.2. Measurement: VCSS, using the item “inflammation”, scored in a 0 to 3 point scale corresponding to the extension of the inflammation, from perimalleolar to above the lower third of the calf. Skin inflammation in the upper extremity was assessed by describing its location.

2.4. Lipodermatosclerosis:

2.4.1. Definition: Lipodermatosclerosis was considered a form of skin induration, as defined in the VCSS: “Presumes venous origin of secondary skin and subcutaneous changes (i.e., chronic edema with fibrosis, hypodermitis). Skin induration includes white atrophy [(“Localized, often circular whitish and atrophic skin areas surrounded by dilated capillaries and sometimes hyperpigmentation; it is a sign of severe chronic venous disease, not to be confused with healed ulcer scars. Scars of healed ulceration may also exhibit atrophic skin with pigmentary changes, but are distinguishable by history of ulceration and appearance from atrophie blanche, and are excluded from this definition”)] [62] and lipodermatosclerosis [(“Localized chronic inflammation and fibrosis of skin and subcutaneous tissues of lower leg, sometimes associated with scarring or contracture of Achilles tendon. Lipodermatosclerosis is sometimes preceded by diffuse inflammatory edema of the skin, which may be painful and is often referred to as hypodermitis. Lipodermatosclerosis must be differentiated from lymphangitis, erysipelas, or cellulitis by their characteristically different local signs and systemic features. Lipodermatosclerosis is a sign of severe chronic venous disease”[62])] [76].

2.4.2. Measurement: VCSS, using the item “induration”, scored in a 0 to 3 point scale, corresponding to the extension of the induration, from perimalleolar to above the lower third of the calf. Skin induration in the upper extremity was assessed by describing its location.

2.5. Venous ulcers:

2.5.1. Definition: “Full-thickness defect of skin, most frequently in the ankle region, that fails to heal spontaneously and is sustained by chronic venous disease” [62].

2.5.2. Measurement: The VCSS items “active ulcer number”, measured in a 0 to 3 point scale ranging from 0 to ≥3 ulcers; “active ulcer duration”, measured in a 0 to 3 point scale ranging from N/A to not healed for >1 year, and “active ulcer size”, measured in a 0 to 3 point scale ranging from NA to > 6 cm.
2.6. Trophic changes:

2.6.1. **Definition**: Trophic changes were described under “atrophy”, which refers to the decrease in the size of the tissue. In superficial cutaneous atrophy, the skin appears thinned and finely wrinkled. Loss of skin markings and epidermal appendages produces a shiny appearance [261]. Lipoatrophy refers to atrophy of the subcutaneous tissue, usually resulting in “deep depressions of the body surface without textural changes of the skin” [261].

2.6.2. **Measurement**: Physical examination (present/absent).

2.7. Altered perfusion:

2.7.1. **Definition**: “Clinical signs of poor perfusion consist of cold, pale, clammy, and mottled skin associated with an increase in capillary refill time. Skin temperature and capillary refill have been advocated as measure of peripheral perfusion” [335].

2.7.2. **Measurement**: Physical examination (present/absent) and thermography. Thermography was performed using the DermaTemp Infrared Surface Skin Scanner. This is a high precision, non-invasive, hand-held infrared thermographic scanner able to detect subtle variations in skin temperature caused by changes in underlying perfusion. Temperature was assessed at three standardized anatomic points (between first and second metacarpal bones – dorsal aspect of the hand-, between the ulnar and radial styloid process, and mid-point between ulnar styloid process and olecranon process for upper extremities; and on the first metatarsal phalangeal joint, below the medial malleolus, and on a mid-point between the medial malleolus and the tibial tuberosity for lower extremities). Values were averaged, and the affected to unaffected limb temperature ratio was calculated from these averaged values.

2.8. Recurrent superficial infections and poor healing:

2.8.1. **Definitions**: a) easy infection of minor peripheral abrasions or cuts; b) abnormal or delayed healing response to peripheral injuries.

2.8.2. **Measurement**: Item on the symptoms questionnaire (yes/no/not sure).

2.9. Easy bruising:

2.9.1. **Definition**: “Bruising without a history of trauma or bruising after minor trauma that would not have caused bruising in the past” [264], specifically related to easy damage to
superficial collateral veins.

2.9.2. **Measurement**: Physical examination (present/absent).

3. **Observable/dilated veins**

Observable/dilated veins were measured using the VCSS [76] and the Vein Prominence Scale [208]. One of the researchers (LA) assessed the presence of all the items that required physical examination. As described above, photographic documentation was obtained and/or a second and independent physical examination was performed by the pediatric dermatologist. Pictures were taken following a standard protocol (i.e., light, distance, settings, positioning of the patients), after obtaining a separate consent.

3.1. **Telangiectasias**:

3.1.1. **Definition**: Small dilated blood vessels (venules, arterioles or capillaries) near the surface of the skin. These vessels can become confluent, but always measure less than 1 mm in caliber [62].

3.1.2. **Measurement**: Physical examination (present/absent).

3.2. **Reticular vein**:

3.2.1. **Definition**: “Dilated bluish subdermal vein, usually 1 mm to < 3 mm in diameter. Usually tortuous. Excludes normal visible veins in persons with thin, transparent skin” [62].

3.2.2. **Measurement**: Physical examination (present/absent).

3.3. **Varicose vein**:

3.3.1. **Definition**: “Subcutaneous dilated vein 3 mm in diameter or larger measured in upright position, may involve saphenous veins, saphenous tributaries, or nonsaphenous superficial leg veins. Varicose veins are usually tortuous, but tubular saphenous veins with demonstrated reflux may be classified as varicose veins. Synonyms include varix, varices, and varicosities”. Also includes corona phlebectatica or ankle flare [62].

3.3.2. **Measurement**: VCSS, using the item “varicose veins”, measured on a 0 to 3 point scale, corresponding to the distribution/severity of varicose veins, from few scattered, including ankle flare to calf and thigh involvement. Varicose veins in the upper extremity were assessed by
describing their location.

3.4. Collateral circulation:

3.4.1. **Definition:** Collateral formation is classically regarded as a mechanism that could potentially compensate obstruction to blood flow in the deep or superficial venous system, acting as a complementary network to bypass flow, thus alleviating the obstruction [39].

3.4.2. **Measurement:** Vein Prominence Scale [208], a 0 to 4 point scale, ranging from “no veins visible” to “veins easily visible under ambient light and >10 cm in length and/or palpable”. (Appendix 5.1)

4. Segmental Edema

The assessment of segmental edema included the measurement of limb circumference, limb volume, and fluid content as determined by bioimpedance spectroscopy. Additional techniques included durometry and a clinical scale to assess pitting edema.

4.1. Limb edema:

4.1.1. **Definition:** “Perceptible increase in volume of fluid in the skin and subcutaneous tissue, which characteristically indents with pressure. Venous edema usually occurs in the ankle region, but may extend to the leg and foot”[62].

4.1.2. **Measurement:**

- Circumferential measurements of the affected/unaffected extremities, using a tension-controlled measuring tape (Gulick II, Country Technology Inc., Wisconsin, US). The tape allows applying constant tension, therefore minimizing measurement error associated with differences in the amount of tension applied. Whenever feasible, measurements were determined while the patient was lying comfortably on the examination table, with his/her arm or leg resting at the sides of the body. In children with diagnosis of upper extremity deep vein thrombosis, the circumference of the mid-arm was measured at the mid-point between the acromion process and olecranon process. In children with diagnosis of lower extremity deep vein thrombosis, the circumference of the mid-calf was assessed at a distance corresponding to the mid-point between the medial malleolus and the tibial tuberosity. The mid-thigh circumference was assessed at the
distance corresponding to the mid-point between the anterior superior iliac spine and the tibial tuberosity. The absolute difference between the affected and unaffected extremity, as well as the ratio of the affected to unaffected extremity were estimated.

- The volume of extremities was determined by assuming the shape of a limb is formed by a series of truncated cone segments. The volume of five truncated cones were added to measure leg volume. The volume the cones was estimated according to circumferential measures taken at six landmarks using a Gulick II tape; the length of the segment or distance between these points was measured using digital calipers, according to the size of the patient (0-150 mm iGaging inside caliper, and a 0-300mm Shimana® and a 0-600mm digital calipers without upper jaws). Landmarks, identified with a mark on the skin to increase accuracy of measurements, were as follows: knuckles at the level of the third metacarpophalangeal joint, first metacarpophalangeal joint, ulnar styloid, mid distance between ulnar styloid and olecranon process, olecranon process, mid distance between the olecranon process and acromium process for upper extremities; first metatarsophalangeal joint, below the medial malleolus, above the medial malleolus, mid distance between the medial malleolus and tibial tuberosity, tibial tuberosity, mid distance between the tibial tuberosity and anterior superior iliac spine for the lower extremities.

The volume of a truncated cone was calculated as follows:

\[ V = \frac{h (C_1^2 + C_1 C_2 + C_2^2)}{12 \pi} \]

where \( V \) was the volume of the segment, \( C_1 \) and \( C_2 \) were the circumferences at the end of the segment, and \( h \) was the length of each measured segment. This method was based on that described by Sander et al [209], who found that the truncated cone or frustum assumption appeared to reproduce the shape of the arm more accurately than the cylinder shape. The frustum method was also used to determine the volume of the leg. Although the disk method is preferred by some authors for LE measurement [213], its relative difficulty as compared to the frustum method made it unsuitable for evaluation of younger patients. The values obtained were normalized by comparing the affected to the contralateral unaffected limb, and expressed as a ratio.

- Ankles were measured using the Figure-of-8 technique [223]. The technique is
described in Chapter 3 (page 65 of this thesis) [268]. The difference in cm between the affected and unaffected extremities and the ratio of the affected to unaffected limb were estimated.

- Bioimpedance Spectroscopy (BIS): Upper extremity and lower extremity tissue resistance was measured using a BIS SFB7 device (ImpediMed, Inc. California, US). Participants were positioned in supine with no pillows, arms at sides with palms facing down, and lower extremities flat and slightly abducted. Younger patients were allowed to sit up during the testing to encourage compliance. Lightly adhesive resting electrodes were placed on the dorsum of the wrists adjacent to the ulnar styloid process, the dorsum of the hands just proximal to the third metacarpophalangeal joint, anterior to the ankle joints between the malleoli, and over the dorsum of the feet over the third metatarsal bone just proximal to the third metatarsophalangeal joint. The data on the measurement was processed using Bioimped Software. The measurement of extracellular fluid content of the affected limb was normalized by comparison to the value of the extracellular content of the contralateral unaffected limb, and expressed as a ratio (unaffected/affected limb resistance). Further details on BIS can be found in Chapter 3, page 64-65.

4.2. Pitting edema:

4.2.1. Definition: “Indentation of the skin that occurs when simple finger pressure is applied” [269].

4.2.2. Measurement:

- 1 to 4 point clinical scale [270] measuring the depth of the indentation left after pressing firmly against the skin for 5 seconds (ranging from 2 to 8 mm).

- Durometer [271]: Hand held digital force gauge (initially a Mecmesin Basic Force gauge, Mecmesin, West Sussex, UK, and later a Chatillon, Ametek Test & Calibration Instruments, Pennsylvania, US) that measures the resistance of tissue to compression. Skin resistance to indentation was measured three times at one anatomic location in both arms (soft tissue on the dorsal surface of the hand between the fourth and fifth metacarpal bones) or legs (midpoint between the medial malleolus and calcaneus tuberosity). Results, expressed in Newtons, were averaged to calculate the affected limb skin resistance to unaffected limb skin resistance ratio.
4.3. Pretibial edema:

4.3.1. **Definition:** Segmental edema localized in the lower leg.

4.3.2. **Measurement:** As described in point 4.2.2.

4.4. Head edema:

4.4.1. **Definition:** Segmental edema localized en the cranial or facial region.

4.4.2. **Measurement:** Physical examination (present/absent).

5. Functional impact of PTS

5.1. Function/Impairment:

5.1.1. **Definition:** According to the International Classification of Functioning, Children and Youth version (ICF-CY) [250], *body functions* are the physiological functions of body systems (including psychological functions), and *impairments* are problems in body function or structure such as a significant deviation or loss.

5.1.2. **Measurement:**

- Self- and proxy-reported strength and endurance were measured using self-reported and proxy-reported questions that compared the strength or endurance of the affected limb to the contralateral unaffected limb, using a modified item scored in a 1 to 13 point scale, ranging from “[Compared to my other leg/arm] I have no endurance/strength in the leg/arm where I had the clot” to “The endurance/strength in the leg/arm where I had the clot is exceptionally better than my other leg/arm” [272].

- The strength of the affected and unaffected limbs was measured using a compact dynamometer, the Commander™ muscle tester (JTECH Medical, Utah, US), in children who were able to comply with the instructions. We then determined the affected to unaffected limb strength ratio.

5.2. Activity and participation/activity limitation and participation restriction:

5.2.1. **Definition:** According to the ICF-CY, “activity is the execution of a task or action by an individual. *Activity limitations* are difficulties an individual may have in executing activities. *Participation* refers to involvement in a life situation. *Participation restrictions* are problems an individual may experience in involvement in life situations”[250].
5.2.2. **Measurement**: We used the parental report version of the Pediatric Outcomes Data Collection Instrument (PODCI) \[273\] in children aged 2-18 years. To decreased burden to the patient, we used the parental version of the tool. The tool contains 5 sub-scales (upper extremity, transfer and basic mobility, sports and physical functioning, pain and comfort, and happiness). A final or global score (0 to 100) and subscale scores (0 to 100) are obtained; 0 represents a “poor outcome/worse health” whereas 100 corresponds to the “best possible outcome/best health”. Normal values in the general population for the global scores are: mean 93.3-95.2; standard deviation 7.2-7.8, depending on the version used.

6. **Other**

6.1. **Pain on palpation:**

6.1.1. **Definition**: “Localized tenderness along the distribution of the deep venous system” \[336\].

6.1.2. **Measurement**: Physical examination (present/absent).

6.2. **Venous claudication:**

6.2.1. **Definition**: “Severe pain and sensation of tightness with vigorous exercise” \[337\].

6.2.2. **Measurement**: Symptoms questionnaire (as reflected by endurance and pain).
Appendix 4.2. Item Generation Survey (REDCap)
Appendix 4.3. Consents and Assent Forms, Study 2a and Study 2b

Research Consent Form for Patients

**Title of Research Project:** Fluid accumulation in pediatric post-thrombotic syndrome

**Investigator(s):**
- **Principal Investigator:**
  - Dr. Leonardo R. Brandao, MD, MSc  
  - Tel: 416-813-7654 (ext. xxx)

- **Co-Investigators:**
  - Dr. Brian Feldman, MD, MSc  
  - Tel: 416-813-7654 (ext. xxx)
  - Dr. Alexander Kiss, PhD  
  - Tel: 416-480-6100 (ext. xxx)
  - Dr. Jennifer Stinson, RN, PhD  
  - Tel: 416-813-7654 (ext. xxx)
  - Dr. Suzan Williams, MD, MSc  
  - Tel: 416-813-7654 (ext. xxx)
  - Dr. Irene Lara-Corrales, MD, MSc  
  - Tel: 416-813-7654 (ext. xxx)
  - Dr. Maria Laura Avila, MD*  
  - Tel: 416-813-7654 (ext. xxx)
  - *(Haematology Fellow, supervisor: Dr. Brian Feldman)

**Research Assistant:**
- Ms. Madeline Montoya, BA (Hons.)  
- Tel: 416-813-7654 (ext. xxx)

**Purpose of the Research:**

The objective of this study is to learn about a potential complication that can present in children who have had a blood clot in a vein. “Blood clots” happen when the blood gets sticky and solid, blocking the vein. The complication we are studying is known as post-thrombotic syndrome. One out of 4 children who have had a blood clot in their legs or arms can have this complication.

One of the most common symptoms that tell us this may be happening is a swollen arm or leg. The swelling is caused by an accumulation of water. Doctors have been using a tape measure to know if a child’s arm or leg is swollen. However, we need to be sure that this is the best way to check for swelling.

To do so, we are using a body water machine called “bioimpedance spectroscopy”, which has been used for decades to measure body fat. This time, we are using it to measure arm and leg swelling. In other words, to know how much water has accumulated.

**Description of the Research:**

At the beginning of the study, we will obtain your height and weight. We will ask you about swelling and other things you may feel in the arm or leg where you had the clot. You will also be asked to fill out a short questionnaire about your symptoms. This will take less than 15 minutes.

After that, we will examine your arms or legs, and measure them using a measuring tape, test the temperature of your skin, and check the presence of swelling by gently pressing the
skin. If you have any lesions in your skin that are related to the clot, we will ask you for your permission to take a picture of the lesion. This will take approximately 20-25 minutes.

We will then place stickers on your arms and legs and measure swelling with the body water machine. This will take 5 minutes. We will also check how strong your arms or legs are with a short exercise involving weights and a pushing device. We will repeat the temperature and body water machine testing right after the exercise.

While we examine you, we will ask your parents to fill out a questionnaire to know if the clot has made some of your everyday activities hard for you to do. This will take them 15-20 minutes.

We will then look through your health record to take some details about your diagnosis, studies performed, presence of catheters in the present or past time, and the medications you received.

If you have any questions, you can reach the researchers by phone at any time.

Please indicate below if you give your permission to store the obtained information for potential use in future research:

Yes  [ ]     No  [ ]

**Potential Harms:**

We know of no harm that taking part in this study could cause you. The machine does not cause any known harm. Like any other material, the stickers could cause a rash, but we know this is not common at all. Taking part in this study will cause no harm to pregnant mothers or their babies.

As in any study, there could be a breach of data confidentiality. We have taken high security measures to prevent this from happening.

**Potential Discomforts or Inconvenience:**

Potential inconvenience relates to the time commitment required to participate (60 minutes more than a regular thrombosis consult).

**Potential Benefits:**

**To individual subjects:**

You may not benefit directly from participating in this study. However, participation will generate information that will help researchers and physicians learn more about this medical problem, which could potentially be of benefit to individual patients, besides their contribution to society.

Because we are learning about bioimpedance, its results will not carry the risk of any unexpected or ‘incidental’ findings to be concerned about.

The results of the bioimpedance will be available to you at your request.

**To society:**

This study will generate information that will help researchers/physicians understand the better ways to measure swelling in children. With the interpretation of the information
collected in this study, we expect to generate new ideas that will be utilized in the proposal of future research studies.

**Alternatives to participation:**
You may choose not to participate in this study.

**Confidentiality:**
We will respect your privacy. No information about who you are will be given to anyone or be published without your permission, unless required by law. For example, the law could make us give information about you if you have an illness that could spread to others, if you or someone else talks about suicide (killing themselves), or if the court orders us to give them the study papers.

SickKids Clinical Research Monitors, employees of the funder or sponsor, or the regulator of the study may see your health record to check on the study. By signing this consent form, you agree to let these people look at your records. We will put a copy of this research consent form in your patient health record and give you a copy as well.

The data produced from this study will be stored in a secure, locked location. Only members of the research team (and maybe those individuals described above) will have access to the data. This could include external research team members. Following completion of the research study the data will be kept as long as required then destroyed as required by SickKids policy. Published study results will not reveal your identity.

**Reimbursement/Community Service hours:**
Reimbursement for parking cost or transportation will be provided in the form of a parking voucher and/or certificate of C$30.

Additionally, a letter for community service hours can be provided to you, if needed, once you have completed the study.

**Participation:**
Your participation in this research study is voluntary. If you choose to take part in this study, you can leave the study at any time. The care you get at Sick Kids will not be affected in any way by whether you take part in this study or not.

Your signing this consent form does not interfere with your legal rights in any way. The staff of the study or the hospital is still responsible, legally and professionally, for what they do.

During this study we may create new tests, new medicines, or other things that may be worth some money. Although we may make money from these findings, we cannot give you any of this money now or in the future because you took part in this study.

**Sponsorship:**
The sponsor/funders of this research study are Dr. Brandao and The Hospital for Sick Children.
Conflict of Interest:
The Principal Investigator, Dr. Brandao, and the other research team members have no conflict of interest to declare.

Consent (participants who can consent for themselves):

“By signing this form, I agree that:
1) You have explained this study to me. You have answered all my questions.
2) You have explained the possible harms and benefits (if any) of this study.
3) I know what I could do instead of taking part in this study. I understand that I have the right not to take part in the study and the right to stop at any time. My decision about taking part in the study will not affect my health care at Sick Kids.
4) I am free now, and in the future, to ask questions about the study.
5) I have been told that my medical records will be kept private except as described to me.
6) I understand that no information about who I am will be given to anyone or be published without first asking my permission.
7) I have read and understood pages 1 to 5 of this consent form. I agree, or consent, to take part in this study.

_________________________________________
Printed Name of Subject & Age

Subject’s signature & date

_________________________________________
Printed Name of person who explained consent

Signature of Person who explained consent & date

_________________________________________
Printed Witness’ name (if the subject/legal guardian does not read English)

Witness’ signature & date

If you have any questions about this study, please contact Madeline Montoya at 4168137654 (ext. xxx) or Laura Avila (ext. xxx). If you have questions about your rights as a subject in a study or injuries during a study, please call the Research Ethics Manager at ext. xxx.”
Research Consent Form for Parents

Title of Research Project: Fluid accumulation in pediatric post-thrombotic syndrome

Investigator(s):
Principal Investigator:
Dr. Leonardo R. Brandao, MD, MSc
Tel: 416-813-7654 (ext. xxx)

Co-Investigators:
Dr. Brian Feldman, MD, MSc
Tel: 416-813-7654 (ext. xxx)
Dr. Alexander Kiss, PhD
Tel: 416-480-6100 (ext. xxx)
Dr. Jennifer Stinson, RN, PhD
Tel: 416-813-7654 (ext. xxx)
Dr. Suzan Williams, MD, MSc
Tel: 416-813-7654 (ext. xxx)
Dr. Irene Lara-Corales, MD, MSc
Tel: 416-813-7654 (ext. xxx)
Dr. Maria Laura Avila, MD*
Tel: 416-813-7654 (ext. xxx)
*(Haematology Fellow, supervisor: Dr. Brian Feldman)

Research Assistant:
Ms. Madeline Montoya, BA (Hons.)
Tel: 416-813-7654 (ext. xxx)

Purpose of the Research:
The objective of this study is to learn about a potential complication that can present in children who have had a blood clot in a vein. “Blood clots” happen when the blood gets sticky and solid, blocking the vein. The complication we are studying is known as post-thrombotic syndrome. One out of 4 children who have had a blood clot in their legs or arms can have this complication.

One of the most common symptoms that tell us this may be happening is a swollen arm or leg. The swelling is caused by an accumulation of water. Doctors have been using a tape measure to know if a child’s arm or leg is swollen. However, we need to be sure that this is the best way to check for swelling.

To do so, we are using a body water machine called “bioimpedance spectroscopy”, which has been used for decades to measure body fat. This time, we are using it to measure arm and leg swelling. In other words, to know how much water has accumulated.

Description of the Research:
At the beginning of the study, we will obtain your child’s height and weight. We will ask you and your child about swelling and other symptoms he/she may feel in the arm or leg affected by the clot. You or your child will be asked to fill out a short questionnaire about these symptoms. This will take less than 15 minutes.

After that, we will examine your child’s arms or legs, and will perform measurements in his/her arms and legs using a measuring tape. We will also check the temperature of the skin and the presence of swelling by gently pressing the skin. If there are any lesions in the skin that are related to the clot, we will ask for your permission to take a picture of the lesion. All this part will take approximately 20-25 minutes.
We will then place stickers in his/her arms and legs and will measure swelling with the body water machine. This will take 5 minutes. We will also check how strong your child's arms or legs are with a short exercise involving weights and a special pushing device. We will take the skin temperature again and repeat body water machine testing right after the exercise.

We will also ask you to fill out a questionnaire to understand if post-thrombotic syndrome is interfering with your child’s daily activities. This will take 15-20 minutes, and you could take the survey while we are examining your child.

We will then look through your child’s health record to take some details about your child’s diagnosis, studies performed, presence of catheters in the past or present time and medications received.

If you have any questions, you can reach the researchers by phone at any time.

Please indicate below if you give your permission to store the obtained information for potential use in future research:

Yes [ ]  No [ ]

**Potential Harms:**

We know of no harm that taking part in this study could cause you. The machine does not cause any known harm. Like any other material, the stickers could cause a rash, but we know this is not common at all. Taking part in this study will cause no harm to pregnant mothers or their babies.

As in any study, there could be a breach of data confidentiality. We have taken high security measures to prevent this from happening.

**Potential Discomforts or Inconvenience:**

Potential inconvenience relates to the time commitment required to participate (60 minutes more than a regular thrombosis consult).

**Potential Benefits:**

**To individual subjects:**

Your child may not benefit directly from participating in this study. However, participation will generate information that is expected to help researchers and physicians learn more about this medical problem, which could potentially be of benefit to individual patients, besides their contribution to society.

Because we are learning about bioimpedance, its results will not carry the risk of any unexpected or ‘incidental’ findings to be concerned about.

The results of the bioimpedance will be available to you at your request.

**To society:**

This study will generate information that will help researchers/physicians understand the mechanisms and better ways measure swelling in children. With the interpretation of the
information collected in this study, we expect to generate new ideas that will be utilized in future research studies.

**Alternatives to participation:**
You may choose not to have your child participate in this study.

**Confidentiality:**
We will respect your child’s privacy. No information about who your child is will be given to anyone or be published without your permission, unless required by law. For example, the law could make us give information about your child if he/she have an illness that could spread to others, if he/she or someone else talks about suicide (killing themselves), or if the court orders us to give them the study papers.

SickKids Clinical Research Monitors, employees of the funder or sponsor, or the regulator of the study may see your child’s health record to check on the study. By signing this consent form, you agree to let these people look at your child’s records. We will put a copy of this research consent form in your child’s patient health record and give you a copy as well.

The data produced from this study will be stored in a secure, locked location. Only members of the research team (and maybe those individuals described above) will have access to the data. This could include external research team members. Following completion of the research study the data will be kept as long as required then destroyed as required by SickKids policy. Published study results will not reveal your child’s identity.

**Reimbursement/Community Service hours:**
Reimbursement for parking cost or transportation will be provided in the form of a parking voucher and/or certificate of C$30.

Additionally, a letter for community service hours can be offer to your child once he/she has completed the study, if needed.

**Participation:**
Your participation in this research study is voluntary. If you choose to have your child to take part in this study, he/she can leave the study at any time. The care your child gets at SickKids will not be affected in any way by whether he or she takes part in this study or not.

If your child is able to read a description of the study, we will provide a simplified form and allow him/her to ask any questions that he/she has before obtaining verbal consent.

Your signing this consent form does not interfere with your child’s legal rights in any way. The staff of the study or the hospital is still responsible, legally and professionally, for what they do.

During this study we may create new tests, new medicines, or other things that may be worth some money. Although we may make money from these findings, we cannot give you any of this money now or in the future because your child took part in this study.

**Sponsorship:**
The sponsor/funders of this research study are Dr. Brandao and The Hospital for Sick Children.
Conflict of Interest:
The Principal Investigator, Dr. Brandao, and the other research team members have no conflict of interest to declare.

Consent (For parent/legal guardians or substitute decision makers consenting for their children):

“By signing this form, I agree that:

1) You have explained this study to me. You have answered all my questions.
2) You have explained the possible harms and benefits (if any) of this study.
3) I know what I could do instead of having my child take part in this study. I understand that I have the right to refuse to let my child take part in the study. I also have the right to take my child out of the study at any time. My decision about my child taking part in the study will not affect my child’s health care at Sick Kids.
4) I am free now, and in the future, to ask questions about the study.
5) I have been told that my child’s medical records will be kept private except as described to me.
6) I understand that no information about my child will be given to anyone or be published without first asking my permission.
7) I have read and understood pages 1 to 4 of this consent form. I agree, or consent, to allow my child to take part in this study.”

_________________________________
Printed Name of Parent/Legal Guardian

Parent/Legal Guardian’s signature & date

_________________________________
Printed Name of person who explained consent

Signature of Person who explained consent & date

_______________________________________
Printed Witness’ name (if the parent/legal guardian does not read English)

Witness’ signature & date

If you have any questions about this study, please contact Madeline Montoya at 4168137654 (ext. xxx) or Laura Avila (ext. xxx). If you have questions about your rights as a subject in a study or injuries during a study, please call the Research Ethics Manager at ext. xxx.
ASSENT FORM

• Title of Study: Fluid accumulation in pediatric post-thrombotic syndrome

• Investigator(s)

Principal Investigator:
Dr. Leonardo R. Brandao, MD, MSc  Tel: 416-813-7654 (ext. xxx)

Co-Investigators:
Dr. Brian Feldman, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Alexander Kiss, PhD  Tel: 416-480-6100 (ext. xxx)
Dr. Jennifer Stinson, RN, PhD  Tel: 416-813-7654 (ext. xxx)
Dr. Suzan Williams, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Irene Lara-Corrales, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Maria Laura Avila, MD*  Tel: 416-813-7654 (ext. xxx)
*(Haematology Fellow, supervisor: Dr. Brian Feldman)

Research Assistant:
Ms. Madeline Montoya, BA (Hons.)  Tel: 416-813-7654 (ext. xxx)

• Why are we doing this study?
You are being asked to take part in this research study because you had a “blood clot”, which happens when blood gets sticky and solid inside a vein. Children with blood clots can have complications. One of these complications, called “post-thrombotic syndrome”, can cause swelling in arms or legs. We want to improve the tests doctors use to check if arms or legs are swollen. At the same time, we want to check if you have any other signs of post-thrombotic syndrome.

This assent form describes the purpose of the study and what you will have to do if you want to take part. Please read this paper carefully and ask as many questions as you like. If you want to talk to the study doctor by yourself, please ask. If this form has any words you do not know, please ask the study doctor what they mean.

Taking part in this study is up to you and your parents. If you do not want to be in the study, it is OK for you to say no. Your study doctor will still take care of you and no one will be mad at you. You will not be punished if you do not want to be in the study.

• What will happen during the study?
A doctor or a research assistant will talk to you, and will ask you what you feel in your arm or leg affected by the clot. You will also be asked to fill out a short questionnaire about your symptoms. We will examine and measure your arms or legs, and test the temperature of your skin. After that, we will put stickers in your arm or leg and take another measure with a special body water machine. This study does not cause any harm. This study does not require you to take any additional medication or have any pokes. Then, we will check how
strong your arms or legs are with a short exercise with weights and a pushing machine. We will check your temperature again and also the amount of water in your arms or legs with the body water machine after the exercise.

Finally, if you have any marks on your skin, we will take pictures of the marks.

• Are there good things and bad things about the study?
  The good thing about participating in this study is that you will help us improve the tests that we need to know if a child has swelling in the arms or legs.

• Who will know about what I did in the study?
  If we feel your health is in danger, we may have to report your results to your doctor. You, the doctors who run the study, and the hospital group that checks on studies will know that you were in the study.

• Can I decide if I want to be in the study?
  Nobody will be angry or mad at you if you do not want to be in the study. You can say yes now and then change your mind later. We are talking to your parent/legal guardians about the study and you should talk to them about it too.

Assent:
The following section must be included at the end of the assent form:

"I was present when _____________________________ read this form and said that he or she agreed, or assented, to take part in this study”.

Printed Name of person who obtained assent _____________________________ Signature & Date _____________________________
Research Ethics Board
Photography consent form (patients)

**Title of Research Project:** Fluid accumulation in pediatric post-thrombotic syndrome

**Investigators:**

**Principal Investigator:**
Dr. Leonardo R. Brandao, MD, MSc  
Tel: 416-813-7654 (ext. xxx)

**Co-Investigators:**
Dr. Brian Feldman, MD, MSc  
Tel: 416-813-7654 (ext. xxx)
Dr. Alexander Kiss, PhD  
Tel: 416-480-6100 (ext. xxx)
Dr. Jennifer Stinson, RN, PhD  
Tel: 416-813-7654 (ext. xxx)
Dr. Suzan Williams, MD, MSc  
Tel: 416-813-7654 (ext. xxx)
Dr. Irene Lara-Corrales, MD, MSc  
Tel: 416-813-7654 (ext. xxx)
Dr. Maria Laura Avila, MD*  
Tel: 416-813-7654 (ext. xxx)
*(Haematology Fellow, supervisor: Dr. Brian Feldman)

**Research Assistant:**
Ms. Madeline Montoya, BA (Hons.)  
Tel: 416-813-7654 (ext. xxx)

**Confidentiality:**

The pictures produced from this study will be stored in a secure, locked location. Only members of the research team (and maybe the SickKids monitor, or employees of the study sponsor or the regulator eg., Health Canada) will have access to them. Following completion of the study the pictures will be kept as long as required in the SickKids “Records Retention and Destruction” policy. They will then be destroyed according to this same policy.

**Consent:**

“By signing this form,

1) I also agree to be photographed during this study. These photographs will be used to create an atlas of skin manifestations seen in children with post-thrombotic syndrome (educational purposes).

2) I understand that I have the right to refuse to take part in this study. I also have the right to withdraw from this part of the study at any time. eg., before or even after the photographs are made. My decision will not affect my health care at SickKids.

3) I am free now, and in the future, to ask questions about the pictures taken.

4) I have been told that my medical records will be kept private. You will give no one information about me, unless the law requires you to.
5) I understand that no information about me (including these pictures) will be given to anyone or be published without first asking my permission.

6) I have read and understood pages 1 to 3 of this consent form. I agree, or consent, to having my picture taken as part of the study.

<table>
<thead>
<tr>
<th>Printed Name of Subject</th>
<th>Subject’s signature &amp; date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Printed Name of person who explained consent</td>
<td>Signature &amp; date</td>
</tr>
<tr>
<td>Printed Witness’ name (subject does not read English)</td>
<td>Witness’ signature &amp; date</td>
</tr>
</tbody>
</table>

In addition, I agree or consent for this photograph(s) to be used for:

1. Other studies on the same topic or
2. Teaching and demonstration at SickKids or
3. Teaching and demonstration at meetings outside SickKids.
4. Not to be used for anything else.

In agreeing to the use of the photograph(s) for other purposes, I have been offered a chance to view photograph(s). I also have the right to withdraw my permission for other uses of the photograph(s) at any time.

<table>
<thead>
<tr>
<th>Printed Name of Subject</th>
<th>Subject’s signature &amp; date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Printed Name of person who explained consent</td>
<td>Signature &amp; date</td>
</tr>
<tr>
<td>Printed Witness’ name (subject does not read English)</td>
<td>Witness’ signature &amp; date</td>
</tr>
</tbody>
</table>
Title of Research Project: Fluid accumulation in pediatric post-thrombotic syndrome

Investigators:
Principal Investigator:
Dr. Leonardo R. Brandao, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Co-Investigators:
Dr. Brian Feldman, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Alexander Kiss, PhD  Tel: 416-480-6100 (ext. xxx)
Dr. Jennifer Stinson, RN, PhD  Tel: 416-813-7654 (ext. xxx)
Dr. Suzan Williams, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Irene Lara-Corrales, MD, MSc  Tel: 416-813-7654 (ext. xxx)
Dr. Maria Laura Avila, MD*  Tel: 416-813-7654 (ext. xxx)
*(Haematology Fellow, supervisor: Dr. Brian Feldman)

Research Assistant:
Ms. Madeline Montoya, BA (Hons.)  Tel: 416-813-7654 (ext. xxx)

Confidentiality:
The pictures of your child produced from this study will be stored in a secure, locked location. Only members of the research team (and maybe the SickKids monitor, or employees of the study sponsor or the regulator eg., Health Canada) will have access to them. Following completion of the study the pictures will be kept as long as required in the SickKids “Records Retention and Destruction” policy. They will then be destroyed according to this same policy.

Consent:
“By signing this form,

1) I also agree to have my child photographed during this study. These photographs will be used to create an atlas of skin manifestations of post-thrombotic syndrome in children.

2) I understand that I have the right to refuse to let my child take part in this study. I also have the right to take my child out of this part of the study at any time. e.g., before or even after the photographs are made. My decision will not affect my child’s health care at SickKids.

3) I am free now, and in the future, to ask questions about the pictures taken.

4) I have been told that my child’s medical records will be kept private. You will give no one information about my child, unless the law requires you to.
5) I understand that no information about my child (including these pictures) will be given to anyone or be published without first asking my permission.

6) I have read and understood pages 1 to 3 of this consent form. I agree, or consent, to have my child photographed as part of the study.

_________________________________
Printed Name of Parent/Legal Guardian

_________________________________
Printed Name of person who explained consent

_________________________________
Printed Witness’ name (subject does not read English)

In addition, I agree or consent for this photograph(s) to be used for:

1. Other studies on the same topic or
2. Teaching and demonstration at SickKids or
3. Teaching and demonstration at meetings outside SickKids.
4. Not to be used for anything else.

In agreeing to the use of the photograph(s) for other purposes, I have been offered a chance to view photograph(s). I also have the right to withdraw my permission for other uses of the photograph(s) at any time.

_________________________________
Printed Name of Parent/Legal Guardian

_________________________________
Printed Name of person who explained consent

_________________________________
Printed Witness’ name (subject does not read English)
Appendix 5.1. Vein Prominence Scale

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No visible veins</td>
</tr>
<tr>
<td>1</td>
<td>Veins visible only in close proximity with direct, bright light and/or Less than or equal to 5 cm in length</td>
</tr>
<tr>
<td>2</td>
<td>Veins visible under ambient light and/or Greater than 5 cm in length</td>
</tr>
<tr>
<td>3</td>
<td>Veins visible under ambient light and Greater than 10 cm in length and/or Sharp edges resembling a pen drawing on the skin</td>
</tr>
<tr>
<td>4</td>
<td>Veins easily visible under ambient light and Greater than 10 cm in length and/or Palpable</td>
</tr>
</tbody>
</table>

Reproduced from Boulden et al[338], with permission

Appendix 5.2. Simplified Scoring System

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No visible veins</td>
</tr>
<tr>
<td>1</td>
<td>Visible veins ≤ 10 cm in length AND no highly contrasting on the skin AND not palpable</td>
</tr>
<tr>
<td>2</td>
<td>Visible veins &gt; 10 cm in length OR highly contrasting on the skin OR palpable vein</td>
</tr>
</tbody>
</table>

Copyright © 2015 The Hospital for Sick Children (SickKids)
Appendix 6.1. Item Reduction Survey, Second Round: Screenshot of the Survey, as Seen by the Respondent

### Symptoms (part 1 of 4)

#### 1- Limb heaviness

**Measurement style:** Self-reported questionnaire

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>Symptom present</td>
<td>40%</td>
<td>94%</td>
</tr>
<tr>
<td>Lower</td>
<td>Symptom present</td>
<td>53%</td>
<td>100%</td>
</tr>
</tbody>
</table>

#### 2- Limb tiredness

**Measurement style:** Self-proxy-reported questionnaire

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>Symptom present</td>
<td>47%</td>
<td>83%</td>
</tr>
<tr>
<td>Lower</td>
<td>Symptom present</td>
<td>41%</td>
<td>93%</td>
</tr>
</tbody>
</table>

---

**Time to apply the tool (minutes)**

0

*The time is reflective of the “real-life” application of the tool to one extremely slow participant.*
Appendix 6.2. Item Reduction Survey, First Round: Detailed Instrument

Post-thrombotic syndrome in children: item selection

Welcome to Round 2 of the pediatric thrombosis expert consensus survey!

Each sign and symptom of post-thrombotic syndrome (PTS) obtained in Round 1 was measured in 140 children with unilateral limb deep vein thrombosis and compared to PTS diagnosis according to an independent expert assessor.

Please select the items (signs and symptoms) that should be kept in the final pediatric PTS index.

Here is a brief tutorial:

TUTORIAL
This screenshot explains what the survey will look like, its different parts, and how to complete it (with a made up example):
1. This is the item and how we measured it; the criterion shown in the table is the optimal cut-off point as per receiver operating curve (ROC) analysis.

2. You can get more data to help with your decision by clicking here.

3. You can minimize the added data by clicking here.

4. Now you will choose to keep or remove the item from the final tool.
Symptoms (part 1 of 4)

1- Limb heaviness  Measurement style: Self-reported questionnaire  LimbCriterionSensitivity  Specificity
Upper  Symptom present 40%  94% Lower  Symptom present53%  100%

Click here for more information
Minimize information
Measurement details: The item was measured in children ≥10 years of age, using a 5-point scale (none-everyday)

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.67</td>
<td>0.47-0.84</td>
<td>0.046</td>
</tr>
</tbody>
</table>

1a- Upper limb
1b- Lower limb

Time to apply the tool (minutes)* *The time is reflective of the “real life” application of the item to one extremity only

2- Limb tiredness Measurement style:
Self/proxy-reported
questionnaireLimbCriterionSensitivity Specificity
Upper Symptom present 47% 88% Lower Symptom present41% 93%

AUC | 95% CI     | p-value |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.76</td>
<td>0.56-0.91</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Keep
Remove
Not sure

Click here for more information
Minimize information
Measurement details: The item was measured in children ≥10 years of age, using a 5-point scale (none-everyday); proxy-report for younger patients was optional.

### AUC 95% CI  p-value

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0.69</strong></td>
<td>0.53-0.83</td>
<td>0.01</td>
</tr>
</tbody>
</table>

### AUC 95% CI  p-value

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0.66</strong></td>
<td>0.51-0.80</td>
<td>0.009</td>
</tr>
</tbody>
</table>

2a- Upper limb  
2b- Lower limb

Time to apply the tool (minutes)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Keep</td>
<td>Remove</td>
</tr>
<tr>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Not sure</td>
<td></td>
</tr>
<tr>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

3- Limb swelling Measurement style:
Self/proxy-reported questionnaire
Limb Criterion
Upper Symptom present 24%  100% Lower Symptom present 42%  94%

Click here for more information
Minimize information
Measurement details: The item was measured using a 5-point scale (none-everyday) in all children, as self- or proxy-reported

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.62</td>
<td>0.50-0.73</td>
<td>0.006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.68</td>
<td>0.56-0.79</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

3a- Upper limb
3b- Lower limb

Time to apply the tool (minutes)

4- Limb pruritus Measurement style:
Self/proxy-reported
questionnaire
Limb Criterion Sensitivity Specificity
Upper Symptom present 8% 98% Lower Symptom present 17% 97%
Measurement details: The item was measured using a 5-point scale (none-everyday), in all children as self- or proxy-reported.

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.53</td>
<td>0.41-0.65</td>
<td>0.34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.57</td>
<td>0.45-0.69</td>
<td>0.05</td>
</tr>
</tbody>
</table>

4a- Upper limb
4b- Lower limb

Time to apply the tool (minutes)
Symptoms (part 2 of 4)

1- Paresthesia (tingling) Measurement style:
Self/proxy-reported questionnaire
LimbCriterion: Sensitivity Specificity
Upper Symptom present 40% 100%
Lower Symptom present 33% 100%

Measurement details: The item was measured in children ≥10 years of age using a 5-point scale (none-everyday); proxy-report for younger patients was optional

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>0.70</td>
<td>0.54-0.83</td>
<td>0.002</td>
</tr>
<tr>
<td>1b</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes)

2- Skin redness/purple or blotchy skin Measurement style: Self/proxy-reported questionnaire
LimbCriterion: Sensitivity Specificity
Upper Symptom present 24% 98%
Lower Symptom present 19% 100%

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.67</td>
<td>0.51-0.80</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

1a- Upper limb
1b- Lower limb

Not sure

08/24/2016 1:42pm www.projectredcap.org
Measurement details: The item was measured using a 5-point scale (none-everyday) in all children, as self- or proxy-reported

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm Skin Redness/Mottled Skin</td>
<td>0.61</td>
<td>0.48-0.72</td>
<td>0.02</td>
</tr>
<tr>
<td>Leg Skin Redness/Mottled Skin</td>
<td>0.60</td>
<td>0.47-0.71</td>
<td>0.004</td>
</tr>
</tbody>
</table>

2a- Upper limb
2b- Lower limb

Time to apply the tool (minutes)

3- Warmer limb Measurement style: Self/proxy-reported questionnaire
Upper Symptom present 20%  98% Lower Symptom present 19% 100%

- ○ Click here for more information
- ○ Minimize information
**Measurement details:** The item was measured using a 5-point scale (none-everyday) in all children, as self- or proxy-reported

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper limb</td>
<td>Symptom present 20%</td>
<td>94%</td>
<td>Lower limb</td>
<td>Symptom present at least once a week</td>
<td>26%</td>
</tr>
</tbody>
</table>

**AUC** | **95% CI** | **p-value**
--- | --- | ---
0.59 | 0.47-0.71 | 0.03

**AUC** | **95% CI** | **p-value**
--- | --- | ---
0.60 | 0.47-0.71 | 0.004

3a- Upper limb | Keep | Remove | Not sure
3b- Lower limb | | | 

**Time to apply the tool (minutes)**

4- Cramps Measurement style: Self-reported questionnaire

Upper Symptom present 20% 94% Lower Symptom present at least once a week 26% 100%

Click here for more information

Minimize information
Measurement details: The item was measured in children ≥10 years of age using a 5-point scale (none-everyday)

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.57</td>
<td>0.37-0.76</td>
<td>0.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.61</td>
<td>0.40-0.79</td>
<td>0.28</td>
</tr>
</tbody>
</table>

4a- Upper limb
4b- Lower limb

Time to apply the tool (minutes)
Symptoms (part 3 of 4)

1- Tightness (limb feels tight) Measurement style:
   Self-reported questionnaire
   Limb Criterion: Sensitivity
   Specificity: Upper Symptom present 10% 100% Lower Symptom present 42% 100%

Measurement details: The item was measured in children ≥10 years of age, using a 5-point scale (none-everyday)

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm</td>
<td>Present</td>
<td>10%</td>
<td>100%</td>
<td>0.55</td>
<td>0.35-0.74</td>
<td>0.32</td>
</tr>
<tr>
<td>Leg</td>
<td>Present</td>
<td>12%</td>
<td>92%</td>
<td>0.71</td>
<td>0.50-0.87</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes)

1a- Upper limb
   Keep
1b- Lower limb
   Not sure

2- Delayed healing of skin lesions Measurement style:
   Self/proxy-reported questionnaire
   Limb Criterion: Sensitivity
   Specificity: Upper Present 10% 100% Lower Present 12% 92%
Measurement details: The item was measured using a question with three response options (yes/no/not sure), as self- or proxy-reported.

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arms</td>
<td>Upper ≤ 6</td>
<td>36%</td>
<td>96%</td>
<td><strong>0.55</strong></td>
<td>0.35-0.74</td>
<td>0.32</td>
</tr>
<tr>
<td>Legs</td>
<td>Lower ≤ 5</td>
<td>36%</td>
<td>94%</td>
<td><strong>0.52</strong></td>
<td>0.33-0.71</td>
<td>0.71</td>
</tr>
</tbody>
</table>

3. Endurance Measurement style: Self/proxy-reported questionnaire
   Limb | Criterion | Sensitivity | Specificity
   Arms | Upper ≤ 6 | 36%         | 96%         | **0.55** | 0.35-0.74 | 0.32 |
   Legs| Lower ≤ 5 | 36%         | 94%         | **0.52** | 0.33-0.71 | 0.71 |

Options:
- **Keep**
- **Remove**
- **Not sure**

- Click here for more information
- Minimize information
Measurement details: Arm and leg endurance was measured by comparing the affected and unaffected limb, using a modified endurance item of the self-perceived fitness questionnaire (1-13 point scale), as self- or proxy-reported.

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Upper ≤5 (strength slightly below the other arm or worse) 8%</th>
<th>Lower ≤6 (strength just below that of the other arm or worse) 25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a- Upper limb</td>
<td>0.66 0.54-0.77</td>
<td>96%</td>
<td>8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b- Lower limb</td>
<td>0.66 0.54-0.77</td>
<td>95%</td>
<td>25%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4- Strength Measurement style: Self/proxy-reported questionnaire

○ Click here for more information
○ Minimize information
Measurement details: Arm and leg strength was assessed by comparing the affected and unaffected limb, using a modified endurance item of the self-perceived fitness questionnaire (1-13 point scale), as self- or proxy-reported.
Symptoms (part 4 of 4)

PAIN

Background: Pain was measured using different scales, which overlapped in some age groups.

- **Visual Analogue Scale (VAS) – proxy report:** 0-3 yo

```
| No pain | Worst possible pain |
```

- **Faces Pain Scale – Revised (FPS-R) – patient self-report:** 4-19 yo

```
[Scale images showing different facial expressions]
```

- **Adolescent Pediatric Pain Tool (VAS with word descriptors) – patient self-report:** 10-19 yo

```
| No Pain | Little Pain | Medium Pain | Large Pain | Worst Possible Pain |
```

Instructions:

- You will be presented with three different potential scale choices
- Please keep in mind that each scale has a different (and sometimes overlapping) age range

1- Pain, Visual Analogue Scale (0-3 yo) Measurement
   style: Proxy report in children 0 to 3 yo
   LimbCriterionSensitivity Specificity Upper Pain
   present 10% 100% Lower Pain present 0% 100%
Measurement details: The item was measured using a visual analogue scale (VAS), applied to proxies of children younger than 4 years of age.

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>Pain present</td>
<td>47%</td>
<td>96%</td>
</tr>
<tr>
<td>Lower</td>
<td>Pain present</td>
<td>50%</td>
<td>80%</td>
</tr>
</tbody>
</table>

**Arm VAS**

- **AUC**: 0.55
- **95% CI**: 0.36-0.73
- **p-value**: 0.32

**Leg VAS**

- **AUC**: 0.50
- **95% CI**: 0.30-0.70
- **p-value**: 1.00

- **1a- Upper limb**: Keep
- **1b- Lower limb**: Remove

2. Pain, Faces Pain Scale-Revised (4-19 yo)

Measurement style: Self-reported scale in children 4 to 19 yo

LimbCriterion: Sensitivity Specificity Upper Pain
present 47% 96% Lower Pain present 50% 80%
Measurement details: We used the Faces Pain Scale – Revised (FPS-R) in children ≥4-19 years old, as self-reported

<table>
<thead>
<tr>
<th>Limb Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Upper Pain present with intensity &gt;1/10</th>
<th>60%</th>
<th>94%</th>
<th>Lower Pain present 67%</th>
<th>100%</th>
</tr>
</thead>
</table>

**AUC | 95% CI | p-value**
---|---|---
Arm FPS-R | 0.72 | 0.56-0.85 | 0.002 |
Leg FPS-R | 0.67 | 0.51-0.81 | 0.01 |

---

2a- Upper limb
2b- Lower limb
3- Pain, Adolescent Pediatric Pain Tool, Visual Analogue Scale (10-19 yo) Measurement style: Self-reported scale in children 10 to 19 yo Limb Criterion Sensitivity Specificity Upper Pain present with intensity >1/10 60% Lower Pain present 67% 100%

- Keep
- Remove
- Not sure

☐ Keep
☐ Remove
☐ Not sure

☐ Click here for more information
☐ Minimize information
Measurement details: We used the Adolescent Pediatric Pain Tool (APPT) in children \(\geq 10\) years old. The APPT assesses pain location (body outline diagram), intensity (visual analogue scale, VAS, with word descriptors) and quality (list of pain descriptors). The results of the VAS are presented here.

<table>
<thead>
<tr>
<th>3a- Upper limb</th>
<th>Keep</th>
<th>Remove</th>
<th>Not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>3b- Lower limb</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes)
The highest possible value is added here.

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.74</td>
<td>0.53-0.89</td>
<td>0.02</td>
</tr>
<tr>
<td>-----</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>0.83</td>
<td>0.63-0.95</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Skin

1. Skin temperature Measurement style: Infrared temperature measurement device
   *Ratio was estimated as affected/unaffected limb
   Limb Criterion: Sensitivity Specificity
   Upper *Ratio <= 0.99 8% 75% Lower *Ratio > 1.00
   56% 65%

Measurement details:
Temperature was measured at three standardized anatomic points. Values were averaged to calculate the affected to unaffected limb temperature ratio

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>*Ratio &lt;= 0.99</td>
<td>8%</td>
<td>75%</td>
</tr>
<tr>
<td>Lower</td>
<td>*Ratio &gt; 1.00</td>
<td>56%</td>
<td>65%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>*Ratio &lt;= 0.99</td>
<td>8%</td>
<td>75%</td>
</tr>
<tr>
<td>Lower</td>
<td>*Ratio &gt; 1.00</td>
<td>56%</td>
<td>65%</td>
</tr>
</tbody>
</table>

AUC | 95% CI | p-value
---|-------|-------
0.51 | 0.39-0.63 | 0.90

AUC | 95% CI | p-value
---|-------|-------
0.53 | 0.4-0.65 | 0.66

Arm Temperature Ratio

Leg Temperature Ratio
1a- Upper limb
1b- Lower limb

Time to apply the tool (minutes)

2- Collaterals Measurement style: Boulden scale
    Limb Criterion Sensitivity Specificity
    Upper Collaterals present 1+ 64%/71% Lower
    Collaterals present 2+, visible/>5cm 14% 88%

Measurement details: Collaterals were measured using the Boulden scale (0-4 point severity scale)

3- Varicose veins Measurement style: Revised venous clinical severity score
    Limb Criterion Sensitivity Specificity
    Upper Present 0%/100% Lower Present 8% 97%

08/24/2016 1:42pm
Measurement details: Varicose veins were measured using the item of the Revised venous clinical severity score (0-3 points: none to severe)

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Upper *Ratio</th>
<th>Lower *Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; = 1.06</td>
<td>67%</td>
<td>60%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;0.98</td>
<td>67%</td>
<td>60%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement details:</th>
<th>Skin durometry Measurement style: Skin resistance to indentation with durometer (force gauge)</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Ratio was estimated as affected/unaffected limb</td>
<td></td>
</tr>
<tr>
<td>LimbCriterionSensitivity Specificity Upper *Ratio</td>
<td></td>
</tr>
<tr>
<td>&lt;= 1.06 67% 60% Lower *Ratio &gt;0.98 67% 60%</td>
<td></td>
</tr>
</tbody>
</table>

3a- Upper limb
3b- Lower limb

Keep  Remove  Not sure

Time to apply the tool (minutes)

4- Skin durometry Measurement style: Skin resistance to indentation with durometer (force gauge)  
*Ratio was estimated as affected/unaffected limb
LimbCriterionSensitivity Specificity Upper *Ratio 🔘 Click here for more information
<= 1.06 67% 60% Lower *Ratio >0.98 67% 60%

Minimize information

AUC  95% CI  p-value
Arm Varicose Veins

0.50  0.38-0.62  1.00

Leg Varicose Veins

0.53  0.10-0.65  0.32
Measurement details: Skin resistance to indentation was measured three times at one standardized anatomical point on each limb. Measurements were averaged to calculate the affected limb skin resistance to unaffected limb skin resistance ratio.

The following skin items (measured using the Revised venous clinical severity score) were NOT OBSERVED during the study.

<table>
<thead>
<tr>
<th></th>
<th>Keep</th>
<th>Remove</th>
<th>Not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>4a- Upper limb</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4b- Lower limb</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes)
Please select the items you would like to KEEP for ARMS

<table>
<thead>
<tr>
<th>Skin pigmentation</th>
<th>Skin induration</th>
<th>Skin inflammation</th>
<th>Skin ulcer</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Please select the items you would like to KEEP for LEGS

<table>
<thead>
<tr>
<th>Skin pigmentation</th>
<th>Skin induration</th>
<th>Skin inflammation</th>
<th>Skin ulcer</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Please note that there is a conflict in your choices

Time to apply the tool (minutes)   ________________________________
Edema

Background: The tools used for objective evaluation of limb edema can be based on the measurement of either limb circumference or limb volume.
- Limb circumference is determined with a measuring tape.
- Limb volume can be assessed with the aid of geometric volume formulas (data are obtained with a measuring tape and a caliper) or using bioimpedance.
- Results of the affected limb are usually compared to those of the contralateral limb.

Instructions:

1. You will first be presented with three different estimation methods for measuring limb edema
2. You will be asked to state your preferred calculation for one of the methods
3. You may be asked which segment you prefer to measure (e.g., thigh and/or calf)
4. Lastly, you will be asked to choose your preferred final method to measure edema in the specified limb

We will start with ARMS:
Method 1: Arm Circumference Measurement style:
Measuring tape
Calculation Criterion
Sensitivity Specificity
Arm Circ. Ratio > 1.0176% 56%  Arm Circ.
Difference >= 0.4 cm 64% 62%

Measurement details:
Arm circumference was measured using a Gulick II anthropometric tape, at a point equidistant from the acromion process to the olecranon process. The absolute difference (in cm) between the affected and unaffected arm and the ratio of the affected to unaffected arm were estimated.

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.62</td>
<td>0.50-0.73</td>
<td>0.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.63</td>
<td>0.51-0.74</td>
<td>0.07</td>
</tr>
</tbody>
</table>
STEP 1: Please choose one of the following calculation methods (you will choose a measurement method later)

Method 2: Arm Volume Measurement style: Geometric measurements
Minimize information

Measurement details: Arm circumference was measured at 6 different anatomical points using a Gulick II tape; the distance between these points was determined with a caliper. Arm volume was mathematically estimated assuming an elliptical truncated cone geometry. The value was normalized by comparison to the value of the contralateral limb, and expressed as a ratio

Method 3: Arm Fluid Content Measurement style: Bioimpedance Spectroscopy (BIS)
Minimize information

AUC  95% CI  p-value
0.65  0.53-0.76  0.03
Measurement details: The principle underlying bioimpedance is that the opposition of tissues to the flow of electrical current is inversely proportional to the fluid volume of the tissue. Impedance at low frequencies reflects the extracellular fluid content alone. The measurement of extracellular fluid content of a limb is normalized by comparison to the value of the contralateral limb, and expressed as a ratio. Larger ratios mean larger volume in the affected limb.

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.62</td>
<td>0.5-0.74</td>
<td>0.08</td>
</tr>
</tbody>
</table>

STEP 2: Please choose a final method to measure ARM edema

Select the method to measure ARM edema you would like to keep

- [ ] arm_circ_kr
- [ ] Arm Vol. Ratio
- [ ] Arm BIS Ratio
- [ ] None
- [ ] Not sure

Time to apply the tool (minutes)

______________________________

We will now move on to the measurement of LEG edema.
Method 1: Leg Circumference Measurement style: Measuring tape

<table>
<thead>
<tr>
<th>CalculationCriterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thigh Circ. Ratio &gt; 1.029</td>
<td>53%</td>
<td>53%</td>
</tr>
<tr>
<td>Thigh Circ. Difference &gt; 1.2 cm</td>
<td>64%</td>
<td>74%</td>
</tr>
<tr>
<td>Calf Circ. Difference &gt; -0.4 cm</td>
<td>100%</td>
<td>29%</td>
</tr>
<tr>
<td>Ankle Circ. Ratio &lt; = 0.99</td>
<td>28%</td>
<td>88%</td>
</tr>
<tr>
<td>Ankle Circ. Difference &lt; -0.5 cm</td>
<td>28%</td>
<td>88%</td>
</tr>
</tbody>
</table>

**Measurement details:** Thigh/calf circumference was measured with a Gulick II tape, at a point equidistant to the apex of the patella and the anterior superior iliac spine apex (thigh) or to the patella and the medial malleolus (calf). Ankles were measured following the figure of 8 technique. The absolute difference [in cm] between the affected and unaffected arm and the ratio of the affected to unaffected arm were estimated.

**STEP 1:** Choose the leg segment that should be measured
- Only thigh
- Only calf
- Only ankle
- Only thigh and calf
- Only thigh and ankle
- Only calf and ankle
- Thigh, calf and ankle

**STEP 2:** Choose the calculation method
- [leg_circ] ratio
- [leg_circ] difference

<table>
<thead>
<tr>
<th>Segments (ratio)</th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thigh</td>
<td>0.71</td>
<td>0.59-0.81</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Calf</td>
<td>0.63</td>
<td>0.51-0.74</td>
<td>0.06</td>
</tr>
<tr>
<td>Ankle</td>
<td>0.55</td>
<td>0.43-0.67</td>
<td>0.48</td>
</tr>
</tbody>
</table>

**Segmenets (difference)**
- Thigh: 0.74, 0.62-0.84, <0.001
- Calf: 0.68, 0.56-0.79, 0.005
- Ankle: 0.54, 0.41-0.68, 0.59
Method 2: Leg Volume Measurement style: Geometric measurements

Measurement Criterion

Sensitivity  Specificity  Leg Volume Ratio >1.02 83% 44%

Method 3: Leg Fluid Content Measurement style: Bioimpedance spectroscopy (BIS)

Measurement Criterion

Sensitivity  Specificity  Leg Bioimpedance (BIS) Ratio >1.0281% 50%

Measurement details: Leg circumference was measured at 6 different anatomical points using a Gulick II tape; the distance between these points was determined with a caliper. Leg volume was mathematically estimated assuming an elliptical truncated cone geometry. The value was normalized by comparison to the value of the contralateral limb, and expressed as a ratio.

\[ \text{volume} = \frac{1}{12\pi} (x^2 + y^2 + h^2) \]

- \( x \)
- \( y \)
- \( h \)

AUC 95%CI p-value

0.61 0.47-0.71 0.18
Measurement details: The principle underlying bioimpedance is that the opposition of tissues to the flow of electrical current is inversely proportional to the fluid volume of the tissue. Impedance at low frequencies reflects the extracellular fluid content alone. The measurement of extracellular fluid content of a limb is normalized by comparison to the value of the contralateral limb, and expressed as a ratio. Larger ratios mean larger volume in the affected limb.

STEP 3: Please choose a final method to measure LEG edema

Select the method to measure LEG edema you would like to keep

- Leg Vol. Ratio
- Leg BIS Ratio
- None
- Not sure

Time to apply the tool (minutes)
The highest possible value is added here
**Other Clinical Signs & Disease Impact**

1- The following items were NOT OBSERVED during the study

<table>
<thead>
<tr>
<th>Pitting edema</th>
<th>Pain on palpation</th>
<th>Both</th>
<th>None</th>
<th>Not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Please select the items you would like to KEEP for ARMS

Please select the items you would like to KEEP for LEGS

<table>
<thead>
<tr>
<th>Time to apply the tool (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>________________________________</td>
</tr>
</tbody>
</table>

2- Limitation in activity & participation

Measurement style: Pediatric Outcomes Data Collection Instrument (PODCI)

LimbCriterionSensitivity Specificity Upper $< 84$ points out of a max score of $100$

Lower $< 91$ points out of $100$

42% 87% Lower $< 91$

43% 88%

Measurement details: We measured activity and participation limitation as per the definitions of the International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY), using the Pediatric Outcomes Data Collection Instruments (PODCI). Normal values in the general population for overall scores are: mean 93.3-95.2; standard deviation 7.9-7.2, depending on the version used.

---

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.63</td>
<td>0.50-0.74</td>
<td>0.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.69</td>
<td>0.56-0.80</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Measurement details: We measured the strength of the affected and the unaffected limb with a compact dynamometer following a standardized protocol. We then determined the affected to the unaffected limb strength ratio.

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.56</td>
<td>0.33-0.77</td>
<td>0.69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.63</td>
<td>0.41-0.81</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>Keep</td>
<td>Remove</td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>3a- Upper limb</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3b- Lower limb</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Estimated maximum amount of time the tool will take to apply with the items you have chosen
Appendix 6.3. Item Reduction Survey, Second Round: Screenshot of the Survey, as Seen by the Respondent

### Symptoms (part 1 of 4)

#### 1- Upper limb heaviness

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>Symptom present</td>
<td>40%</td>
<td>94%</td>
</tr>
</tbody>
</table>

- **Upper limb votes:** keep 66%, remove 20%, not sure 15%
- **Item KEPT for lower limbs (86% votes)**

<table>
<thead>
<tr>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Time to apply the tool for arms (minutes)*

*Four minutes is the starting point with all the items already included for the assessment of arms*

#### 2- Upper limb tiredness

<table>
<thead>
<tr>
<th>Limb</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>Symptom present</td>
<td>47%</td>
<td>86%</td>
</tr>
</tbody>
</table>

**Upper limb votes:** keep 85%, remove 20%, not sure 15%
**Item KEPT (75% votes) for lower limbs**

<table>
<thead>
<tr>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Time to apply the tool (minutes) for arms*

*must provide value*
Welcome to the last round of the pediatric thrombosis expert consensus survey!

In this page, you can see a summary of the responses obtained in the previous round (items voted "keep" or "remove" by at least 70% of respondents).

The 18 items that did not achieve a 70% endorsement are re-circulated for a final voting in this survey. Items that do not obtain 70% endorsement for either keep or remove will be eliminated from the tool.

It will take you less than 10 min to answer the survey.

These are the items included so far. Each one had at least 70% votes for "keep".
The following items were excluded on the previous round. Each one had at least 70% votes for "remove".

<table>
<thead>
<tr>
<th>Upper limbs</th>
<th>Lower limbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delayed healing</td>
<td>Delayed healing</td>
</tr>
<tr>
<td>Strength (reported and tested)</td>
<td>Strength (tested)</td>
</tr>
<tr>
<td>Skin temperature</td>
<td>Skin temperature</td>
</tr>
<tr>
<td>Skin induration, inflammation,</td>
<td>Skin inflammation</td>
</tr>
<tr>
<td>pigmentation</td>
<td>Pain on palpation</td>
</tr>
<tr>
<td>Ulcers</td>
<td>Durometry</td>
</tr>
<tr>
<td>Pitting edema</td>
<td></td>
</tr>
<tr>
<td>Pain on palpation</td>
<td></td>
</tr>
<tr>
<td>Durometry</td>
<td></td>
</tr>
<tr>
<td>Varicose veins</td>
<td></td>
</tr>
</tbody>
</table>

This screenshot is just a reminder of how to complete the survey (with a made up example):
Click "Next Page" to start

**Symptoms (part 1 of 4)**

1- Upper limb heaviness LimbCriterionSensitivity Specificity Upper Symptom present 40% 94%

○ Click here for measurement style and ROC curve
○ Minimize information
Measurement details: The item was measured in children ≥10 years of age, using a 5-point scale (none-everyday)

- Upper limb votes: keep 65%, remove 20%, not sure 15%
- Item KEPT for lower limbs (85% votes)

Upper limb heaviness

Time to apply the tool for arms (minutes)* *Four minutes is the starting point with all the items already included for the assessment of arms

2- Upper limb tiredness LimbCriterionSensitivity Specificity Upper Symptom present 47% 88% Click here for measurement style and ROC curve Minimize information
Upper limb votes: keep 65%, remove 20%, not sure 15%
Item KEPT (75% votes) for lower limbs

Tired upper limb

Time to apply the tool (minutes) for arms

3- Limb pruritus
LimbCriterionSensitivity Specificity Upper Symptom present 8% 98% Lower Symptom present 17% 97%

AUC  95% CI  p-value
0.69  0.53-0.83  0.01
Measurement details: The item was measured using a 5-point scale (none-everyday), in all children as self- or proxy-reported.

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.53</td>
<td>0.41-0.65</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Upper limb votes: keep 25%, remove 65%, not sure 10%
Lower limb votes: keep 30%, remove 60%, not sure 10%

3a- Upper limb pruritus  
3b- Lower limb pruritus

Time to apply the tool (minutes) for arms

Time to apply the tool (minutes) for legs *13 min is the starting point with all the items already included for legs (patient+clinician reported)

4- Paresthesia (tingling) LimbCriterionSensitivity Specificity Upper Symptom present 40% 100% Lower Symptom present 33% 100%

08/24/2016 2:01pm  
www.projectredcap.org
Upper limb votes: keep 65%, remove 20%, not sure 15%
Lower limb votes: keep 55%, remove 30%, not sure 15%

4a- Upper limb tingling
4b- Lower limb tingling

Time to apply the tool (minutes) for arms
Time to apply the tool (minutes) for legs

Measurement details: The item was measured in children ≥10 years of age using a 5-point scale (none-everyday); proxy-report for younger patients was optional.
### Symptoms (part 2 of 4)

1. Skin redness/purple or blotchy skin
   - Limb criterion: Sensitivity 24%  Specificity 98%
   - Upper limb: Symptom present 24%  Specificity 98%
   - Lower limb: Symptom present 19%  Specificity 100%

   ![Arm Skin Redness/Mottled Skin](image1)
   ![Leg Skin Redness/Mottled Skin](image2)

   **Measurement details:** The item was measured using a 5-point scale (none-everyday) in all children, as self- or proxy-reported

   **AUC** | **95% CI** | **p-value**
   --- | --- | ---
   0.61 | 0.48-0.72 | 0.02

   **Upper limb votes:**
   - Keep 60%, remove 30%, not sure 10%
   - 1a- Upper limb red/blotchy skin
   - 1b- Lower limb red/blotchy skin

   **Time to apply the tool (minutes) for arms**

   **Time to apply the tool (minutes) for legs**

2. Warmer limb
   - Limb criterion: Sensitivity 20%  Specificity 98%
   - Upper limb: Symptom present 20%  Specificity 98%
   - Lower limb: Symptom present 19%  Specificity 100%

   ![AUC](image3) | ![AUC](image4)

   **AUC** | **95% CI** | **p-value**
   --- | --- | ---
   0.60 | 0.47-0.71 | 0.004

   **Lower limb votes:**
   - Keep 60%, remove 30%, not sure 10%

   2- Warmer limb

   ![Click here for measurement style and ROC curves](image5)
   ![Minimize information](image6)
Measurement details: The item was measured using a 5-point scale (none-everyday) in all children, as self- or proxy-reported

Upper limb votes: keep 35%, remove 50%, not sure 15%
Lower limb votes: keep 35%, remove 50%, not sure 15%

2a- Warmer upper limb
2b- Warmer lower limb

Time to apply the tool (minutes) for arms
Time to apply the tool (minutes) for legs

3- Cramps LimbCriterionSensitivity  Specificity
Upper  Symptom present 20%  94% Lower  Symptom present at least once a week 26%  100%

AUC  95% CI  p-value
0.59  0.47-0.71  0.03

AUC  95% CI  p-value
0.60  0.47-0.71  0.004

Upper limb votes: keep 35%, remove 50%, not sure 15%
Lower limb votes: keep 35%, remove 50%, not sure 15%

2a- Warmer upper limb
2b- Warmer lower limb

Keep
Remove

○
○

○
○

○
○

○
○

Click here for measurement style and ROC curves
Minimize information
Measurement details: The item was measured in children $\geq 10$ years of age using a 5-point scale (none-everyday)

Upper limb votes: keep 25%, remove 55%, not sure 20%
Lower limb votes: keep 30%, remove 55%, not sure 15%

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm Cramps</td>
<td>0.57</td>
<td>0.37-0.76</td>
<td>0.33</td>
</tr>
<tr>
<td>Leg Cramps</td>
<td>0.61</td>
<td>0.40-0.79</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes) for arms
Time to apply the tool (minutes) for legs
Symptoms (part 3 of 4)

1- Tightness (limb feels tight)
   LimbCriterion: Sensitivity Specificity Upper
   Symptom present 10% 100% Lower Symptom present 42%
   100%

   ○ Click here for measurement style and ROC curves
   ○ Minimize information

Measurement details: The item was measured in children ≥10 years of age, using a 5-point scale (none-everyday)

![](image)

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0.55</strong></td>
<td><strong>0.35-0.74</strong></td>
<td>0.32</td>
</tr>
</tbody>
</table>

Upper limb votes: keep 15%, remove 60%, not sure 25%
Lower limb votes: keep 65%, remove 20%, not sure 15%

<table>
<thead>
<tr>
<th>1a- Upper limb tightness</th>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td>1b- Lower limb tightness</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes) for arms

Time to apply the tool (minutes) for legs
2- Lower limb strength (reported)
LimbCriterionSensitivity  Specificity  Lower  $< = 6$
(strength just below that of the other leg or
worse)25% 95%

Measurement details: Leg strength was assessed by
comparing the affected and unaffected limb, using a
modified endurance item of the self-perceived fitness
questionnaire (1-13 point scale), as self- or proxy-reported

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.61</td>
<td>0.48-0.72</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Lower limb votes: keep 30%, remove 50%, not sure 20%
Item REMOVED (70% votes) for upper limbs

Lower limb strength

Time to apply the tool (minutes) for legs

3- Pain, Visual Analogue Scale (0-3 yo)
LimbCriterionSensitivity  Specificity  Upper  Pain
present 10%  100% Lower  Pain present 0% 100%

Click here for measurement style and ROC curves

Minimize information
Measurement details: The item was measured using a visual analogue scale (VAS), applied to proxies of children younger than 4 years of age.

Upper limb votes: keep 30%, remove 55%, not sure 15%
Lower limb votes: keep 30%, remove 55%, not sure 15%

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm VAS</td>
<td>0.55</td>
<td>0.36-0.73</td>
<td>0.32</td>
</tr>
<tr>
<td>Leg VAS</td>
<td>0.50</td>
<td>0.30-0.70</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Upper limb votes: 3a- Upper limb VAS Keep, 3b- Lower limb VAS Keep
Lower limb votes: 3a- Upper limb VAS Remove, 3b- Lower limb VAS Remove

Pain has already been counted in, as it has been voted as part of the tool for older children.
Skin

1- Lower limb collaterals
Limb Criterion Sensitivity Specificity
Lower Collaterals present 2+, visible>5cm 14% 88%

Measurement details: Collaterals were measured using the Boulden scale (0-4 point severity scale)

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimize information</td>
<td>14%</td>
<td>88%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Limb</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower</td>
<td>8%</td>
<td>97%</td>
</tr>
</tbody>
</table>

Lower limb votes: keep 50%, remove 35%, not sure 15%
Item KEPT (90% votes) for upper limbs

Lower limb collaterals

Time to apply the tool (minutes) for legs

2- Lower limb varicose veins Limb
Criterion Sensitivity Specificity
Lower Present 8% 97%

<table>
<thead>
<tr>
<th>Limb</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimize information</td>
<td>8%</td>
<td>97%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Limb</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower</td>
<td>8%</td>
<td>97%</td>
</tr>
</tbody>
</table>

AUC 95% CI p-value
0.51 0.39-0.64 0.81
Measurement details: Varicose veins were measured using the item of the Revised venous clinical severity score (0-3 points = none to severe)

<table>
<thead>
<tr>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.53</td>
<td>0.10-0.65</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Lower limb votes: keep 45%, remove 50%, not sure 5%
Item REMOVED (70% votes) for upper limbs

<table>
<thead>
<tr>
<th>Lower limb varicose veins</th>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes) for legs

The following skin items (measured using the Revised venous clinical severity score) were NOT OBSERVED during the study.

All items were REMOVED for upper limbs

<table>
<thead>
<tr>
<th>Skin Pigmentation in Lower Limbs (60% keep, 40% remove)</th>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Skin Induration in Lower Limbs (45% keep, 55% remove)</th>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes) for legs
Edema

Background: The tools used for objective evaluation of limb edema can be based on the measurement of either limb circumference or limb volume.
- Limb circumference is determined with a measuring tape.
- Limb volume can be assessed with the aid of geometric volume formulas (data are obtained with a measuring tape and a caliper) or using bioimpedance.
- Results of the affected limb are usually compared to those of the contralateral limb.

1- Measurement of ARM edema

<table>
<thead>
<tr>
<th>Arm Circumference Calculation</th>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arm Circ. Ratio &gt;1.0176%</td>
<td>56%</td>
<td>Arm Circ.</td>
<td></td>
</tr>
<tr>
<td>Difference&gt;=0.4 cm</td>
<td>64%</td>
<td>62%</td>
<td></td>
</tr>
</tbody>
</table>
Measurement details:
Arm circumference was measured using a Gulick II anthropometric tape, at a point equidistant from the acromion process to the olecranon process. The absolute difference (in cm) between the affected and unaffected arm and the ratio of the affected to unaffected arm were estimated.
Summary of votes for arm measurement

- Arm Circumference Difference (45% votes)
- Arm Circumference Ratio (25% votes)
- Arm Volume Ratio (15% votes)
- None (5% votes)

Please select the method you would like to keep for the measurement of ARMS

- Arm Circumference Difference (45% votes)
- Arm Circumference Ratio (25% votes)
- Arm Volume Ratio (15% votes)
- None (5% votes)

Time to apply the tool (minutes) for arms

2- Measurement of LEG edema
## Leg Circumference Calculation

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thigh Circ. Ratio &gt; 1.0293%</td>
<td>53% Thigh Circ. Difference &gt; 1.2 cm</td>
<td>64% 74% Calf Circ. Ratio &gt; 0.99 94% 35% Calf Circ. Difference &gt; -0.4 cm 100% 29% Ankle Circ. Ratio &lt; = 0.99 28% 88% Ankle Circ. Difference &lt; -0.5 cm 28% 88%</td>
</tr>
</tbody>
</table>

**Measurement details:** Thigh/calf circumference was measured with a Gulick II tape, at a point equidistant to the apex of the patella and the anterior superior iliac spine apex (thigh) or to the patella and the medial malleolus (calf). Ankles were measured following the figure of 8 technique. The absolute difference (in cm) between the affected and unaffected arm and the ratio of the affected to unaffected arm were estimated.

![Graphs and images showing measurement details.](image-url)

<table>
<thead>
<tr>
<th>Segments (ratio)</th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thigh</td>
<td>0.71</td>
<td>0.59-0.81</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Calf</td>
<td>0.63</td>
<td>0.51-0.74</td>
<td>0.06</td>
</tr>
<tr>
<td>Ankle</td>
<td>0.55</td>
<td>0.43-0.67</td>
<td>0.48</td>
</tr>
</tbody>
</table>
Summary of measurements for legs

- Thigh & calf difference (45% votes)
- Thigh & calf ratio (20% votes)
- Thigh, calf & ankle ratio (10% votes)
- Thigh, calf & ankle difference (5% votes)
- Only thigh ratio (5% votes)
- Only thigh difference (5% votes)
- Only calf difference (5% votes)
- Not sure (5% votes)

Please select the method you would like to keep for the measurement of LEGS

- Thigh & calf difference (45% votes)
- Thigh & calf ratio (20% votes)
- Thigh, calf & ankle ratio (10% votes)
- Thigh, calf & ankle difference (5% votes)
- Only thigh ratio (5% votes)
- Only thigh difference (5% votes)
- Only calf difference (5% votes)

Time to apply the tool (minutes) for legs

_______________________________
Other Clinical Signs & Disease Impact

1- The following item was NOT OBSERVED during the study

<table>
<thead>
<tr>
<th>Pitting edema in lower limbs</th>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td>(40% keep, 55% remove, 5% not sure)</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes) for legs

2- Limitation in activity & participation

<table>
<thead>
<tr>
<th>LimbCriterion</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Upper&lt; =84</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limb</td>
<td></td>
<td></td>
<td>Minimize information</td>
</tr>
<tr>
<td>Criterion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Click here for measurement style and ROC curve

Measurement details: We measured activity and participation limitation as per the definitions of the International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY), using the Pediatric Outcomes Data Collection Instruments (PODCI). Normal values in the general population for overall scores are: mean 93.3-95.2; standard deviation 7.8-7.2, depending on the version used

AUC 95% CI p-value
0.63 0.50-0.74 0.09

Upper limb votes: keep 65%, remove 25%, not sure 10%

Item KEPT (75% votes) for lower limbs

<table>
<thead>
<tr>
<th>Upper limb PODCI</th>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Time to apply the tool (minutes) for arms

Upper limb votes: keep 65%, remove 25%, not sure 10%

Item KEPT (75% votes) for lower limbs

<table>
<thead>
<tr>
<th>Upper limb PODCI</th>
<th>Keep</th>
<th>Remove</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
Estimated maximum amount of time the tool will take to apply for UPPER limbs

________________________

Estimated maximum amount of time the tool will take to apply for LOWER limbs

________________________
Appendix 7.1. Evolution of Test Validity Theory

The earliest definitions of validity emerged along with educational and psychological testing, and were influenced by statistical advances of the time [293]. At the beginning of the 20th century, soon after Pearson published his formula for correlation coefficients, validity was thought of in terms of the correlation between test scores and some criterion [339]. The concept became more general with time, and by the 1920s and 30s, a test was considered valid for anything it correlated with [293, 340]. A more theoretical definition proposed in the 1930s was that validity was the degree or extent to which a test “measures what it purports to measure” [341].

In the 1940s, there was growing criticism of relying on the correlational approach to assess test validity and the adequacy of the criteria used for validation was questioned [293, 294], driving the validity theory to evolve. Operationalization and content became relevant [342], and the concept of content validity was thus introduced.

In the 1950s, four categories of validity were described in the Technical Recommendations and Standards: predictive validity, status validity, content validity, and congruent validity [343]. In the 1954 version of the Standards, status and congruent validity were renamed concurrent and construct validity, respectively. In subsequent editions, predictive and concurrent validity were combined as “criterion-related validity”, giving rise to the three probably best known types [317], aspects [318], or categories [319] (appearing in that order) of validity: criterion, construct, and content-related validity [293].

The listing of construct validity in the 1954 version of the Standards was the result of the influence of philosophy of science, introduced by Cronbach and Meehl [294]. These authors shifted the concept of validity from correlation, the dominant view up until then, to explanation [293]. The definition these authors provided in 1955 is of particular interest: “Construct validation is involved whenever a test is to be interpreted as a measure of some attribute or quality which is not ‘operationally defined’.” [344].

Attributes that are not operationally defined (described in section 7.2.1) are frequent in education and psychology, given the complexity and level of abstraction of the constructs these fields study.
Under the concept developed by Cronbach and Meehl, the identity of such constructs can be defined by a system of laws and relationships between/among the construct and observable terms, according to theory. This theoretical network then allows the operationalization of the construct [301], supporting the meaning of scores [293]. Campbell and Fiske later added that the measure of a construct should be strongly related to other similar measures of the same construct, and weakly related to measures of different constructs [294, 345], a popular concept in measurement.

Modern theorists consider that this theoretical network is an “overly complicated” framework of validity [293]. Furthermore, the idea that the meaning of a term is fixed and implicitly given by its position in a theoretical network is currently the subject of great controversy. To give an example, the meaning of length is not defined by its relation with weight [140, 293].
Appendix 7.2. Representation of Post-thrombotic Syndrome

The triangle represents all possible states of the vascular system of patients who had a deep vein thrombosis (DVT, see text). These states are defined by the combination of signs and symptoms of post-thrombotic syndrome (PTS). Potential cut-off scores representing the transition between relevant states are shown (diseased/non-diseased; treatment cut-off).