GLOBAL ASSESSMENT IN PEDIATRIC SURGERY: AN EVIDENCE-BASED, WEIGHTED, CAPACITY ASSESSMENT TOOL FOR CHILDREN'S SURGERY IN LOW-RESOURCE SETTINGS

Yasmine Yousef

Department of Surgical and Interventional Sciences, Faculty of Medicine and Health Sciences

McGill University, Montreal

December 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of PhD in Surgical and Interventional Sciences

DEDICATION

This thesis is dedicated to my late father, Ibrahim. To him I owe my love of research and my staunch belief that all of us, no matter our place of birth, deserve access to medical and surgical care. To my beautiful daughter who was lucky enough to be born with a congenital anomaly in a high-income country. To my husband, my rock, my support, who inspires me and who has pushed me through the finish line. Finally, to my mother, she epitomizes all that a mother should be, the woman who has taught me what true endurance, tenacity, and courage are, my unwavering pillar throughout my entire life.

Table Of Contents

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

BACKGROUND

Pediatric surgical care in low- and middle-income countries (L-MICs) is hindered by limited resources, infrastructure, training, and systemic inefficiencies, contributing to high child morbidity and mortality rates. Existing capacity assessment tools are insufficient for comprehensively evaluating pediatric surgical capabilities in resource-constrained settings. The Global Assessment for Pediatric Surgery (GAPS) tool was developed to bridge this gap, providing an evidence-based instrument designed for L-MICs to assess surgical capacity and facilitate global surgical partnerships.

METHODS

GAPS was developed through a systematic review, expert consultations, and pilot testing phases. Version 1 of the tool, comprising 139 items, was expanded to 168 items following expert feedback. GAPS Version 2 was pilot tested in 65 institutions across various LMICs to ensure feasibility and assess item performance. Using construct validation, items were evaluated for their capacity to discriminate between basic and advanced surgical care centers. The refined GAPS Version 3 consisted of 64 questions across five key domains: human resources, material resources, accessibility, outcomes, and education. A capacity-based weighting system was then incorporated to enhance predictive validity based on external health metrics such as human development index (HDI), under-five mortality rate (U5MR), neonatal mortality rate (NNMR), deaths due to injury and deaths due to congenital anomalies.

RESULTS

Pilot testing demonstrated that 64 items effectively distinguished between basic and advanced surgical care facilities. The tool's weighted scoring system correlated significantly with key health indicators such as HDI ($p < 0.001$), U5MR ($p < 0.001$), and NNMR ($p < 0.001$). This highlighted GAPS' capability to serve as a comprehensive measure of pediatric surgical capacity and its potential to inform resource allocation and strategic partnerships.

CONCLUSION

The GAPS tool is a validated and methodologically robust instrument for evaluating pediatric surgical capacity in L-MICs. It fills a crucial need for objective assessments that inform quality improvement initiatives and foster effective collaborations. Future plans include

developing a free, accessible application that would facilitate multi-user assessments within institutions to refine the methodological limitations of the tool and the tools temporal stability. Such advancements will make GAPS an essential tool for driving strategic improvements in global pediatric surgical care and monitoring outcomes during and after partnerships with high-income countries or other L-MICs.

ABREGE

CONTEXTE

Les soins chirurgicaux pédiatriques dans les pays à faible et moyen revenu (PFMR) sont entravés par des ressources limitées, des infrastructures insuffisantes, un manque de formation et des inefficacités systémiques, contribuant ainsi à des taux élevés de morbidité et de mortalité infantile. Les outils d'évaluation de la capacité existants ne suffisent pas à évaluer de manière exhaustive les capacités chirurgicales pédiatriques dans les environnements à ressources limitées. L'outil Global Assessment for Pediatric Surgery (GAPS) a été développé pour combler cette lacune, offrant un instrument basé sur des preuves, conçu pour les PFMR afin d'évaluer la capacité chirurgicale et de faciliter les partenariats chirurgicaux mondiaux.

METHODES

Le développement de GAPS a été réalisé à travers une revue systématique, des consultations d'experts et des phases de tests pilotes. La version 1 de l'outil, comprenant 139 items, a été élargie à 168 items suite aux retours des experts. La version 2 de GAPS a été testée en pilote dans 65 institutions à travers divers PFMR pour garantir sa faisabilité et évaluer la performance des items. Grâce à la validation de construit, les items ont été évalués pour leur capacité à discriminer entre les centres de soins chirurgicaux de base et les centres de soins avancés. La version affinée GAPS Version 3 comprenait 64 questions réparties en cinq domaines clés : ressources humaines, ressources matérielles, accessibilité, résultats et éducation. Un système de pondération basé sur la capacité a ensuite été intégré pour améliorer la validité prédictive en se basant sur des métriques externes de santé telles que l'indice de développement humain, le taux de mortalité des moins de cinq ans, le taux de mortalité néonatale, les décès dus aux blessures et les décès dus aux anomalies congénitales.

RESULTATS

Les tests pilotes ont démontré que 64 items distinguaient efficacement les installations chirurgicales de base et avancées. Le système de notation pondérée de l'outil a montré une corrélation significative avec des indicateurs clés de santé tels que l'indice de développement humain ($p < 0,001$), le taux de mortalité des moins de cinq ans ($p < 0,001$) et le taux de mortalité néonatale ($p < 0,001$). Cela a mis en évidence la capacité de GAPS à servir de mesure globale de

la capacité chirurgicale pédiatrique et son potentiel à éclairer l'allocation des ressources et les partenariats stratégiques.

CONCLUSION

L'outil GAPS est un instrument validé et méthodologiquement solide pour évaluer la capacité chirurgicale pédiatrique dans les PFMR. Il répond à un besoin crucial d'évaluations objectives qui informent les initiatives d'amélioration de la qualité et favorisent des collaborations efficaces. Les plans futurs incluent le développement d'une application gratuite et accessible qui faciliterait les évaluations multi-utilisateurs au sein des institutions pour affiner les limites méthodologiques de l'outil et sa stabilité temporelle. Ces avancées feront de GAPS un outil essentiel pour stimuler les améliorations stratégiques dans les soins chirurgicaux pédiatriques mondiaux et surveiller les résultats avant, pendant et après les partenariats avec les pays à revenu élevé ou d'autres PFMR.

ACKNOWLEDGMENTS

This thesis represents the culmination of a journey that would not have been possible without the invaluable support, guidance, and contributions of many individuals and groups.

First and foremost, I extend my deepest gratitude to my supervisors, Dr. Dan Poenaru, Dr. Robert Baird, and Dr. Jean-Martin Laberge. Their unwavering support, mentorship, and insightful feedback have been instrumental in the success of this project.

I am also profoundly grateful to my thesis chair, Dr. Julio Fiore, for his steadfast support and guidance throughout this process. His thoughtful advice and encouragement have been pivotal in navigating the complexities of this work, and I deeply value his dedication.

The collection of materials, thesis design, and pilot data testing were made possible through the extensive assistance of several individuals. I am particularly thankful to Dr. Sherif Emil, Dr. Jean-Martin Laberge, Dr. Pramod Puligandla, Dr. Kenneth Shaw, and Dr. Robin Petroze for their critical contributions to the design of the thesis. Additionally, I am deeply appreciative of the following individuals, whose invaluable assistance in data collection and pilot studies was essential to the completion of this work: Kevyn Armand, Daniel Agbo-Panzo, Marilyn Butler, Beda Espineda, Kenzy Abdel-Hamid, Sonia Anchouche, Ai-Xuan Holterman, George W. Galiwango, Nathalie MacKinnon, Ritesh Shrestha, Desigen Reddy, William Harkness, Phyllis Kisa, Rashmi Kumar, Ibrahima Fall, Justin Onen, Ananda K. Lamahewage, Mike Ganey, Tessa Concepcion, Peter Donkor, Zaitun Bokhary, Anette S. Jacobsen, Reju Thomas, Brouh Yapo, Adiyasuren Jamiyanjav, Marcia Matias, Benno Ure, Vanda Amado, Bahati Robert, Lubna Samad, Roumanatou Bankolé, Sarah Cairo, Etienne St-Louis, Laura F. Goodman, Doulia M. Hamad, Emily R. Smith, Mohamed Abdelmalak, Zipporah Gathuy, Faye Evans, Maryam Ghavami Adel, Ki K. Bertille, Milind Chitnis, Leecarlo Millano, Peter Nthumba, Sergio d'Agostino, Bruno Cigliano, Luis Zea-Salazar, Doruk Ozgediz, and Elena Guadagno. Their collective contributions have significantly enriched the scope and rigor of this research, and I am truly grateful for their time and expertise.

Finally, I could not have become a pediatric surgeon or completed this thesis without the unwavering support and guidance of Dr Sherif Emil, to whom I am deeply indebted.

LIST OF ABBREVIATIONS

LIC: Low-income country

MIC: Middle-income country

LMIC: low middle-income country

UMIC: upper middle-income country

HIC: High income country

L-MIC: Low- and middle-income countries

GAPS: Global Assessment in Pediatric Surgery

HDI: Human Development Index

U5MR: Under-five mortality rate

NNMR: Neonatal mortality rate

CONTRIBUTIONS TO ORIGINAL KNOWLEDGE

This thesis makes a significant contribution to the field of global pediatric surgery by developing and validating the Global Assessment for Pediatric Surgery (GAPS) tool. As discussed in this body of work, unlike existing tools, GAPS provides a comprehensive, evidence-based approach for assessing pediatric surgical capacity in L-MICs, addressing a critical gap in global surgery initiatives. The research offers original insights through its rigorous multi-phase methodology. This thesis establishes a new standard for evaluating pediatric surgical capacity and provides an objective measure for strategic resource allocation and international partnerships aimed at improving pediatric surgical outcomes and pediatric surgical infrastructure in L-MICs.

CONTRIBUTIONS OF AUTHORS

I, Yasmine Yousef, was the primary author for all chapters in this thesis, including the manuscripts contained within. I was responsible for conceptualizing the research questions, conducting the analyses, interpreting the findings, and drafting the chapters and manuscripts. I synthesized the literature, designed the studies, and prepared the final documents for submission and presentation. Contributions of all manuscript authors are detailed below per manuscript.

Manuscript 1: Yasmine Yousef, Robert Baird, Emily R. Smith and Dan Poenaru contributed to study conception and design. Yasmine Yousef and Elena Guadagno developed the search strategy. Yasmine Yousef and Etienne St-Louis reviewed all manuscripts. The first draft of the manuscript was written by Yasmine Yousef and all authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

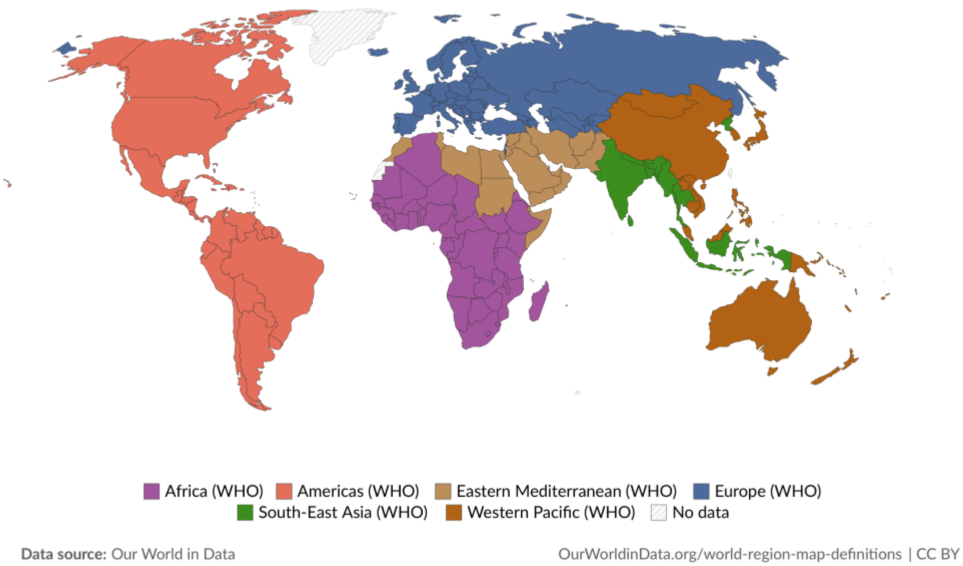
Manuscript 2: Yasmine Yousef, Dan Poenaru, and Jean-Martin Laberge contributed to the study conception and design. Material preparation, data collection and analysis were performed by Yasmine Yousef, Sarah Cairo, Laura F. Goodman, Doulia M. Hamad, Jean-Martin Laberge, and Dan Poenaru. The first draft of the manuscript was written by Yasmine Yousef and all authors commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

Manuscript 3: Yasmine Yousef participated in conceptualization, methodology, data curation, and analysis. Dan Poenaru, Emmanuel Ameh participated in conceptualization and methodology. Luc Malemo Kalisya participated in data collection and methodology. The first draft of the manuscript was written by Yasmine Yousef, and all authors reviewed the manuscript.

CHAPTER 1: INTRODUCTION

There has been a profound surge in academic output within the realm of global surgery, notably since the seminal 2015 Lancet Commission on Global Surgery (LCoGS) article, "*Global Surgery 2030: Evidence and solutions for achieving health, welfare, and economic development*" [1, 2]. This landmark publication has acted as a catalyst for research in the field of global surgery. Since the publication of LCoGS, the publication of and interest in pediatric global surgery-focused literature has risen steeply and comprised nearly one-fifth of all global surgery articles published in the period from January 2015 to February 2016 [2]. A significant portion of this body of work has been disseminated through the *World Journal of Surgery*, primarily by authors affiliated with High-Income Countries (HICs) yet focusing on the African region as delineated by the World Health Organization (WHO)(Figure 1)[2, 3].

Figure 1.World regions according to World Health Organization (WHO)



Moreover, the thematic focus of this burgeoning body of literature has revealed a predominant emphasis on health system strengthening and surgical epidemiology, accounting for over half articles published [2]. Within the extensive subset of 747 studies dedicated to healthcare delivery and management, a nuanced distribution of specific domains is evident: 53% of these studies focus on infrastructure and resources, while 48% delve into issues of quality and safety [2]. This distribution underscores an intentional focus on foundational and operational aspects of healthcare systems, highlighting both resource allocation and the necessity of maintaining high standards of care.

Building on the significant focus on pediatric surgery within the global surgery discourse, it's crucial to contextualize the urgency of addressing pediatric surgical needs, particularly in low- and middle-income countries (L-MICs). These regions exhibit a demographic profile heavily skewed towards youth, with median ages ranging between 14 and 30 years across Africa (Figure 2, Figure 3)[4].

Figure 2. Median Age Across the World, 2023

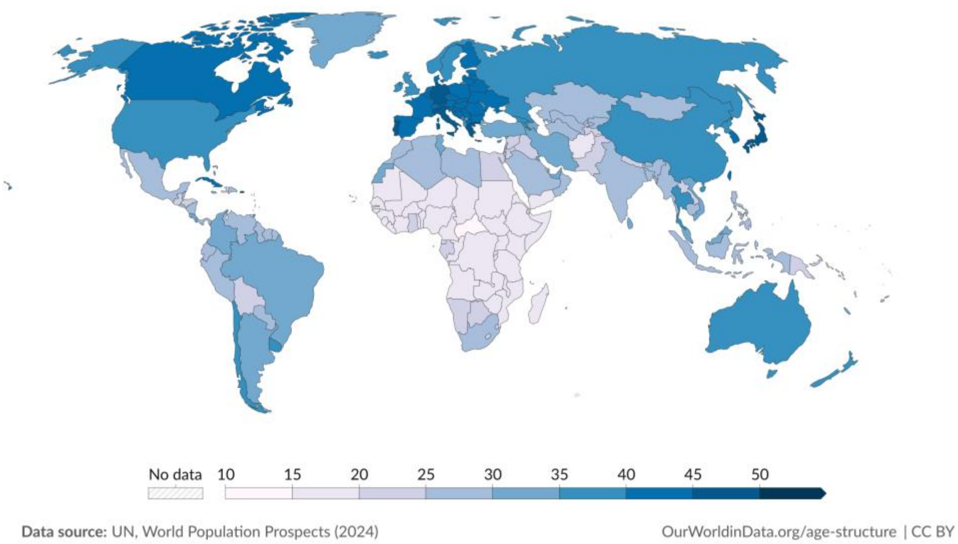
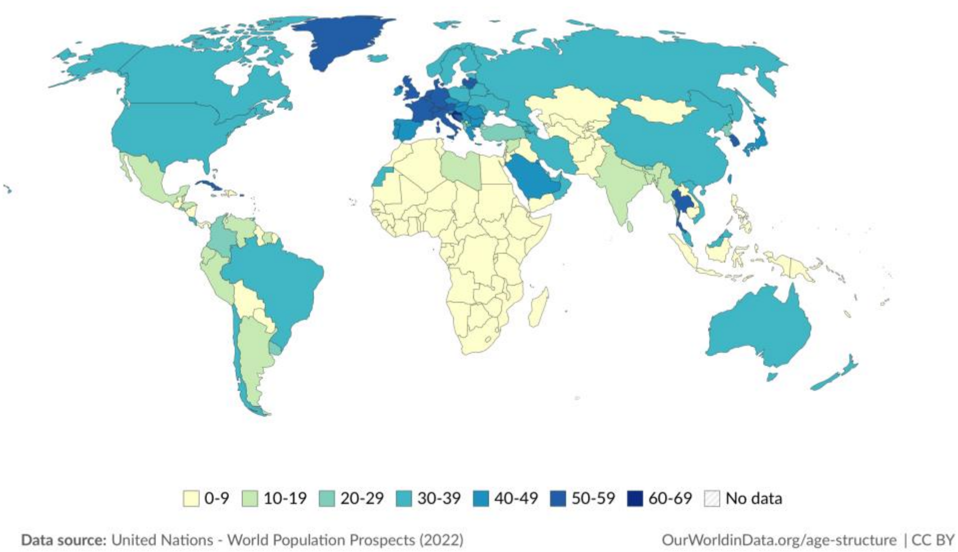


Figure 3. Ten-year Age Group with the Largest Population, 2021



Furthermore, projections indicate that children under 15 years old are expected to represent more than 30% of the African continent's population by 2050, amounting to over 786 million children (Figure 4 and Figure 5)[4].

Figure 4. Population by age group, Africa (United Nations)

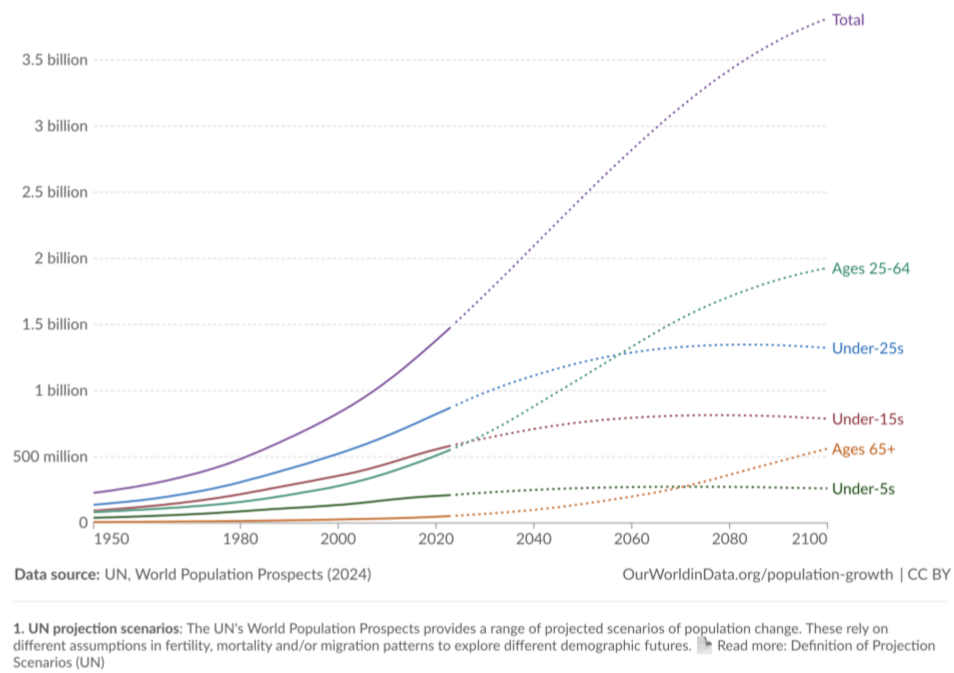
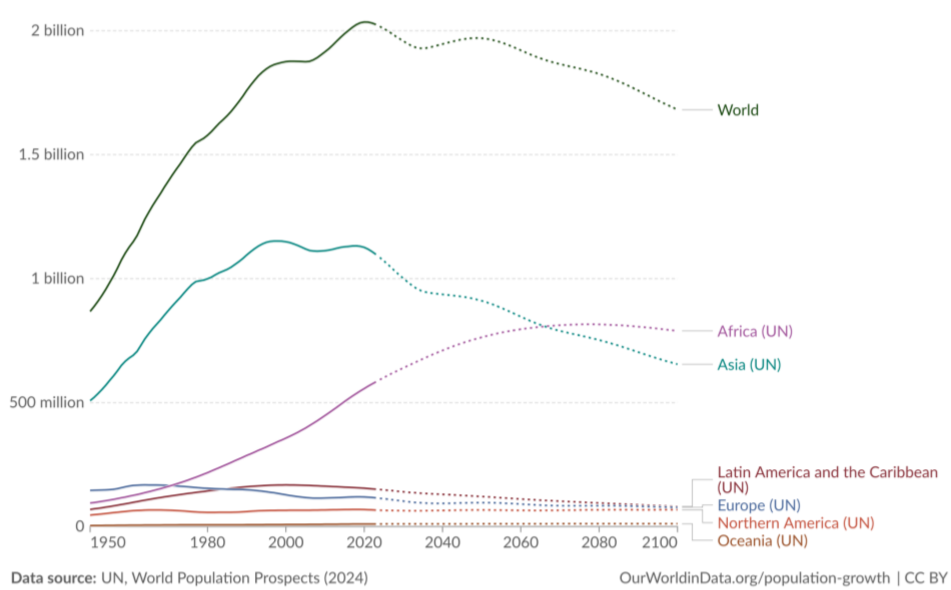
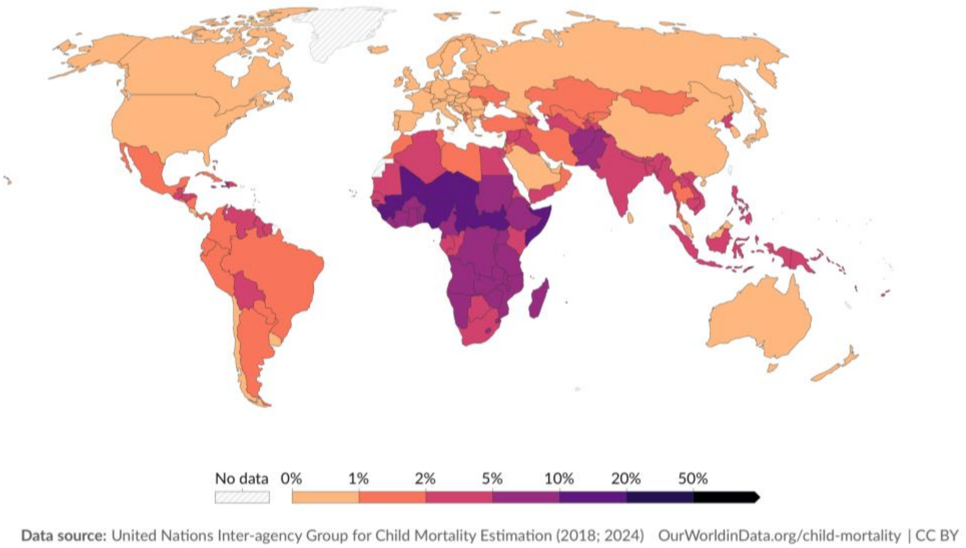


Figure 5. Children under age 15, by world region



This demographic trend underscores the critical need for pediatric surgical interventions in L-MICs, especially when considering the disparity in youth mortality rates—up to 15 times higher in L-MICs compared to HICs (Figure 6)[5].

Figure 6. Youth (before age 15) Mortality Rate, 2022



A study by Bickler et al. in 2003 highlighted this disparity, estimating that by the age of 15, up to 85% of children in L-MICs may require treatment for surgical conditions [6]. This staggering statistic not only illustrates the vast unmet need for pediatric surgical care in these settings but also emphasizes the potential for significant health outcomes improvement through targeted interventions. As the population of children in L-MICs continues to grow, the development and implementation of capacity assessment tools (CATs) for pediatric surgery become increasingly essential to effectively address these critical healthcare gaps.

Acknowledging the stark demographic challenges and the high prevalence of surgically treatable conditions among children in L-MICs, it's essential to also consider the broader implications of these surgical conditions on the overall disease burden. Research conducted by Smith et al. shines a light on this aspect by evaluating the impact of delayed surgical access on pediatric disease burden in Somaliland[7]. Their findings reveal that the mean averted disability-adjusted life years (DALYs) per child amounted to 8.4 across all conditions studied, with this figure surging to 22.1 when focusing solely on general surgery procedures[7]. This comparison becomes even more striking when juxtaposed with interventions in other health domains; for instance, antiretroviral therapy to prevent mother-to-child transmission of HIV is estimated to avert 8.6 DALYs compared to no treatment, underscoring the profound potential impact of enhanced surgical care[7].

The implications of these findings are twofold. Firstly, they quantify the significant health benefits that can be achieved through timely surgical interventions, offering a compelling case for the prioritization of surgical capacity development in L-MICs. Secondly, by comparing the averted DALYs associated with surgical interventions to those achieved through established public health interventions (like antiretroviral therapy), Smith et al. demonstrated the relative value of surgical interventions and highlight how integrating surgical care into broader health system strengthening efforts in L-MICs could have a transformative effect on public health outcomes and be cost efficient[7, 8].

The potential for surgical interventions to reduce pediatric surgical disease burden in L-MICs, is mirrored in the increasing engagement in international pediatric surgical partnerships aimed at addressing surgically treatable conditions in children within these regions[9]. A scoping review focusing on such partnerships in Sub-Saharan Africa (SSA) identified cardiac surgery and general pediatric surgery as the fields with the highest volume of published collaborative efforts[9]. Partnerships most frequently involved HICs, notably the United States of America, the United Kingdom, France, and Canada, collaborating with L-MICs – Kenya and Ethiopia being the most frequently cited[9].

Global Surgical Partnership Models

Numerous different models of partnerships have evolved between HIC and L-MICs[9-15]. While these models vary in specifics, they typically fall into one of four categories[13]:

1. *Collaborations between foreign medical institutions or charities and local institutions*
2. *Periodic medical missions or outreach*
3. *International health electives for surgical residents*
4. *Training of surgeons at foreign medical institutions*

In a 2014 review of international partnerships in pediatric surgery in SSA, Ekenze et al. identified 31 articles discussing these different models of programs and missions. These articles were largely descriptive and lacked objective outcome measures for the reported initiatives. The review noted a predominance of partnerships led by the United States, followed by the United Kingdom, Canada, Germany, and Italy. Common challenges highlighted included poorly defined targets, ethical concerns around supervision and patient protection, inadequate audits, misallocation of already scarce resources, and insufficient follow-up. Additionally, assessing long-term outcomes was problematic due to the ambiguous duration of missions, lack of sustainability, and challenges in defining measurable outcomes. The reviewed articles, consistent with other reports on medical missions and partnerships, suggested a limited impact on long-term outcomes. The authors suggested that effective planning should address objective definition, target setting, ethical considerations, funding, and exit strategies. Furthermore, they highlighted the importance of actively involving the host institution to provide contextual expertise, facilitating a better understanding of the local disease burden, healthcare systems, and societal and cultural complexities [16].

In recent years, the strategy of global health efforts has evolved from short-term initiatives with limited long-term impact towards more sustainable models that prioritize local community needs[17]. A systematic review by Woods et al. aimed to categorize and critically evaluate past partnerships to support future capacity-building efforts [9]. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews (PRISMA-ScR) guidelines, the authors reviewed articles published in multiple languages between January 2009 and December 2019[9]. Their analysis included 63 articles, and they defined 18 distinct categories of partnership models[9].

1. *Academic Partnerships:* Partnership between two universities; a university and an academic teaching hospital; or an academic society and a university or academic teaching hospital.
    - Focused on pediatric surgical education for surgical trainees and practicing surgeons in SSA, along with supporting other activities such as research.
2. *Fellow Exchange Program:* Pediatric surgery fellows from a country outside of SSA and pediatric surgery fellows from a country in SSA switch places and train at the other fellows' institution for a short period of time (four to six weeks).
    - The goal is for a mutually beneficial educational experience.
3. *Government Partnership:* Partnership between the government of a country in SSA and a Non-Governmental Organization (NGO), hospital, or university in a

foreign country. Alternatively, a partnership between a foreign country's government and an NGO, hospital, or university in SSA.

    - The government provides resources with the aim of scaling up pediatric surgical care in the country in SSA.

4. *International Hospital Partnership:* A partnership between a hospital in SSA and a hospital or NGO in another country with the aim of supporting pediatric surgical activity at the hospital in SSA.
    - NGO's provide financial support or coordinate follow-up & community health care for patients. Foreign hospitals send teams of pediatric surgical providers to the hospitals in SSA and provide opportunities for providers from SSA to receive pediatric surgical training at foreign institutions.
5. *International Patient Transfer:* Pediatric patients from a country in SSA are transferred to a foreign country for complete surgical care.
    - This is either the only means for local patients to receive any surgical care, or for complex patients requiring specialty surgical care that is not available in the patients' country of origin.
6. *International Surgery Rotation:* Medical students, general surgery residents, or pediatric surgery fellows from a country outside SSA complete clinical rotations at a partner institution in SSA and can receive elective credit for the rotation.
7. *International Surgical Society/Symposium:* An international society of pediatric surgeons from multiple countries around the world (including SSA) that organizes regular meetings and activities, or a symposium meeting in which pediatric surgeons from around the world (including SSA) gather without being organized into a formal society.
8. *Long-term Surgical Outreach:* Surgical activity performed by a partner country organization in a host country in SSA for a continuous period longer than one year.
9. *Long-term Surgical Volunteers:* A single expatriate pediatric surgeon practices surgery in a country in SSA for a continuous period longer than one year.
10. *Military Surgical Mission:* Surgery outreach performed by foreign military personnel in a country in SSA for a defined period.

- Must include operations on local civilians in addition to war-related surgeries

11. *Non-Governmental Organization Partnership:* Partnership between an NGO and a hospital, healthcare system, NGO, or university.
    - The NGO may be based in a country in SSA or a country outside of SSA.
12. *Residency/Fellowship Training Program:* A general surgery residency program that includes training in pediatric surgery or a pediatric surgery fellowship program in SSA developed in collaboration with surgical faculty and/or organizations in a country outside SSA.
13. *Short-term Rotating Teams:* Visiting faculty and healthcare personnel go to a single facility in SSA for a short period of time (one or more weeks) on a rotational schedule so that visiting team members are present continuously throughout the year.
    - This can either be for education (visiting surgical faculty) or to have a continuous presence of foreign healthcare providers present to perform surgical care.
14. *Short-term Surgical Trip:* Surgeons and healthcare personnel from a country outside of SSA travel to a country in SSA to provide surgery for local patients and/or education to local healthcare providers for relatively short period of time (less than one year).
15. *Surgical Camp:* A team from the partner country and a team from the host country go on a joint surgical outreach to a defined area of need within the host country to perform operations over a short period of time (one to two weeks).
    - May or may not focus on a single surgical condition or procedure.
16. *Surgical Ships:* Fully equipped hospital ship visits countries in SSA and provides surgery to the local population.
17. *Surgical Simulation Workshop:* A training experience that teaches providers from SSA how to perform a specific procedure using dummies and/or simulation software.
    - Workshops may include didactic lectures on surgical content.
18. *Telemedicine:* Collaboration between surgeons in SSA and in a foreign country

by transmission of patient information through email with consultations between surgeons over the phone, email, or video, or live, joint assessments of surgical patients via video conference.

The most common partnership model identified was short-term surgical trips, documented in 29 articles (46%), followed by partnerships with NGOs, mentioned in 28 articles (44%). Of the articles reviewed, 36 (57%) included components of education or training for the host country. Short-term trips varied in duration from three days to three months, with 10 days and two weeks being the most common lengths. Among the 48 articles where outcomes were deemed appropriate, 21 (45%) reported including post-operative follow-up care, and 31 (66%) documented post-operative patient outcomes, primarily focusing on mortality (71%) and post-operative complications (55%). Data indicate that surgeries conducted through humanitarian platforms may experience higher complication rates and poorer clinical outcomes compared to those performed in HIC hospitals [18]. A significant shortcoming identified was the lack of follow-up, which future partnerships in SSA must address, although the authors did not provide an objective method for measuring outcomes. Of the 36 articles that incorporated educational activities within their partnerships, 19 (53%) reported some form of outcome measure. The most frequently cited outcome, found in only seven articles, was the ability of surgeons and trainees in the host country to perform surgeries independently. To ensure the effectiveness of training in international partnerships, it is crucial to design and report specific educational outcome measures. The independent surgical capability of host country providers should serve as a key benchmark for all partnership efforts [9].

Recommended practices for short-term medical missions have previously been published by Roche et al and are detailed in Table 1 [19].

Table 1. Recommended Practices for Short-Term Medical Missions

WHO health system building block	*Subcategory*	*Recommendation*
Service delivery	Patient safety and quality of care	Uphold the standard of care
		Practice within your skill set and provide appropriate supervision of trainees

		Develop an ethical framework
		Conduct patient selection in a particular way
		Use interpreters
		Establish emergency plans, limit the number of patients the VMT sees or hours they work, or operate on patients in a particular order
	Patient autonomy	Get informed consent from patients
		Consider cultural and/or language barriers to informed consent
	Continuity of care	Consider the feasibility of follow-up care after the VMT's departure
		Send patients requiring additional care to the local healthcare system
	Minimizing impact on regular service delivery	Avoid disrupting the regular flow of service delivery
	Integration with local health services and sustainability	Consider the sustainability of STMM work
		Integrate/coordinate STMM care with existing health services
		Use local health services when possible (e.g. for imaging studies)
Human resources	–	Include local healthcare providers in the STMM
Health information	–	Leave behind documentation of the care the patient received
		Acquire patients' existing health records, if possible
Medical products and technology	–	Donate supplies and equipment in a particular way
		Consider care continuity needs related to medicines and technologies

		Distribute/dispose of medications in a particular way
Health financing	–	Consider patient costs
		Do not burden host institutions with costs
Leadership and governance	Collaborative partnerships	Engage in collaborative goal setting with in-country stakeholders
	Obtaining legal authorization	Research local laws on practicing medicine and/or register with appropriate health authorities
	Other accountability measures	Assess patient outcomes
		Recommends a specific accountability measure for STMMs
Other recommendations	Pre-mission planning	Consider VMT members' personal safety, immunizations, passport and visa requirements, what to pack, and local transport options
		Select site based on certain criteria and/or conduct a pre-mission site visit/needs assessment
		Partner with an in-country organization, such as an NGO
	VMT characteristics	Research the destination country and/or undergo pre-departure training
		Respect local culture/deliver culturally sensitive care
		Select VMT members based on certain criteria
Table derived from Roche et al *Table 2* [19].		

Despite the apparent enthusiasm for and investment in these partnerships, there is a conspicuous gap in the literature concerning their objectives and the efficacy of resource management within these collaborative frameworks. To date, there have been no comprehensive studies investigating the optimization of resource allocation in these cross-border surgical partnerships, nor have there been objective tools developed explicitly to quantitatively measure the improvements and outcomes of such endeavors. This lack of evaluative mechanisms and objective measures poses a significant challenge, especially given the substantial financial and manpower investments involved in establishing and maintaining these international collaborations. Without clear metrics for assessing the impact of these partnerships, the financial and temporal investments remain difficult to justify, highlighting an urgent need for the development of objective assessment tools to gauge the effectiveness and value of these vital international health initiatives.

Established in 2016, the Global Initiative for Children's Surgery (GICS) united the experience of surgical and anesthetic care providers from L-MICs with the resources and capabilities of HIC partners[10, 20]. Collaborative efforts involving academic pediatric surgeons from Nigeria, Canada, the United States, and the United Kingdom set the stage for a global dialogue dedicated to enhancing the accessibility and quality of safe surgical care for children in low-resource settings. GICS aspired to ensure that every child worldwide with a surgical need had access to the necessary resources to optimize their care. To this end, GICS defined and promoted essential resources for pediatric surgery in under-resourced areas, engaging a broad spectrum of surgical care providers. To achieve this goal the initiative forged alliances with multiple professional organizations, including the American Society of Anesthesiologists (ASA), the American Pediatric Surgical Association (APSA), the British Association of Paediatric Surgeons (BAPS), the Canadian Association of Paediatric Surgeons (CAPS), the Pan-African Pediatric Surgical Association (PAPSA), the World Federation of Societies of Anaesthesiologists (WFSA), the Canadian Anesthesiologists' Society (CAS), and the International Society for Pediatric Neurosurgery (ISPN). Since its inception, GICS has grown into one of the largest global coalitions of children's surgical providers, boasting over 260 members from more than 40 countries. The initiative's founding principles and objectives were designed to be L-MIC centric and flexible, responding to the dynamic nature of global surgery [10, 20, 21]:

1. Honest, open, respectful, and unbiased delivery of our shared goals
2. Diverse, inclusive, and equitable structure and membership in a world without borders.
3. Treat all people in a way that demonstrates their high value as individuals, regardless of colour, race, ethnicity, gender, background, disability, sexuality, politics, or beliefs.
4. Advocate for the basic human right to health for all children and families worldwide and for the essential role of surgical, anesthetic, and nursing care, using a sustainable

approach.

5. Catalyze, advocate for, and promote increased commitment and investment in all aspects of children's surgical, nursing, and anesthetic care, particularly in areas with limited resources.
6. Elevate the voices of children and families affected by children's surgical conditions, especially those who are disproportionately affected in L-MIC, in conflicts and in disaster areas.
7. Alleviate health inequity by reducing disparities and supporting provider of all aspects of children's surgical, anesthetic, nursing, and perioperative care, especially those challenged by limited resources in L-MICs.

These guidelines were developed through priority setting and working group discussions at the first GICS meeting in London, UK in 2016. Once challenges were identified, a second meeting took place in Washington, USA and a third in Valore, India, in 2016 and 2018 respectively, facilitating in-person interactions between HIC and L-MIC participants. Infrastructure, service delivery, training, and research emerged as the major thematic areas. The Optimal Resources for Children's Surgery (OReCS) document, which describes service delivery by hospital level, and the Global Assessment for Pediatric Surgery (GAPS), the subject of this thesis, aimed at evaluating existing gaps in pediatric surgical capacity, were identified as landmark projects to address these themes[10, 22-29].

Classically, north-south collaborations were dominated by the HIC partner through funding, education, and initiation of the partnership. More recently organizations such as GICS and the Global Pediatric Surgical Network (GPSN) have provided a formal context for HIC and L-MIC partners to set common goals for partnerships aimed at addressing the pertinent difficulties of the LMIC partner and working together to achieve predefined objectives[10, 11, 20, 23, 30]. A survey published in 2020 distributed to surgeons from L-MICs assessed the nature and perception of these global surgical collaborations and their benefits[12]. Clinical and educational partnerships were ranked most important by respondents, although they were simultaneously also thought to be least sustainable. It is important to note that the article did not provide a definition of sustainable. Over three quarters of respondents reported expressing their needs before HIC arrival, but less than half of HIC partners understood the overall environment well upon arrival in L-MICs. There was no mention of the use of needs assessment tools, or any objective measures used to identify needs. Almost all survey respondents felt that a formal process of collaboration and a structured partnership would benefit all parties, specifically noting that a pre-arrival checklist would be beneficial for HIC institutions and surgeons to understand the reality of an L-MIC environment. When L-MIC partners were asked about the benefits of partnerships, the provision of otherwise unavailable medical care was reported by only 21% of respondents. Most respondents reported benefits in opportunities for trainee education (95%), forums for faculty research or educational exchange/collaboration (63%), and access to the latest technology and techniques (53%). When asked what respondents believe the largest benefits for their HIC partners are, the opportunity to enrich trainee education was cited as number one (68%), closely followed by giving back to the global community (58%), and recognition in academic and professional circles (63%). The authors underscored the importance of developing a tool that would initiate communication and critically evaluate individual needs and encourage transparent communication. They further emphasized the importance of longitudinal follow-up and continued reflection on the status of the partnership, obstacles, and goals. However strong the partnership, success cannot be achieved without clear outcome measures to serve as a predefined definition of success. The end goal is a self-propagating relationship that can withstand challenges, with the long-term partnership structure outlasting any of the individuals involved[12].

Recognizing the challenges outlined above, particularly the need for objective measures to evaluate the effectiveness of international pediatric surgical partnerships, we embarked on the development of the GAPS. This tool was envisaged as an objective, validated capacity assessment tool specifically designed to accurately gauge pediatric surgical capacity in low-resource settings. It aims to empower stakeholders in global pediatric surgery with the ability to precisely identify their needs, thereby facilitating the acquisition and optimal utilization of focused assistance, be it in the form of resources, manpower, or funding.

To construct a capacity assessment instrument tailored for low-resource settings that provide pediatric surgical services, it was imperative to employ a surrogate marker of outcomes. This necessity arose from the paucity of available institutional outcome data or validated benchmarking procedures for children's surgery. Consequently, the hospital level was utilized as a surrogate indicator of its capacity to manage pediatric surgical cases. Levels of healthcare facilities were based on the "Essential Surgery" volume of Disease Control Priorities, 3rd Edition

(DCP3) classification and the GICS OReCS [10, 20, 24-26, 31-33]. They include 1st level hospital, 2nd level hospital, 3rd level hospital and national children's hospital, as pre-defined by respective countries governmental agencies.

Our intent with GAPS was to create a comprehensive scoring system accompanied by a personalized recommendation report, highlighting potential areas for enhancement. By doing so, GAPS would serve not only as a critical instrument for establishing clear, long-term partnerships aimed at elevating the quality of surgical care but also as a pioneering model for the development of analogous tools across various surgical subspecialties. Additionally, GAPS is intended to function as a dynamic quality maintenance tool. It will enable the monitoring of progress and advancements achieved through North-South partnerships, offering a platform for ongoing reassessment and recalibration of these collaborations. This approach is designed to ensure that partnerships remain acutely aligned with the evolving, unbiased needs of low-resource centers, thereby maximizing the impact of global efforts to improve pediatric surgical care.

Capacity Assessment Tool

Developing a healthcare tool involves a systematic process that transforms conceptual health constructs into measurable instruments. The process undertaken for the development of the GAPS was based on two main sources: DeVellis et al's, *Scale development: theory and applications* and Streiner et al's, *Health measurement scales*[34, 35]. After careful reflection, the methodological process undertaken for GAPS included: devising items, scaling responses, selecting items, refining items, and evaluating the performance of the items [34, 35].

Devising The Items

Devising the items is the first step involved in the creation of a capacity assessment tool. The goal is to generate a pool of items that represent the construct of interest. In this case the construct of interest was the ability to undertake pediatric surgical care in low-resource centers. Because this is a vague construct we choose level of center designation, basic surgical care (BSC) or advanced surgical care (ASC) as a surrogate for providing pediatric surgical care in low-resource settings.

This began with a conceptualization of what construct the tool is intended to measure. It required a thorough literature review to understand the theoretical underpinnings of the construct, in this case the construct was the ability to deliver safe pediatric surgical care. The goal was to ensure content validity – the degree to which the items of a tool are representative of the construct to be measured.

To devise the items in GAPS, we conducted a systematic review of current capacity assessment tool (Manuscript 1). This literature review allowed us to identify previously used items and scales as well as theoretical models of the construct and misgivings in existing tools. The findings of the systematic reviews guided the item generation process by ensuring that items covered all relevant aspects of the construct, adhering to a coherent theoretical rationale. Coupled with expert consultation, the list of items was refined. In our cases we opted to conduct a small expert consultation panel of eight stakeholders to assess item interpretation and relevance. Experts provided insight into the construct's most critical aspects, suggested additional items, and ensured the items were understandable and relevant. Item wording and format must be clear and concise with the formatting adapted to the item in question. The culmination of these efforts is the development of an initial pool of items that is deliberately larger than the intended final scale. This pool should reflect the construct's breadth, incorporating a diverse range of items that cover all identified dimensions of the construct. The goal at this stage was to be inclusive, with the understanding that subsequent phases will refine this initial item pool through empirical testing and analysis.

Devising items is a critical and complex phase in the development of a healthcare tool. It blends theoretical insight with empirical evidence and stakeholder engagement to create a robust item pool that accurately represents the construct of interest. This meticulous process lays the foundation for the subsequent steps of scale development, aiming for a tool that is both scientifically rigorous and practically useful in healthcare settings.

Scaling Responses

Scaling responses is a determinant factor in how participants interact with the questionnaire. This step involves deciding on the format and nature of responses for each item in the tool and ensuring that the scale effectively captures the gradations of the construct being measured. Effective scaling transforms qualitative aspects of health-related constructs into quantifiable data, enabling precise analysis and interpretation. The choice of response format is influenced by the construct's nature, the item's content, and the target population's characteristics. Common format includes Likert scales, visual analog scales, nominal scales, or ordinal scales. Likert scales are often used for attitudes, beliefs, and subjective evaluations, offering options like "strongly disagree" to "strongly agree." They are valuable for capturing the intensity of respondents' feelings or agreement levels. Visual analog scales are most often used for measuring subjective experiences with continuous variables, such as pain intensity. Nominal scales are appropriate for categorical data without inherent order, such as types of healthcare services used. Ordinal scales capture ordered categories, such as frequency of events, without assuming equal intervals between categories.

Furthermore, the scale should be sensitive enough to detect nuances within the construct without overwhelming respondents. A balance between sensitivity and burden is essential to ensure high-quality data while maintaining respondent engagement. The number of response options depends on the desired precision and the respondents' ability to discriminate between options. More options can increase sensitivity but may also add complexity for respondents. Providing clear, descriptive labels for each point on the scale, termed anchoring labels, can aid in consistency of responses. These labels can be numerical, verbal, or both, tailored to the construct and population. In addition, neutral options can be included for questions where lack of knowledge is informative. Pretesting the chosen scaling method within an expert panel or target population is crucial and can reveal potential issues with the scale interpretation or item ambiguity.

Scaling responses is a nuanced process that directly impacts a healthcare tool's ability to yield meaningful, actionable data. It requires careful consideration of the construct, the target population, and the research objectives. By thoughtfully designing response options and scales, developers can enhance the precision, reliability, and validity of health measurement tools, facilitating their utility in both clinical and research settings.

Refining The Items

Refining the items is a critical stage in the development of a healthcare tool, where the initial pool of items generated is scrutinized, tested, and adjusted to ensure each item contributes effectively to the tool's overall validity and reliability. This phase aims to polish and hone the questionnaire, making it a precise instrument for measuring the construct of interest. Refining items involves a series of steps designed to evaluate and improve the items based on empirical data and expert judgment.

The first step of refining the items consist of a thorough review of each item for clarity, relevance, and conciseness. To ensure that the item is understandable by the target population, language is simplified. Each item is evaluated to ensure that it is essential for capturing the construct. Items that are redundant and may contribute to respondent fatigue without adding value are removed.

Pilot testing with a small, representative sample from the target population is a useful step in the refinement process. It helps identify issues not evident during the item generation process. Pilot testing can reveal misinterpretation of item wording or response options, item bias where certain items may not be equally applicable across different subgroups within the population, and length issues to ensure the questionnaire is not too burdensome for respondents.

Reliability

Empirical data collected during pilot testing are analyzed to inform further item refinement. Item analysis examines how well an item correlates with a predefined factor; suggesting how closely aligned each item is with the construct being measured. Internal consistency and reliability are classically measured by Cronbach's alpha to assesses how well the items measure a single, unified construct. Items that detract from the scale's reliability may need to be revised or discarded. While Cronbach's alpha has been traditionally used to assess internal consistency, Item Response Theory (IRT) offers a sophisticated alternative that can provide deeper insights into item performance and the underlying construct.

Unlike classical reliability approaches, IRT pays attention to the properties of the individual items composing a scale. In classical measurement theory, a scale's reliability is often increased by redundancy—adding more items. Reliability is enhanced in IRT approaches not by redundancy but (where possible) by identifying better items. IRT approaches view reliability as fundamentally about the individual items, whereas classical approaches tend to view it from the perspective of the scale as a whole. IRT explicitly examines what level of the attribute being measured most strongly influences an item. When applied to multi-response items, IRT is conducted through the framework of a Graded Response Model (GRM). This model provides information about how each of the multiple-response options relates to ability. For a good item, each response option should occupy a more-or-less distinct portion of the ability continuum. IRT models can generate Item Characteristic Curves (ICCs), which plot the probability of endorsing an item across various levels of the latent trait. Examining ICCs helps in assessing whether items discriminate well between individuals with different levels of the trait. Items with flat curves do not discriminate well and may be candidates for removal or revision. "Item Discrimination" reflects how well an item can distinguish between respondents with slightly different levels of the latent trait. High discrimination values indicate that an item is effective in differentiating among levels of the trait. Items with low discrimination may be too easy or too hard for most respondents, contributing little to the measurement of the construct. A discrimination value of 1.5 indicates that the item is moderately good at discriminating between different levels of the latent trait. The higher the discrimination parameter, the steeper the slope of the ICC, and the more sensitive the item is at distinguishing between items with slightly different levels of the latent trait. In the context of IRT, "Item Difficulty" refers to the location on the latent trait where the item has a 50% chance of being endorsed. This parameter helps identify items that are appropriately challenging for the target population. Items that are too difficult or too easy for the majority of the sample might not provide useful information about the construct and could be refined or excluded. Moreover, IRT models can handle missing responses more flexibly, making them suitable for scales with optional items or in situations where not all respondents answer all items.

Furthermore, iterative feedback from expert review provides insight into the content validity and theoretical alignment of items. Respondent feedback, collected during the pilot testing phase, offers direct input from the target population on item clarity, relevance, and sensitivity. After iterative revisions, the item pool undergoes a final review to ensure that the tool accurately represents the construct, covering all necessary dimensions without unnecessary duplication. The tool should also balance comprehensiveness with brevity, to maintain respondent engagement without sacrificing the depth of measurement. Moreover, the tool must meet the anticipated analytical needs, ensuring that the scale will support the intended statistical analyses for research or clinical application.

Test-Retest Reliability and Inter-Rater Reliability

Test-Rest Reliability assesses the stability of a tool over time by administering the same scale to the same respondents at multiple different time points. If there is a strong correlation between the sets of responses, the tool is thought to have good temporal stability.

Inter-rater reliability measures the extent to which different raters give consistent estimates of the same phenomenon.

Validity

<u>Content Validity</u>

Content validity ensures that the scale covers all relevant facets of the construct of interest. This is usually defined by the theoretical framework and the literature surrounding the area of study. Therefore, content validity often relies on expert consensus and opinions about the content of the items. This approach focuses solely on whether the items are appropriate and comprehensive according to the experts.

<u>Construct Validity</u>

Involves testing theoretical predictions about how the measure should relate to other variables and constructs. It is concerned with whether a test accurately measures the theoretical construct it claims to measure. Construct validation is notoriously complex to achieve and requires an ongoing process since it assesses both the theory and the measures at the same time. Cronbach and Meehl's seminal article on construct validation described three steps: (1) describing a set of theoretical concepts and their relationship to one another, (2) developing scales to measure the hypothetical constructs, and (3) testing the relationships among the constructs[36]. Construct validation by extreme groups, or *discriminative validation*, is done by giving an instrument to two groups, one with the trait under study and the other without. The assumption is that the group with the studied trait should score higher than the group without on the scale being tested. This can only be done if there is a predetermined truth, or 'gold-standard' that has already been established. Thereafter, the theory of 'bootstrapping' can be employed and if the new scale gives more accurate predictions or explains further findings, or achieve better inter-rater observer agreement, it can replace the original criterion. Other methods to assess construct validation are through *convergent validation*; the new index of study should correlate with other measures of the same construct. The construct should not correlate with dissimilar unrelated ones referred to as *divergent validation*. Both *convergent validation* and *divergent validation* are encompassed in *criterion validity* and discussed in the next section (Criterion Validity) [34, 35]

Criterion Validity

Criterion validity involves testing theoretical predictions about how the measure should relate to other variables and constructs. Techniques include factor analysis to examine the underlying structure of the scale and hypothesis testing for *convergent* and *divergent validity*.[34, 35]

Convergent validity refers to the degree to which two measures that are supposed to be measuring the same construct are related. High convergent validity implies that different methods of measuring the same concept yield similar results. *Divergent* (or discriminant) *validity* is the degree to which a measure does not correlate strongly with other measures that are theoretically different. High divergent validity ensures that a measure is not mistakenly capturing unrelated constructs. In essence, convergent validity ensures similarity between related measures, whereas divergent validity confirms the distinction between dissimilar constructs.[34, 35]

CHAPTER 3: A REVIEW OF CAPACITY ASSESSMENT TOOLS

PREFACE

As detailed in the introduction, the field of global health has traditionally focused on addressing individual diseases, usually communicable diseases, in L-MICs, with limited emphasis on the organization and delivery of surgical and anesthesia care[1]. The LCoGS identified six core surgical indications suitable for monitoring and evaluation at delivery sites in order to reduce the global burden of surgical disease: access to timely essential surgery, specialist surgical workforce density, surgical volume, peri-operative mortality, and protection against impoverishing and catastrophic expenditures for surgical care[1]. Efforts to strengthen the quality of global surgical care, including specifically global pediatric surgery, have resulted in a proliferation of partnerships between HIC and L-MIC institutions[14, 30, 37-40]. However, specific metrics to develop, guide, and evaluate the efficacy of these partnerships have not been clearly defined[38].

Several tools have been developed in HICs to identify and monitor performance indicators for quality improvement and maintenance in surgical care. In North America, the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) has been particularly effective in tracking surgical outcomes, including a pediatric component [37, 38]. The World Health Organization created the Tool for Situational Analysis to Assess Emergency and Essential Surgical Care (WHO TSA) [39]. Rhee et al. have utilized the Agency for Healthcare Research and Quality's Pediatric Indicators to define resources and stratify care [40, 41]. In L-MICs, tools specifically designed for quality assessment, such as the Surgeons OverSeas Assessment of Surgical Needs: Personnel, Infrastructure, Procedures, Equipment, and Supplies (SOSAS PIPES), have been employed. This tool has also been adapted for pediatric surgery, resulting in Pediatric Surgery Personnel, Infrastructure, Procedures, Equipment, and Supplies (Pedi-PIPES) [42, 43].

Following guidelines outlined by DeVellis and Streiner’s for development of health measurement tools, a thorough literature review was the first step in tool development[34, 35]. We opted to proceed with a systematic review to identify and evaluate the already existing quality assessment tools for pediatric surgery. The aim of this systematic review was to critically evaluate CATs and quality indicator tools, particularly surveys or questionnaires, designed to assess the needs and quality of pediatric surgery in institutions with established pediatric surgical capacity.

The development of the Global Assessment in Pediatric Surgery, as outlined in the following chapters, represents a critical and much-needed advancement in addressing the current gap in monitoring the outcomes of global surgical partnerships. In the absence of objective assessment tools, the evaluation of these partnerships remains largely reliant on subjective judgments. Such subjective evaluations are often fraught with complex ethical dilemmas, financial conflicts, academic ambiguities, and, in some cases, colonial undertones [30, 39-45]. Establishing a standardized, objective approach is essential to mitigate these issues and provide a more reliable measure of the impact of global surgical collaborations.

The findings of Manuscript 1 were presented at the World Federation of Associations of Pediatric Surgeons (WOFAPS) in October 2016 in Washington D.C., USA

Manuscript 1 has been published in print under the following reference[28]:
Yousef Y, St-Louis E, Baird R, Smith ER, Guadagno E, St-Vil D, Poenaru D. A systematic review of capacity assessment tools in pediatric surgery: Global Assessment in Pediatric Surgery (GAPS) Phase I. Journal of Pediatric Surgery. 2019 Apr;54(4):831-837. doi: 10.1016/j.jpedsurg.2018.11.005.

MANUSCRIPT 1: A SYSTEMATIC REVIEW OF CAPACITY ASSESSMENT TOOLS IN PEDIATRIC SURGERY: GLOBAL ASSESSMENT IN PEDIATRIC SURGERY (GAPS) PHASE I.

Yasmine Yousef[1,2], Etienne St-Louis[1], Robert Baird[1], Emily R. Smith[3], Elena Guadagno[1], Dickens St-Vil[2], Dan Poenaru[1]

[1] McGill University Health Center, Division of Pediatric General and Thoracic Surgery, Montreal Children's Hospital

[2] Centre Hospitalier de l'Université de Montréal, Hôpital Sainte-Justine, Département de chirurgie générale pédiatrique

[3] Duke Global Health Institute

Keywords: pediatric surgery, global surgery, capacity assessment, quality improvement, health-systems, systematic review

Funding :

Fonds de Recherche du Québec - Santé (FRQS)

Canadian Institutes of Health Research

Acknowledgements:

We would like to thank Dr. Sherif Emil, Dr. Jean-Martin Laberge, Dr. Pramod Puligandla, Dr. Kenneth Shaw and Dr. Robin Petroze for their input with this work.

Abstract

<u>Background</u>

The Lancet Commission on Global Surgery highlighted global surgical need but offered little insight into the specific surgical challenges of children in low-resource settings. Efforts to strengthen the quality of global pediatric surgical care have resulted in a proliferation of partnerships between low-and middle-income countries (LMICs) and high-income countries (HICs). Standardized tools able to reliably measure gaps in delivery and quality of care are important aids for these partnerships. We undertook a systematic review (SR) of capacity assessment tools (CATs) focused on needs assessment in pediatric surgery.

Methods

A comprehensive search strategy of multiple electronic databases was conducted per PRISMA guidelines without linguistic or temporal restrictions. CATs were selected according to pre-defined inclusion criteria. Articles were assessed by two independent reviewers. Methodological quality of studies was appraised using the COSMIN checklist with 4-point scale.

Results

The search strategy generated 16,641 original publications, of which three CATs were deemed eligible. Eligible tools were either excessively detailed or oversimplified. None used weighted scores to identify finer granularity between institutions. No CATs comprehensively included measures of resources, outcomes, accessibility/impact and training.

Discussion

The results of this study identify the need for a CAT capable of objectively measuring key aspects of surgical capacity and performance in a weighted tool designed for pediatric surgical centers in LMICs.

Type of Study: Systematic Review

Keywords: Pediatric surgery, global surgery, capacity assessment, quality improvement, health-systems, systematic review

Level of Evidence: II

Introduction

Global health has traditionally focused on addressing individual communicable and non-communicable (i.e. heart disease, diabetes) diseases in low-and middle-income countries (LMICs), with limited emphasis on the organization and delivery of surgical and anesthesia care [1–4]. The Lancet Commission on Global Surgery together with several other publications, reports and declarations have emphasized the significance of surgical burden of disease within global health and the need to improve the delivery of surgical care [2–6]. However, these reports made little reference to the specific and pressing challenges facing children's surgery [1]. Considering that children comprise more than 50% of the population in the least developed regions of the world, we can infer that the avertable surgical burden in children in LMICs is colossal [1,7]. Thus, developing strategies to improve the quality of pediatric surgical and trauma care in low-resource settings can significantly decrease child morbidity and mortality and promote economic development [8].

Recent efforts to strengthen capacity of global surgical care of children have resulted in a proliferation of various partnerships models between LMICs and high-income countries (HICs) in almost all sub-specialties of children's surgery [9–16]. However, the quality of these partnerships is highly variable, and the goals are often unclear or unspecified [9,12]. In order to guide these partnerships in identifying priorities for capacity improvement, as well as assist in the monitoring of improvements, it is critical to develop standardized capacity assessment tools that can reliably measure existing gaps in care [9].

The objective of this systematic review was to identify and assess all pediatric surgery capacity assessment tools in the aim of identifying one tool best suited to address the needs of centers providing pediatric surgery in LMICs.

Methods

We conducted a systematic review (SR) of pediatric surgical capacity assessment tools to ensure that all pediatric literature was evaluated. This SR took place between July 19, 2016 and November 1, 2016. It followed the 'Preferred reporting items for systematic reviews and meta-analysis (PRISMA)' guidelines [17] and was prospectively registered in PROSPERO (CRD42016042069).

Search Strategy

The search strategy (Appendix A) was created in collaboration with a clinical librarian (EG) and peer-reviewed by a second independent clinical librarian. The search included no language restrictions but was limited to the pediatric population and human only studies, from inception to June 13, 2016. The following databases were included: Medline (Ovid), Embase (Ovid), Cochrane (Wiley), Africa-Wide Information (Ebsco), Allied & Complementary Medicine (Ovid), Biosis (Ovid), CINAHL (Ebsco), Global Health (Ovid), Latin-American and Caribbean Center on Health Sciences Information, PubMed (NLM) and Web of Science (Thomson Reuters). The systematic literature search identified articles that included variations of the terms "pediatric surgery", combined with variations of measurement tools and global/international concepts, found as text-words in the Title/Abstract/Keyword fields as well in the Medical Subject Headings (MesH). See Appendix A for the detailed search strategy (Medline search was extrapolated to all other databases).

Inclusion Criteria

Inclusion criteria were as follows: the tool must have self-identified as a surgical capacity assessment tool, the study must have reported on institutions with surgical capacities treating children up to 18 years of age, the tool must have been comprised of questionnaires or surveys, and the study must have included as one of its objectives the development of a measurement instrument, its evaluation, or its validation. Exclusion criteria included: diagnostic or screening instruments, prognostic studies (i.e. prediction models), instruments evaluating outcomes (i.e. complications, mortality, etc.), and comprehensive surgical capacity assessments not focused on pediatric surgery or trauma.

Critical Appraisal

Two reviewers (YY, ESL), first independently assessed eligibility of studies based on titles and abstracts. A second round of review by YY and ESL analyzed the selected full-text articles and their respective reference lists for eligibility based on the inclusion criteria. Any disagreements between the reviewers regarding a study's inclusion were evaluated by a senior author (DP).

The methodological quality of the included studies was assessed using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist with 4-point scale [18–21]. The COSMIN checklist is a critical appraisal tool for evaluating the methodological quality of studies of health measurement instruments. Assessment of the methodological quality was performed independently by two reviewers (YY, ESL), and consensus was achieved through discussion whenever required.

Data Abstraction

Information was extracted from the included articles by two re- viewers in an independent fashion using a formal data extraction tool (Appendix B). If available, the following data items were extracted from all studies selected for inclusion: general characteristics of the instruments (i.e. construct, sub-scales, number of items, version, etc.), characteristics of the institutions in which the measurement properties were assessed, and results of the measurement properties. Items not available were noted. Items were separated into four categories (resources, out- come, accessibility/impact, and training) in order to provide a comprehensive overview of the different types of questions included in each tool.

The data extracted was reviewed by a senior author. Discrepancies in data extraction were reviewed by all three reviewers together and changes were made based on a consensus between the majority (two) of the reviewers. Pilot data extraction of the first 30% of full texts included was performed to standardize the process of extraction.

Results

The search strategy of the systematic review identified 19,868 records. After removal of duplicates, 16,641 titles and/or abstracts were retained for eligibility, of which 15 full text articles were reviewed in detail. A total of 12 tools did not fit the inclusion criteria and were excluded; 10 studies were deemed to be prognostic tools or predication models and 2 tools were not specific to children. After final review, 3 capacity assessment tools were included in the systematic review (PRISMA Diagram: Fig. 1). The three tools are the Pediatric Surgery Personnel, Infrastructure, Procedures, Equipment, Supplies (Pedi-PIPES); the Children's Surgical

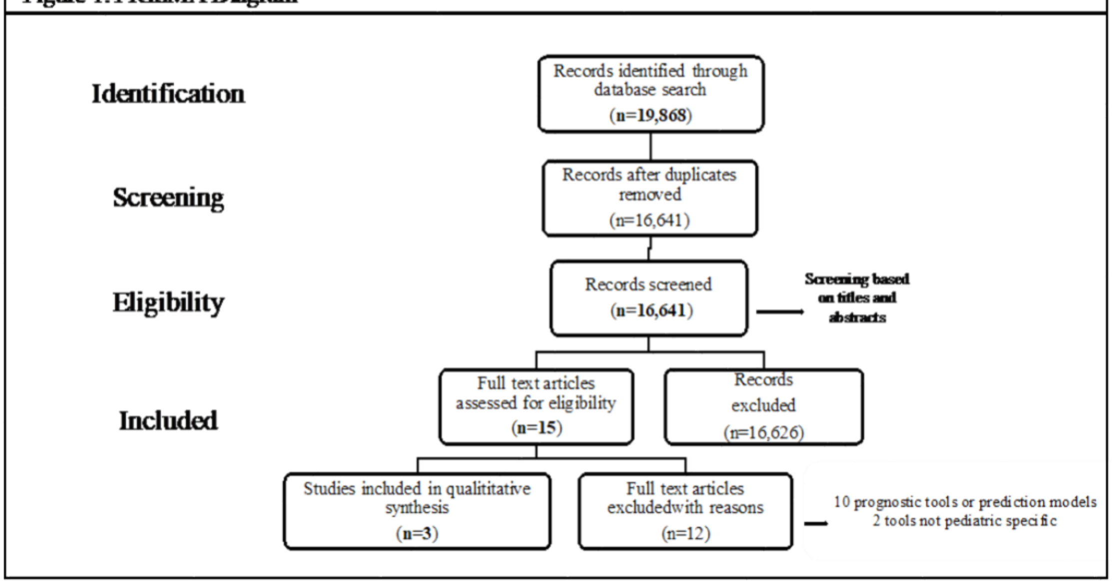
Center Designations with Scope of Practice (CSCDSP); and the Checklist for a Children's Trauma Room (CCTR).

Table 1 provides a synopsis of the reference articles and capacity assessment tools included in this SR. Both Pedi-PIPES and CSCDSP's aims were to assess pediatric surgical capacity. The goal of the CCTR was to provide a list of equipment and supplies to prevent “disastrous” omission when stocking the trauma room. The number of items in each tool varied from 27 in the CSCDSP, 118 in Pedi-PIPES, to 165 in CCTR [22–24]. Unlike the other two tools, Pedi-PIPES focused exclusively on assessment of centers in LMICs [24]. All tools were available exclusively in English [22–24]. Funding for the development of Pedi-PIPES was provided by SurgeonOverSeas (SOSAS) and the Children's Hospital Association and the American College of Surgeons for CSCDSP [24]. There were no funding details available for the CCTR [22].

All tools were created by groups working in HICs; Pedi-PIPES also had a contribution from pediatric surgeons in LMICs and was first implemented in several West African countries (i.e. Nigeria, Niger, Benin, Togo, Ghana, Burkina Faso, Côte d'Ivoire, Liberia, Sierra Leone, Guinea, Gambia, and Senegal) [22–24].

All tools were previously based on capacity assessment tools aimed for adult surgical centers [22–24]. For example, the CCTR was based on a similar tool meant for adult trauma rooms at the Parkland Hospital Emergency Department in Dallas, Texas [22]. Pedi-PIPES is a modified version of the SOSAS PIPES [24]. The SOSAS PIPES tool itself was developed by

Figure 1. PRISMA Diagram



SOSAS in an attempt to simplify the World Health Organization Tool for Situational Analysis to Assess Emergency and Essential Surgical Care (WHO TSA), and enable easy comparison between institutions and over time [25,26]. The SOSAS PIPES modifications were made based on consensus within a small group of pediatric surgeons (four American and five African pediatric surgeons) [24]. Though the Children's Hospital Association and the American College of Surgeons created the CSCDSP based on a literature review and expert consensus, the WHO TSA factored heavily in its development [23]. The method of creation of the CCTR was not detailed by the authors [22].

Neither the CSCDSP nor the CCTR included any index enabling longitudinal comparison between institutions [22,23]. Pedi-PIPES is divided into four sections (personnel, procedures, equipment and supplies) and has an associated score, calculated as follows [24]: points are allocated in each of the four sections and allotted to each data item equally, depending on whether or not the item is “always available” (1 point) or “not always available” (0 points), to yield a total score for each section [24]. The total scores for all sections are summed to yield a Total Pedi-PIPES score [24]. The PIPES-index is calculated by dividing the Total Pedi-PIPES score by the number of items (118) and then multiplied by 10 [24]. This index does not have a maximum score, and there is no score weighting [24].

Concerning the subjects addressed within each tool, all tools focused disproportionately on resources followed by accessibility [22–24]. Out- come of surgical procedures was only addressed by the CSCDSP [23]. No tool addressed training [22–24]. When comparing pediatric and neonatal specific components of the tools, neonatal items were substantially underrepresented with a maximum of 11% of items referring exclusively to neonates in the CSCDSP tool [22–24]. Moreover, only 17% of questions in Pedi-PIPES were pediatric specific. Further details are presented in Fig. 2. Advantages and limitations of each tool are detailed in Table 2.

To our knowledge, no tools have been interrogated for internal consistency, inter- and intra-rater reliability, or construct validity. All tools scored poorly on all aspects of the COSMIN Checklist with 4-point scale including: internal consistency, reliability, measurement error, structural validity, hypothesis testing, criterion validation, and responsiveness [22–24].

Table 1. Description of Reference articles and capacity assessment tools												
	Reference article	Year of creation	Cited	Language of Tool	Funding	Aim of assessment tool	Classification of economy where tool was created	Based on previously published surgical capacity assessment tools		Method of creation	Weighted index
								Adults	Children		
Pedi-PIPES	Okoye, Mekam T., et al. "A pilot survey of pediatric surgical capacity in West Africa." *World journal of surgery* 39.3 (2015): 669-676.[22]	2015	6	English	Surgeon OverSeas	Assess capacity to provide Essential and Emergency Surgical Care to children in LMICs	LMIC	SOSAS PIPES (WHO Situational Analysis Tool)	-	Expert consensus panel editing the SOSAS PIPES tool for adults	No
Children's Surgical Center Designations with Scope of Practice	Oldham, Keith T. "Optimal Resources for Children's Surgical Care." *Journal of the American College of Surgeons* 220.5 (2015): 970-971.[20]	2013	16	English	Children's Hospital Association and the American College of Surgeons	To optimize the delivery of children's surgical care and develop consensus recommendations	HIC	WHO Situational Analysis Tool	-	Literature review and expert opinion	No
Checklist for a Children's Trauma Room	Morse, TS, JA Haller, and B. Othersen. "Checklist for a Children's Trauma Room." *Journal of Trauma-Injury Infection and Critical Care* 16.10 (1976): 763-765.[29]	1976	2	English	Unknown	Prevent disastrous omissions when the trauma room is being stocked	HIC	Parkland Hospital Emergency Department (Dallas)	-	Unknown	No

Figure 2. Number data points in each CAT stratified by subject

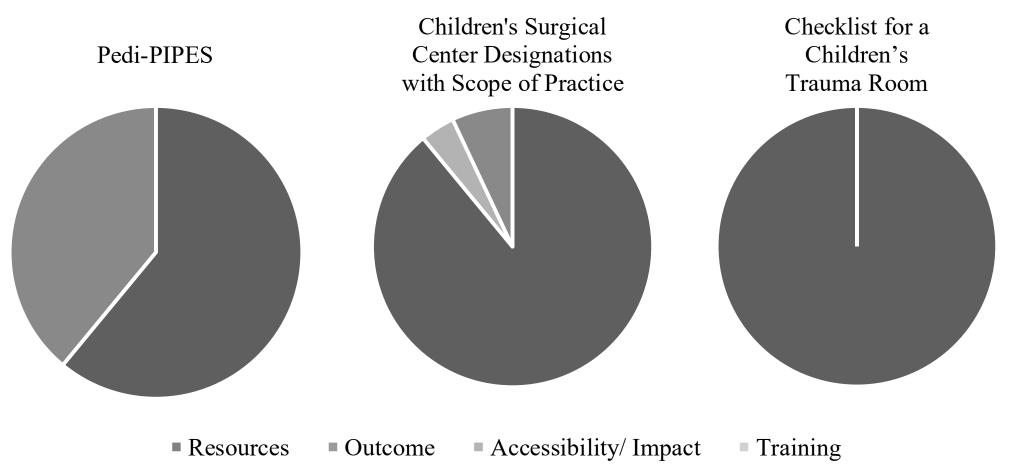


Table 2. Advantages and Disadvantages of Selected Capacity Assessment Tools		
	Advantages	**Limitations**
Pedi-PIPES	Developed for LMICs Developed for children Developed by professionals in LMICs and HICs Implemented in several west African countries Includes an associated score Addressed resources and accessibility	Available only in English Based on tool for adults Developed using expert consensus only No weighted index for comparison Previously based on capacity assessment tools aimed for adult surgical centers Did not address training Inadequate emphasis on neonatal specific questions Not interrogated for internal consistency, inter- / intra-rater reliability, or construct validity
Children's Surgical Center Designations with Scope of Practice	Developed for children Based on literature review and expert consensus Concise Addressed resources, outcomes, and accessibility	Not developed for LMICs Available only in English Developed by professionals in HICs only No weighted index for comparison Previously based on capacity assessment tools aimed for adult surgical centers Did not address training Inadequate emphasis on neonatal specific questions Not interrogated for internal consistency, inter- / intra-rater reliability, or construct validity
Checklist for a Children's Trauma Room	Developed for children Addressed resources only	Not developed for LMICs Available only in English Developed by professionals in HICs only No weighted index for comparison Previously based on capacity assessment tools aimed for adult surgical centers Did not address training No emphasis on neonatal specific questions Not interrogated for internal consistency, inter- / intra-rater reliability, or construct validity

Discussion

This systematic review identified three capacity assessment tools aimed at evaluating surgical capacity in pediatric centers in both HICs and LMICs (Table 1). There have been no published attempts to date to validate these tools. Moreover, of the three tools only Pedi-PIPES has a quantitative component [24]. However, the Pedi-PIPES index is a non-weighted sum of the tool's data points based on a dichotomous evaluation of each item, thus missing any granularity between “avail- able” and “not available” [24]. Neither the Pedi-PIPES tool, nor its index, have undergone formal validation [24]. All three tools are disproportionately focused on resource assessment with little reference to im- pact, accessibility, or outcome [22–24]. The selection of items in these tools were not based on patient outcomes and thus, no association can be drawn between these capacity assessment tools and improved surgical outcomes [22–24]. Furthermore, no tool has addressed training capacity, quality assurance measures (i.e. morbidity and mortality rounds, tumor board meetings, trauma meetings, etc.) or research endeavors [22–24]. Thus, existing pediatric capacity assessment tools all suffer from various shortcomings. Certain capacity assessment tools include an exhaustive list of all possible material resources while other tools include only the most basic resources carried by most primary health-care facilities (i.e. sutures, cautery). To avoid oversimplification or excessive detail data points need to be carefully selected to include a variety of resources covering both basic and complex care. At present, capacity assessment tools are ill-suited for the scope of pediatric surgical practice and poorly adapted to the limitations of LMICs.

Guidelines created by the American College of Surgeons, British As- sociation of Pediatric Surgery, and the Global Initiative for Children's Surgery have identified resources for pediatric surgical care stratified by the level of care facility [27–29]. However, none of these guidelines provide a quantitative method of evaluating or comparing centers.

Current literature has focused on the descriptive assessment of surgical capacity in different LMICs [7]. However, literature aimed at developing methods to improve this capacity is lacking, even though investing in surgical care for children is highly cost-effective in several cases and pro- vides societal benefits [30]. The past 10 years have seen an exponential in- crease in attempts to quantify, analyze, and scale up surgical capacity in LMICs for the general population [1,3,9,31–35]. This is evidenced by the number of tools and guidelines detailing necessary resources for optimal surgical care across pediatric and adult populations. To evaluate recent efforts in determining global surgical capacity, assessment tools in use and potential areas of study, two SRs have been published to examine the different types of adult surgical CATs and their development [36,37].

WHO TSA was the first capacity and quality assessment tool with the aim of identifying and comparing lacunae in surgical capacity [26]. Developed in 2007 by the WHO Global Initiative for Emergency and Essential Surgical Care, the WHO TSA is composed of 256 data items based mostly on resources [25,26,38]. Kwon et al. developed a method to calculate an index based on the WHO TSA tool, however, reliability of the quantitative analysis was poor due to high response variability [25,26,37,39]. The goal of SOSAS PIPES was to create a surgical CAT, simpler than the TSA, with a quantitative analytic dimension allowing com- parison between institutions as well as within the same institution over time [25]. SOSAS PIPES, the Lancet Commission on Global Surgery Humanitarian Assessment Tool, and the Harvard Humanitarian Initiative are all tools, based on the WHO TSA, used to qualitatively assess surgical capacity in LMIC [1,32,37,40–48]. Combined, these tools have been used to assess surgical capacity in over 30 countries [37]. However, none of these tools have been used to evaluate evolution of surgical capacity in an institution or success of capacity building initiatives.

The global burden of surgical disease is heavily skewed towards LMICs [1]. The burden of surgical disease in children in LMICs is difficult to assess due to poor quality epidemiological data [7]. Most of our cur- rent assessments are derived from hospital data or national estimates based on under-powered populations [7]. Children below 17 years of age make up 35–48% of the population in LMICs, suggesting that the burden of surgical disease is very significant in this population [1,7,49–51]. Though existing data may not be validated, it is believed to severely underestimate the burden of surgical disease [1,3,7]. Due to this need, partnerships between institutions in HICs and LMICs have been developed to scale up the pediatric surgical workforce. How- ever, there are no methods to adequately assess health systems and follow improvements made from partnerships [7,52].

Current tools are comprised of a checklist-style assessment of mate- rial and human resources as well as infrastructure but disregard other important issues such as, accessibility, impact, training and outcome (Table 2). Accessibility to surgical care in LMICs is a complex issue but must be accounted for in any tool aimed at detailing surgical capacity, yet only Pedi-PIPES substantially addresses the issue of accessibility [1,24]. Closely linked to accessibility is impact; referring to the effect the institution has on the population it serves (i.e. averted disability adjusted life years, catchment area, procedures performed, etc.). Education in LMICs has not been a focus of any optimal resource guidelines on children's surgery [22–24,27–29]. Resident involvement in surgical care has shown to decrease morbidity and mortality as well as serve as a valid method of increasing surgical capacity in LMICs [7,53–57]. The link between improved outcomes and the presence of post- graduate surgical programs supports the addition of training in guide- lines and tools that hope to increase surgical capacity. Outcome is arguably the most important indicator of surgical quality. Unfortunately, outcome information of children's surgery in LMICs is not readily avail- able due to poor epidemiological data and precarious record keeping practices [7]. At present no CAT reflects on or integrates outcomes of surgical care in children. Thus, there is a significant need for a tool that incorporates data points on resources, accessibility/impact, outcome and training, and features a quantitative analytical component capable of generating a weighted score for each variable.

This study serves as the basis for the development of the Global Assessment of Pediatric Surgery (GAPS), an objective, evidence-based CAT specifically designed for pediatric surgical centers in low- and middle- income countries (LMICs). Based on the findings of this systematic review, we will create an evidenced-based CAT that incorporates the advantages of the current tools while consciously addressing their limitations (Table 2). We plan to establish construct validity by proving that the items in GAPS successfully differentiate between level of care and will create a weighted tool based on outcomes. We hope the GAPS will serve as a feasible method for use in assessment of health care facilities and in prioritizing and monitoring global surgical capacity development efforts.

Conclusion

As partnerships between LMIC and HIC surgical centers continue to multiply, there is a growing need to establish these partnerships on clearly defined and articulated goals, themselves based on the objectively determined needs of the LMIC partners. Thus, to maximize the impact and effectiveness of these partnerships a comprehensive tool is needed to identify the quality of provision of pediatric surgical care in host institutions, and identify potential deficiencies. This implementation step will be the aim of the next phase of our study, the development and validation of The Global Assessment of Pediatric Surgery (GAPS). GAPS, a validated outcomes-based pediatric surgical capacity tool which will serve as an objective measure of needs, while simultaneously identifying gaps, providing the framework for increasing care capacity and monitoring improvements made through international partnerships.

Acknowledgements

We would like to thank Dr Sherif Emil, Dr Jean-Martin Laberge, Dr Pramod Puligandla, Dr Kenneth Shaw and Dr Robin Petroze for their input with this work.

Funding Sources

Dr Yasmine Yousef has received financial support for this work from the Canadian Institute of Health Research (CIHR); the 'Fondation de Recherche Québec en Santé' (FRQS); and the Montreal Children's Hospital Pediatric Surgery Research and Education Fund.

References

1. Meara JG, Leather AJ, Hagander L, Alkire BC, Alonso N, Ameh EA, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. Lancet. 2015; 386:569-624
2. Dare AJ, Grimes CE, Gillies R, Greenberg SL, Hagander L, Meara JG, et al. Global surgery: defining an emerging global health field. Lancet. 2014; 384:2245-7
3. Meara JG, Hagander L, Leather AJ. Surgery and global health: a Lancet Commission. Lancet. 2014; 383:12-3
4. Bickler SW, Spiegel D. Improving surgical care in low- and middle-income countries: a pivotal role for the World Health Organization. World J Surg. 2010; 34:386-90
5. Makasa EM. Letter to global health agency leaders on the importance of surgical indicators. Lancet. 2014; 384:1748
6. Huber B. Finding surgery's place on the global health agenda. Lancet. 2015; 385:1821-2
7. Greenberg SL, Ng-Kamstra JS, Ameh EA, Ozgediz DE, Poenaru D, Bickler SW. An investment in knowledge: Research in global pediatric surgery for the 21st century. Semin Pediatr Surg. 2016; 25:51-60
8. Peltzer K. Injury and social determinants among in-school adolescents in six African countries. Inj Prev. 2008; 14:381-8
9. Ibrahim GM, Cadotte DW, Bernstein M. A framework for the monitoring and evaluation of international surgical initiatives in low- and middle-income countries. PLoS One. 2015; 10:e0120368
10. Butler MW, Ozgediz D, Poenaru D, Ameh E, Andrawes S, Azzie G, et al. The Global Paediatric Surgery Network: a model of subspecialty collaboration within global surgery. World J Surg. 2015; 39:335-42
11. Krishnaswami S, Nwomeh BC, Ameh EA. The pediatric surgery workforce in low- and middle-income countries: problems and priorities. Semin Pediatr Surg. 2016; 25:32-42
12. Butler MW. Developing pediatric surgery in low- and middle-income countries: An evaluation of contemporary education and care delivery models. Semin Pediatr Surg. 2016; 25:43-50

13. Baird R, Poenaru D, Ganey M, Hansen E, Emil S. Partnership in fellowship: Comparative analysis of pediatric surgical training and evaluation of a fellow exchange between Canada and Kenya. J Pediatr Surg. 2016; 51:1704-10
14. Patel PB, Hoyler M, Maine R, Hughes CD, Hagander L, Meara JG. An opportunity for diagonal development in global surgery: cleft lip and palate care in resource-limited settings. Plast Surg Int. 2012; 2012:892437
15. Blair GK, Duffy D, Birabwa-Male D, Sekabira J, Reimer E, Koyle M, et al. Pediatric surgical camps as one model of global surgical partnership: a way forward. J Pediatr Surg. 2014; 49:786-90
16. Lakhoo K, Msuya D. Global health: A lasting partnership in paediatric surgery. Afr J Paediatr Surg. 2015; 12:114-8
17. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009; 62:1006-12
18. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007; 60:34-42
19. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010; 10:22
20. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010; 63:737-45
21. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010; 19:539-49
22. Morse T, JA Haller, B. Othersen. Checklist for a children's trauma room. J Trauma. 1976; 16:763-5

23. Task Force for Children's Surgical C. Optimal resources for children's surgical care in the United States. J Am Coll Surg. 2014; 218:479-87, 87 e1-4
24. Okoye MT, Ameh EA, Kushner AL, Nwomeh BC. A pilot survey of pediatric surgical capacity in West Africa. World J Surg. 2015; 39:669-76
25. Groen RS, Kamara TB, Dixon-Cole R, Kwon S, Kingham TP, Kushner AL. A tool and index to assess surgical capacity in low income countries: an initial implementation in Sierra Leone. World J Surg. 2012; 36:1970-7
26. Osen H, Chang D, Choo S, Perry H, Hesse A, Abantanga F, et al. Validation of the World Health Organization tool for situational analysis to assess emergency and essential surgical care at district hospitals in Ghana. World J Surg. 2011; 35:500-4
27. . American College of Surgeons: Optimal Resources for Children's Surgical Care American College of Surgeons, 2015
28. . Optimal Resources for Children's Surgery, Global Initiative for Children's Surgery, 2016
29. Alshehri A, Lo A, Baird R. An analysis of early nonmortality outcome prediction in esophageal atresia. Journal of pediatric surgery. 2012; 47:881-4
30. Saxton AT, Poenaru D, Ozgediz D, Ameh EA, Farmer D, Smith ER, et al. Economic Analysis of Children's Surgical Care in Low- and Middle-Income Countries: A Systematic Review and Analysis. PLoS One. 2016; 11:e0165480
31. Macpherson L, Collins M. Training responsibly to improve global surgical and anaesthesia capacity through institutional health partnerships: a case study. Trop Doct. 2016;
32. Chao TE, Burdic M, Ganjawalla K, Derbew M, Keshian C, Meara J, et al. Survey of surgery and anesthesia infrastructure in Ethiopia. World J Surg. 2012; 36:2545-53
33. Cook M, Howard BM, Yu A, Grey D, Hofmann PB, Moren AM, et al. A Consortium Approach to Surgical Education in a Developing Country: Educational Needs Assessment. JAMA Surg. 2015; 150:1074-8
34. Ozgediz D, Wang J, Jayaraman S, Ayzengart A, Jamshidi R, Lipnick M, et al. Surgical training and global health: initial results of a 5-year partnership with a surgical training program in a low-income country. Arch Surg. 2008; 143:860-5; discussion 5
35. Lipnick M, Mijumbi C, Dubowitz G, Kaggwa S, Goetz L, Mabweijano J, et al. Surgery and anesthesia capacity-building in resource-poor settings: description of an ongoing academic partnership in Uganda. World J Surg. 2013; 37:488-97

36. Carlson LC, Lin JA, Ameh EA, Mulwafu W, Donkor P, Derbew M, et al. Moving from data collection to application: a systematic literature review of surgical capacity assessments and their applications. World J Surg. 2015; 39:813-21

37. Blair KJ PL, Shaw PL, Shapiro MB, Nwomeh BC, Swaroop M. Surgical and Trauma Care in Low- and Middle-Income Countries: A Review of Capacity Assessments. Journal of Surgical Research. 2016;

38. . WHO Global Initiative for Emergency and Essential Surgical Care, 2016

39. Kwon S, Kingham TP, Kamara TB, Sherman L, Natuzzi E, Mock C, et al. Development of a surgical capacity index: opportunities for assessment and improvement. World J Surg. 2012; 36:232-9

40. Solis C, Leon P, Sanchez N, Burdic M, Johnson L, Warren H, et al. Nicaraguan surgical and anesthesia infrastructure: survey of Ministry of Health hospitals. World J Surg. 2013; 37:2109-21

41. Notrica MR, Evans FM, Knowlton LM, Kelly McQueen KA. Rwandan surgical and anesthesia infrastructure: a survey of district hospitals. World J Surg. 2011; 35:1770-80

42. Linden AF, Sekidde FS, Galukande M, Knowlton LM, Chackungal S, McQueen KA. Challenges of surgery in developing countries: a survey of surgical and anesthesia capacity in Uganda's public hospitals. World J Surg. 2012; 36:1056-65

43. Lebrun DG, Saavedra-Pozo I, Agreda-Flores F, Burdic ML, Notrica MR, McQueen KA. Surgical and anesthesia capacity in Bolivian public hospitals: results from a national hospital survey. World J Surg. 2012; 36:2559-66

44. Lebrun DG, Dhar D, Sarkar MI, Imran TM, Kazi SN, McQueen KA. Measuring global surgical disparities: a survey of surgical and anesthesia infrastructure in Bangladesh. World J Surg. 2013; 37:24-31

45. Knowlton LM, Chackungal S, Dahn B, LeBrun D, Nickerson J, McQueen K. Liberian surgical and anesthesia infrastructure: a survey of county hospitals. World J Surg. 2013; 37:721-9

46. Sun LY, Laberge JM, Yousef Y, Baird R. The Esophageal Anastomotic Stricture Index (EASI) for the management of esophageal atresia. Journal of pediatric surgery. 2015; 50:107-10

47. . Qualitative Hospital Assessment Tool The Lancet Commission on Global Surgery. Global Surgery 2030, 2015
48. Baker T, Lugazia E, Eriksen J, Mwafongo V, Irestedt L, Konrad D. Emergency and critical care services in Tanzania: a survey of ten hospitals. BMC Health Serv Res. 2013; 13:140
49. Butler EK, Tran TM, Fuller AT, Brammell A, Vissoci JR, de Andrade L, et al. Quantifying the pediatric surgical need in Uganda: results of a nationwide cross-sectional, household survey. Pediatr Surg Int. 2016; 32:1075-85
50. Butler EK, Tran TM, Nagarajan N, Canner J, Fuller AT, Kushner A, et al. Epidemiology of pediatric surgical needs in low-income countries. PLoS One. 2017; 12:e0170968
51. Smith ER, Vissoci JRN, Rocha TAH, Tran TM, Fuller AT, Butler EK, et al. Geospatial analysis of unmet pediatric surgical need in Uganda. J Pediatr Surg. 2017; 52:1691-8
52. Ozgediz D, Poenaru D. The burden of pediatric surgical conditions in low and middle income countries: a call to action. J Pediatr Surg. 2012; 47:2305-11
53. Chatterjee A, Pyfer B, Chen L, Czerniecki B, Tchou J, Fisher C. Resident and Fellow Participation in Breast Surgery: An American College of Surgeons NSQIP Clinical Outcomes Analysis. J Am Coll Surg. 2015; 221:988-94
54. Ejaz A, Spolverato G, Kim Y, Wolfgang CL, Hirose K, Weiss M, et al. The impact of resident involvement on surgical outcomes among patients undergoing hepatic and pancreatic resections. Surgery. 2015; 158:323-30
55. Ferraris VA, Harris JW, Martin JT, Saha SP, Endean ED. Impact of Residents on Surgical Outcomes in High-Complexity Procedures. J Am Coll Surg. 2016; 222:545-55
56. Kiran RP, Ahmed Ali U, Coffey JC, Vogel JD, Pokala N, Fazio VW. Impact of resident participation in surgical operations on postoperative outcomes: National Surgical Quality Improvement Program. Ann Surg. 2012; 256:469-75
57. Rickard J. Systematic Review of Postgraduate Surgical Education in Low- and Middle-Income Countries. World J Surg. 2016; 40:1324-35

Appendices

<u>Appendix A</u>

Medline Search Strategy

Medline [Ovid] (June 13, 2016)

Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

#	Searches	Results
1	exp Specialties, Surgical/	180962
2	exp Surgical Procedures, Operative/	2685912
3	su.fs.	1723842
4	(surger* or surgic* or procedure* or operati*).tw,kf.	2559499
5	Surgery Department, Hospital/	4067
6	perioperative care/	10168
7	preoperative period/	3623
8	exp Postoperative Complications/ or Postoperative Care/ or exp Postoperative Period/	537781
9	(preop* or pre-op* or presurg* or pre-surg* or perop* or periop* or peri-op* or perisurg* or peri-surg* or intraop* or intra-op* or postop* or postsurg* or (post adj2 op*)).tw,kf.	718951
10	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	4687635
11	exp pediatrics/ or exp adolescent/ or exp child/ or exp infant/ or Adolescent Medicine/	3136125
12	(newborn* or new-born* or neonat* or neo-nat* or infan* or child* or adolesc* or paediatr* or pediatr* or baby* or babies* or toddler* or kid or	2056805

	kids or boy* or girl* or juvenile* or teen* or youth* or pubescen* or preadolesc* or prepubesc* or preteen or tween).tw,kf.	
13	(paediatr* or pediatr*).jw.	505316
14	11 or 12 or 13	3767887
15	10 and 14	862085
16	st.fs.	602856
17	Data Collection/	86097
18	Checklist/	3565
19	exp "surveys and questionnaires"/	788223
20	exp Forecasting/	75736
21	exp "Outcome and Process Assessment"/	846604
22	"Quality of Health Care"/	62182
23	limit 22 to yr="1968-1979"	6349
24	exp quality indicators, health care/	15699
25	Quality Assurance, Health Care/	52034
26	exp quality improvement/	11410
27	exp health services research/	139705
28	Health Services/	22691
29	limit 28 to yr="1974-1979"	2628
30	Research/	192971
31	limit 30 to yr="1966-1979"	16970
32	Health Resources/	10141

33	(qualit* adj2 (assurance* or improv* or indicator* or improv* or assess* or measure* or control*)).tw,kf.	193487
34	16 or 17 or 18 or 19 or 20 or 21 or 23 or 24 or 25 or 26 or 27 or 29 or 31 or 32 or 33	2401134
35	15 and 34	165512
36	(((data adj1 (collect* or aggregat*)) or aggregat* or (initiative* or standard* or measurement* or checklist* or check-list* or survey* or questionnair* or tool or tools or interview* or index or scale* or indicator*)) adj3 (surger* or surgic* or procedure* or operati* or preop* or pre-op* or presurg* or pre-surg* or perop* or periop* or peri-op* or perisurg* or peri-surg* or intraop* or intra-op* or postop* or postsurg* or (post adj2 op*))).ab.	77135
37	(((data adj1 (collect* or aggregat*)) or (initiative* or standard* or measurement* or checklist* or check-list* or survey* or questionnair* or tool or tools or interview* or index or scale or indicator*)) and (surger* or surgic* or procedure* or operati* or preop* or pre-op* or presurg* or pre-surg* or perop* or periop* or peri-op* or perisurg* or peri-surg* or intraop* or intra-op* or postop* or postsurg* or (post adj2 op*))).ti,kf.	21876
38	36 or 37	95407
39	14 and 38	17105
40	35 or 39	175978
41	Global Health/	35190
42	exp "united states dept. of health and human services"/	75785
43	united states agency for international development/	55
44	exp world health organization/	29797
45	international agencies/	7710

46	exp internationality/	145152
47	exp societies, hospital/	4877
48	Hospitals/ or Societies/	82699
49	limit 48 to yr="1966-1977"	8754
50	exp societies, medical/	63467
51	Advisory Committees/	7959
52	(global* or international* or world* or humanitarian*).kf.	17739
53	(global* or international* or world* or humanitarian* or agency or agencies or association* or society or societies or nation* or federation* or iNGO or NGO or organization* or task force* or committee* or academy or commission).ti. /freq=2	44456
54	(global* or international* or world* or humanitarian* or agency or agencies or association* or society or societies* or nation* or federation* or iNGO or NGO or organization* or task force* or committee* or academy or commission).ab. /freq=5	73051
55	((WHO or lancet or unicef) adj (commission or index or global* or survey* or situational or tool* or checklist* or check-list* or surg* or indicator* or safe*)).tw,kf.	1168
56	41 or 42 or 43 or 44 or 45 or 46 or 47 or 49 or 50 or 51 or 52 or 53 or 54 or 55	442675
57	40 and 56	4526
58	Animals/ not (Animals/ and Humans/)	4230832
59	57 not 58	4521
60	remove duplicates from 59	4403

Please contact the author for the full Search strategy in all databases.

Appendix B.

Data Extraction Sheet

1. Tool Name
2. Article Information
    A. Author
    B. Title
    C. Journal
    D. Year Published
    E. DOI
3. Tool specifics
    A. Language of tool
    B. Funding agency of tool
    C. Aim of assessment tool
    D. Name of assessment tool
    E. Country of assessment tool
    F. Classification of economy where assessment tool was created/validated
    G. Tool based on previously published surgical capacity assessment tool
        1. Original capacity assessment tool
4. Details of Tool
    H. Number of items
        1. Resources
        2. Outcome
        3. Accessibility/Impact
        4. Training
    B. Method of development
    C. Score
        1. Calculation of score
    D. Year of tool created
5. Consistency & Reliability
    E. Cronbach's Alpha
    F. Intraclass correlation coefficient

G. Percentage of missing items

H. Handling of missing items

I. Administered in how many different institutions

    1. Administered in how many secondary or primary centers

    2. Administered in how many tertiary centers

F. Administered in how many countries

G. Method with which pilot institutions were chosen

H. Instrument administered at least twice in the same institutions

I. Administrators independent

J. Time interval between administrations

K. Test conditions similar for both administrations

6. Validity

L. Reassessment of relevance of all items

M. Confirmation of face validity

N. Expertise of the people involved in formulation of tool

O. Criterion used or employed considered as a reasonable gold standard

7. Hypothesis Testing

P. Hypotheses regarding correlations formulated a priori

Q. Expected direction of correlations

Chapter 4: Creation And Piloting Of The GAPS Tool

Preface

The systematic review (Manuscript 1) identified three capacity assessment tools (CATs) designed to evaluate the surgical capacity of pediatric centers: the *Pediatric Surgery Personnel, Infrastructure, Procedures, Equipment, Supplies* (Pedi-PIPES); the *Children's Surgical Center Designations with Scope of Practice* (CSCDSP); and the *Checklist for a Children's Trauma Room* (CCTR). All these tools were adapted from other CATs evaluating adult surgical centers. Pedi-PIPES is the only tool that focuses exclusively on the assessment of centers in L-MICs. Neither the CSCDSP nor the CCTR include scores enabling longitudinal comparison of institutions. Though Pedi-PIPES can generate a score, it is a non-weighted sum of the tool's data points based on a dichotomous evaluation of each item - resulting in limited interpretation and use. Moreover, all three tools focus disproportionately on resources (human and material) and accessibility, omitting outcomes, quality assurance measures, research endeavors, and training. No tool comprehensively evaluates the full scope of pediatric surgical care or the existing gaps in pediatric surgical capacity in L-MICs. Beyond the findings of Manuscript 1, GAPS was developed based on ten other CATs and surgical resource guidelines (detailed in Manuscript 1 and Appendix B of Manuscript 2)[27, 28].

Following the publication of both Manuscript 1 and Manuscript 2, Nuss et al. introduced the Children's Surgical Assessment Tool (CSAT), designed to evaluate the capacity of district and tertiary hospitals in Nigeria for pediatric surgical services[27, 28, 46]. This tool assesses hospital infrastructure, workforce, service delivery, financing, and training capacity specific to pediatric surgery. The foundation for CSAT can be traced back to Lin et al. in 2020, who developed the initial surgical assessment tool using a systematic review and the Delphi method [47]. Lin's tool incorporated 15 different surgical tools, comprising 216 unique elements within the domains of infrastructure, service delivery, and workforce. Their final version included 169 items, expanding to cover infrastructure, service delivery, workforce, information management, and financing.

Building on this foundation, Nuss et al. adapted CSAT following the OReCS guidelines[20, 22, 24-26, 32, 46, 48]. To refine the tool, an expert panel engaged in a three-round modified Delphi process to achieve consensus. The tool's validity was evaluated using the Content

Validity Index (CVI), a quantitative measure assessing how relevant and representative individual items are of the content being measured. Experts rate items based on relevance, and the CVI is calculated as the proportion of items deemed highly relevant. A CVI closer to 1.0 signifies high content validity.

In this study, 18 experts participated, including 14 from Rwanda and 4 from the USA. The panel consisted of 6 pediatric surgeons, 5 pediatric anesthesiologists, 1 general surgeon, 2 surgical subspecialists, 1 pediatrician, and 3 health systems researchers. After three Delphi rounds, the final survey consisted of 171 items, with an average CVI of 0.84.[46]

The CSAT collects comprehensive data across various areas: facility characteristics broken down by patient age; access and referral systems; infrastructure (including emergency rooms, pharmacies, radiology, pathology, and blood supply); service delivery, detailing pediatric surgery volumes in general surgery, trauma, orthopedics, and plastic surgery; quality and safety measures (such as training, outcome tracking, electronic medical records, and registries); operating equipment; workforce; financing; and training and research capabilities.[46]

However, it is important to note that the CSAT has not been validated with clinical outcomes, such as mortality or morbidity. Additionally, there is no scoring system or index associated with the tool to quantify the hospital's capacity in relation to patient outcomes. This lack of validation with outcomes limits its application for directly predicting or improving patient care results.

The CATs identified in Manuscript 1 were disproportionately focused on resource assessment with little reference to accessibility, outcomes, education, or quality assurance measures. Moreover, there have been no published attempts to date to validate any capacity assessment tools aimed at pediatric surgery in L-MICs.

Manuscript 2 details the second phase of the iterative development of GAPS. Developing the tool began with expert consultation, grounded in evidence gathered from Manuscript 1. Following this, international pilot testing assessed the tool's feasibility and applicability across diverse settings, ensuring its adaptability to varying environments. The third stage focused on construct validation and item analysis, where items were evaluated for their ability to differentiate effectively between levels of care. Finally, instrument refinement retained only those items that could be successfully answered in L-MICs and demonstrated a significant ability to distinguish among various levels of hospital care.

DeVellis et al. emphasize on the importance of theoretical and empirical foundations in item development[34]. Informing item development using professional insights ensures that initial items are conceptually relevant. Moreover, expert consultation should be carried out early in scale development so that initial item generation draws from evidence-based criteria to promote content validity. International testing follows the recommendation of both DeVellis and Streiner, who stress the importance of evaluating feasibility and validity across diverse populations—an essential step for creating tools with wide-ranging applicability[34, 35]. As per both DeVellis et al. and Streiner et al., pilot testing is imperative for pretesting scales in various contexts to uncover potential issues in feasibility and relevance, thereby ensuring the tool's adaptability to different settings[34, 35]. Similarly, Streiner et al. highlight that testing in multiple environments can reveal culturally specific challenges, enhancing the scale's relevance and effectiveness in international contexts[35].

In order to proceed with construct validity, we chose to focus on whether or not GAPS was able to identify between level of surgical care. We selected this metric as a theoretical surrogate for surgical capacity, particularly in the context of L-MICs where reliable objective measures for surgical capacity or outcomes are very often unavailable. When assessing construct validity items must demonstrate the ability to discriminate effectively across constructs of interest—in this case, levels of care. Construct validation is a core component in DeVellis et al.'s framework, as it assesses whether items measure the intended concepts[34]. Streiner et al. further elaborate on this, highlighting that effective item analysis should reveal how well each item contributes to the overall construct, ensuring that items align with the scale's theoretical framework[35].

This final phase follows DeVellis et al.'s principle of iterative refinement, retaining only the items that contribute meaningfully to differentiation in target settings, particularly L-MICs. DeVellis et al. suggest using empirical testing to retain items that demonstrate utility and discriminative power, eliminating those that do not perform well across intended settings, and enhancing both reliability and validity [34]. Such refinement to ensure that the scale remains practical, accessible, and interpretable, especially in L-MICs where resource constraints may influence item interpretation and response[35].

Manuscript 2 was presented at the annual meeting of the Canadian Association of Pediatric Surgeons (CAPS) in September 2018 in Toronto, Canada.

Manuscript 2 has been published under the following reference[27]: Yousef Y, Cairo S, St-Louis E, Goodman LF, Hamad D, Baird R, Smith ER, Emil S, Laberge JM, Abdelmalak M, Gathuy Z, Evans F, Adel MG, Bertille KK, Chitnis M, Millano L, Nthumba P, d'Agostino S, Cigliano B, Zea-Salazar L, Ameh E, Ozgediz D, Guadagno E, Poenaru D. GAPS Phase II: Development and Pilot Results of the Global Assessment in Pediatric Surgery, an Evidence-Based Pediatric Surgical Capacity Assessment Tool for Low-Resource Settings. Pediatric Surgery International. 2024 Jun 19. 40, 158 (2024). https://doi.org/10.1007/s00383-024-05741-w

MANUSCRIPT 2: GAPS PHASE II: DEVELOPMENT AND PILOT RESULTS OF THE GLOBAL ASSESSMENT IN PEDIATRIC SURGERY, AN EVIDENCE-BASED PEDIATRIC SURGICAL CAPACITY ASSESSMENT TOOL FOR LOW-RESOURCE SETTINGS

Yasmine Yousef[1], Sarah Cairo[2], Etienne St-Louis[1], Laura F. Goodman[3], Doulia M. Hamad[4], Robert Baird[5], Emily R. Smith[6], Sherif Emil[1], Jean-Martin Laberge[1], Mohamed Abdelmalak[7,8], Zipporah Gathuy[9], Faye Evans[10], Maryam Ghavami Adel[11], Ki K. Bertille[12], Milind Chitnis[13], Leecarlo Millano[14], Peter Nthumba[15], Sergio d'Agostino [16], Bruno Cigliano [16], Luis Zea-Salazar [17], Emmanuel Ameh[18], Doruk Ozgediz[19], Elena Guadagno[1], Dan Poenaru [1]

[1]Harvey E. Beardmore Division of Pediatric Surgery, The Montreal Children's Hospital, McGill University Health Centre
[2]University of California San Francisco, Department of Surgery, San Francisco, CA, USA
[3]Children's Hospital of Orange County and University of California Irvine, Orange, CA, USA
[4] Faculty of Medicine, McGill University, Montreal, QC, CANADA
[5] Division of Pediatric General Surgery BC Children's Hospital, University of British Columbia, Vancouver, BC, CANADA
[6] Department of Emergency Medicine and Center for Global Surgery and Health Equity, Duke University, Durham, NC, USA
[7] Nile of Hope Hospital, Alexandria, EGYPT
[8] Alexandria University Children's Hospital, Alexandria, EGYPT
[9] The Nairobi Hospital, Nairobi, KENYA
[10] Boston Children's Hospital and Harvard Medical School, Boston, MA, USA
[11] Valiasr Hospital, Tehran University of Medical Sciences, Tehran, IRAN
[12] Charles de Gaulle Pediatric Academic Hospital, Ouagadougou, BURKINA FASO
[13] Frere Hospital, Walter Sisulu University, East London, SOUTH AFRICA
[14] Faculty of Medicine and Nursing, Krida Wacana University, Jakarta, INDONESIA
[15] AIC Kijabe Hospital, Kenya and Department of Plastic Surgery, Vanderbilt University Medical Center, TN, USA
[16] Surgery for Children – ONLUS, Vicenza, ITALY
[17] OMNI HOSPITAL, Guayaquil, ECUADOR

[18] Division of Paediatric Surgery, Department of Surgery, National Hospital, Abuja, NIGERIA
[19] Department of Surgery, University of California, San Francisco, San Francisco, California, USA

Keywords: pediatric surgery, children's surgery, global surgery, global health, low-income country, middle-income country

Funding :
Fonds de Recherche du Québec - Santé (FRQS)
Canadian Institutes of Health Research

Acknowledgements:
Special thanks to the following: Kevyn Armand, Daniel Agbo-Panzo, Marilyn Butler, Beda Espineda, Kenzy Abdel-Hamid, Sonia Anchouche, Ai-Xuan Holterman, George W. Galiwango, Nathalie MacKinnon, Ritesh Shrestha, Desigen Reddy, William Harkness, Phyllis Kisa, Rashmi Kumar, Ibrahima Fall, Justin Onen, Ananda K. Lamahewage, Mike Ganey, Tessa Concepcion, Peter Donkor, Zaitun Bokhary, Anette S. Jacobsen, Reju Thomas, Brouh Yapo, Adiyasuren Jamiyanjav, Marcia Matias, Benno Ure, Vanda Amado, Bahati Robert, Lubna Samad, Roumanatou Bankolé, Pramod Puligandla, and Robin Petroze.

Abstract

<u>*Purpose*</u>:

Pediatric surgical care in low- and middle-income countries is often hindered by systemic gaps in healthcare resources, infrastructure, training, and organization. This study aims to develop and validate the Global Assessment of Pediatric Surgery (GAPS) to appraise pediatric surgical capacity and discriminate between levels of care across diverse healthcare settings.

<u>*Methods*</u>:

The GAPS Version 1 was constructed through a synthesis of existing assessment tools and expert panel consultation. The resultant GAPS Version 2 underwent international pilot testing. Construct validation categorized institutions into providing Basic or Advanced Surgical Care. GAPS was further refined to Version 3 to include only questions with a > 75% response rate and those that significantly discriminated between Basic or Advanced Surgical settings.

<u>*Results*</u>:

GAPS Version 1 included 139 items, which, after expert panel feedback, was expanded to 168 items in Version 2. Pilot testing, in 65 institutions yielded a high response rate. Of the 168 questions in GAPS Version 2, 64 significantly discriminated between Basic and Advanced Surgical Care. The refined GAPS Version 3 tool comprises 64 questions: Human Resources (9), Material Resources (39), Outcomes (3), Accessibility (3), and Education (10).

<u>*Conclusion*</u>:

The GAPS Version 3 tool presents a validated instrument for evaluating pediatric surgical capabilities in low-resource settings.

Introduction

The critical issues of accessibility and quality of surgical care in low-resource centers were brought to the forefront of global health by the Lancet Commission on Global Surgery and several other non-governmental groups [1–5]. Support for global efforts aimed at increasing availability of pediatric surgical resources often comes from “North–South” partnerships [1, 2]. The objectives of these collaborations are infrequently stated, potentially resulting in inefficient resource allocation and limited impact [6–10].

To guide such partnerships in identifying priorities and monitoring quality improvement, standardized tools are required which measure existing gaps in care and identify the most urgent needs [11]. A 2019 systematic review published by our group found no capacity assessment tools (CAT), which comprehensively evaluated the full scope of pediatric surgical care or the existing gaps in pediatric surgical capacity in low-resource settings [12]. We sought to develop a validated surgical CAT, the Global Assessment in Pediatric Surgery (GAPS). The creation of GAPS includes the following three phases.

- Phase I - systematic review and item pool generation: consisted of the aforementioned systematic review to identify current relevant tools available [12].
- Phase II - instrument development and preliminary validation: the subject of this article, consisted of four components:
    1. Initial Item Development Expert Consultation: initial tool development based on evidence found in Phase I
    2. International pilot testing: feasibility and applicability testing in multiple settings.
    3. Construct Validation and Item Analysis: Assessing the ability of items to discriminate between levels of care
    4. Instrument Refinement: retaining only items that can successfully be answered in LMICs and can differentiate between levels of hospital care
- Phase III - Outcome Based Weighting for Construct Validity: addition of a weighted scoring system based on outcomes to enhance the tools predictive validity.

Methods

<u>*Definitions*</u>

Levels of healthcare facilities were based on the *"Essential Surgery"* volume of *Disease Control Priorities, 3rd Edition* (DCP3) classification and the *Global Initiative for Children's Surgery: Optimal Resources for Children's Surgery* (GICS ORECS). They include first-level hospital, second- level hospital, third-level hospital, and national children's hospital [13–15], as predefined by governmental agencies.

Definitions of low-income country (LIC), lower-middle- income country (LMIC), upper-middle-income (UMIC), and high-income country (HIC) were based on The World Bank's 2018 designation of gross-national income per capita [16].

Component 1: GAPS INITIAL ITEM DEVELOPMENT AND EXPERT CONSULTATION ITEM DEVELOPMENT

Using the insights of our systematic review [12], we synthesized the items from existing pediatric surgical CATs. These were organized into five domains: human resources (HR), material resources (MR), accessibility (AC), outcomes (OU), and education (ED). Responses to each item and their scales were developed based on the level of hospital using data from DCP3 and the GICS ORECS [13, 15]. This preliminary iteration, GAPS Version 1, was reviewed by a panel of eight authors to ensure its comprehensibility, ambiguity, and alignment with the unique challenges faced in low-resource settings.

EXPERT CONSULTATION: FROM DELPHI TO PANEL

Our initial strategy to refine GAPS Version 1 involved a modified two-round Delphi process. However, responses rendered the identification of overarching themes challenging. Consequently, we pivoted to a single-round expert consultation panel. Experts were asked to comment on the relevance, utility, and clarity of each data point and the pertinence of its inclusion in the tool. Their collective feedback was used in the evolution from GAPS Version 1 to Version 2.

Component 2: INTERNATIONAL PILOT TESTING

To ensure that GAPS Version 2 was feasible and universally applicable, we implemented an in-person pilot testing phase between February 2017 and March 2018. Each survey was completed in person on paper with the aid of a local senior medical representative. Countries were chosen based on their geographical location, ease of access, safety for travel, and number of accessible institutions.

Component 3: CONSTRUCT VALIDATION AND ITEM ANALYSIS CONSTRUCT VALIDATION: DISCRIMINATING LEVELS OF CARE

To further refine GAPS, participating institutions from the pilot phase were stratified into two categories: basic care and advanced care centers [17]. "Basic surgical care" (BSC) was defined as a first- or second-level hospital. "Advanced surgical care" (ASC) was defined as a third- level or national children's hospital.

Responses to GAPS were compared to predefined "ideal" responses per hospital level, previously identified by our expert panel. Responses for every question were separated into BSC and ASC. These responses were then compared with an objective assessment of the hospital's level to determine whether the question matched the hospital's actual designation.

The analysis led to the preservation of questions that accurately represented the resource constraints in low- resource settings and successfully distinguished between the two care levels.

ITEM ANALYSIS

The primary endpoint was to determine if items in GAPS were able to differentiate between centers providing BSC and ASC. Questions in each section were evaluated for their ability to discriminate between BSC and ASC, using Fisher's Exact Test. P-values < 0.05 were deemed statistically significant. A Bonferroni Correction was not applied to avoid the potential exclusion of relevant questions from the GAPS tool analysis. The data were analyzed using STATA / MP 13.0 (StataCorp, College Station, TX).

Component 4: INSTRUMENT REFINING

Informed by the insights from the pilot testing and construct validity phases, GAPS was refined to include questions that were significantly able to discriminate between BSC and ASC ($p < 0.05$) and had a response rate over 75% during pilot testing.

Reliability assessment was done using a graded response model (GRM) within the item response theory (IRT) framework to estimate item discrimination and difficulty.

STUDY APPROVAL

This study was approved by the Pediatric Research Ethics Board of the McGill University Health Centre Research Institute (27-2502).

Results

Component 1: GAPS INITIAL ITEM DEVELOPMENT AND EXPERT CONSULTATION ITEM DEVELOPMENT

GAPS Version 1 (Appendix A) was based on three surgical CATs for children, five surgical CATs for adults, and five surgical resource guidelines (Appendix B). These tools and guidelines were evaluated by the first author and ordinal or categorical items were created based on commonalities between tools to address human resources, material resources, accessibility, outcome, and education. GAPS Version 1 consisted of 139 items separated into five sections - 19 HR, 71 MR, 22 OU, 17 AC, and 10 ED.

EXPERT CONSULTATION

A panel of 37 experts in the field of global children's surgery agreed to participate in the single-round expert consultation panel. Experts included in the panel originated from 28 countries (Figure 1). Most participants were from LMICs (14/37, 38%), followed by LICs (9/37, 24%), HICs (8/37, 22%) and UMICs (6/37, 16%). Participants were most often pediatric general surgeons (25/37, 68%). Other pediatric specialties represented on the panel included anesthesia (5/37, 14%), neurosurgery (2/37, 5%), plastic surgery (2/37, 5%), urology (1/37, 3%), oral maxillofacial surgery (1/37, 3%) and critical care (1/37, 3%).

GAPS Version 1 was well-received by the expert panel (Appendix C). The panelists concurred with the tool's sectional division. However, they voiced concerns over the insufficient detail concerning human resources, specifically anesthesia and nursing training. There was a shared desire for more granularity on the origins of funding, material sourcing, and protocols for single-use instruments. A significant challenge highlighted was the difficulty in capturing outcome data, given that many institutions lack robust outcome monitoring. While GAPS was seen as a promising initiative, there was consensus on the need for enhanced comprehensiveness, specificity, and clarity.

Based on the opinions and notes of the panel, GAPS Version 1 was revised to yield GAPS Version 2 (Appendix D), composed of 168 items divided as such: 26 HR, 86 MR, 14 AC, 27 OU, 15 ED.

Using the notes from the expert panel, ORECS and DCP3, responses to each question were edited to reflect ideal responses per level of hospital as described in the methods.

Component 2: INTERNATIONAL PILOT TESTING

GAPS Version 2 was piloted in 65 institutions across eight countries (Figure 1) at all hospital levels: 27 (42%) 1st level hospitals, 23 (35.9%) 2nd level hospitals, 11 (17.2%) 3rd level hospitals, and 3 (4.7%) national children's hospitals.

Figure 1. Countries of Expert Panel Member Institutions and GAPS Version 2 Tool Pilot Testing Countries

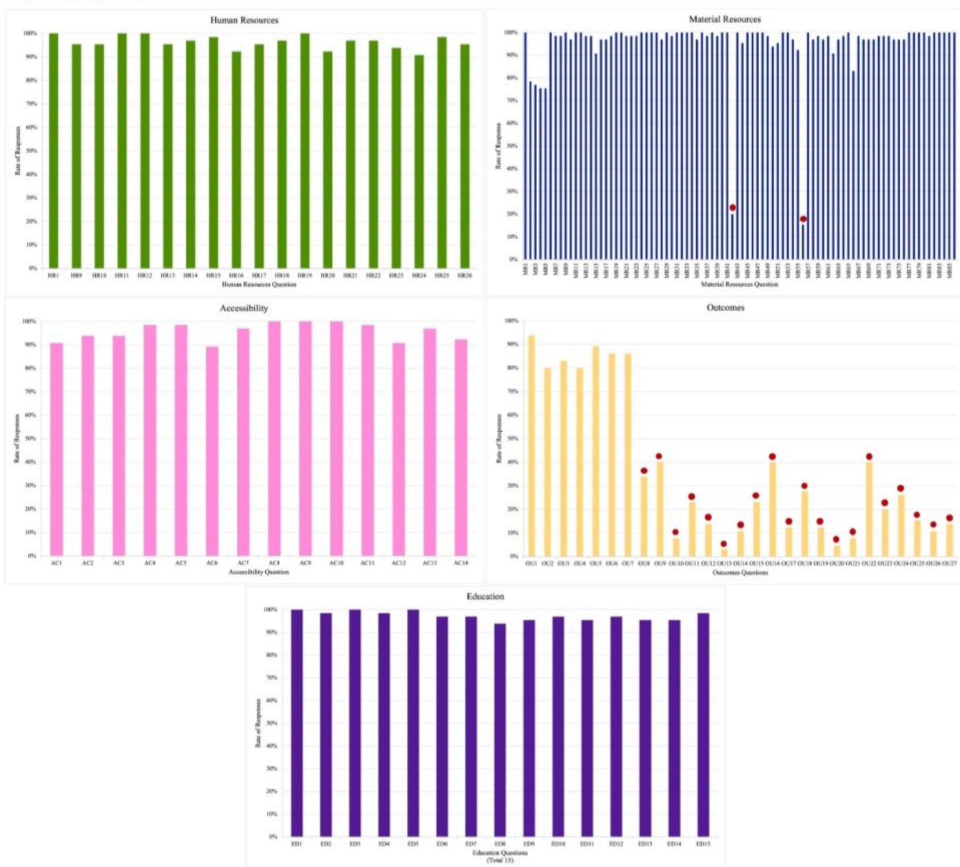
GAPS Version 2 Pilot Testing Countries: Algeria, Democratic Republic of Congo (DRC), Egypt, Mongolia, Somaliland, Thailand, Uganda, and Vietnam
Expert Panel Member Countries: Brazil, Burkina Faso, Canada, Côte D'Ivoire, Ecuador, Egypt, Germany, Ghana, Guinea, India, Indonesia, Iran, Italy, Kenya, Mongolia, Mozambique, Nepal, Nigeria, Pakistan, Philippines, Senegal, Singapore, South Africa, Sri Lanka, Tanzania, United Kingdom, United States of America, and Uganda

Most countries included in the pilot study were LIC (Somaliland, Democratic Republic of Congo (DRC), Uganda), followed by LMIC (Mongolia, Egypt, and Vietnam), and lastly UMIC (Algeria and Thailand). The DRC comprised 63.1% of the institutions, followed by Mongolia (12.3%). Hospitals were most frequently public (69.2%), followed by faith-based (24.6%), non-governmental (3.1%) and private organizations (3.1%). Table 1.

Table 1. Details on countries and institutions of pilot centers			
	Basic Care N=50 (%)	Advanced Care N=15 (%)	Total N=65 (%)
Hospital Type			
Public	33 (66)	12 (80)	45 (69.2)
Non-governmental organization	1 (2)	1 (6.7)	2 (3.1)
Faith-Based	15 (30)	1 (6.7)	16 (24.6)
Private	1 (2)	1 (6.7)	2 (3.1)
Country			
Algeria	3 (6)	0 (0)	3 (4.6)
Democratic Republic of Congo	37 (74)	4 (26.7)	41 (63.1)
Egypt	0	4 (26.7)	4 (6.2)
Mongolia	7 (14)	1 (6.7)	8 (12.3)
Somaliland	0	1 (6.7)	1 (1.5)
Thailand	0	1 (6.7)	1 (1.5)
Uganda	2 (4)	2 (13.3)	4 (6.2)
Vietnam	1 (2)	2 (13.3)	3 (4.6)
Income level			
Low-income	39 (78)	7 (46.7)	46 (70.8)
Low-middle income	8 (16)	7 (46.7)	15 (23.1)
Upper-middle income	3 (6)	1 (6.7)	4 (6.2)

The response rate for all instrument sections was over 90% (HR 96%, MR 95%, AC 96%, ED 97%) except for the outcomes section (36%) (Figure 2). Regarding MR, two questions had a response rate below 30%: the availability of pediatric endoscopes (colonoscope, gastroscope, rigid esophagoscope, cystoscope, bronchoscope) and the availability of sterilized surgical instrument sets for emergent cricothyroidotomy. Representatives from partner institutions were less responsive to questions regarding rates of: surgical site infection, 30-day follow-up, mortality related to anesthesia or surgery, continuous respiratory monitoring, and depth of anesthesia. When asked if centers had a method of measuring outcomes, 28 institutions (28/65, 43%) answered "No" and four (4/65, 6%) did not answer. Of the 33 centers (33/65, 51%) that had a method of measuring outcomes over time, none were able to answer all questions in the outcomes section – particularly regarding mortality rate.

Figure 2. Rate of Responses of Questions in Pilot Study of GAPS Version 2



This figure details the cumulative response rate of each question in the GAPS Version 2 during the pilot study. Human resources is comprised of 25 questions. However, HR 2 – 8 are not represented as they are dependent on the answer to question HR1 and thus the response rate is not interpretable. of Material resources is comprised of 86 questions. Accessibility is comprised of 14 questions. Outcomes is comprised of 27 questions. Education is comprised of 15 questions. A red circle above the question identifies those that had a lower than 75% response rate. The full questions can be found in Appendix D.

Component 3: CONSTRUCT VALIDATION AND ITEM ANALYSIS

GAPS Version 2 was piloted in 50 BSC and 15 ASC facilities. Overall, 38% (64/168) of questions significantly discriminated between BSC and ASC.

In HR, 9/26 (35%) questions significantly discriminated between BSC and ASC. Questions that could discriminate between levels of care were centered around the presence of pediatric surgical subspecialists (general surgery/urology, orthopedic surgery, otolaryngology, cardiac surgery) and medical subspecialties (pediatrics, pediatric anesthesia, intensive care, emergency) (Table 2).

Table 2. Questions Able to Discriminate between Level of Care: Human Resources

		Basic Care N=50 (%)	Advanced Care N=15 (%)	P-value*
HUMAN RESOURCES				
Pediatric Surgical Specialties				
HR2	General / Urology	7 (14)	9 (60)	*0.001*
HR3	Orthopedic	3 (6)	4 (26.7)	*0.044*
HR4	Otolaryngology	1 (2)	3 (20)	*0.036*
HR6	Cardiac	0	3 (20)	*0.010*
Pediatric Specialties				
HR10	Pediatrician	12 (25.5)	10 (76.9)	*0.002*
HR11	General Anesthesia (> 12months)	12 (25)	10 (66.7)	*0.005*
HR15	Intensive Care Unit	6 (24)	8 (61.5)	*0.035*
HR19	Emergency Room	4 (11.1)	6 (42.9)	*0.020*
MATERIAL RESOURCES				
Infrastructure				
MR1	Dedicated Children's Hospital	2 (4)	4 (26.7)	*0.022*
MR3	FAST	10 (26.3)	8 (66.7)	*0.017*
MR6	Maternal Unit	47 (94)	11 (73.3)	*0.044*
MR7	Cesarean Section Availability	49 (98)	10 (71.4)	*0.007*
MR8	Beds > 10	4 (8.2)	7 (46.7)	*0.002*
MR9	Dedicated ORs > 2	3 (6)	6 (80)	*0.003*
MR10	Available ORs > 2	2 (4.1)	6 (42.9)	*0.001*
MR13	Prolonged Mechanical Ventilation	26 (52)	13 (92.9)	*0.005*
MR19	PACU	28 (56)	15 (100)	*0.001*

MR20	PACU for Children	1 (2)	6 (40)	*< 0.001*
MR21	PICU	5 (10)	8 (57.1)	*0.001*
MR22	PICU beds Available to Children	26 (53)	15 (100)	*< 0.001*
MR23	Internet	13 (26.5)	9 (60)	*0.028*
MR26	Electricity	25 (50)	14 (93.3)	*0.002*
MR33	CT Scan	7 (14)	7 (46.7)	*0.012*
MR34	Anesthesia Machine	25 (50)	14 (93.3)	*0.002*
Equipment				
MR35	Neonatal Incubators	18 (36.7)	12 (85.7)	*0.002*
MR36	Neonatal Ventilators	9 (18)	12 (80)	*< 0.001*
MR37	Pediatric Ventilators	11 (22.5)	13 (86.7)	*< 0.001*
MR39	Supplemental Oxygen	23 (46)	12 (85.7)	*0.013*
MR44	Defibrillator	2 (4.2)	5 (35.7)	*0.005*
MR45	Electrocautery	23 (46)	14 (93)	*0.001*
MR46	Laryngoscope handle and pediatric sized blades *(Miller ≤ 2 or Macintosh ≤ 3)*	30 (60)	14 (93.3)	*0.025*
MR48	Pediatric endotracheal tubes *(2.5 - 6 mm)*	25 (50)	13 (86.7)	*0.016*
MR49	Pediatric oral airways *(000-4)*	27 (54)	12 (85.7)	*0.036*
MR50	Pediatric Nasal Airways	22 (45.8)	11 (84.6)	*0.026*
MR53	Sterilized surgical instrument sets for open abdominal, thoracic, cardiac, neurosurgery and orthopedic surgery	24 (48)	12 (80)	*0.039*
MR55	Sterilized surgical instrument sets for emergency thoracotomy	9 (19.7)	9 (64.3)	*0.003*
MR57	Functioning autoclave (sterilizer)	33 (66)	15 (100)	*0.007*
MR61	Capnograph	3 (6)	6 (42.9)	*0.002*
MR62	Electric warming blanket	7 (15.2)	10 (76.9)	*< 0.001*
MR63	Depth of anesthesia monitor	5 (10)	6 (46.2)	*0.007*
Supplies				
MR73	Pediatric chest tubes (<20Fr)	13 (26.5)	9 (60)	*0.028*
MR74	Functioning chest drainage system	17 (35.4)	10 (66.7)	*0.041*
MR77	Anticonvulsants	14 (28)	10 (66.7)	*0.013*
MR79	Hypertonic Saline	17 (34)	10 (66.7)	*0.036*
MR81	Parenteral Nutrition	18 (36.7)	11 (73.3)	*0.018*

MR84	Paralytics	23 (46)	14 (93.3)	*0.001*
MR86	Vasopressors	33 (66)	14 (93.3)	*0.049*
EDUCATION				
Academic Affiliation				
ED1	Affiliated with a University	10 (20)	11 (73.3)	*< 0.001*
ED2	Surgical Residents Involved in Care	8 (16)	9 (64.3)	*0.001*
ED3	Fellows Involved in Care	1 (2)	8 (53.3)	*< 0.001*
ED4	Residency Training Program	3 (6)	11 (78.6)	*< 0.001*
ED5	Fellowship Training Program	0	6 (40)	*< 0.001*
ED6	Anesthesiology Training Program	2 (4)	5 (38.5)	*0.003*
Continuing Professional Development				
ED9	Participation in continuing medical education	15 (30.6)	10 (76.9)	*0.004*
ED10	Research Involvement	8 (16.7)	10 (71.4)	*< 0.001*
ED13	Tumor Boards	3 (6.3)	7 (50)	*0.001*
Clinical Service				
ED15	Clinical Service dedicated to Children's Surgery	2 (4.1)	9 (60)	*<0.001*
**Fisher Exact Test*				
FAST = Focused Assessment with Sonography in Trauma, OR = Operating Room, PACU= Pediatric Anesthesia Care Unit, PICU = Pediatric Intensive Care Unit, CT = Computerized tomography, mm=millimeteres, Fr = French				

MR were divided into Infrastructure, Equipment, and Supplies. Half of the questions related to infrastructure (16/34; 47.1%) and equipment (16/29; 55.2%), and a third of questions referring to supplies (7/23; 30.4%) significantly discriminated between the two levels of care (Table 2).

OU and AC were the least discriminating sections in GAPS Version 2. Three questions in OU (3/27; 11.1%) demonstrated a difference BSC and ASC; (1) *30-day follow-up rate > 50%* (p = 0.011), (2) *Continuous monitoring of airway and ventilation > 50% of the time* (p = 0.022), (3) *Adequacy of ventilation monitored with capnography > 50% of the time* (p = 0.001). Similarly, only three questions in AC (3/14; 21.4%) could discern between levels of care; (1*) Catchment area > 1 million* (p = 0.006), (2) *Presence of minimally invasive surgery* (p = 0.001), (3) *ASA Class > 3* (p = 0.013).

ED was the most successful section with 66.7% (10/15) of questions significantly discriminating between levels of care (Table 2).

Component 4: INSTRUMENT REFINEMENT

Based on the above findings, GAPS Version 2 was revised to include only items that were able to significantly differentiate between BSC and ASC. Consequently, the GAPS Version 3 (Appendix E) includes 64 items in the following distribution: 9 HR, 39 MR, 3 OU, 3 AC, and 10 ED. (Table 3).

Table 3. Evolution of Item refinement per section in GAPS Version 1, Version 2, and Version 3

Item Sections	Version 1	Version 2	Version 3
Human Resources	19	26	9
Material Resources	74	89	39
Infrastructure	29	34	16
Supplies	22	23	7
Equipment	20	32	16
- Conditional questions	3	3	-
Accessibility	17	14	3
Outcome	22	27	3
Education	10	15	10
Total questions including conditional questions	*142*	*171*	*64*
Total questions	**139**	**168**	**64**

GAPS Version 3 was analyzed in a GRM to better understand the data's structure and assess tool's internal reliability (Appendix F). Forty items (40/64, 63%) showed very high discrimination (> 1.5): 6 HR (6/9, 67%), 23 MR (26/39, 67%), 7 ED (7/10, 70%), 1 AC (1/3, 33%) none in OU. Fourteen items (14/64, 22%) showed high discrimination (1 - 1.5): 1 HR (1/9, 11%), 7 MR (7/39, 18%), 3 ED (3/10, 30%), 2 AC (2/3, 67%) and 1 OU (1/3, 33%). Seven items (7/64, 11%) showed moderate discrimination (0.5 - 1): 1 HR (1/9, 11%), 4 MR (4/39, 10%), 2 OU (2/3, 67%) and none in ED or AC. Only three items showed poor discrimination: 1 HR (HR2: -0.07) and 2 MR (MR4: -2.96 and MR3: -1.26). The item in HR that showed poor discrimination was "*The presence of pediatric general surgeons/ urologist*", was retained in GAPS Version 3 as it was felt to be a relevant question for the tool. The items in MR showing poor discrimination both had negative coefficients with an absolute integer > 1.5 suggesting high discrimination. Thus, both items were kept in the tool. The difficulty parameters, representing the thresholds for each response category, also showed variability across items.

Discussion

This publication describes the development, piloting, and construction validation of the GAPS, an evidence-based pediatric surgical CAT designed for low-resource settings. A systematic review, published by our group, describing existing CATs and surgical resource guidelines [12] served as the basis for the development of GAPS. A 37-person multidisciplinary expert panel collaborated to establish consensus about the data items in the tool.

The development of GAPS was an intensive endeavor, highlighting the complexities associated with research in low-resources settings. The observed lower response rate for two questions in the MR section pertaining to the availability of specialized equipment, such as pediatric endoscopes and sterilized surgical instrument sets for emergent cricothyroidotomy, could be due to several factors. Foremost, the scarcity of these resources in low-resource settings may result in fewer respondents possessing pertinent information or practical experience. Moreover, respondents may have less familiarity with these specific resources, especially if they are not routinely used in their clinical practice. Ultimately, the inherent complexity or perceived infrequency of these items could lead to less attention or prioritization of these questions, further contributing to the lower response rate. The limited response rate of the OU and AC sections is likely attributable to the persistent challenges in data col- lection inherent to research in the developing world [18, 19]. Low- and middle-income countries often lack the infrastructure for consistent outcome reporting, complicating proof of intervention for quality improvement [20, 21]. Improving outcome reporting in low-resource settings can be achieved through several strategies [22]. The most often attempted remains capacity building and training to enhance data quality and consistency [14]. More recently, leveraging technology for real-time data entry and initiating sustainable data

collection practices through collaborative efforts have gained prevalence in low-resource environments [23]. Furthermore, it is essential to provide context-specific adjustments and streamlined data collection tools. Outcome reporting in HICs is often burdensome due to the multitude of variables in data collection tools such as the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) with over 130 variables. Dimick demonstrated that five variables sufficiently risk-adjust procedure specific quality measures for core general surgery operations [24]. Subsequently, Ullrich identified a minimal dataset for predicting pediatric surgical perioperative mortality in low- resource settings [19].

Comparatively, ED was the section most successful at discriminating between the levels of surgical care. Unlike outcomes, there is an aspect of positive validation that accompanies the presence of training programs and external collaborations that may lead to a reporting bias [25, 26]. Historically, individual partnerships between low and high resource institutions have been the most common method of furthering the quality of care and access to surgical services [27]. The global focus on enhancing medical education may accentuate the disparity in educational resources between basic and advanced care settings, potentially introducing reporting bias due to the direct observation and reporting of educational activities.

As described, GAPS successfully discerns between the levels of surgical care, the first CAT tool of its kind to do so. Complex data does not necessarily translate to granularity and can come at the expense of utility [28]. There exist simplified models for robust scoring systems and quality-outcome databases such as the American Society of Anesthesiologists scoring system and the ACS NSQIP, respectively [28, 29]. To facilitate and streamline the use of GAPS, we abridged GAPS to include only items that significantly discriminated between levels of care (GAPS Version 3; Appendix C). The GRM IRT further supports the choice of items in GAPS Version 3.

Though GAPS is a step forward in the development of an all-encompassing evidence-based CAT, it is not without its limitations. The pilot study and validation were confined to a narrow set of countries, with a significant emphasis on the DRC, due to the availability of local advocates and our operational reach. Despite these constraints, the diversity of participating countries and income levels suggests our results have a broad applicability. Perhaps, the most important limitation of GAPS is that the assessment of internal consistency of responses was not feasible. Furthermore, as suggested by Stewart et al., previous assessments of surgical capacity have not translated to surgical output [30]. However, Stewart et al.'s study, focused on 25 essential surgical resources as proxies for broader surgical capacity, primarily emphasizing material resources [30]. We concur that traditional measures of surgical capacity, focusing on material resources, are insufficient surrogates for surgical output. Consequently, in devising the GAPS tool, we incorporated elements beyond material resources. The integration of further capacity metrics—human resources, accessibility, outcomes, and education—in GAPS represents our effort to transcend an inventory approach. However, though GAPS may be more comprehensive, its ability to directly measure surgical output remains untested. Accurately measuring surgical output necessitates a comprehensive evaluation of surgical procedures and their respective outcomes, a topic that extends beyond the scope of this current paper. This limitation reflects a crucial area for future investigations. Nevertheless, it is essential to initially address factors such as human and material resources, accessibility, education, and basic outcomes to create a foundation for objectively evaluating surgical output. Our hope is that GAPS will serve as a valuable tool in future research aimed at assessing surgical output in addition to surgical capacity.

Conclusion

GAPS was developed, piloted, and validated to distinguish pediatric surgical capacity in low-resource settings, with particular attention to discriminating BSC from ASC. GAPS Version 1, comprising 139 items, was expanded to 168 items in Version 2 after a diverse 37-member expert panel from across income strata and pediatric specialties provided feedback. The international pilot study of GAPS Version 2 indicated barriers in outcome data reporting, underscoring persistent challenges in data collection within low-resource settings. Construct validation showed that 38% of the questions effectively discriminated between levels of care, leading to a refined Version 3 with 64 questions.

GAPS has the potential to provide a framework for increasing care capacity and monitoring improvements made through international partnerships to improve the overall pediatric surgical care in low-resource settings.

Appendices

Appendix A.

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

GENERAL INFORMATION					
Date (DD/MM/YYYY):					
Country / City:					
Institution Name:					
Evaluator's Name:					
Evaluator's Position					
Type of Hospital:	☐Public	☐Private	☐NGO	☐Faith-based	
Which level of Healthcare facility do you consider yourself to be?	☐Community Facility	☐1st Level Hospital	☐2nd Level Hospital	☐3rd Level Hospital	☐National Children's Hospital

***Community facility and primary health center**: Health care facility usually located within a rural community. Provides basic health care services of very general nature. Proper triage is a key expectation. (ie. Community health center, Primary health center, Comprehensive health center)*
***1st level Hospital**: Services available are of a general nature and not specialized. Proper triage is a key expectation. (ie. General hospital, District hospital)*
***2nd level hospital:** Clinical services are specialized by function, but not all specialties and limited specialized technical equipment are available. (ie. Regional hospital, Provincial hospital, General mission hospital)*
***3rd level hospital:** Highly specialized hospital, with similarly specialized staff and technical equipment to support level of care provided. Clinical services are highly differentiated by function. Teaching is provided. Both children and adults are treated within the institution (ie. Academic/teaching/university hospital, National hospital, Central hospital, niche / specialized mission hospital)*
National Children's Hospital: Highly specialized hospital, dedicated to the care of children. Highly specialized staff and technical equipment in all children's specialties are available. Complex multidisciplinary and chronic care. Specialized teaching and research in all children's specialties provided. (i.e. Children's Hospital)

Please answer the following questions according the state of the institution in the past month.

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

HUMAN RESOURCES:					
1. Surgical capability	☐ None	☐ Non-surgeon with children's surgical expertise *(Health care professional (both MD and non-MD) that preform pediatric / neonatal surgery)*	☐ Surgeon with children's /neonatal expertise *(Surgeon with expertise in pediatric/neonatal surgery however no official /formal subspecialty pediatric surgical training)*	☐ Children's Surgeon (*Surgeon with official/ formal specialized training in children's surgery of > 1 year)*	
Pediatric General Surgeon / Urologist *Surgeon with formal specialized training in pediatric surgery of > 1 year*			☐ No	☐ Yes	☐ Occasionally
Pediatric Orthopedic Surgeon *Surgeon with formal specialized training in pediatric orthopedic surgery of > 1 year*			☐ No	☐ Yes	☐ Occasionally
Pediatric Otolaryngologist/ Maxillofacial Surgeon *Surgeon with formal specialized training in pediatric plastic surgery of > 1 year*			☐ No	☐ Yes	☐ Occasionally
Pediatric Plastic Surgeon *Surgeon with formal specialized training in pediatric plastic surgery of > 1 year*			☐ No	☐ Yes	☐ Occasionally
Pediatric Cardiac Surgeon *Surgeon with formal specialized training in pediatric cardiac surgery of > 1 year*			☐ No	☐ Yes	☐ Occasionally
Pediatric Neurosurgeon *Surgeon with formal specialized training in pediatric neurosurgery of > 1 year*			☐ No	☐ Yes	☐ Occasionally
2. Pediatric care capability	☐ None	☐ General Practitioner	☐ Pediatrician ☐ Full-time ☐ Part-time/Occasional *Medical doctor with formal specialization of > 1 year in pediatrics /neonatology*		
3. General Anesthesia Capacity for Children > 12 months old	☐ None	☐ Non-physician anesthesia provider *Healthcare professional who is not a medical doctor that performs pediatric anesthesia independently (ie. nurse anesthetist, respiratory therapist, clinical officer anesthetist)*	☐ Anesthesiologist with pediatric experience *Anesthesiologist with experience in pediatric anesthesia but no official/formal subspecialty pediatric anesthesiology training*	☐ Pediatric Anesthesiologist *Anesthesiologist with official/formal specialization of > 1 year in pediatric anesthesia*	
4. General Anesthesia Capacity for Neonates and Infants < 12 months old	☐ None	☐ Non-physician anesthesia provider *Healthcare professional who is not a medical doctor that preforms neonatal anesthesia independently (ie. Nurse*	☐ Anesthesiologist with pediatric experience *Anesthesiologist with experience in neonatal anesthesia however no*	☐ Pediatric Anesthesiologist *Anesthesiologist with official/formal specialization of > 1 year in neonatal anesthesia*	

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

		anesthetists, respiratory therapist, clinical officer anesthetist)	*official/formal subspecialty neonatal anesthesiology training.*	
5. Critical Care Capability for children and neonates	☐ None	☐ General Practitioner with intensive care experience *Medical doctor working in PICU / NICU with no official/formal specialization in pediatric/ neonatology intensive care*	☐ Pediatrician with experience *Pediatrician working in PICU / NICU with no official/formal specialization in pediatric/ neonatology intensive care*	☐ Pediatric intensivist *Pediatrician or intensivist working in PICU / NICU with formal specialization > 1 year in pediatric /neonatology intensive care*
6. Imaging Reading Capability	☐ None	☐ Physician	☐ General Radiologist ☐ On-site ☐ Distance *Radiologist with informal pediatric/neonatal expertise*	☐ Pediatric radiologist ☐ On-site ☐ Distance *Radiologist with formal specialization in pediatric/neonatal radiology*
7. Pathology Capability	☐ None	☐ General Pathologist *Medical doctor with > 1 year formal training in pathology*	☐ Pediatric Pathologist ☐ On-site ☐ Distance *Pathologist with official/formal pediatric/neonatal specialization of > 1 year*	
8. Emergency Services for children/neonates	☐ None	☐ Physician	☐ Emergency physician *Emergency physician with informal pediatric expertise*	☐ Pediatric Emergency Physician *Emergency physician with > 1 year of specialization in pediatrics*
9. What is your nurse-to-patient ratio on a surgical ward > 50% of the time?	☐ 1 nurse to >10 patients	☐ 1 nurse to 7-10 patients	☐ 1 nurse to < 6 patients	
10. Do > 50% of nurses that care for children in your institution have a bachelor's degree in nursing?	☐ No	☐ Yes		
11. Are > 50% of your nurses that care for children pediatric nurses? (*Nurse with pediatric expertise for > 1 year or formal training in pediatrics of > 6 months)*	☐ No	☐ Yes		
12. Perfusionist *a healthcare professional with formal training in using heart–lung machine during cardiac surgery or surgeries requiring cardiopulmonary bypass*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
13. Child-life/Family support Staff *Professionals who work with children and families in hospitals to help them cope with the challenges of hospitalization, illness, and disability.*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

MATERIAL RESOURCES					
Infrastructure					
1. Is this a dedicated children's hospital? *(hospital treating children only)*	☐ No	☐ Yes			
2. Is there a maternal unit attached or associated to your institution? *(a unit treating obstetrical patients)*	☐ No	☐ Yes			
3. How many hospital beds are dedicated to children's surgery?	☐ None	☐ < 5	☐ 6-10	☐ 11-15	☐ >15
4. How many operating rooms are dedicated to children's surgery?	☐ None	☐ 1-2	☐ 3-4	☐ >5	
5. How many operating rooms have the capacity to be operational after 17h00?	☐ None	☐ 1	☐ 2	☐ >2	
6. What is the maximum number of operating rooms that can be running simultaneously at any one time over a 24hour period?	☐ None	☐ 1	☐ 2	☐ >2	
7. If you need access to blood, what is the average amount of time it takes for the blood to be available? *(minutes)*	☐ Not applicable	☐ < 30	☐ 30 – 60	☐ 60 - 120	☐ > 120
8. How is medical information recorded in your institution?	☐ No medical records kept	☐ Paper only	☐ Electronic only	☐ Paper and Electronic	
9. Do you use a trauma database, and if so how often is it available?	☐ No database / Never used	☐ < 50%	☐ > 50%	☐ Always	
10. How often is a physician dedicated to the emergency department in a 24hour period?	☐ Never	☐ < 50%	☐ > 50%	☐ Always	
11. Is there an area dedicated to a Post-operative Recovery Unit (PARU) or Post-operative Anesthesia Care Unit (PACU)	☐ No	☐ Yes			
12. How often is a physician or nurse dedicated to the PARU or PACU in a 24hour period?	☐ Never	☐ < 50%	☐ > 50%	☐ Always	
13. Is there an area in the PARU / PACU dedicated to children only?	☐ No	☐ Yes			
14. Is there an area dedicated to critically ill patients? *(ie. Intensive care unit (ICU), Pediatric-ICU (PICU), neonatal-ICU(NICU))*	☐ No area dedicated to treating critically ill pediatric / neonatal patients	☐ Adult ICU treating pediatric / neonatal patients	☐ Same area for both PICU and adult ICU, but separate from NICU	☐ Separate area for both PICU and NICU	
15. How many beds in the ICU/PICU/NICU are dedicated to pediatric/neonatal patients?	☐ None	☐ 1-4	☐ 5-9	☐ > 10	

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

16. How often is a physician dedicated to the PICU/NICU in a 24-hour period?	☐ None	☐ < 50%	☐ > 50%	☐ Always
17. What is the nurse to-patient-ratio in the PICU/NICU?	☐ Not Applicable	☐ 1 nurse to > 10 patients	☐ 1 nurse to 5-9 patients	☐ 1 nurse to < 4 patients
Are the following resources available to you? Please answer according to the past month.				
18. Internet (strong enough to load a webpage)	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
19. Running water	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
20. Functioning telephone line or cell phone	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
21. External (or public) electricity	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
22. Functioning back-up generator	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
23. Functioning incinerator	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
24. Laboratory/Biochemical services	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
25. Blood bank services	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
26. Functioning x-ray machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
27. Functioning ultrasound machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
28. Functioning CT scan	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
29. Functioning anesthesia machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
Equipment				
30. How many functional neonatal incubators do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5
Are they available/working all the time?	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
31. How many functional neonatal ventilators do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5
Are they available/working all the time?	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
32. How many functional pediatric ventilators do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5
Are they available/working all the time?	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*	
33. What do you use for oxygen delivery > 50% of the time?	☐ Nothing	☐ Compressed (Cylinder)	☐ Piped oxygen	☐ Oxygen concentrator

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

34. In what percentage of operating rooms do you have access to a pulse oximeter?	☐ None	☐ < 50 %	☐ > 50%	☐ 100%
35. What suction source is used in more > 50% of cases?	☐ None	☐ Manual	☐ Electric	☐ Vacuum
36. Please check all pediatric-sized endoscopes available at your institution> 50 % of the time.	☐ Colonoscope/ Gastroscope	☐ Rigid Esophagoscop e	☐ Cystoscope	☐ Bronchoscope
37. What do you use as your primary operating room lighting source in > 50% of cases?	☐ Nothing	☐ Standing lamp	☐ Headlamp	☐ Ceiling light

Is the following equipment available at your institution? *Please answer according to the past month.*			
38. Electrocautery	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
39. Laryngoscope *(Miller ≤ 3 or Macintosh ≤ 4)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
40. Resuscitator bag valve and mask *(< 550ml bag with < size 3 mask)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
41. Pediatric endotracheal tubes *(< 6 mm)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
42. Functioning weighing scale for infants/neonates	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
43. Functioning blood pressure measuring equipment *(ie. Sphygmomanometer, electronic machine, etc.*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
44. Sterilized surgical instrument sets for open abdominal, thoracic, cardiac, neurosurgery and orthopedic surgery	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
45. Sterilized surgical instrument sets for laparoscopic surgery	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
46. Sterilized surgical instrument sets for emergency thoracotomy	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
47. Functioning autoclave *(sterilizer)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
48. Functioning apnea monitor	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
49. Cardiopulmonary bypass machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
Supplies			
Do you have access to the following supplies? Please answer according to the past month.			
50. Disinfectant hand wash	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
51. Sterilizing pre-operative solutions *(ie. Chlorhexidine, iodine, etc…)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
52. Pediatric nasogastric tubes *(< 10 Fr)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
53. Blood transfusion set	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

54. Small sutures *(≤ 4-0 or 0.15mm)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
55. Pediatric urinary catheter *(< 10 Fr)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
56. Functioning sharps disposal container	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
57. Eye protection *(goggles, safety glasses)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
58. Pediatric chest tubes *(<20 Fr)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
59. Functioning chest drainage system	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
60. Chest board/immobilization board	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
61. Pediatric central venous catheters *(< 12 cm or < 6 Fr)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
Do you have access to the following medications? Please answer according to the past month.			
62. Phenytoin	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
63. Dextrose 50%	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
64. Hypertonic saline (3%)	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
65. Local analgesic *(ie. Lidocaine, Marcaine/Bupivacaine, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
66. Parenteral Nutrition (*amino-acids, Intralipid)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
67. Intravenous narcotics *(ie. morphine, Dilaudid, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
68. Broad spectrum antibiotics *(ie. > 3rd generation cephalosporin, combination of penicillin+carbapenem, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
69. Paralytics *(ie. Rocuronium, succinylcholine, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
70. Sedatives *(ie. ketamine, midazolam, propofol, etomidate, etc..)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*
71. Vasopressors *(ie. norepinephrine, epinephrine, dopamine, dobutamine, vasopressin, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days per month)*

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

OUTCOME					
1. What is the surgical site infection rate at your institution for clean cases?	☐<1%	☐1-5%	☐6-10%	☐>10%	
2. What is the 30-day follow-up rate after major procedures?	☐None	☐< 50%	☐> 50%	☐100%	
What is the peri-operative mortality rate at your institution for the following conditions? (up to discharge…)					
Urgent					
3. Neonatal General Surgery/Urology Procedures: *Gastroschisis repair, esophageal atresia repair, neonatal intestinal obstruction, necrotizing enterocolitis,*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
4. Pediatric General Surgery/Urology Procedures: *Trauma laparotomy/thoracotomy, Intussusception &, typhoid perforation laparotomy*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
5. Pediatric Neurosurgery Procedures: *Trauma craniotomy*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
6. Pediatric Orthopedic Procedures: *Drainage of Septic Arthritis*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
7. Pediatric Plastic Surgical Procedures: *Fasciotomy*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
8. Pediatric Cardiac Procedures: *Jatene Procedure (arterial switch for transposition of the great arteries)*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
9. Pediatric Otolaryngology/ Maxillofacial Procedures:	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
Semi-Urgent					
10. Neonatal General Surgery/Urology Procedures: *Anorectal malformation repair, bladder exstrophy repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
11. Pediatric General Surgery/Urology Procedures: *Appendectomy*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
12. Pediatric Neurosurgery Procedures: *Hydrocephalus decompression (i.e. shunt), chiari decompression, myelomeningocele repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
13. Pediatric Orthopedic Procedure: *Amputation, open reduction-internal fixation for long bone fractures, debridement of osteomyelitis*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
14. Pediatric Plastic Surgery Procedures: *Escharotomy*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

15. Pediatric Cardiac Surgery Procedures: *Atrial/ventricular septal defect repair, tetralogy of fallot repair, coarctation of the aorta repair, patent ductus arteriosus repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
16. Pediatric Otolaryngologist/ Maxillofacial Procedures: *Craniofacial reconstruction*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
Not Urgent					
17. Pediatric General Surgery/Urology Procedures: *Pediatric inguinal hernia repair, Intestinal stoma closure, orchidopexy, hypospadias repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
18. Pediatric Neurosurgery Procedures: *Meningocele/lipomeningocele repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
19. Pediatric Orthopedic Surgery: *Screw fixation for slipped capital femoral epiphysis; club foot repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
20. Pediatric Plastic Surgery Procedures: *Contracture release / skin grafting, left Lip/palate Repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
21. Pediatric Cardiac Surgery Procedures:	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
22. Pediatric Otolaryngologist/ Maxillofacial Procedures:	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

ACCESSIBILITY/IMPACT					
1. How many people does your institution serve? *(in millions)*	☐ < 1	☐ 1-10	☐ 11-20	☐ > 20	
2. What percentage of patients can reach your institution within 2 hours of travel?	☐ None	☐ < 50%	☐ > 50%	☐ 100%	
3. What percentage of pediatric patients arrive by ambulance?	☐ None	☐ < 50%	☐ > 50%	☐ 100%	
4. Do you preform unplanned or emergency surgeries?	☐ No	☐ Yes			
5. Do you have the capacity to perform surgery 24 hours a day?	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*		
6. What is your average wait time for children's surgical care for elective surgery? *(months)*	☐ < 6	☐ 6 – 12	☐ 12 – 24	☐ > 24	
7. Does your institution perform minimally invasive surgery? (*ie. laparoscopy, thoracoscopy, arthroscopy etc…)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*		
Do you have access to each of the following? Please answer according to the past month.					
8. General anesthesia	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*		
9. IV sedation *(ie. ketamine, midazolam, etc...)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*		
10. Spinal anesthesia	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*		
11. Regional/local anesthesia	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*		
12. What percentage of the hospital budget is allocated to children's surgery?	☐ None	☐ < 50%	☐ >50%	☐ 100%	
13. What percentage of pediatric surgical patients have health insurance?	☐ None	☐ < 50%	☐ >50%	☐ 100%	
14. What is the highest ASA class of children operated at your institution?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
15. What is the ASA class of the majority (> 50%) of children operated at your institution?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
16. What is the wound classification of the majority of surgeries performed at your institution?	☐ Class I: Clean *An uninfected operative wound in which no inflammation is encountered and the respiratory, alimentary, genital, or uninfected urinary tract is not entered.*	☐ Class II: Clean – Contaminated *An operative wound in which the respiratory, alimentary, genital, or urinary tracts are entered under controlled conditions and without unusual contamination or evidence of infection. (i.e. biliary tract, appendix, vagina, and oropharynx)*	☐ Class III: Contaminated *Open, fresh, accidental wounds. In addition, operations with major breaks in sterile technique or gross spillage from the gastrointestinal tract, and incisions in which acute, non-purulent inflammation is encountered.*	☐ Class IV: Dirty *Old traumatic wounds with retained devitalized tissue and those that involve existing clinical infection or perforated viscera.*	
17. How often is the WHO Surgical Safety Checklist used pre-operatively?	☐ Never	☐ < 50%	☐ > 50%	☐ 100%	

Global Assessment in Pediatric Surgery (GAPS): VERSION 1

EDUCATION			
1. Are surgical residents/registrars involved in the care of pediatric surgical patients? *A surgical resident/registrar is a medical doctor that is pursuing official/formal specialization in surgery of > 1 years*	☐ No	☐ Yes	☐ Occasionally *(> 10 days per month)*
2. Are pediatric surgical fellows involved in the care of pediatric surgical patients? *A surgical fellow is a medical doctor that has completed official/formal residency training program in surgery (> 1 years) or completed junior residency training in surgery, and is pursuing a specialization or senior residency training in children's surgery*	☐ No		☐ Yes
3. Does your institution have an accredited pediatric surgical fellowship?	☐ No		☐ Yes
4. How many ongoing research projects does the department of surgery have?	☐ None	☐ 1-5	☐ > 5
5. How many papers have been published in peer-reviewed journals by the department of surgery at your institution in the past year?	☐ None	☐ 1-5	☐ > 5
6. Are morbidity and mortality rounds regularly held and attended at your institution?	☐ No	☐ Yes	☐ Occasionally *(> 3 times per year)*
7. Are oncological cases regularly discussed at an interdisciplinary tumor board?	☐ No	☐ Yes	☐ Occasionally *(> 3 times per year)*
8. Is there a regular meeting to discuss the trauma patients?	☐ No	☐ Yes	☐ Occasionally *(> 3 times per year)*
9. Do at least 50% of physicians in your department attend a minimum of 1-2 continuing medical education events per year? *(i.e. conferences, lectures, courses)*	☐ No		☐ Yes
10. Is there a clinical service dedicated solely to children's surgery?	☐ No	☐ Yes	☐ Occasionally *(operational > 10 days per month)*

Appendix B.

Appendix B. Documents detailing optimal resources for surgical care in both pediatric and adult populations.
Capacity Assessment Tool for Children
Pedi-PIPES
Children's Surgical Center Designations with Scope of Practice
Checklist for a Children's Trauma Room
Capacity Assessment Tool for Adults
WHO Tool for Situational Analysis to Assess Emergency and Essential Surgical Care (TSAAEESC)
SOSAS PIPES
Lancet Commission on Global Surgery: Hospital Assessment Tool (LCoGS HAT)
Harvard Humanitarian Initiative (HHI)
WHO Emergency and Critical Care Tool (EaCC)
Surgical Resource Guidelines
WHO Guidelines for Essential Trauma Care
Resources for Optimal Care of the Injured Patient
Global Initiative for Children’s Surgery (GICS): Optimal Resources for Children’s Surgery (OReCS)
British Association of Pediatric Surgery (BAPS): Standards for Children’s Surgery (SCC)
American Pediatric Surgical Association (APSA): Optimal Resources for Children’s Surgical Care (ORCSC)

Appendix C.

Appendix C. Summary of feedback from expert panel

Comprehensiveness:

- ❑Inclusion of politics, more sub-specialties, and LMIC-specific training.

Specificity:

- ❑More details regarding human resources, skill stratification, and resource sourcing.

Outcome Monitoring:

- ❑Emphasis on anesthesia outcomes
- ❑Presence of patient registry.

Education:

- ❑Address cultural biases
- ❑Anesthesia/nursing training
- ❑Mentorship

Overall Structure:

- ❑Improve clarity and organization

Appendix D.

Global Assessment in Pediatric Surgery (GAPS) – VERSION 2

GENERAL INFORMATION:					
Date (DD/MM/YYYY):					
Country / City:					
Institution Name:					
Evaluator's Name/Postion:					
Type of Hospital:	☐ Public/Government	☐ Private	☐ NGO		☐ Faith-based
Which level of Healthcare facility do you consider yourself to be?	☐ Community Facility	☐ 1st Level Hospital	☐ 2nd Level Hospital	☐ 3rd Level Hospital	☐ National Children's Hospital

Community facility and primary health center*: Health care facility usually located within a rural community. Provides basic health care services of very general nature. Proper triage is a key expectation. (ie. Community health center, Primary health center, Comprehensive health center)*

1st level Hospital*: Services available are of a general nature and not specialized. Proper triage is a key expectation. (ie. General hospital, District hospital)*

2nd level hospital: *Clinical services are specialized by function, but not all specialties and limited specialized technical equipment are available. (ie. Regional hospital, Provincial hospital, General mission hospital)*

3rd level hospital: *Highly specialized hospital, with similarly specialized staff and technical equipment to support level of care provided. Clinical services are highly differentiated by function. Teaching is provided. Both children and adults are treated within the institution (ie. Academic/teaching/university hospital, National hospital, Central hospital, niche / specialized mission hospital)*

National Children's Hospital: Highly specialized hospital, dedicated to the care of children. Highly specialized staff and technical equipment in all children's specialties are available. Complex multidisciplinary and chronic care. Specialized teaching and research in all children's specialties provided. (i.e. Children's Hospital)

Please answer the following questions according the state of the institution in the past month.
For all questions, please choose a single answer unless otherwise specified.

HUMAN RESOURCES:				
1. Surgical capability *Please check all that apply*	☐ None	☐ Non-surgeon with children's surgical expertise *All health care professionals that perform pediatric / neonatal surgery*	☐ Surgeon with children's /neonatal expertise *Trained surgeon with expertise in pediatric or neonatal surgery however no formal subspecialty pediatric surgical training*	☐ Children's Surgeon *Trained surgeon with formal specialized training in children's surgery of > 1 year*

To be answered only if Children's Surgeon was selected in the Q1.				
2. Pediatric General Surgeon / Urologist *Trained surgeon with formal specialized training in pediatric surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally	
3. Pediatric Orthopedic Surgeon *Trained surgeon with formal specialized training in pediatric orthopedic surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally	
4. Pediatric Otolaryngologist/ Maxillofacial Surgeon *Trained surgeon with formal specialized training in pediatric otolaryngology/ maxillofacial surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally	
5. Pediatric Plastic Surgeon Trained surgeon with formal specialized training in pediatric plastic surgery of > 1 year	☐ No	☐ Yes	☐ Occasionally	
6. Pediatric Cardiac Surgeon *Trained surgeon with formal specialized training in pediatric cardiac surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally	
7. Pediatric Neurosurgeon *Trained surgeon with formal specialized training in pediatric neurosurgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally	
8. Pediatric Ophthalmologist *Trained surgeon with formal specialized training in pediatric ophthalmology of > 1 year*	☐ No	☐ Yes	☐ Occasionally	
9. What percentage of children / neonates are treated surgically by a formally trained children's surgeon?	☐ None	☐ < 50%	☐ > 50%	☐ Always
10. Pediatric care capability *Please check all that apply*	☐ None	☐ General Practitioner	☐ Pediatrician (☐ Full-time ☐ Part-time/Occasional) *Medical doctor with formal specialization of > 1 year in pediatrics/neonatology*	

1

Global Assessment in Pediatric Surgery (GAPS) – VERSION 2

11. General Anesthesia Capacity for Children > 12 months old *Please check all that apply*	☐ None	☐ Non-physician anesthesia provider *Health-care professional (not a physician) who performs pediatric anesthesia independently and has NOT completed formal training in anesthesia*	☐ Non-physician anesthesia provider with formal training *Health-care professional (not a physician) who performs pediatric anesthesia independently and has completed formal training in anesthesia (ie. nurse anesthetist, respiratory therapist, clinical officer anesthetist)*	☐ Anesthesiologist with pediatric experience *Trained anesthesiologist with experience in pediatric anesthesia but no formal subspecialty pediatric anesthesiology training.*	☐ Pediatric Anesthesiologist *Trained anesthesiologist with formal specialization of a minimum of 1 year in pediatric anesthesia.*
12. General Anesthesia Capacity for Neonates and Infants < 12 months old *Please check all that apply*	☐ None	☐ Non-physician anesthesia provider *Health-care professional (not a physician) who performs neonatal anesthesia independently and has NOT completed formal training in anesthesia*	☐ Non-physician anesthesia provider with formal training *Health-care professional (not a physician) who performs neonatal anesthesia independently and has completed formal training in anesthesia (ie. nurse anesthetist, respiratory therapist, clinical officer anesthetist)*	☐ Anesthesiologist with neonatal experience *Trained anesthesiologist with experience in neonatal anesthesia but no formal subspecialty pediatric anesthesiology training.*	☐ Pediatric Anesthesiologist *Trained anesthesiologist with formal specialization of a minimum of 1 year in pediatric anesthesia*
13. If anesthesia is provided by non-physician personal, are these providers supervised by physician anesthesiologists?		☐ No	☐ Yes	☐ Occasionally	☐ All anesthesia provided by physician anesthesiologist
14. Is an anesthesia provider immediately present with the patient throughout cases of general anesthesia?		☐ Not Applicable	☐ No	☐ Yes	☐ Occasionally
15. Critical care capability for children and neonates *PICU = Pediatric Intensive Care Unit, NICU = Neonatal Intensive Care Unit*	☐ None	☐ General Practitioner with intensive care experience *Physician working in PICU / NICU with no formal specialization in pediatric/ neonatology intensive care*		☐ Pediatrician with experience *Pediatrician working in PICU / NICU with no formal specialization in pediatric/ neonatology intensive care*	☐ Pediatric intensivist *Physician with formal specialization of a minimum of 1 year in pediatric /neonatology intensive care*
16. How often is a physician dedicated to the PICU/NICU in a 24-hour period?		☐ None	☐ < 50%	☐ > 50%	☐ Always
17. Imaging Reading Capability *Please check all that apply*	☐ None	☐ Physician	☐ General Radiologist ☐ On-site ☐ Distance *Radiologist with informal pediatric/neonatal expertise*		☐ Pediatric radiologist ☐ On-site ☐ Distance *Radiologist with formal specialization in pediatric/neonatal radiology*
18. Pathology Capability *Please check all that apply*	☐ None		☐ General Pathologist ☐ On-site ☐ Distance *Medical doctor with > 1 year formal training in pathology*		☐ Pediatric Pathologist ☐ On-site ☐ Distance *Pathologist with official/formal pediatric/neonatal specialization of > 1 year*
19. Emergency Services for children/neonates *Please check all that apply*	☐ None	☐ Non-Physician *Healthcare professional (not a physician)*	☐ Physician	☐ Emergency physician *Trained emergency physician with no formal pediatric expertise*	☐ Pediatric Emergency Physician *Trained emergency physician with > 6 months of formal training in pediatrics*
20. How often is a physician dedicated to the emergency present in the emergency department in a 24hour period?		☐ Never	☐ < 50%	☐ > 50%	☐ Always
21. Are > 50% of your nurses that care for children pediatric nurses? *Nurse with pediatric expertise for > 1 year or formal training in pediatrics of > 6 months*			☐ No	☐ Yes	
22. Do > 50% of nurses that care for children in your institution have a bachelor's degree in nursing?			☐ No	☐ Yes	
23. What is your nurse-to-patient ratio on a surgical ward > 50% of the time?		☐ 1 nurse to < 7 patients	☐ 1 nurse to 7-10 patients	☐ 1 nurse to > 10 patients	
24. What is the nurse-to patient-ratio in the PICU-NICU? *PICU = Pediatric Intensive Care Unit, NICU = Neonatal Intensive Care Unit*		☐ Not Applicable	☐ 1 nurse to < 5 patients	☐ 1 nurse to 5-9 patients	☐ 1 nurse to > 9 patients
25. Perfusionist *A healthcare professional with formal training in using heart–lung machine during cardiac surgery or surgeries requiring cardiopulmonary bypass*		☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	

26. Child-life/Family support Staff *Professionals who work with children and families in hospitals to help them cope with the challenges of hospitalization, illness, and disability.*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*

RESOURCES

Infrastructure

1. Is this a dedicated children's hospital? *(hospital treating children only)*	☐ No	☐ Yes			
2. Is there a designated trauma bay at your institution?	☐ No	☐ Yes			
3. Does your institution have the capacity to do FAST ultrasounds and X-rays in the emergency department?	☐ No	☐ Yes			
4. Does your institution have a dedicated trauma team?	☐ No	☐ Yes			
5. Do you have trauma equipment available in your emergency room? *(chest tube, interosseuse line, endotracheal tube, etc.)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
6. Is there a maternal unit attached or associated to your institution? *(a unit treating obstetrical patients)*	☐ No	☐ Yes			
7. Does your institution have capacity to perform cesarean sections?	☐ No	☐ Yes			
8. How many hospital beds are dedicated to children's surgery?	☐ None	☐ < 5	☐ 6-10	☐ 11-15	☐ >15
9. How many operating rooms are dedicated to children's surgery?	☐ None	☐ 1-2	☐ 3-4	☐ >5	
10. How many operating rooms are available to children's surgery?	☐ None	☐ 1-2	☐ 3-4	☐ >5	
11. How many operating rooms have the capacity to be operational after 17h00?	☐ None	☐ 1	☐ 2	☐ >2	
12. What is the maximum number of operating rooms that can be running simultaneously at any one time over a 24hour period?	☐ None	☐ 1	☐ 2	☐ >2	
13. Is your institution able to provide prolonged mechanical ventilation in the operating room?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
14. Where are blood products available for your institution?	☐ Not available	☐ Staff donations as needed	☐ Off-site blood bank	☐ On-site blood bank	
15. How is medical information recorded in your institution?	☐ No medical records kept	☐ Medical records kept by patient	☐ Medical records kept by institution ☐ Paper only ☐ Electronic only ☐ Paper and Electronic		
16. Do you use a trauma database, and if so how often is it available?	☐ No database / Never used	☐ < 50%	☐ > 50%	☐ Always	
17. How often is the WHO Surgical Safety Checklist used pre-operatively?	☐ Never	☐ < 50%	☐ > 50%	☐ 100%	
18. Are the details of each anesthetic documented in an anesthesia record?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
19. Is there an area dedicated to a Post-operative Recovery Unit (PARU) or Post-operative Anesthesia Care Unit (PACU)?	☐ No	☐ Yes			
20. Is there an area in the PARU / PACU dedicated to children only?	☐ No	☐ Yes			
21. Is there an area dedicated to critically ill patients? *(ie. Intensive care unit (ICU), Pediatric-ICU (PICU), neonatal-ICU(NICU))*	☐ No area dedicated to treating critically ill pediatric / neonatal patients	☐ Adult ICU treating pediatric / neonatal patients	☐ Same area for both PICU and adult ICU, but separate from NICU	☐ Separate area for both PICU and NICU	
22. How many ICU/PICU/NICU beds are available to pediatric/neonatal patients?	☐ None	☐ 1-4	☐ 5-9	☐ > 10	
Are the following resources available to you? Please answer according to the past month.					
23. Internet *(strong enough to load a webpage)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
24. Running water	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
25. Functioning telephone line or cell phone	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
26. External (or public) electricity	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
27. Functioning back-up generator	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
28. Functioning incinerator (enclosed furnace for burning biological waste)	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		

Global Assessment in Pediatric Surgery (GAPS) – VERSION 2

29. Laboratory services	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
30. Blood bank services	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
31. Functioning x-ray machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
32. Functioning ultrasound machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
33. Functioning CT scan	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
34. Functioning anesthesia machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
Equipment				
35. How many functional neonatal incubators do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5
a) Are they available/working all the time?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
36. How many functional ventilators for use in neonates do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5
a) Are they available/working all the time?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
37. How many functional ventilators for use in children do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5
a) Are they available/working all the time?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
38. What do you use for oxygen delivery > 50% of the time in the operating room?	☐ Nothing	☐ Compressed (Cylinder)	☐ Piped oxygen	☐ Oxygen concentrator
39. Is supplemental oxygen available to children undergoing general anesthesia?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
40. What percentage of operating rooms have a dedicated pulse oximeter?	☐ None	☐ < 50 %	☐ > 50%	☐ 100%
41. What suction source is used in more > 50% of cases?	☐ None	☐ Manual	☐ Electric	☐ Vacuum
42. Please check all pediatric-sized endoscopes available at your institution> 50 % of the time.	☐ Colonoscope/ Gastroscope	☐ Rigid Esophagoscope	☐ Cystoscope	☐ Bronchoscope
43. What do you use as your primary operating room lighting source in > 50% of cases?	☐ Nothing	☐ Standing lamp	☐ Headlamp	☐ Ceiling light
44. Is there a defibrillator available in each operating room?	☐ None	☐ Available in some operating rooms	☐ Available in all operating rooms	
Is the following equipment available at your institution*? Please answer according to the past month.*				
45. Electrocautery	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
46. Functioning Laryngoscope Handle and Pediatric sized blades *(Miller ≤ 2 or Macintosh ≤ 3)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
47. Resuscitator bag valve and mask *(< 550ml bag with < size 3 mask)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
48. Pediatric endotracheal tubes *(2.5 - 6 mm)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
49. Pediatric oral airways *(000-4)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
50. Pediatric nasal airways	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
51. Functioning weighing scale for infants/neonates	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
52. Pediatric sized functioning blood pressure measuring equipment *(ie. Sphygmomanometer, electronic machine, etc.*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
53. Sterilized surgical instrument sets for open abdominal, thoracic, cardiac, neurosurgery and orthopedic surgery	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
54. Sterilized surgical instrument sets for laparoscopic surgery	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
55. Sterilized surgical instrument sets for emergency thoracotomy	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
56. Sterilized surgical instrument sets for emergency cricothyroidotomy	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	
57. Functioning autoclave *(sterilizer)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*	

Global Assessment in Pediatric Surgery (GAPS) – VERSION 2

58. Functioning apnea monitor	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
59. Cardiopulmonary bypass machine	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
60. Pulse oximeter appropriate for use with neonates/infants	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
61. Capnograph	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
62. Electric warming blanket	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
63. Depth of anesthesia monitor *(ie. BIS)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
Supplies			
Do you have access to the following supplies? Please answer according to the past month.			
64. Disinfectant hand wash	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
65. Sterilizing pre-operative solutions *(ie. Chlorhexidine, iodine, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
66. Pre-formed Silos	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
67. Pediatric nasogastric tubes *(< 10 Fr)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
68. Blood transfusion set	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
69. Small sutures *(≤ 4-0 or 0.15mm)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
70. Pediatric urinary catheter *(< 10 Fr)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
71. Functioning sharps disposal container	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
72. Eye protection *(goggles, safety glasses)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
73. Pediatric chest tubes *(<20 Fr)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
74. Functioning chest drainage system	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
75. Chest board/immobilization board	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
76. Pediatric central venous catheters *(< 12 cm or < 6 Fr)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
Do you have access to the following medications? Please answer according to the past month.			
77. Phenytoin	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
78. Dextrose 50%	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
79. Hypertonic saline (3%)	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
80. Local analgesic *(ie. Lidocaine, Marcaine/Bupivacaine, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
81. Parenteral Nutrition *(ie. amino-acids, Intralipid)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
82. Intravenous narcotics *(ie. morphine, Dilaudid, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
83. Broad spectrum antibiotics *(ie. > 3rd generation cephalosporin, combination of penicillin+carbapenem, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
84. Paralytics *(ie. Rocuronium, succinylcholine, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
85. Sedatives *(ie. ketamine, midazolam, propofol, etomidate, etc..)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
86. Vasopressors *(ie. norepinephrine, epinephrine, dopamine, dobutamine, vasopressin, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*

OUTCOME					
1. Does your institution have a method of measuring outcomes over time?		☐No		☐Yes	
Please answer the questions 2-7 in this section if you answered yes to Q1 above. If you answered no move onto Q8.					
2. What is the surgical site infection rate at your institution for clean cases?		☐<1%	☐1-5%	☐6-10%	☐>10%
3. What is the 30-day follow-up rate after major procedures?		☐None	☐< 50%	☐> 50%	☐100%
4. What is the mortality rate at discharge related to anesthesia?	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%

Global Assessment in Pediatric Surgery (GAPS) – VERSION 2

5. How often is the airway and ventilation continuously monitored while providing anesthesia?		Never	< 50%	> 50%	Always
6. How often is adequacy of ventilation monitored with capnography?		Never	< 50%	> 50%	Always
7. Is depth of anesthesia regularly assessed by depth of anesthesia monitor?		No	Yes	Occasionally	
What is the mortality rate at discharge at your institution for the following procedures?					
Urgent					
8. Neonatal General Surgery/Urology Procedures: *Gastroschisis repair, esophageal atresia repair, neonatal intestinal obstruction, necrotizing enterocolitis,*	Not Applicable	<1%	1-5%	6-10%	>10%
9. Pediatric General Surgery/Urology Procedures: *Trauma laparotomy/thoracotomy, Intussusception &, typhoid perforation laparotomy*	Not Applicable	<1%	1-5%	6-10%	>10%
10. Pediatric Neurosurgery Procedures: *Trauma craniotomy*	Not Applicable	<1%	1-5%	6-10%	>10%
11. Pediatric Orthopedic Procedures: *Drainage of Septic Arthritis*	Not Applicable	<1%	1-5%	6-10%	>10%
12. Pediatric Plastic Surgical Procedures: *Fasciotomy*	Not Applicable	<1%	1-5%	6-10%	>10%
13. Pediatric Cardiac Procedures: *Jatene Procedure (arterial switch for transposition of the great arteries)*	Not Applicable	<1%	1-5%	6-10%	>10%
14. Pediatric Otolaryngology/ Maxillofacial Procedures: *Cricothyroidotomy, tracheotomy*	Not Applicable	<1%	1-5%	6-10%	>10%
Semi-Urgent					
15. Neonatal General Surgery/Urology Procedures: *Anorectal malformation repair, bladder exstrophy repair*	Not Applicable	<1%	1-5%	6-10%	>10%
16. Pediatric General Surgery/Urology Procedures: *Appendectomy*	Not Applicable	<1%	1-5%	6-10%	>10%
17. Pediatric Neurosurgery Procedures: *Hydrocephalus decompression (i.e. shunt), chiari decompression, myelomeningocele repair*	Not Applicable	<1%	1-5%	6-10%	>10%
18. Pediatric Orthopedic Procedure: *Amputation, open reduction-internal fixation for long bone fractures, debridement of osteomyelitis*	Not Applicable	<1%	1-5%	6-10%	>10%
19. Pediatric Plastic Surgery Procedures: *Escharotomy*	Not Applicable	<1%	1-5%	6-10%	>10%
20. Pediatric Cardiac Surgery Procedures: *Unstable Atrial/ventricular septal defect repair, unstable tetralogy of fallot repair, unstable patent ductus arteriosus repair, coarctation of the aorta repair*	Not Applicable	<1%	1-5%	6-10%	>10%
21. Pediatric Otolaryngologist/ Maxillofacial Procedures: *Craniofacial reconstruction*	Not Applicable	<1%	1-5%	6-10%	>10%
Not Urgent					
22. Pediatric General Surgery/Urology Procedures: *Pediatric inguinal hernia repair, Intestinal stoma closure, orchidopexy, hypospadias repair*	Not Applicable	<1%	1-5%	6-10%	>10%
23. Pediatric Neurosurgery Procedures: *Meningocele/lipomeningocele repair*	Not Applicable	<1%	1-5%	6-10%	>10%

Global Assessment in Pediatric Surgery (GAPS) – VERSION 2

24. Pediatric Orthopedic Surgery: *Screw fixation for slipped capital femoral epiphysis; club foot repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
25. Pediatric Plastic Surgery Procedures: *Contracture release / skin grafting, left Lip/palate Repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
26. Pediatric Cardiac Surgery Procedures: *Stable Atrial/ventricular septal defect repair, stable tetralogy of fallot repair, stable patent ductus arteriosus repair*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%
27. Pediatric Otolaryngologist/ Maxillofacial Procedures: *Tonsilectomy, adenoidectomy, cervical mass excision*	☐Not Applicable	☐<1%	☐1-5%	☐6-10%	☐>10%

ACCESSIBILITY					
1. How many people does your institution serve? *(in millions)*	☐< 1	☐1-10	☐11-20	☐> 20	
2. What percentage of patients can reach your institution within 2 hours of travel?	☐None	☐Unknown	☐< 50%	☐> 50%	☐100%
3. What percentage of pediatric patients arrive by ambulance?	☐None	☐Unknown	☐< 50%	☐> 50%	☐100%
4. Do you perform unplanned or emergency surgeries?	☐No	☐Yes			
5. Do you have the capacity to perform surgery 24 hours a day?	☐No	☐Yes	☐Occasionally *(> 10 days/month)*		
6. What is your average wait time for children's surgical care for elective surgery? *(months)*	☐< 6	☐6 – 12	☐12 – 24	☐> 24	
7. Does your institution perform minimally invasive surgery? *(ie. laparoscopy, thoracoscopy, arthroscopy etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*		
Do you have access to each of the following? Please answer according to the past month.					
8. General anesthesia	☐No	☐Yes	☐Occasionally *(> 10 days/month)*		
9. IV sedation *(ie. ketamine, midazolam, etc…)*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*		
10. Spinal anesthesia of any form	☐No	☐Yes	☐Occasionally *(> 10 days/month)*		
11. Regional blocks	☐No	☐Yes	☐Occasionally *(> 10 days/month)*		
12. What percentage of the hospital budget is allocated to children's surgery?	☐None	☐Unknown	☐< 50%	☐>50%	☐100%
13. What percentage of pediatric surgical patients have health insurance?	☐None	☐Unknown	☐< 50%	☐>50%	☐100%
14. What is the highest ASA class of children operated at your institution?	☐1	☐2	☐3	☐4	☐5

EDUCATION			
1. Is your institution affiliated with a university or medical school	☐No	☐Yes	
2. Are surgical residents/registrars involved in the care of pediatric surgical patients? *A surgical resident/registrar is a medical doctor that is pursuing formal specialization in surgery of > 1 year*	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
3. Are surgeons undergoing formal training in children's surgery involved in the care of pediatric surgical patients?	☐No	☐Yes	☐Occasionally *(> 10 days/month)*
4. Does your institution have a formal surgical residency/registrar training program?	☐No	☐Yes	
5. Does your institution have any formal fellowship in children's surgical specialties?	☐No	☐Yes	
6. Does your institution have a formal anesthesiology training program?	☐No	☐Yes	
7. Is your institution involved in a formal training program for surgical trainees?	☐No	☐Yes	
8. Is maintenance of competence / professional development available for the care of children? *(check all that apply)*	☐No	☐Yes ☐Surgeons ☐Anesthesia ☐Nurses ☐Other ancillary staff	
9. Do at least 50% of health-care professionals involved in the care of children attend a minimum of 1-2 continuing medical education events per year? *(i.e. conferences, lectures, courses)*	☐No	☐Yes	

Global Assessment in Pediatric Surgery (GAPS) – VERSION 2

10. How many ongoing research projects does the department of surgery have?	☐None	☐1-5	☐> 5
11. How many papers have been published in peer-reviewed journals by the department of surgery at your institution in the past year?	☐None	☐1-5	☐> 5
12. Are morbidity and mortality rounds regularly held and attended at your institution?	☐No	☐Yes	☐Occasionally *(> 3 times/year)*
13. Are oncological cases regularly discussed at an interdisciplinary tumor board?	☐No	☐Yes	☐Occasionally *(> 3 times/year)*
14. Is there a regular meeting to discuss care of pediatric trauma patients?	☐No	☐Yes	☐Occasionally *(> 3 times/year)*
15. Is there a clinical service dedicated solely to children's surgery?	☐No	☐Yes	☐Occasionally *(operational > 10 days/month)*

Appendix E.

Global Assessment in Pediatric Surgery (GAPS) – VERSION 3

GENERAL INFORMATION:

Date (DD/MM/YYYY):					
Country / City:					
Institution Name:					
Evaluator's Name/Postion:					
Type of Hospital:	☐ Public/Government	☐ Private	☐ NGO		☐ Faith-based
Which level of Healthcare facility do you consider yourself to be?	☐ Community Facility	☐ 1st Level Hospital	☐ 2nd Level Hospital	☐ 3rd Level Hospital	☐ National Children's Hospital

Community facility and primary health center*: Health care facility usually located within a rural community. Provides basic health care services of very general nature. Proper triage is a key expectation. (ie. Community health center, Primary health center, Comprehensive health center)*
1st level Hospital*: Services available are of a general nature and not specialized. Proper triage is a key expectation. (ie. General hospital, District hospital)*
2nd level hospital: *Clinical services are specialized by function, but not all specialties and limited specialized technical equipment are available. (ie. Regional hospital, Provincial hospital, General mission hospital)*
3rd level hospital: *Highly specialized hospital, with similarly specialized staff and technical equipment to support level of care provided. Clinical services are highly differentiated by function. Teaching is provided. Both children and adults are treated within the institution (ie. Academic/teaching/university hospital, National hospital, Central hospital, niche / specialized mission hospital)*
National Children's Hospital: Highly specialized hospital, dedicated to the care of children. Highly specialized staff and technical equipment in all children's specialties are available. Complex multidisciplinary and chronic care. Specialized teaching and research in all children's specialties provided. (i.e. Children's Hospital)

Please answer the following questions according the state of the institution in the past month.S
For all questions, please choose a single answer unless otherwise specified.

HUMAN RESOURCES:

1. Surgical capability *Please check all that apply*
   - ☐ None
   - ☐ Non-surgeon with children's surgical expertise *All health care professionals that perform pediatric / neonatal surgery*
   - ☐ Surgeon with children's /neonatal expertise *Trained surgeon with expertise in pediatric or neonatal surgery however no formal subspecialty pediatric surgical training*
   - ☐ Children's Surgeon *Trained surgeon with formal specialized training in children's surgery of > 1 year*

To be answered only if Children's Surgeon was selected in the Q1.

2. Pediatric General Surgeon / Urologist *Trained surgeon with formal specialized training in pediatric surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally
3. Pediatric Orthopedic Surgeon *Trained surgeon with formal specialized training in pediatric orthopedic surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally
4. Pediatric Otolaryngologist/ Maxillofacial Surgeon *Trained surgeon with formal specialized training in pediatric otolaryngology/ maxillofacial surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally
5. Pediatric Cardiac Surgeon *Trained surgeon with formal specialized training in pediatric cardiac surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally

6. Pediatric care capability *Please check all that apply*
   - ☐ None
   - ☐ General Practitioner
   - ☐ Pediatrician
     - ☐ Full-time
     - ☐ Part-time/Occasional

     Medical doctor with formal specialization of > 1 year in pediatrics/neonatology

7. General Anesthesia Capacity for Children > 12 months old *Please check all that apply*
   - ☐ None
   - ☐ Non-physician anesthesia provider *Health-care professional (not a physician) who performs pediatric anesthesia independently and has NOT completed formal training in anesthesia*
   - ☐ Non-physician anesthesia provider with formal training *Health-care professional (not a physician) who performs pediatric anesthesia independently and has completed formal training in anesthesia (ie. nurse anesthetist, respiratory therapist, clinical officer anesthetist)*
   - ☐ Anesthesiologist with pediatric experience *Trained anesthesiologist with experience in pediatric anesthesia but no formal subspecialty pediatric anesthesiology training.*
   - ☐ Pediatric Anesthesiologist *Trained anesthesiologist with formal specialization of a minimum of 1 year in pediatric anesthesia.*

1

Global Assessment in Pediatric Surgery (GAPS) – VERSION 3

8. Critical care capability for children and neonates *PICU = Pediatric Intensive Care Unit, NICU = Neonatal Intensive Care Unit*	☐ None	☐ General Practitioner with intensive care experience *Physician working in PICU / NICU with no formal specialization in pediatric/ neonatology intensive care*	☐ Pediatrician with experience *Pediatrician working in PICU / NICU with no formal specialization in pediatric/ neonatology intensive care*	☐ Pediatric intensivist *Physician with formal specialization of a minimum of 1 year in pediatric /neonatology intensive care*	
9. Emergency Services for children/neonates *Please check all that apply*	☐ None	☐ Non-Physician *Healthcare professional (not a physician)*	☐ Physician	☐ Emergency physician *Trained emergency physician with no formal pediatric expertise*	☐ Pediatric Emergency Physician *Trained emergency physician with > 6 months of formal training in pediatrics*

RESOURCES					
Infrastructure					
1. Is this a dedicated children's hospital? *(hospital treating children only)*	☐ No	☐ Yes			
2. Does your institution have the capacity to do FAST ultrasounds and X-rays in the emergency department?	☐ No	☐ Yes			
3. Is there a maternal unit attached or associated to your institution? *(a unit treating obstetrical patients)*	☐ No	☐ Yes			
4. Does your institution have capacity to perform cesarean sections?	☐ No	☐ Yes			
5. How many hospital beds are dedicated to children's surgery?	☐ None	☐ < 5	☐ 6-10	☐ 11-15	☐ >15
6. How many operating rooms are dedicated to children's surgery?	☐ None	☐ 1-2	☐ 3-4	☐ >5	
7. How many operating rooms are available to children's surgery?	☐ None	☐ 1-2	☐ 3-4	☐ >5	
8. Is your institution able to provide prolonged mechanical ventilation in the operating room?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
9. Is there an area dedicated to a Post-operative Recovery Unit (PARU) or Post-operative Anesthesia Care Unit (PACU)?	☐ No	☐ Yes			
10. Is there an area in the PARU / PACU dedicated to children only?	☐ No	☐ Yes			
11. Is there an area dedicated to critically ill patients? *(ie. Intensive care unit (ICU), Pediatric-ICU (PICU), neonatal-ICU(NICU))*	☐ No area dedicated to treating critically ill pediatric / neonatal patients	☐ Adult ICU treating pediatric / neonatal patients	☐ Same area for both PICU and adult ICU, but separate from NICU	☐ Separate area for both PICU and NICU	
12. How many ICU/PICU/NICU beds are available to pediatric/neonatal patients?	☐ None	☐ 1-4	☐ 5-9	☐ > 10	
Are the following resources available to you? Please answer according to the past month.					
13. Internet *(strong enough to load a webpage)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
14. External (or public) electricity	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
15. Functioning CT scan	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
16. Functioning anesthesia machine	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
Equipment					
17. How many functional neonatal incubators do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5	
18. How many functional ventilators for use in neonates do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5	
19. How many functional ventilators for use in children do you have access to?	☐ None	☐ 1-2	☐ 3-4	☐ > 5	
20. Is supplemental oxygen available to children undergoing general anesthesia?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
21. Is there a defibrillator available in each operating room?	☐ None	☐ Available in some operating rooms	☐ Available in all operating rooms		
Is the following equipment available at your institution*? Please answer according to the past month.*					
22. Electrocautery	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
23. Functioning Laryngoscope Handle and Pediatric sized blades *(Miller ≤ 2 or Macintosh ≤ 3)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
24. Pediatric endotracheal tubes *(2.5 - 6 mm)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		

Global Assessment in Pediatric Surgery (GAPS) – VERSION 3

25. Pediatric oral airways *(000-4)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
26. Pediatric nasal airways	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
27. Sterilized surgical instrument sets for open abdominal, thoracic, cardiac, neurosurgery and orthopedic surgery	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
28. Sterilized surgical instrument sets for emergency thoracotomy	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
29. Functioning autoclave *(sterilizer)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
30. Capnograph	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
31. Electric warming blanket	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
32. Depth of anesthesia monitor *(ie. BIS)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
Supplies			
Do you have access to the following supplies? Please answer according to the past month.			
33. Pediatric chest tubes *(<20 Fr)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
34. Functioning chest drainage system	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
Do you have access to the following medications? Please answer according to the past month.			
35. Anticonvulsants	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
36. Hypertonic saline (3%)	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
37. Parenteral Nutrition *(ie. amino-acids, Intralipid)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
38. Paralytics *(ie. Rocuronium, succinylcholine, etc…)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
39. Vasopressors *(ie. norepinephrine, epinephrine, dopamine, dobutamine, vasopressin, etc…)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*

OUTCOME					
1. What is the 30-day follow-up rate after major procedures?	☐ None	☐ < 50%	☐ > 50%	☐ 100%	
2. How often is the airway and ventilation continuously monitored while providing anesthesia?		☐ Never	☐ < 50%	☐ > 50%	☐ Always
3. How often is adequacy of ventilation monitored with capnography?		☐ Never	☐ < 50%	☐ > 50%	☐ Always

ACCESSIBILITY					
1. How many people does your institution serve? *(in millions)*	☐ < 1	☐ 1-10	☐ 11-20	☐ > 20	
2. Does your institution perform minimally invasive surgery? *(ie. laparoscopy, thoracoscopy, arthroscopy etc…)*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*		
3. What is the highest ASA class of children operated at your institution?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

EDUCATION			
1. Is your institution affiliated with a university or medical school	☐ No	☐ Yes	
2. Are surgical residents/registrars involved in the care of pediatric surgical patients? *A surgical resident/registrar is a medical doctor that is pursuing formal specialization in surgery of > 1 year*	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
3. Are surgeons undergoing formal training in children's surgery involved in the care of pediatric surgical patients?	☐ No	☐ Yes	☐ Occasionally *(> 10 days/month)*
4. Does your institution have a formal surgical residency/registrar training program?	☐ No	☐ Yes	
5. Does your institution have any formal fellowship in children's surgical specialties?	☐ No	☐ Yes	
6. Does your institution have a formal anesthesiology training program?	☐ No	☐ Yes	

Global Assessment in Pediatric Surgery (GAPS) – VERSION 3

7. Do at least 50% of health-care professionals involved in the care of children attend a minimum of 1-2 continuing medical education events per year? *(i.e. conferences, lectures, courses)*	☐No	☐Yes	
8. How many ongoing research projects does the department of surgery have?	☐None	☐1-5	☐> 5
9. Are oncological cases regularly discussed at an interdisciplinary tumor board?	☐No	☐Yes	☐Occasionally *(> 3 times/year)*
10. Is there a clinical service dedicated solely to children's surgery?	☐No	☐Yes	☐Occasionally *(operational > 10 days/month)*

Appendix F.

Appendix F. Discrimination Parameters Across Items			
	Discrimination		
	Coefficient	P value	95% Confidence Interval
Accessibility			
AC2	1.28	0.001	0.549 - 2.014
AC1	1.34	< 0.001	0.611 - 2.077
AC3	1.65	< 0.001	0.895 - 2.415
Education			
ED9	1.05	0.003	0.365 - 1.738
ED1	1.30	0.001	0.499 - 2.101
ED7	1.46	0.002	0.546 - 2.375
ED2	1.50	< 0.001	0.666 - 2.336
ED4	1.53	0.001	0.588 - 2.469
ED8	1.74	< 0.001	0.841 - 2.643
ED5	2.10	0.004	0.679 - 3.520
ED6	2.21	0.003	0.758 - 3.660
ED3	2.26	< 0.001	1.072 - 3.448
ED10	3.27	0.001	1.372 - 5.175
Human Resources			
HR2	-0.07	0.928	-1.594 - 1.453
HR3	0.66	0.221	-0.398 - 0.725
HR4	1.22	0.083	-0.160 - 2.607
HR8	1.69	<0.001	0.941 - 2.446
HR5	1.94	0.13	-0.054 - 4.413
HR1	2.02	<0.001	1.112 - 2.943
HR6	2.52	<0.001	1.207 - 3.828
HR7	2.76	<0.001	1.674 - 3.837
HR9	3	<0.001	1.747 - 4.255
Material Resources			
MR4	-2.96	0.017	5.392 - 0.533
MR3	-1.26	0.005	(-2.141) – (-0.375)
MR9	0.69	0.050	0.0004 – 1.377
MR35	0.72	0.016	0.132 – 1.312
MR36	0.74	0.012	0.161 – 1.321
MR32	0.99	0.014	0.204 – 1.772
MR2	1.00	0.009	0.245 – 1.761

MR21	1.11	0.001	0.473 – 1.743
MR12	1.15	<0.001	0.559 – 1.747
MR29	1.21	0.014	0.244 – 2.171
MR27	1.23	0.003	0.425 – 2.042
MR5	1.34	<0.001	0.667 – 2.013
MR11	1.44	<0.001	0.734 – 2.137
MR34	1.54	0.001	0.663 – 2.427
MR33	1.63	<0.001	0.741 – 2.529
MR8	1.77	0.001	0.683 – 2.865
MR37	1.78	0.001	0.747 – 2.813
MR15	1.80	0.001	0.742 – 2.848
MR30	1.80	0.002	0.661 – 2.942
MR31	1.89	0.001	0.759 – 3.023
MR26	2.03	0.001	0.849 – 3.213
MR6	2.10	<0.001	1.164 – 3.039
MR13	2.28	<0.001	1.110 – 3.451
MR1	2.32	0.004	0.736 – 3.895
MR20	2.40	<0.001	1.104 – 3.685
MR7	2.44	<0.001	1.414 – 3.470
MR22	2.45	<0.001	1.072 – 3.833
MR28	2.46	0.001	1.046 – 3.879
MR17	2.52	<0.001	1.404 – 3.643
MR14	3	<0.001	1.360 – 4.638
MR23	3.02	0.001	1.265 – 4.770
MR19	3.07	<0.001	1.686 – 4.461
MR18	3.10	<0.001	1.673 – 4.535
MR24	3.39	<0.001	1.584 – 5.198
MR25	3.69	<0.001	1.691 – 5.689
MR16	4.17	<0.001	1.897 – 6.451
MR39	4.24	0.002	1.526 – 6.963
MR38	5.04	< 0.001	2.262 – 7.817
MR10	5.17	0.063	(-0.273769) – 10.604
Outcomes			
OU1	0.90	0.014	0.184 – 1.621
OU3	0.97	0.008	0.252 – 1.685
OU2	1.31	0.001	0.513 – 2.108

References

1. Meara JG, Leather AJ, Hagander L, Alkire BC, Alonso N, Ameh EA, Bickler SW, Conteh L, Dare AJ, Davies J, Merisier ED, El-Halabi S, Farmer PE, Gawande A, Gillies R, Greenberg SL, Grimes CE, Gruen RL, Ismail EA, Kamara TB, Lavy C, Lundeg G, Mkandawire NC, Raykar NP, Riesel JN, Rodas E, Rose J, Roy N, Shrime MG, Sullivan R, Verguet S, Watters D, Weiser TG, Wilson IH, Yamey G , Yip W (2016) Global surgery 2030: Evidence and solutions for achieving health, welfare, and economic development. Int J Obstet Anesth 25:75-78. https://10.1016/j.ijoa.2015.09.006.ORG/
2. Meara JG, Hagander L , Leather AJM (2014) Surgery and global health: A lancet commission. Lancet 383:12-13. https://10.1016/S0140-6736(13)62345-4.ORG/
3. Makasa EM (2014) Letter to global health agency leaders on the importance of surgical indicators. Lancet 384:1748. https://10.1016/S0140-6736(14)62012-2.ORG/
4. Huber B (2015) Finding surgery's place on the global health agenda. Lancet 385:1821-1822. https://10.1016/S0140-6736(15)60761-9.ORG/
5. Dare AJ, Grimes CE, Gillies R, Greenberg SL, Hagander L, Meara JG , Leather AJ (2014) Global surgery: Defining an emerging global health field. Lancet 384:2245-2247. https://10.1016/S0140-6736(14)60237-3.ORG/
6. Wright NJ , Global PaedSurg Research C (2019) Management and outcomes of gastrointestinal congenital anomalies in low, middle and high income countries: Protocol for a multicentre, international, prospective cohort study. BMJ Open 9:e030452. https://10.1136/bmjopen-2019-030452.ORG/
7. Okoye MT, Ameh EA, Kushner AL , Nwomeh BC (2015) A pilot survey of pediatric surgical capacity in west africa. World J Surg 39:669-676. https://10.1007/s00268-014-2868-5.ORG/
8. Niconchuk JA , Newton MW (2022) Global pediatric surgery and anesthesia inequities: How do we have a global effort? Curr Opin Anaesthesiol 35:351-356. https://10.1097/ACO.0000000000001122.ORG/
9. Krishnaswami S, Nwomeh BC , Ameh EA (2016) The pediatric surgery workforce in low- and middle-income countries: Problems and priorities. Semin Pediatr Surg 25:32-42. https://10.1053/j.sempedsurg.2015.09.007.ORG/

10. Butler MW (2016) Developing pediatric surgery in low- and middle-income countries: An evaluation of contemporary education and care delivery models. Semin Pediatr Surg 25:43-50. https://10.1053/j.sempedsurg.2015.09.008.ORG/

11. Wimmer S, Truche P, Guadagno E, Ameh E, Samad L, Makasa EMM, Greenberg S, Meara JG, van Dijk TH , Poenaru D (2023) Assessing the inclusion of children's surgical care in national surgical, obstetric and anaesthesia plans: A policy content analysis. BMJ Open 13:e051248. https://10.1136/bmjopen-2021-051248.ORG/

12. Yousef Y, St-Louis E, Baird R, Smith ER, Guadagno E, St-Vil D , Poenaru D (2019) A systematic review of capacity assessment tools in pediatric surgery: Global assessment in pediatric surgery (gaps) phase i. J Pediatr Surg 54:831-837. https://10.1016/j.jpedsurg.2018.11.005.ORG/

13. Mock CN, Donkor P, Gawande A, Jamison DT, Kruk ME, Debas HT , Group DCPESA (2015) Essential surgery: Key messages from disease control priorities, 3rd edition. Lancet 385:2209-2219. https://10.1016/S0140-6736(15)60091-5.ORG/

14. Goodman LF, St-Louis E, Yousef Y, Cheung M, Ure B, Ozgediz D, Ameh EA, Bickler S, Poenaru D, Oldham K, Farmer D, Lakhoo K , Collaborators G (2018) The global initiative for children's surgery: Optimal resources for improving care. Eur J Pediatr Surg 28:51-59. https://10.1055/s-0037-1604399.ORG/

15. Global Initiative for Children's S (2019) Optimal resources for children's surgical care: Executive summary. World J Surg 43:978-980. https://10.1007/s00268-018-04888-7.ORG/

16. World bank country and lending groups. World Bank Group. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

17. Debas HT, Donkor P, Gawande A, Jamison DT, Kruk ME and Mock CN (2015) In: Essential surgery: Disease control priorities, third edition (volume 1), The International Bank for Reconstruction and Development / The World Bank Washington (DC),

18. Greenberg SL, Ng-Kamstra JS, Ameh EA, Ozgediz DE, Poenaru D , Bickler SW (2016) An investment in knowledge: Research in global pediatric surgery for the 21st century. Semin Pediatr Surg 25:51-60. https://10.1053/j.sempedsurg.2015.09.009.ORG/

19. Ullrich SJ, Kisa P, Muzira A, Kakembo N, Nabukenya M, Tumukunde J, Sekabira J, Chang DC , Ozgediz D (2022) Pediatric surgical quality improvement in low- and middle-income

countries: What data to collect? Surgery 171:1067-1072. https://10.1016/j.surg.2021.09.010.ORG/
20. Riviello R , Scott JW (2016) Closing the data gaps for surgical care delivery in lmics. Lancet Glob Health 4:e138-139. https://10.1016/S2214-109X(16)00042-5.ORG/
21. Ng-Kamstra JS, Greenberg SL, Kotagal M, Palmqvist CL, Lai FY, Bollam R, Meara JG , Gruen RL (2015) Use and definitions of perioperative mortality rates in low-income and middle-income countries: A systematic review. Lancet 385 Suppl 2:S29. https://10.1016/S0140-6736(15)60824-8.ORG/
22. Vicente-Crespo M, Agunbiade O, Eyers J, Thorogood M , Fonn S (2020) Institutionalizing research capacity strengthening in lmics: A systematic review and meta-synthesis. AAS Open Res 3:43. https://10.12688/aasopenres.13116.3.ORG/
23. Knight SR, Ots R, Maimbo M, Drake TM, Fairfield CJ , Harrison EM (2019) Systematic review of the use of big data to improve surgery in low- and middle-income countries. Br J Surg 106:e62-e72. https://10.1002/bjs.11052.ORG/
24. Dimick JB, Osborne NH, Hall BL, Ko CY , Birkmeyer JD (2010) Risk adjustment for comparing hospital quality with surgery: How many variables are needed? J Am Coll Surg 210:503-508. https://10.1016/j.jamcollsurg.2010.01.018.ORG/
25. Baird R, Kisa P, Muzira A, Wesonga AS, Sekabira J, Butterworth S, MacNeily AE, Duffy D , Blair G (2022) Training low-middle-income (lmic) pediatric surgeons in a high-income country (hic) program. Pediatr Surg Int 38:269-276. https://10.1007/s00383-021-05027-5.ORG/
26. Pulvirenti R, Gortan M, Cumba D, Gamba P , Tognon C (2022) Pediatric surgery and anesthesia in low-middle income countries: Current situation and ethical challenges. Front Pediatr 10:908699. https://10.3389/fped.2022.908699.ORG/
27. Cunningham AJ, Stephens CQ, Ameh EA, Mshelbwala P, Nwomeh B , Krishnaswami S (2019) Ethics in global pediatric surgery: Existing dilemmas and emerging challenges. World J Surg 43:1466-1473. https://10.1007/s00268-019-04975-3.ORG/
28. Anderson GA, Bohnen J, Spence R, Ilcisin L, Ladha K , Chang D (2018) Data improvement through simplification: Implications for low-resource settings. World J Surg 42:2725-2731. https://10.1007/s00268-018-4535-8.ORG/

29. Farmer DL (2012) Nsqip lite: A potential tool for global comparative effectiveness evaluations. Arch Surg 147:803-804. https://10.1001/archsurg.2012.899.ORG/

30. Stewart BT, Gyedu A, Gaskill C, Boakye G, Quansah R, Donkor P, Volmink J , Mock C (2018) Exploring the relationship between surgical capacity and output in ghana: Current capacity assessments may not tell the whole story. World J Surg 42:3065-3074. https://10.1007/s00268-018-4589-7.ORG/

CHAPTER 5: VALIDATION OF THE GAPS TOOL

PREFACE

Manuscript 2 describes the development and validation of the GAPS, an assessment tool specifically designed to evaluate pediatric surgical capacity in L-MICs. Pediatric surgical care in these regions is often hampered by significant gaps in infrastructure, healthcare resources, and specialized training, contributing to disparities in surgical outcomes[7, 14, 22, 49-55]. Addressing this need, the GAPS tool was created to systematically measure and prioritize critical resource and care gaps in diverse L-MIC settings. Initially, GAPS Version 1 consisted of 139 items across domains such as human resources, material resources, accessibility, outcomes, and education. Through expert consultation and pilot testing in 65 institutions across multiple countries, the tool was refined to its current form, GAPS Version 3, which includes 64 questions shown to effectively differentiate between basic and advanced surgical care settings. The tool's validation process demonstrated its reliability in assessing hospital capacity and its potential to inform improvements in pediatric surgical care across L-MICs. The study underscores the ongoing challenges in data collection within low-resource settings and the importance of adapting assessment tools to local contexts. GAPS thus represents a structured, evidence-based approach to guide resource allocation and track advancements in pediatric surgical capacity where they are most needed.

Following the development of the GAPS tool in Manuscript 2, the subsequent phase, detailed in Manuscript 3, focuses on refining the tool through the addition of capacity-based weighting. Manuscript 2 established the GAPS tool as an evidence-based framework that effectively identifies and measures gaps in pediatric surgical care by focusing on key domains such as human resources, material resources, outcomes, accessibility, and education. This unweighted approach allowed a preliminary yet impactful differentiation between basic and advanced surgical care settings.

In developing the GAPS tool's weighting strategy, the decision to apply equal weight to each item reflects the principles articulated by Streiner et al and DeVellis et al [34, 35]. Both authors emphasize that differential weighting in extensive scales—with over 40 items—offers limited predictive enhancement but significantly increases computational complexity, often without improving interpretive clarity. Within the GAPS tool, each domain—human resources, material resources, education, accessibility, and outcomes—is composed of homogenous items where weights naturally fall within a narrow range, minimizing the potential impact of differential weighting. According to Streiner et al., in homogenous scales, the benefit of weighting becomes marginal. Within each of the five subsections of GAPS (human resources, material resources, education, accessibility, and outcome), the scale is comprised of homogenous items, thus the weights will fall within a narrow range rendering the impact of differential weighting minimal. However, because each of the five subsections themselves represent distinctly different aspects of pediatric surgical capacity, it was methodologically consistent to assign a weighted score to each subsection. This approach acknowledges the distinct importance each domain brings to the overall capacity assessment, while minimizing the added complexity of differential weighting and maintaining a clear, structured foundation for effective scoring.

The model's overall fit was evaluated through the likelihood ratio chi-squared test, yielding a significant fit (LR $\chi^2(5) = 65.05$, $p < 0.0001$), while the pseudo R^2 value (0.5078) indicated a moderate relationship between the predictors and the ordinal outcome, effectively supporting the rationale for equal weighting. In calculating the subsection scores, substantial ranges emerged: human resources (0–1,080), material resources (0–4,680), accessibility (54–360), education (0–1,200), and outcomes (0–360), reflecting the scale's complexity and the varied resource needs across domains. Importantly, outcomes contributed most significantly to the model, aligning with the priority of patient impact, while material resources were the only statistically significant predictor. This approach underlines DeVellis et al's perspective that equal weighting avoids unnecessary complexity while simultaneously allowing each domain to be represented equitably, thus aligning with the scale's purpose without compromising analytical rigor[34].

In Manuscript 3, the GAPS tool matures by incorporating a weighting system that scales each domain according to its contribution to overall surgical capacity. This nuanced approach captures the multifaceted demands of pediatric surgical services in L-MICs, enabling the GAPS tool to more accurately reflect a hospital's capacity relative to its specific limitations across the five domains examined. The weighted GAPS tool not only supports a more comprehensive assessment but also enhances its utility as a metric for objectively directing targeted interventions and partnerships within L-MIC settings. Through this refined scoring, Phase III aligns GAPS with complex global health objectives, supporting a precise evaluation of pediatric surgical capacity that can guide strategic improvements across diverse resource-limited environments.

Manuscript 3 has been published under the following reference[56]: Yousef Y, Ameh E, Kalisya LM, Poenaru D. GAPS Phase III: incorporation of capacity-based weighting in the global assessment for pediatric surgery. Pediatr Surg Int. 2024 Nov 6;40(1):292. doi: 10.1007/s00383-024-05870-2. PMID: 39503855.

MANUSCRIPT 3: GAPS PHASE III: INCORPORATION OF CAPACITY BASED WEIGHTING IN THE GLOBAL ASSESSMENT FOR PEDIATRIC SURGERY

Yasmine Yousef[1, 2, 3], Emmanuel Ameh[4], Luc Malemo Kalisya[5], Dan Poenaru[1, 3]

[1]Harvey E. Beardmore Division of Pediatric Surgery, The Montreal Children's Hospital, McGill University Health Centre

[2] Joe Di Maggio Children's Hospital, Memorial Health System, Hollywood, FL, USA

[3] Faculty of Medicine, McGill University, Montreal, QC, CANADA

[4] Division of Paediatric Surgery, Department of Surgery, National Hospital, Abuja, Nigeria

[5] Department of Community Health and Development, Great Lakes University of Kisumu, Kisumu, Kenya.

KEYWORDS: pediatric surgery, children's surgery, global surgery, global health, low-income country, middle-income country, capacity assessment

FUNDING :

Fonds de Recherche du Québec - Santé (FRQS)

Canadian Institutes of Health Research

CONFLICT OF INTEREST STATEMENT BY AUTHORS:

Yasmine Yousef – NONE

Emmanuel Ameh – NONE

Luc Malemo Kalisya – NONE

Dan Poenaru – NONE

Abstract

<u>*Introduction*</u>

The Global Assessment for Pediatric Surgery (GAPS) tool was developed to enhance pediatric surgical care in Low- and Middle-Income Countries. This study presents the addition of a capacity-based weighting system to the GAPS tool.

<u>*Methods*</u>

GAPS, developed through a multi-phase process including systematic review, international testing, item analysis, and refinement, assesses 64 items across five domains: human resources, material resources, education, accessibility, and outcomes. This new weighting system differentially weighs each domain. The GAPS Score was evaluated using pilot study data, focusing on hospital and country income levels, human development index, under-five mortality rate, neonatal mortality rate, deaths due to injury and deaths due to congenital anomalies. Analysis involved the Kruskal-Wallis test and linear regression. Benchmark values for the GAPS overall score and subsection scores were identified.

<u>*Results*</u>

The GAPS score's capacity-based weighting system effectively discriminated between levels of hospital ($p=0.0001$) and country income level ($p=0.002$). The GAPS scores showed significant associations with human development index ($p < 0.001$) and key health indicators such as under-five mortality rates ($p < 0.001$), neonatal mortality rate ($p < 0.001$), and deaths due to injury ($p < 0.001$). Benchmark scores for the GAPS overall score and the subsection scores included most institutions within their respective hospital level.

<u>*Conclusions*</u>

The GAPS tool and score, enhanced with the capacity-based weighting system, marks progress in assessing pediatric surgical capacity in resource-limited settings. By mirroring the complex reality of hospital functionality in low-resource centers, it provides a refined mechanism for fostering effective partnerships and data-driven strategic interventions.

Introduction

The Global Surgery movement, spearheaded by key publications from the Lancet Commission on Global Surgery and allied entities, has underscored the urgency of addressing the accessibility and quality of surgical care in low- and middle-income countries [1-5].

In low- and middle-income countries, where children make up to half the population, surgical conditions significantly impact the global disease burden [6-8]. Estimates suggest that by age 15, up to 85% of these children will need treatment for a surgically addressable condition. Moreover, the delayed disability adjusted life years (DALY) related to surgical procedures was recently reported to be 8.4 averted DALY / child, a figure comparable to other public health interventions [7]. Addressing these challenges through a multifaceted approach—including infrastructure, training, interdisciplinary collaboration, and sociocultural sensitivity—is pivotal for reducing pediatric morbidity and mortality in resource-limited settings.

Despite advances, quantification of the pediatric surgical disease burden in low- and middle-income countries is fraught with inaccuracies [7, 9-12]. Current North-South partnerships aiming to bolster pediatric surgical resources in low- and middle-income countries have, at times, been hampered by undefined objectives, resulting in suboptimal resource use and impact [13-15]. There is a pressing need for precise tools to guide these partnerships effectively.
In this context, our research team's systematic review in 2019 delineated the landscape of existing capacity assessment tools (CATs) for pediatric surgical services (1). The review revealed a need for a more encompassing tool that factors in not only local resources and access to care, but also outcomes, quality measures, and capacity-building elements. To address these gaps, we had initiated the development of the Global Assessment for Pediatric Surgery (GAPS), a comprehensive, standardized instrument designed explicitly for low- and middle-income settings [16].

We recently published the GAPS initial development, international pilot testing, construct validation, and item refinement [16]. The entire project is comprised of three phases: Phase I, a systematic review to identify and generate a pool of items from existing relevant CATs [17]; Phase II, detailed in the preceding article, including initial tool development, international pilot testing, construct validation, and refinement to ensure relevance and discriminative capacity for low- and middle-income contexts [16]; and Phase III, the focus of this current work, introducing capacity-based weighting to GAPS, optimizing it as a metric for assessing and enhancing pediatric surgical capacity in low- and middle-income countries.

The current phase III is therefore critical toward operationalizing a nuanced, evidence-based tool that aligns with global health objectives and enables effective, targeted interventions in pediatric surgical care.

Methods

<u>*Determining GAPS score weights*</u>

Each item in the GAPS tool was assigned an equal weight of one point, as previous research has shown that giving different weights to similar items offers little advantage when dealing with large assessment tools [18, 19]. However, because the GAPS tool is divided into five distinct subsections, a specific weight was assigned to each subsection.

Items with ordinal responses were scored by dividing one by the total number of possible responses. To ensure the calculations resulted in whole numbers, the weight of each item was multiplied by a factor of 120, referred to as 'Q'. The scores for the items in each subsection were then summed to produce a subsection score, referred to as 'E', for each of the five subsections: 'E_{HR}, E_{MR}, E_{AC}, E_{ED}, E_{OU}'.

A multiple ordinal logistic regression was then conducted, with the hospital level as the outcome, and the subsection scores ($E_{Subsection}$) as the predictor variables. Each regression coefficient was multiplied by the corresponding subsection score ($E_{Subsection}$) to calculate the final GAPS score. If any data were missing, those items were assigned a score of zero. The final GAPS score was then rounded to the nearest whole number.

This method was chosen for its balance between simplicity and capturing meaningful variability across items and subsections. As described by DeVellis et al and Streiner et al, assigning equal weights to individual items minimizes unnecessary complexity and reduces the risk of overfitting, which can arise from arbitrarily weighting similar items [18, 20]. The application of specific weights to the five distinct subsections reflects the structure of the GAPS tool, ensuring that each area is appropriately represented.

Ordinal logistic regression was used to account for the ranked nature of the hospital-level outcome, allowing for a model that effectively captures how subsection scores contribute to the overall assessment. This approach avoids incorrect assumptions about linearity, providing a straightforward yet precise method. Treating missing values as zero and rounding final scores to the nearest integer ensures both statistical accuracy and ease of interpretation.

This approach offers a balance between statistical rigor and practical usability. It captures the unique contributions of each subsection without overcomplicating the model or introducing unnecessary parameters. Ordinal logistic regression was chosen for its appropriateness in modeling an ordinal outcome, ensuring reliability and robustness. The final GAPS score is thus statistically sound and easy to interpret, making it a practical tool for use in clinical settings. Following is a formulaic representation of the analyses:

Score of each item * 120 = $Q1 : Q64$

Sum($Q1{:}Q9$)=E_{HR} ; Sum ($Q10{:}Q48$)=E_{MR} ; Sum($Q49{:}\ Q51$)=E_{AC} ; Sum($Q52{:}Q54$)=E_{ED} ; Sum($Q55{:}Q64$)=E_{OU}

GAPS Score = (E_{HR}*Coeff$_{\mathrm{HR}}$) + (E_{MR}*Coeff$_{\mathrm{MR}}$) + (E_{AC}*Coeff$_{\mathrm{AC}}$) + (E_{ED}*Coeff$_{\mathrm{ED}}$) + (E_{OU}*Coeff$_{\mathrm{OU}}$)

Correlation of GAPS score with capacity and outcomes

Data for these analyses was retrieved from the pilot study phase [57]. The tool was piloted in person between February 2017 and March 2018. Each survey was completed with the aid of a local senior medical representative. Countries were chosen pragmatically based on their geographical localization, ease of access, safety for travel, and number of accessible institutions.

The primary outcome was level of hospital care (1st level, 2nd level, 3rd level and national children's hospital). Level of hospital was predefined by the corresponding countries ministry of health. Secondary outcomes included country level of income as per World Bank 2018 and 2017 [58, 59], human development index [60], under-five mortality rate (per 1,000 live births) (U5MR), and neonatal mortality rate (NNMR), deaths due to congenital anomalies (CA) per 1,000 births (0 – 27 days, 1 – 59 months, and 0 – 4 years), and deaths due to injuries (INJ) per 1,000 births (0 – 27 days, 1 – 59 months, and 0 – 4 years).

Analyses consisted of the Kruskal-Wallis Test and linear regression, with $p < 0.05$ considered significant. All analyses were performed in STATA/MP 13.0 (StataCorp, College Station, TX).

GAPS Score Benchmarks

Benchmark GAPS scores for each hospital level were derived from the 75th and 90th percentiles, then analyzed via dot plots to establish cut-offs maximizing data point inclusion.

Results

Determining GAPS score weights

The GAPS pilot encompassed 65 institutions in eight countries, including 28 1st level (43%), 23 2nd level (35%), 11 3rd level (17%), and three national children's hospitals (5%). Most countries included in the pilot study were Low-Income Countries (LIC: Somaliland, Democratic Republic of Congo (DRC), Uganda). The DRC comprised 63% of institutions. Hospitals were most frequently public (45/65; 69%), followed by faith-based (15/65; 25%). (**Table 1**).

Table 1. Descriptive Statistics of Pilot Data by Level of Hospital

	1st Level (n=28)	2nd Level (n=23)	3rd Level (n=11)	National Children's Hospital (n=3)
Level of Income				
LIC	26	14	6	0
LMIC	1	7	4	3
UMIC	1	2	1	0
Country				
Somaliland	0	0	1	0
Democratic Republic of Congo	26	12	3	0
Uganda	0	2	2	0
Mongolia	1	6	1	0
Algeria	1	2	0	0
Thailand	0	0	1	0
Vietnam	0	1	0	2
Egypt	0	0	3	1
Type of Hospital				
Public / Government	18	16	8	3
Non-Governmental Organization	0	1	1	0
Faith-based	9	6	1	0
Private	1	0	1	0

GAPS Score [median (IQR)]	3 (2)	6 (1)	11 (4)	16 (5)
GAPS Subsection Scores [median (IQR)]				
Human Resources	135 (100)	290 (300)	480 (450)	1020 (400)
Material Resources	1005 (770)	2380 (1040)	3180 (1000)	4260 (1360)
Accessibility	78 (60)	78 (96)	138 (258)	360 (84)
Education	0 (120)	120 (180)	660 (540)	900 (840)
Outcomes	80 (80)	80 (40)	120 (120)	280 (80)

LIC: Low-Income Country, LMIC: Low-Middle Income Country, UMIC: Upper-Middle Income Country, IQR: Inter-quartile range

The GAPS Score equation (**Appendix A**) is detailed below:

$$\boldsymbol{GAPS\ Score} = (E_{HR} * 0.003297) + (E_{MR} * 0.0020106) + (E_{AC} * 0.0008398) + (E_{ED} * 0.0026368) + (E_{OU} * 0.0042668)$$

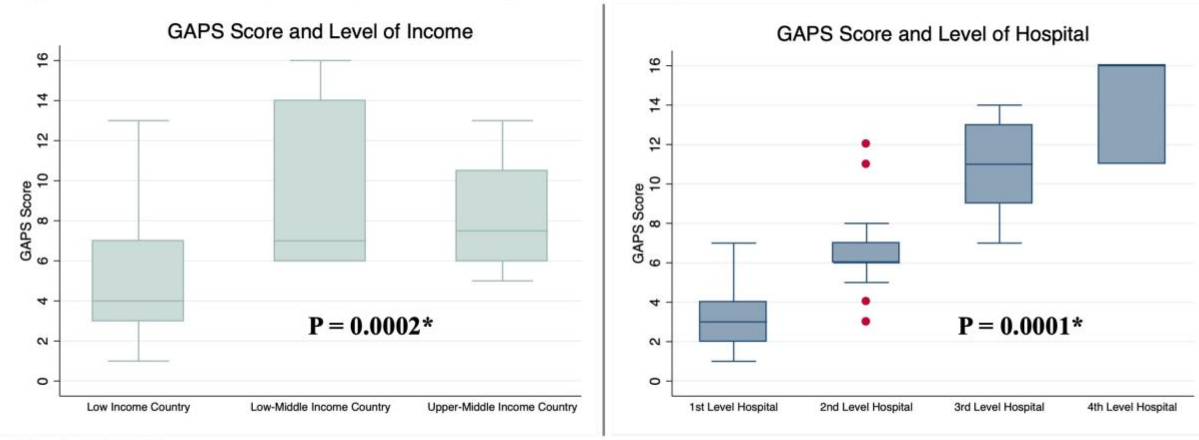
The resultant possible range of values are 0.0453492 – 17.577264. The score is then rounded to the nearest integer, resulting in the final ***GAPS Score*** (range of values from 0 – 18).

Correlation of GAPS score with capacity and outcomes

Capacity

The range of values for the GAPS score in our pilot data for the 65 institutions was 1 - 16 with a median of 6 and an interquartile range (IQR) of 5. Median GAPS scores significantly increased with level of hospital (p=0.0001) (**Figure 1)**. Furthermore, GAPS score significantly increased with level of income (p = 0.0001, **Figure 1**). When evaluating hospital type, median GAPS score was lowest in public / government hospital (6) and faith-based hospital (6), followed by private (7.5), and non-governmental organizations (10.5).

Figure 1. GAPS Score by Country Level of Hospital and Country Income Level



**Kruskal-Wallis Test*

All subsection scores progressively and significantly rose with each hospital level. Human resources and material resources subsection scores significantly increased with level of income (p = 0.0001, p = 0.0004, respectfully). Hospital type was only significant in the human resources subsection score (p = 0.0491). (**Appendix B**).

Furthermore, a linear regression of GAPS Score with country income was significant (p < 0.001) between LIC and Low- and Middle-Income Countries (LMIC), and between 1st level and all other hospital levels. (**Table 2**).

Table 2. Linear Regression Analysis of GAPS Score to Human development Index, Under 5 Mortality Rate, and Neonatal Mortality Rate.

	GAPS Score		
	Regression Coefficient	P value	95% Confidence Interval
Human Development Index	0.02	< 0.001	0.01 – 0.02
Under 5 Mortality Rate (per 1,000 live births)	- 4.48	< 0.001	(- 6.10) – (- 2.85)
Neonatal Mortality Rate (per 1,000 live births)	- 1.21	< 0.001	(- 1.66) – (- 0.75)
Deaths due to Congenital Anomalies (per 1,000 live births)			
0 – 27 days	0.02	0.015	0.00 – 0.03
1 – 59 months	- 0.00	0.889	(-)0.03 – 0.03
0 – 4 years	0.01	0.385	(-) 0.02 – 0.05
Deaths due to Injury (per 1,000 live births)			
0 – 27 days	- 0.02	< 0.001	(-) 0.03 – (-) 0.15

1 – 59 months	- 0.22	< 0.001	(-)0.30 – (-) 0.14
0 – 4 years	- 0.24	< 0.001	(-)0.32 – (-) 0.15

Outcomes

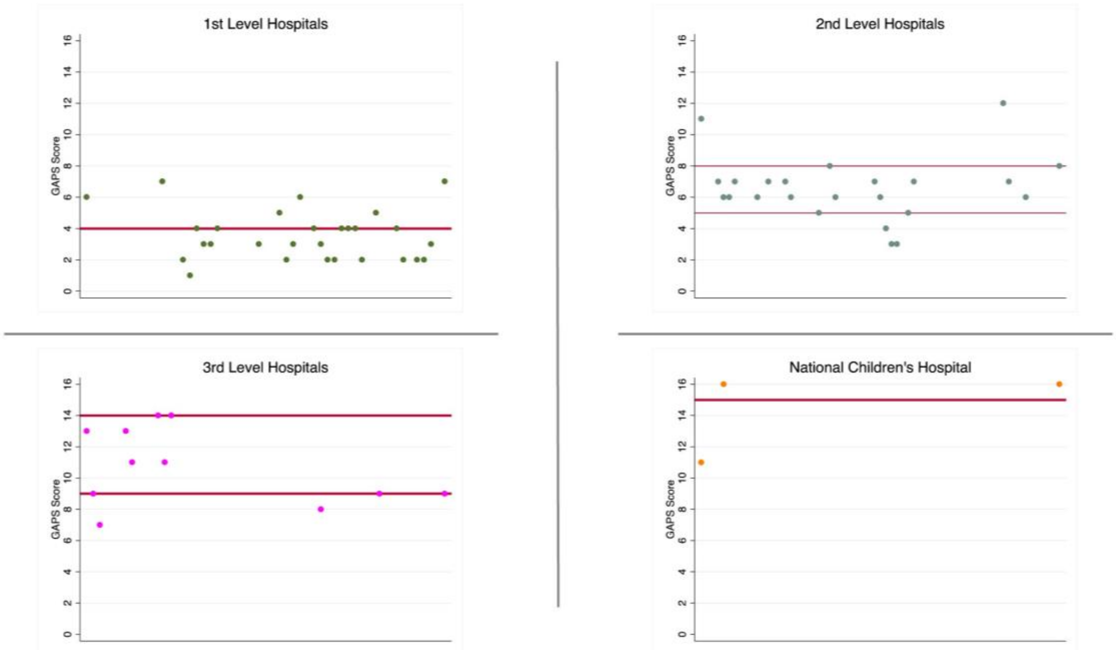
The GAPS score showed a significant association with HDI ($p < 0.001$), U5MR ($p < 0.001$) and NNMR ($p<0.001$). (**Table 2, Appendix C**). Linear regression analysis of GAPS score to U5MR identified that for each unit increase in GAPS score the expected value of U5MR decreases by 4.5 units. Furthermore, a substantial proportion (33%) of the variability in U5MR is explained by the GAPS score. Similarly, for each unit increase in GAPS score the expected value of NNMR decreases by 1.2 units and explains a large proportion (31%) of the variability within NNMR. Regarding HDI, each unit increase in GAPS score is associated with an increase in HDI of 0.02 units.

The GAPS score showed a significant increase as the number of deaths due to injuries declined across all age groups (0 – 27 days, $p < 0.001$; 1 – 59 months, $p < 0.001$; 0 – 4 years, $p< 0.001$) (**Table 2**). However, no significant relationship was found between the GAPS score and deaths due to congenital anomalies. While the p-value for deaths due to congenital anomalies in neonates (0 – 27 days) was statistically significant, the corresponding 95% confidence interval included zero, indicating that this finding does not reflect a statistically significant effect (**Table 2**).

GAPS Score Benchmarks

Dot graphs (**Figure 2**) show the number of institutions captured by the GAPS score benchmarks for each level of hospital (**Table 3**). The GAPS benchmarks correctly captured 78% (21/27) of hospital identified as 1st level, 78% (18/23) of 2nd level, (9/11) 82% of 3rd level, and (2/3) 67% of National Children's Hospitals.

Figure 2. Benchmark values for GAPS Score per Level of Hospital



Each point represents a single institution. The red longitudinal lines demarcate the lowest and highest benchmark value for each level of hospital. GAPS ≤ 4 for 1st level hospitals, 5 – 8 for 2nd level hospitals, 9 – 14 for 3rd level hospitals, and ≥ 15 for national children's hospitals.

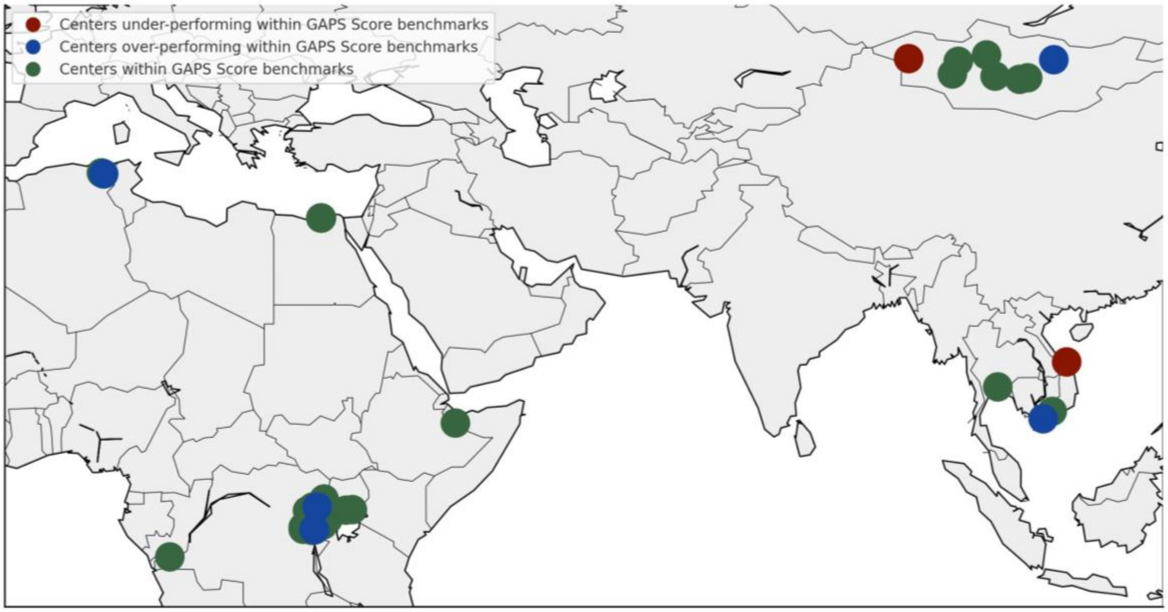
Table 3. Benchmarks Values for GAPS Score by Level of Hospital

	GAPS Score	Human Resources	Material Resources	Accessibility	Education	Outcomes
1st Level Hospital	≤ 4	≤ 200	≤ 2000	≤ 105	≤ 120	≤ 80
2nd Level Hospital	5 – 8	201 – 560	2001 – 3560	106 – 200	121 - 400	81 - 150
3rd Level Hospital	9 – 14	561 – 800	3561 – 4090	201 – 359	401 – 799	151 – 300
National Children's Hospital	≥ 15	≥ 801	≥ 4091	≥ 360	≥800	≥ 301

In the 1st level hospital group, six hospitals scored above the benchmark (≤ 4) (**Figure 2**): centers 12, 23, 40, 43, 54 and 64 (**Appendix D, Figure 3**) located in Mongolia, Algeria, and the DRC (**Figure 4**).

Figure 3. Radar Plots of Centers that Out-Performed and Under-Performed on the GAPS Score with Respect to the Level of Hospital

Figure 4. GAPS Score Performance Mapping of Healthcare Centers in Pilot Study



Each point represents a single institution.

In the 2nd level hospital group, five hospitals scored outside the benchmark GAPS scores (5 – 8): three below and two above (**Figure 2**). Despite identifying as 2nd level hospitals, centers 34, 35, and 36 located in the DRC scored below benchmark values (**Figure 3**). Center 1, located in Vietnam and Center 55, located in the DRC scored within benchmarks for 3rd level hospitals (**Figure 2 and 3, Appendix D**).

For 3rd level hospitals, two hospitals scored below the benchmark (< 9) and none above (>14) (**Figure 2**). Center 10, located in Mongolia, scored below benchmarks for 3rd level hospitals in all sections. Center 44, located in the DRC, scored above benchmark levels for all section except education. (**Figure 2 and 3, Appendix D**).

For National Children's Hospitals, one hospital scored below the benchmark (**Figure 2**). Center 2, located in Vietnam, scored below benchmarks for all sections except human resources (**Figure 2 and 3, Appendix D).**

Discussion

The GAPS tool, based on a 2019 systematic review of CATs and surgical guidelines, was developed through multiple iterations with a 37-person expert panel, international testing, and item refinement. It includes 64 questions across five domains: Human Resources, Material Resources, Outcomes, Accessibility, and Education [16, 17].

The introduction of a capacity-based weighting system to the GAPS tool represents a methodological advancement compared to previously published CATs [17]. In a study of the Service Availability and Readiness Assessment across six countries, O'Neill et al highlight the importance of tailored weighting in health assessment tools to improve the accuracy of capacity evaluations [24]. Assigning differential weights to the various domains of pediatric surgical capacity allows the GAPS tool to be more closely aligned with the multifaceted reality of hospital functionality. This approach of capacity-based weighting is consistent with the multi-dimensional assessment frameworks proposed by Cometto et al, highlighting the necessity of encapsulating the intricacies of health systems within evaluation instruments [25].

The GAPS weighting system marks a significant departure from the equal weighting paradigm used in homogeneous item scales such as the Surgeons OverSeas Pediatric - Personnel, Infrastructure, Procedure, Equipment, and Supplies (Pedi-PIPES) [26]. Unlike Pedi-PIPES, which applies equal weighting across all its items, GAPS employs a dual approach: it uses equal weighting for items within each subscale but adopts differential weighting for the subscales themselves. This method allows for a more accurate representation of hospital capacity, recognizing the differential impact of individual domains on overall surgical capability [19, 20]. This approach aligns with DeVellis' concept that differential weighting enhances the sensitivity of tools in heterogeneous settings [18].

The resultant GAPS score revealed significant correlations with hospital levels and income categories, as well as delineating differences across different hospital types. This score serves not merely as a gauge of capacity, but as a mirror reflecting the intricate interplay between a facility's resources and the socioeconomic milieu in which it operates. The study's findings underscore a positive correlation between GAPS scores and hospital levels, a relationship consonant with the stratified nature of healthcare systems [24]. Similarly, the association between GAPS scores and income levels reinforces the influence of economic status on healthcare quality and availability [27].

In terms of outcomes, the GAPS score's significant association with health indicators such as U5MR and NNMR substantiates its validity as a performance measure [10, 28]. Furthermore, the score correlates with the number of child deaths due injury, accentuating its potential as an indicator of pediatric surgical outcomes. While we found no significant correlation between the GAPS score and deaths from congenital anomalies, we attribute this to the limited sample size of the participating institutions. The score therefore reinforces the connection between health system inputs and health outcomes [9, 29-31].

The establishment of benchmark GAPS scores delineates a clear framework for assessing pediatric surgical capacity across hospital levels, reflecting a critical step towards standardized performance metrics in low- and middle-income countries. These benchmarks provide a tangible target for institutions striving to enhance their surgical services. In our analysis, the benchmarks effectively discriminated between different hospital levels, corroborating the tiers of healthcare delivery described by Hsia et al., who emphasize the importance of clear standards in evaluating service capacity [32]. The performance analysis against these benchmarks revealed that most hospitals conformed to expected levels further supporting the weighting of the GAPS score, however outliers provided offer valuable insights. Certain first-level hospitals, notably in Mongolia and Algeria, surpassed their benchmark scores. While this may suggest overqualification for their designated level or misalignment in hospital categorization, further investigation is needed to determine whether these results reflect successful local policies, improved resource allocation, health system improvements or other contextual factors [27]. Conversely, some second-level hospitals, particularly in the DRC, fell short of their benchmarks, potentially due to systemic challenges such as resource constraints, inadequate training, or governance issues, as highlighted by Grimes et al. [33]. These outliers underscore the variability within categories and illustrate that while benchmarks are effective for setting targets and evaluating performance, they must be interpreted within the specific context of each hospital's health ecosystem and resource availability. The GAPS score provides a valuable starting point for this analysis, but deeper investigation into the drivers of outperformance and underperformance is necessary.

The cross-sectional analysis of hospital performance using the GAPS score highlights notable disparities, with some institutions out- or under-performing relative to their designated levels. Hospitals that exceeded their expected benchmarks, such as those in Mongolia, may reflect the influence of local factors such as policy initiatives, improved resource allocation, or other health system improvements. However, this study does not directly assess the impact of policy changes, and further research is required to validate these potential explanations. Conversely, hospitals scoring below their expected level, particularly 2nd level hospitals in the DRC, may be facing wider structural issues such as inadequate resources, training, or governance

[34]. This mirrors previous findings of the DRC's significant barriers to increasing surgical capacity [35-37]. Understanding the underlying factors contributing to these performance variations is crucial for guiding future policy changes and optimizing international aid. However, the present study focuses primarily on the utility of the GAPS score as a framework for assessing surgical capacity, and the drivers behind outliers remain speculative. Future research should investigate these mechanisms in greater depth to inform targeted capacity-building efforts.

Compared to other CATs explored in our systematic review, GAPS uniquely integrates a more granular, multidimensional analysis of pediatric surgical services [17]. Unlike prior tools focusing inventorying material resources, GAPS's inclusive framework ensures that both tangible and previously intangible facets of surgical capacity are accounted for, providing a more holistic evaluation.

GAPS's capacity-based weighting system, not found in other CATs, reflects the complexity of health service delivery in low- and middle-income countries. It stands as a unique contribution to the field and a potential to guide to refined resource allocation, training initiatives, and policy reforms for pediatric surgery [38].

The methodological rigor underpinning the development and validation of the GAPS tool constitutes a significant strength of this research. The systematic approach, informed by a comprehensive review and subsequent refinement phases, ensures that the tool is both inclusive and reflective of the multifaceted nature of surgical capacity in low-resource settings. The iterative process of development provided a robust construct validation, enhancing the tool's credibility and applicability [16, 39]. Nonetheless, the reliance on pilot data from limited geographical settings may impact the generalizability of the findings. However, any tool's responsiveness to changes over time remains to be evaluated [40]. Future research could address these limitations by expanding the validation process to include a broader range of countries, potentially employing machine learning techniques to refine the weighting system for simplicity and efficacy [41]. Furthermore, the reliance on surveys completed with the aid of local senior medical representatives is a clear source of reporting bias. Moreover, our study relies on internal data and pilot testing for validation. External validation with a broader range of institutions in different settings would be necessary to ensure the tool's applicability and reliability across diverse global health contexts. Additionally, the study does not encompass long-term follow-up data to evaluate temporal changes. Incorporating longitudinal studies would offer valuable perspectives on the tool's responsiveness to evolving surgical capacities, thereby enhancing its effectiveness in ongoing quality enhancement [42-44]. Another important limitation is the lack of integration of surgical output in the GAPS tool. Stewart et al previously noted that assessments of surgical capacity have not adequately correlated with actual surgical output[45]. We concur that material-focused CATs are inadequate in predicting surgical output. To mitigate this, the GAPS tool incorporates additional dimensions (human resources, accessibility, outcomes, and education) moving beyond an inventory-based approach. While GAPS offers a more panoramic view, its effectiveness in directly quantifying surgical output has not been validated. A thorough evaluation of surgical procedures and outcomes is necessary for an accurate assessment of surgical output, an area beyond the scope of this study but vital for future research. We envision GAPS as a pivotal tool in future research, enhancing the evaluation of both surgical capacity and output.

Conclusion

GAPS quantifies pediatric surgical services in low-resource settings, offering a structured method for identifying healthcare disparities and prioritizing needs. The capacity-based weighting system offers a granular assessment of hospital capabilities, thus enabling interested parties to transcend traditional resource inventories and engage in more refined evaluation of service readiness. This specificity is crucial in guiding partnerships, ensuring collaborations are driven by precisely identified needs, thereby enhancing their efficiency and impact. Moreover, the GAPS tool facilitates targeted interventions by providing benchmarks reflective of service levels. The GAPS tool serves as an instrumental resource in the concerted effort to improve pediatric surgical outcomes through data-driven, context-specific strategies.

Acknowledgements :

Special thanks to the following: Kevyn Armand, Daniel Agbo-Panzo, Marilyn Butler, Beda Espineda, Kenzy Abdel-Hamid, Sonia Anchouche, Ai-Xuan Holterman, George W. Galiwango, Nathalie MacKinnon, Ritesh Shrestha, Desigen Reddy, William Harkness, Phyllis Kisa, Rashmi Kumar, Ibrahima Fall, Justin Onen, Ananda K. Lamahewage, Mike Ganey, Tessa Concepcion, Peter Donkor, Zaitun Bokhary, Anette S. Jacobsen, Reju Thomas, Brouh Yapo, Adiyasuren Jamiyanjav, Marcia Matias, Benno Ure, Vanda Amado, Bahati Robert, Lubna Samad and Roumanatou Bankolé, Sarah Cairo, Etienne St-Louis, Laura F. Goodman, Doulia M. Hamad, Robert Baird, Emily R. Smith, Sherif Emil, Jean-Martin Laberge, Robin Petroze, Mohamed

Abdelmalak, Zipporah Gathuy, Faye Evans, Maryam Ghavami Adel, Ki K. Bertille, Milind Chitnis, Leecarlo Millano, Peter Nthumba, Sergio d'Agostino, Bruno Cigliano, Luis Zea-Salazar, Doruk Ozgediz, and Elena Guadagno.

Appendices

Appendix A

Appendix A. Ordinal Logistic Regression for Weights of Each Subsection of GAPS: Human Resources, Material Resources, Outcomes, Accessibility, Education			
	Coefficient	P value	95% Confidence Interval
Human Resources	0.003297	0.147	-0.001 – 0.008
Material Resources	0.0020106	0.002	0.001 – 0.003
Outcomes	0.0042668	0.880	-0.010 – 0.012
Accessibility	0.0008398	0.475	-0.007 – 0.016
Education	0.0026368	0.162	-0.001 – 0.006

The model's overall fit was significant (LR $chi^2(5) = 67.96$, $p < 0.0001$), suggesting that the overall model was able to distinguish well between different levels of hospital care. The Pseudo R^2 value was 0.5237, indicating a moderate relationship between the predictors and the ordered outcome.

Appendix B

Appendix B. GAPS Subsection Scores by Country Income, Level of Hospital, and Type of Hospital										
	Human Resources		**Material Resources**		**Accessibility**		**Education**		**Outcome**	
	HR Score Median (IQR)	p-value*	MR Score Median (IQR)	p-value*	AC Score Median (IQR)	p-value*	ED Score Median (IQR)	p-value*	OU Score Median (IQR)	p-value*
Level of Hospital										
1	135 (100)	0.0001	1005 (770)	0.0001	78 (60)	0.0006	0 (120)	0.0001	80 (40)	0.006
2	290 (300)		2380 (1040)		78 (96)		120 (180)		80 (40)	
3	480 (450)		3180 (1000)		138 (258)		660 (540)		120 (120)	
4	1020 (400)		4260 (1360)		360 (84)		900 (840)		280 (80)	
Level of Income										
LIC	170 (90)	0.0001	1495 (1590)	0.0004	78 (60)	0.051	120 (300)	0.209	80 (40)	0.054
LMIC	580 (490)		2640 (1320)		192 (330)		180 (840)		120 (80)	
UMIC	515 (190)		2430 (1350)		108 (102)		120 (270)		220 (200)	
Hospital Type										
Public / Government	210 (310)	0.0491	1900 (1610)	0.2465	93 (96)	0.6979	120 (300)	0.2963	120 (80)	0.4314

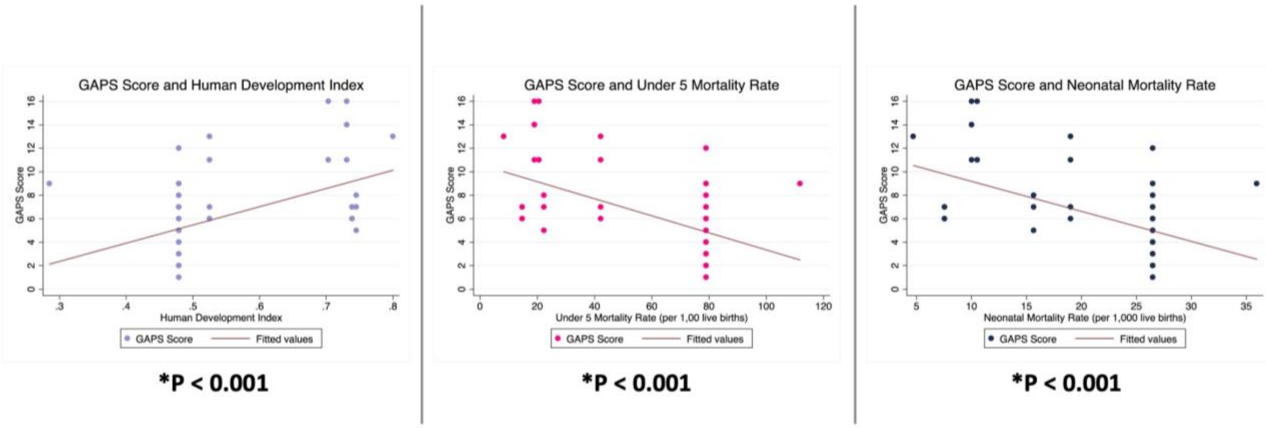
Non-Governmental Organization	485 (130)		3250 (620)		123 (6)		630 (180)		140 (40)	
Faith-based	165 (65)		1690 (1530)		96 (87)		120 (330)		80 (40)	
Private	430 (580)		2510 (1740)		219 (282)		240 (480)		120 (0)	

** Kruskal-Wallis Test*

LIC: Low Income Country, LMIC: Low-Middle Income Country, UMIC: Upper-Middle Income Country, IQR: Inter-quartile range

Appendix C

Appendix C. GAPS Score and Under 5 Mortality Rate, Neonatal Mortality Rate, and Human Development Index



Each point represents a single institution.
** Linear Regression Analysis*

Appendix D

Appendix D. Centers that fall outside GAPS Score Benchmarks per level of Hospital						
Center	GAPS Score	Human Resources	Material Resources	Accessibility	Education	Outcomes
1st Level Hospital						
Above benchmark						
12	6	270*	2080*	150*	0	200*
23	7	660*	2210*	168*	120	0
40	5	130	1540	54	240*	160*
43	6	160	1780	54	420*	80
54	5	170	1900	78	0	160*
64	7	130	2640*	228*	120	80
2nd Level Hospital						

Below benchmark						
34	4	140°	1450°	174	60°	80°
35	3	170°	1070°	114	0°	80°
36	3	210	1030°	54°	180	0
Above Benchmark						
1	11	720*	3420	300*	600*	0°
55	12	550	3560	126	720*	160*
3rd Level Hospital						
Below benchmark						
10	7	270°	2420°	150°	120°	120°
44	8	200°	2640°	138°	720	80°
National Children's Hospital						
Below benchmark						
2	11	1020	3320°	276°	360°	0°

**Above subsection benchmark for level of hospital ° Below subsection benchmark for level of hospital. See Table 3 for GAPS Score and Subsection score benchmark values.*

Appendix E

Appendix F. GAPS Subsection Scores Percentiles per Level of Hospital

	10th	25th	50th	75th	90th
Human Resources					
1st Level Hospital	70	70	135	170	200
2nd Level Hospital	140	170	290	470	560
3rd Level Hospital	200	270	480	720	840
National Children's Hospital	680	680	1020	1080	1080
Material Resources					
1st Level Hospital	620	820	1005	1590	2080
2nd Level Hospital	1450	1600	2380	2640	2900
3rd Level Hospital	2600	2640	3180	3640	3820
National Children's Hospital	3320	3320	4260	4680	4680
Accessibility					
1st Level Hospital	54	54	78	114	150
2nd Level Hospital	24	54	78	150	216

3rd Level Hospital	84	102	138	360	360
National Children's Hospital	276	276	360	360	360
Education					
1st Level Hospital	0	0	0	120	300
2nd Level Hospital	0	60	120	240	360
3rd Level Hospital	480	480	660	1020	1080
National Children's Hospital	360	360	900	1200	1200
Outcomes					
1st Level Hospital	0	40	80	120	160
2nd Level Hospital	40	80	80	120	160
3rd Level Hospital	80	120	120	240	320
National Children's Hospital	240	240	280	320	320

References

1. Abdul-Mumin A, Anyomih TTK, Owusu SA, Wright N, Decker J, Niemeier K, Benavidez G, Abantanga FA, Smith ER , Tabiri S (2020) Burden of neonatal surgical conditions in northern ghana. World J Surg 44:3-11. https://10.1007/s00268-019-05210-9.ORG/
2. Butler MW (2016) Developing pediatric surgery in low- and middle-income countries: An evaluation of contemporary education and care delivery models. Semin Pediatr Surg 25:43-50. https://10.1053/j.sempedsurg.2015.09.008.ORG/
3. Concepcion TL, Dahir S, Mohamed M, Hiltbrunn K, Ismail EA, Poenaru D, Rice HE, Smith ER , Global Initiative for Children's S (2020) Barriers to surgical care among children in somaliland: An application of the three delays framework. World J Surg 44:1712-1718. https://10.1007/s00268-020-05414-4.ORG/
4. Dare AJ, Grimes CE, Gillies R, Greenberg SL, Hagander L, Meara JG , Leather AJ (2014) Global surgery: Defining an emerging global health field. Lancet 384:2245-2247. https://10.1016/S0140-6736(14)60237-3.ORG/
5. Huber B (2015) Finding surgery's place on the global health agenda. Lancet 385:1821-1822. https://10.1016/S0140-6736(15)60761-9.ORG/
6. Makasa EM (2014) Letter to global health agency leaders on the importance of surgical indicators. Lancet 384:1748. https://10.1016/S0140-6736(14)62012-2.ORG/
7. Meara JG, Hagander L , Leather AJM (2014) Surgery and global health: A lancet commission. Lancet 383:12-13. https://10.1016/S0140-6736(13)62345-4.ORG/
8. Meara JG, Leather AJ, Hagander L, Alkire BC, Alonso N, Ameh EA, Bickler SW, Conteh L, Dare AJ, Davies J, Merisier ED, El-Halabi S, Farmer PE, Gawande A, Gillies R, Greenberg SL, Grimes CE, Gruen RL, Ismail EA, Kamara TB, Lavy C, Lundeg G, Mkandawire NC, Raykar NP, Riesel JN, Rodas E, Rose J, Roy N, Shrime MG, Sullivan R, Verguet S, Watters D, Weiser TG, Wilson IH, Yamey G , Yip W (2016) Global surgery 2030: Evidence and solutions for achieving health, welfare, and economic development. Int J Obstet Anesth 25:75-78. https://10.1016/j.ijoa.2015.09.006.ORG/
9. Yousef Y, Lee A, Ayele F , Poenaru D (2019) Delayed access to care and unmet burden of pediatric surgical disease in resource-constrained african countries. J Pediatr Surg 54:845-853. https://10.1016/j.jpedsurg.2018.06.018.ORG/

10. Smith ER, Concepcion TL, Shrime M, Niemeier K, Mohamed M, Dahir S, Ismail EA, Poenaru D, Rice HE , Global Initiative for Children's S (2020) Waiting too long: The contribution of delayed surgical access to pediatric disease burden in somaliland. World J Surg 44:656-664. https://10.1007/s00268-019-05239-w.ORG/
11. Bickler SW , Rode H (2002) Surgical services for children in developing countries. Bull World Health Organ 80:829-835. https://www.ncbi.nlm.nih.gov/pubmed/12471405
12. Smith ER, Concepcion TL, Niemeier KJ , Ademuyiwa AO (2019) Is global pediatric surgery a good investment? World J Surg 43:1450-1455. https://10.1007/s00268-018-4867-4.ORG/
13. Truche P, Smith ER, Ademuyiwa A, Buda A, Nabukenya MT, Kaseje N, Ameh EA, Greenberg S, Evans F, Bickler S, Meara JG , Rice HE (2022) Defining surgical workforce density targets to meet child and neonatal mortality rate targets in the age of the sustainable development goals: A global cross-sectional study. World J Surg 46:2262-2269. https://10.1007/s00268-022-06626-6.ORG/
14. Hamad D, Yousef Y, Caminsky NG, Guadagno E, Tran VA, Laberge JM, Emil S , Poenaru D (2020) Defining the critical pediatric surgical workforce density for improving surgical outcomes: A global study. J Pediatr Surg 55:493-512. https://10.1016/j.jpedsurg.2019.11.001.ORG/
15. Vigliotti VS, Concepcion T, Mohamed M, Dahir S, Ismail EA, Poenaru D, Rice HE , Smith ER (2022) Modeling the scale-up of surgical services for children with surgically treatable congenital conditions in somaliland. World J Surg 46:2489-2497. https://10.1007/s00268-022-06651-5.ORG/
16. Y Y, S C, E S-L, L.F G, D.M H, R B, E.R S, S E, JM L, M A, Z G, F E, M.G A, K.K B, M C, L M, P N, S dA, B C, L Z-S, E A, D O, E G , D P (2023) Development and pilot results of the global assessment in pediatric surgery (gaps): An evidence-based pediatric surgical capacity assessment tool for low-resource settings: Phase ii. World Journal of Surgery,
17. Yousef Y, St-Louis E, Baird R, Smith ER, Guadagno E, St-Vil D , Poenaru D (2019) A systematic review of capacity assessment tools in pediatric surgery: Global assessment in pediatric surgery (gaps) phase i. J Pediatr Surg 54:831-837. https://10.1016/j.jpedsurg.2018.11.005.ORG/
18. DeVellis RF , Thorpe CT (2022) Scale development : Theory and applications. Thousand Oaks, California

19. L. Streiner D, R. Norman G , Cairney J (2023) Health measurement scales. New York
20. Streiner DL (2003) Starting at the beginning: An introduction to coefficient alpha and internal consistency. J Pers Assess 80:99-103. https://10.1207/S15327752JPA8001_18.ORG/
21. (2018) World bank country and lending groups. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups
22. World bank country and lending groups. World Bank Group. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups
23. (2023) Human development index (hdi). United Nations Development Programme. https://hdr.undp.org/data-center/human-development-index#/indicies/HDI
24. O'Neill K, Takane M, Sheffel A, Abou-Zahr C , Boerma T (2013) Monitoring service delivery for universal health coverage: The service availability and readiness assessment. Bull World Health Organ 91:923-931. https://10.2471/BLT.12.116798.ORG/
25. Cometto G, Ford N, Pfaffman-Zambruni J, Akl EA, Lehmann U, McPake B, Ballard M, Kok M, Najafizada M, Olaniran A, Ajuebor O, Perry HB, Scott K, Albers B, Shlonsky A , Taylor D (2018) Health policy and system support to optimise community health worker programmes: An abridged who guideline. Lancet Glob Health 6:e1397-e1404. https://10.1016/S2214-109X(18)30482-0.ORG/
26. Okoye MT, Ameh EA, Kushner AL , Nwomeh BC (2015) A pilot survey of pediatric surgical capacity in west africa. World J Surg 39:669-676. https://10.1007/s00268-014-2868-5.ORG/
27. Anand S , Barnighausen T (2004) Human resources and health outcomes: Cross-country econometric study. Lancet 364:1603-1609. https://10.1016/S0140-6736(04)17313-3.ORG/
28. Roa L, Jumbam DT, Makasa E , Meara JG (2019) Global surgery and the sustainable development goals. Br J Surg 106:e44-e52. https://10.1002/bjs.11044.ORG/
29. Landrum K, Cotache-Condor CF, Liu Y, Truche P, Robinson J, Thompson N, Granzin R, Ameh E, Bickler S, Samad L, Meara J, Rice HE, Smith ER , Global Initiative for Children's S (2021) Global and regional overview of the inclusion of paediatric surgery in the national health plans of 124 countries: An ecological study. BMJ Open 11:e045981. https://10.1136/bmjopen-2020-045981.ORG/

30. Saxton AT, Poenaru D, Ozgediz D, Ameh EA, Farmer D, Smith ER , Rice HE (2016) Economic analysis of children's surgical care in low- and middle-income countries: A systematic review and analysis. PLoS One 11:e0165480. https://10.1371/journal.pone.0165480.ORG/
31. Seyi-Olajide JO, Anderson JE, Kaseje N, Ozgediz D, Gathuya Z, Poenaru D, Johnson W, Bickler SW, Farmer DL, Lakhoo K, Oldham K, Ameh EA , Global Initiative for Children's S (2021) Inclusion of children's surgery in national surgical plans and child health programmes: The need and roadmap from global initiative for children's surgery. Pediatr Surg Int 37:529-537. https://10.1007/s00383-020-04813-x.ORG/
32. Hsia RY, Mbembati NA, Macfarlane S , Kruk ME (2012) Access to emergency and surgical care in sub-saharan africa: The infrastructure gap. Health Policy Plan 27:234-244. https://10.1093/heapol/czr023.ORG/
33. Grimes CE, Law RS, Borgstein ES, Mkandawire NC , Lavy CB (2012) Systematic review of met and unmet need of surgical disease in rural sub-saharan africa. World J Surg 36:8-23. https://10.1007/s00268-011-1330-1.ORG/
34. Farmer PE , Kim JY (2008) Surgery and global health: A view from beyond the or. World J Surg 32:533-536. https://10.1007/s00268-008-9525-9.ORG/
35. Cairo SB, Kalisya LM, Bigabwa R , Rothstein DH (2018) Characterizing pediatric surgical capacity in the eastern democratic republic of congo: Results of a pilot study. Pediatr Surg Int 34:343-351. https://10.1007/s00383-017-4215-z.ORG/
36. Cairo SB, Pu Q, Malemo Kalisya L, Fadhili Bake J, Zaidi R, Poenaru D , Rothstein DH (2020) Geospatial mapping of pediatric surgical capacity in north kivu, democratic republic of congo. World J Surg 44:3620-3628. https://10.1007/s00268-020-05680-2.ORG/
37. Kalisya LM, Yap A, Mitume B, Salmon C, Karafuli K, Poenaru D , Onyango R (2023) Determinants of access to essential surgery in the democratic republic of congo. J Surg Res 291:480-487. https://10.1016/j.jss.2023.07.014.ORG/
38. Price R, Makasa E , Hollands M (2015) World health assembly resolution wha68.15: "Strengthening emergency and essential surgical care and anesthesia as a component of universal health coverage"-addressing the public health gaps arising from lack of safe, affordable and accessible surgical and anesthetic services. World J Surg 39:2115-2125. https://10.1007/s00268-015-3153-y.ORG/

39. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM , de Vet HC (2007) Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60:34-42. https://10.1016/j.jclinepi.2006.03.012.ORG/
40. Fayers PM , Machin D (2016) Quality of life : The assessment, analysis, and reporting of patient-reported outcomes. Chichester, West Sussex, UK ; Hoboken, NJ
41. Obermeyer Z , Emanuel EJ (2016) Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med 375:1216-1219. https://10.1056/NEJMp1606181.ORG/
42. Debas H, Alatise OI, Balch CM, Brennan M, Cusack J, Donkor P, Jaffe BM, Mazariegos GV, Mock C, Mutiibwa D, Numann P, Nyagatuba JKM, O'Neill JA, Jr., Tarpley JL, Tesfaye S, Tefera G , Tuttle TM (2020) Academic partnerships in global surgery: An overview american surgical association working group on academic global surgery. Ann Surg 271:460-469. https://10.1097/SLA.0000000000003640.ORG/
43. Philipo GS, Nagraj S, Bokhary ZM , Lakhoo K (2020) Lessons from developing, implementing and sustaining a participatory partnership for children's surgical care in tanzania. BMJ Glob Health 5:e002118. https://10.1136/bmjgh-2019-002118.ORG/
44. Naluyimbazi R, Nimanya S , Kisa P (2023) Anatomy and lessons of partnerships in global pediatric surgery. Semin Pediatr Surg 32:151353. https://10.1016/j.sempedsurg.2023.151353.ORG/
45. Stewart BT, Gyedu A, Gaskill C, Boakye G, Quansah R, Donkor P, Volmink J , Mock C (2018) Exploring the relationship between surgical capacity and output in ghana: Current capacity assessments may not tell the whole story. World J Surg 42:3065-3074. https://10.1007/s00268-018-4589-7.ORG/

CHAPTER 6: DISCUSSION

THE EVOLUTION OF THE GAPS TOOL

The concept of a CAT for pediatric surgery in low-resource settings originated from the shared concerns expressed by numerous pediatric surgical groups in both L-MICs[9, 15-20, 23, 39, 44, 45, 51, 52]. These groups highlighted that their collaborative efforts were not meeting intended goals and raised issues related to underlying colonial dynamics and the challenges associated with volunteer-based initiatives[39, 40, 44, 45, 61]. The development of the GAPS tool followed a rigorous, structured methodological framework, inspired by the work of DeVellis et al. and Streiner et al., and was specifically adapted to reflect the conditions in low-resource settings. This development process encompassed three distinct phases[34, 35].

Manuscript 1

Phase I - Systematic Review and Item Pool Generation: This initial phase involved conducting a systematic review to identify existing relevant tools that could inform the creation of a comprehensive item pool for GAPS.

Manuscript 2

Phase II - Instrument Development and Preliminary Validation: As the core subject of this study, Phase II involved four critical components:

1. Initial Item Development and Expert Consultation: Items were developed based on the evidence gathered in Phase I, with input from an expert panel to ensure their relevance.
2. International Pilot Testing: The feasibility and applicability of the tool were tested across multiple settings, providing insights into its practical implementation.
3. Construct Validation and Item Analysis: This component evaluated the items' ability to differentiate effectively between varying levels of surgical care.
4. Instrument Refinement: Only items that were feasible to answer in LMICs and could reliably distinguish between hospital care levels were retained.

Manuscript 3

Phase III - Outcome-Based Weighting for Construct Validity: In this final phase, a weighted scoring system was introduced, grounded in outcome measures, to enhance the tool's predictive validity and ensure its utility in assessing and advancing pediatric surgical capacity.

The process began with a systematic review to identify whether an existing comprehensive tool could fulfill the need for a multidimensional assessment of capacity, incorporating not only resources but also education, accessibility, and outcomes. Manuscript 1 presents this systematic review, aimed at evaluating the limitations and scope of current capacity assessment tools used for pediatric surgical care in L-MICs. Despite the significant unmet surgical needs of children in these regions, tools such as Pedi-PIPES, CSCDSP, and CCTR primarily focus on resource availability and omit essential metrics related to outcomes, accessibility, and training. These omissions restrict the usefulness of these tools in driving comprehensive improvements in pediatric surgical care. To fill these gaps, the GAPS tool was developed as an outcomes-oriented, quantitatively driven instrument tailored to the specific requirements of L-MICs. This review underscored the necessity for a holistic, standardized approach to assessing pediatric surgical capacity, aiming to create an objective and reproducible framework for enhancing surgical care in low-resource settings. GAPS was designed to address the shortcomings of existing models, providing a robust structure for assessing and advancing pediatric surgical capacity through partnerships between high- and low-resource settings.

The systematic review identified 19,868 records, narrowing down to 16,641 titles and abstracts screened for eligibility, with 15 full-text articles reviewed in detail. Ultimately, three CATs were included: Pedi-PIPES, Children's Surgical Center Designations with Scope of Practice (CSCDSP), Checklist for a Children's Trauma Room (CCTR). These tools were assessed for their ability to evaluate pediatric surgical capacity, particularly in L-MICs.

1. **Pediatric Surgery Personnel, Infrastructure, Procedures, Equipment, and Supplies (Pedi-PIPES):** This tool, developed by SOSAS and based on the WHO TSA with input from pediatric surgeons in both HICs and L-MICs, focuses on pediatric surgical capacity in L-MICs. It assesses personnel, infrastructure, procedures, equipment, and supplies. Pedi-PIPES contains 118 items, using a binary evaluation of availability for each item, thereby lacking granularity. Although Pedi-PIPES incorporates accessibility, it remains limited in addressing training and outcomes. The tool incorporates a quantitative dimension; the Pedi-PIPES index is a non-weighted sum of its data points.
2. **Children's Surgical Center Designations with Scope of Practice (CSCDSP):** The CSCDSP tool aims to evaluate pediatric surgical centers by categorizing facilities based on their scope of practice. It contains 27 items and focuses on assessing resources at the

facility level. However, it lacks a quantitative index and does not assess training or surgical outcomes, making it more resource oriented.

3. **Checklist for a Children's Trauma Room (CCTR):** The CCTR is an equipment checklist designed for pediatric trauma rooms, containing 165 items. It is based on a similar tool for adult trauma rooms and aims to prevent omissions in stocking trauma rooms with critical supplies. CCTR does not address training, accessibility, or outcomes and is purely focused on material resources.

None of the three tools have undergone formal validation following established guidelines for tool development for internal consistency, inter- and intra-rater reliability, construct validity, criterion validity or content validity[34, 35]. Moreover, all tools are resource-heavy with insufficient attention to training, accessibility, and outcome measurement. Neonatal-specific components are particularly underrepresented across all the tools.

Most notably, none of the tools have been evaluated against outcome metrics to establish a relationship between their results and pediatric surgical outcomes in L-MICs. Despite its widespread use, Pedi-PIPES has not undergone formal validation and places a disproportionate focus on resources. It is used exclusively as a capacity survey tool. In a recent study by Gajewski et al., Pedi-PIPES was applied in 67 district-level hospitals across Malawi, Tanzania, and Zambia. The study concluded that access to safe pediatric surgical care is compromised in these regions. However, no research has yet examined the correlation between the Pedi-PIPES index and the ability to perform "safe" surgical care. The Pedi-PIPES index has not been directly linked to specific morbidity or mortality rates or national health indicators, thus it cannot be definitively concluded that a low Pedi-PIPES index signifies reduced access to safe pediatric surgical care[55].

Similarly, a study by Okoye et al. makes analogous assumptions regarding the results of the Pedi-PIPES survey and the capacity of health facilities to deliver essential and emergency surgical care. Pedi-PIPES data from 41 hospitals in SSA, primarily in Nigeria, were collected and compared to the countries' Gross Domestic Product (GDP). The authors concluded that, despite Nigeria's stronger economic indicators, there were no significant advantages in the delivery of essential and safe surgical care compared to other countries. However, the relationship between the ability to deliver safe surgical care and the Pedi-PIPES index has not been formally investigated, and the study thus extrapolated their conclusions without identifying an association with the Pedi-PIPES index and the ability to deliver essential and emergency surgical care [62].

Wright et al. applied a heavily modified version of the Pedi-PIPES tool in their landmark PaedSurg Africa study [63]. The hospital survey conducted by Wright's group collected data on personnel, infrastructure, procedures, and anesthesia/resuscitation. Unlike the original Pedi-PIPES, each variable was rated on a three-point ordinal scale: always, sometimes, or never available, which was later modified to indicate low, medium, or high availability, respectively. The study's primary outcome was all-cause in-hospital mortality for discharged patients, or 30-day post-primary intervention mortality for those still admitted. Secondary outcomes focused on morbidity, including surgical site infection, wound dehiscence, unplanned re-interventions, length of hospital stay, and 30-day post-primary intervention mortality. However, during the analysis, the researchers reverted to the original dichotomous stratification of Pedi-PIPES without explaining the switch between these methodologies. Additionally, rather than using the Pedi-PIPES index, the authors opted to divide each of the four domains into three subgroups to maintain data granularity. Unfortunately, they did not provide clear criteria or rationale for defining these subgroups. This lack of explanation raises concerns about whether the chosen stratification compromised the level of detail in assessing resource availability. This study presented a valuable opportunity to validate Pedi-PIPES by linking it to morbidity and mortality outcomes in the congenital anomalies under investigation. Despite this opportunity, such validation was not undertaken. Establishing a correlation between Pedi-PIPES scores and actual clinical outcomes, such as morbidity and mortality rates, would have provided crucial evidence regarding Pedi-PIPES predictive value and its utility in assessing the capacity to deliver safe pediatric surgical care. The absence of this validation leaves the connection between resource availability, as measured by Pedi-PIPES, and clinical outcomes unexamined, limiting the broader applicability of the tool in resource-limited settings.

The necessity of evaluating the outcomes of surgical partnerships between HIC and L-MIC has been consistently emphasized in the literature, with numerous publications identifying critical components for successful collaborations [9, 16, 38]. While much attention has been directed towards defining "effective" or "successful" partnerships, the criteria for these terms remain ambiguous. Without clearly defined metrics, it becomes challenging to objectively measure the quality of these partnerships [15, 38, 41]. Often, the sustainability of global surgical partnerships is used as a primary indicator of success [15, 38, 41, 64], yet sustainability does not inherently reflect improvements in surgical capacity or surgical outcomes.

Binda et al. proposed a framework for the evaluation of sustainable global surgical partnerships, utilizing a Delphi process to identify six core pillars: stakeholder engagement, multidisciplinary collaboration, context-relevant education and training, bilateral authorship, multisource funding, and outcome measurement [38]. While outcome measurement was recognized as a crucial pillar, the outcomes addressed in their framework primarily pertained to the educational aspects of these partnerships. Moreover, their final framework included a statement recommending the inclusion of "a critical analysis of potential harms and negative impacts of the partnership." However, no specific tool or method for conducting such an analysis was provided. The authors envisioned their framework as a pre-implementation checklist for assessing the readiness of prospective global surgical partnerships. While this tool is undoubtedly essential, continuous monitoring of partnerships and the development of robust metrics to evaluate their success are equally important to ascertain the value of sustainability. After all, why should institutions on either side invest time, energy, and resources in a partnership that, while sustainable – depending on the definition of such, may not yield meaningful benefits for the institutions involved or, more critically, for the patient population served?

As described in Manuscript 2, Phase I thus informed the creation of a tool based on five core domains: human resources, material resources, accessibility, outcomes, and education. The initial version of GAPS consisted of 139 items, which were refined through feedback from a 37-member panel of pediatric surgery specialists, resulting in a revised tool of 168 items. Subsequently, GAPS Version 2 underwent international pilot testing across 65 institutions in eight diverse countries to assess its feasibility and applicability in real-world settings. This phase revealed high response rates for most sections but noted persistent challenges in capturing outcome data due to limited reporting infrastructure within many of the participating L-MICs. To ensure the tool's validity, construct validation and item analysis was undertaken, in which institutions were categorized by level of care—basic or advanced—and items were evaluated based on their ability to differentiate effectively between these care levels. The analysis led to the retention of questions that could discriminate between the institutional care levels with high accuracy, culminating in a refined set of 64 questions that demonstrated a high response rate.

Finally, instrument refinement produced GAPS Version 3, a streamlined tool containing 64 items across the original five domains: human resources, material resources, accessibility, education, and outcomes. This version was specifically tailored to focus on items that best distinguished care levels, making it an intentional and comprehensive framework for assessing pediatric surgical capacity in resource-constrained settings. Through its design, GAPS aims to address the gaps identified in prior tools by providing a robust, standardized means to evaluate pediatric surgical capacity. This approach supports targeted improvements and fosters international partnerships for optimizing pediatric surgical care in L-MICs, addressing in an evidence-based objective fashion a critical need for a comprehensive, reproducible framework for assessing surgical capacity in diverse healthcare contexts.

In an effort to go beyond a checklist style CAT and create an objective method of comparison between and within institutions, Manuscript 3 outlines the incorporation of capacity-based weighting in the GAPS tool. Incorporating a weighted score to GAPS ensures that each domain contributes meaningfully to the overall assessment score, based on its unique impact on surgical capacity and outcomes. The development process involved a multi-step approach: assigning weights to each item within its domain using logistic regression to base the weighted scores on health indicators such as the under-five mortality rate, neonatal mortality rate, and the Human Development Index, which collectively underscored the tool's ability to reflect pediatric health outcomes accurately. The results showed that the GAPS score could effectively distinguish between hospital levels, with higher scores correlating with greater resource availability and higher-income regions. This improved granularity allowed the establishment of benchmark scores for each hospital level, providing clear performance standards and identifying capacity gaps across institutions. By capturing these differences, the weighted GAPS tool offers a more robust framework for evaluating pediatric surgical services in resource-limited settings, supporting global health initiatives that aim to strengthen pediatric surgery in L-MICs through evidence-based partnerships and interventions.

Implications for Global Pediatric Surgical Capacity

Despite concerted efforts over the past decades, disease burden, access to care, and outcomes for pediatric surgical disease in L-MICs remain imprecisely defined and likely underestimated [28, 49, 61, 65-68]. Figure 7 illustrates the causes of death among children under 5-years old in low-income countries[69]. Figure 8 represents the causes of deaths globally in children < 14 years-old[70]. Both charts highlight the shift over time from communicable to non-communicable diseases, with congenital anomalies comprising a significant portion of cause of mortality[49, 65, 71-73]. Banu et al conducted a review to assess the burden of surgical congenital anomalies with emphasis on those that present within the first 8000 days of life[73]. They demonstrated that children in lower Sustainable Development Index nations experience higher mortality rates from comparable causes[70, 73]. Furthermore, many of the current measures used to evaluate disease burden neither account for the physical, psychological, and financial impacts on families caring for sick or injured children [74, 75] - nor do we know the impact of social determinants of health on surgical outcomes in L-MICs. Moreover, there exists specific challenges related to providing resources for pediatric surgical services in L-MICs [7, 8, 51, 53, 76]. The avertable surgical burden of disease in children is substantial with surgical conditions accounting for 30% of the global burden of disease, and children compromising approximately 50% of the population in the least developed regions of the world [7, 8, 51, 53, 77]. Developing strategies to improve the quality of pediatric surgical care in low-resource settings may significantly decrease child morbidity and mortality. Most importantly a focus on improving child surgical care protects an important segment of society and may contribute to an economically productive future [78].

Figure 7. Change in Child Mortality in Low-Income Countries

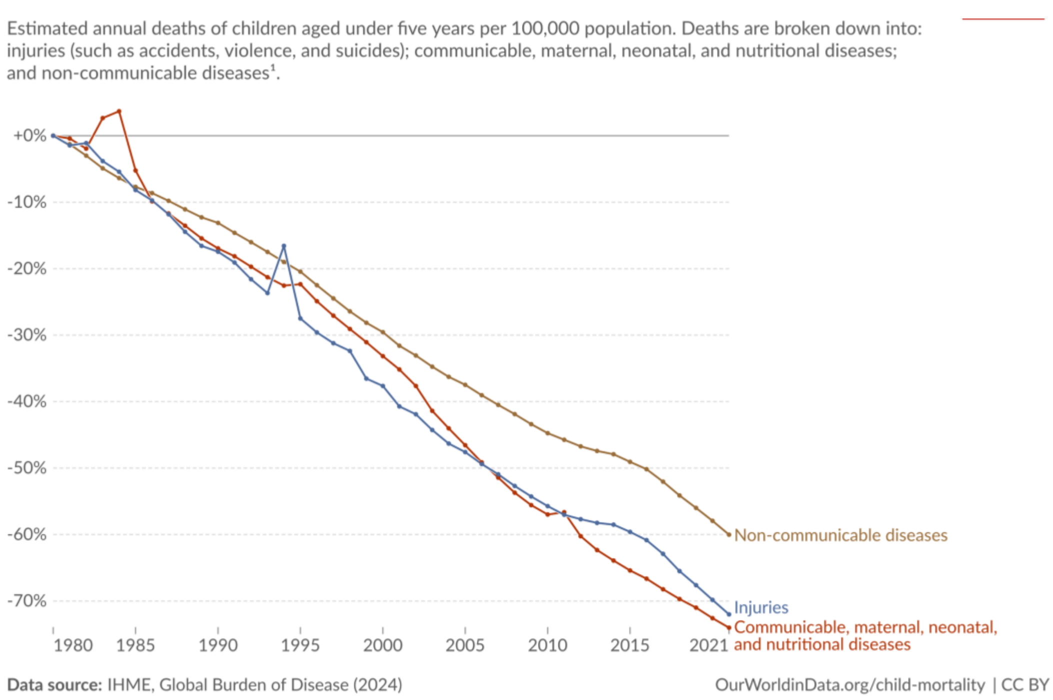
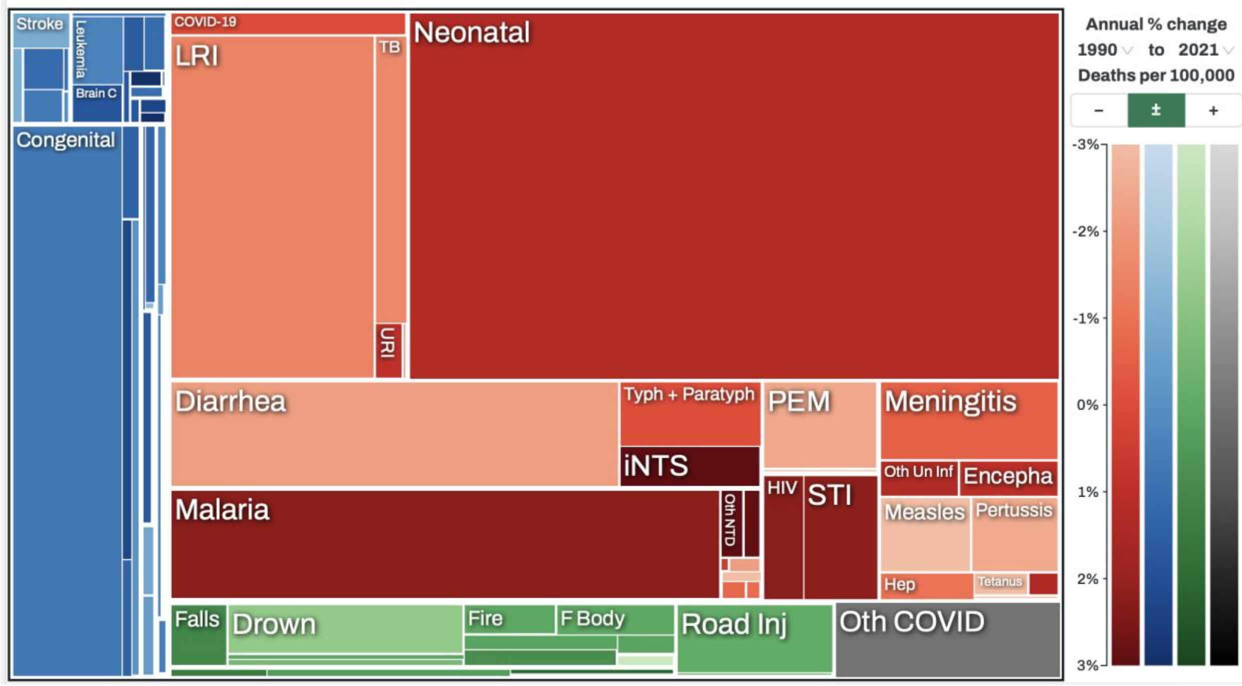


Figure 8. Causes of death globally in children < 14 years-old



In addition to specialized surgical equipment, dedicated infrastructure, advanced training, and patient follow-up, optimal pediatric surgical care mandates an integrated approach encompassing maternal healthcare support, ethical frameworks, preventative strategies, interdisciplinary collaboration, and a nuanced understanding of various sociocultural determinants [45, 50, 61, 77]. While tools like Pedi-PIPES focus on procedural, infrastructure, personnel, equipment, and supplies, GAPS extends this scope by incorporating necessary outcome measures and capacity-building elements [79]. Compared to the CSCDSP, GAPS offers a more complete assessment of surgical capacity by delineating a dynamic and quantitative evaluation of service capacity [22, 28, 80]. Beyond the scope of the items, GAPS introduces a novel capacity-based weighting system that acknowledges the differential impact of various domains on pediatric surgical outcomes, an aspect not found in other CATs. This quantitative feature of GAPS reflects the nuanced complexity of health service delivery in L-MICs, underscoring its unique contribution to the field and its potential to guide appropriate resource allocation, training initiatives, and policy reforms for pediatric surgery—a crucial need [81].

The GAPS tool serves as a standardized approach to evaluating pediatric surgical capacity across different healthcare facilities by providing a uniform method for identifying existing gaps in infrastructure, human resources, accessibility, education and outcomes in pediatric surgical settings. This is a dilemma that has plagued multiple HIC to L-MIC relationships in almost all subspecialties [10, 20]. Lippa et al published-on initiative and challenges within the global neurosurgical community and created a tool to serve as a guideline for clinical partnerships the Engagements and assets, Capacity, Operative autonomy, Sustainability, and scalability (ECOSystem) [82]. They comment that few published research studies centered on global pediatric neurosurgery report a proper needs assessment from their partnership [82, 83]. They emphasize how crucial it is to distinguish between the needs identified by neurosurgeons, those outlined by ministry experts, and comparisons with neighboring countries [82]. The ECOSystem of care involves 5 tiers: Tiers 0 (foundation), 1 (essential), 2 (complexity), 3 (autonomy), and 4 (final) [82]. Tier 0 involves establishing a new connection to plan a sustainable collaboration [82]. Tier 1 lists essential items and educational efforts to manage essential clinical needs in neurosurgery safely [82]. In a survey distributed to surgeons from L-MICs to assess the nature and perception of global surgery collaborations almost all participants felt a formal process of collaboration and a structured partnership would benefit all parties in assessing needs [14]. Only 77% of respondents reported expressing their needs before HIC team arrival [14]. Kisa et al analyzed the different models of pediatric global surgical partnerships and identified that clear objectives to partnerships are instrumental to success [39]. These objectives should ideally be driven by local hosts site needs and priorities though identify objective goals can be difficult with an instrument to assess the starting point and follow-up. the ability to identify these objectives [39].

Guilfoyle et al conducted a systematic review what makes global healthcare partnerships successful and determined that of 26 collaborations, only two had more than a decade of collaboration and had achieved a “self-sustaining program”; the definition of which was not clear [41]. Lack of funding was identified as a major barrier to sustainability. With the detailed assessment of human resources, material resources, education, accessibility, and outcomes, GAPS can help policymakers and healthcare organizations allocate resources more effectively. The tool's weighted scoring system highlights the areas needing the most attention, ensuring that resources, including financial, are directed toward the most impactful aspects of pediatric surgical care. Guilfoyle et al underscored that funding should be secured prior to program development in anticipation of a long-term project and proposed a list of guidelines to improve the chances of a successful and sustainable collaboration [41].

Training in L-MICs represents a promising opportunity within the evolving landscape of artificial intelligence, remote learning, and simulation-based medical education. Despite its potential, simulation-based medical education has not been extensively implemented in L-MICs. The GAPS tool highlights the critical role of education and training as key components of surgical capacity. By incorporating educational resources as an assessment domain, GAPS promotes the creation of training programs tailored to the unique needs of pediatric surgery in resource-limited settings, ultimately contributing to better outcomes through well-trained surgical teams. For instance, a recent study conducted over four years in Myanmar implemented 11 simulation-based medical education courses, which demonstrated significant improvements in all self-rated confidence measures on pre- and post-course Likert scales[84]. This example illustrates a cost-effective initiative that can enhance education in an L-MIC once the need has been identified. The GAPS tool effectively identifies such educational needs and provides a means to measure improvements following interventions beyond basic self-rated confidence assessments.

In addition, the ethics of global surgery needs to be considered when discussing global surgical interventions [45]. Grant et al published a scoping review on ethics in global surgery and identifying gaps in the current literature. They identified four domains in the literature on the ethics of global surgery [44]:

1. DOMAIN 1: Clinical Care and Delivery
    a. Global surgical missions can strain local resources, disrupt continuity of care, and present ethical challenges in patient selection. Variations in care standards and issues with cultural awareness and informed consent can compromise patient outcomes.
2. DOMAIN 2: Education, Exchange of Trainees, and Certification
    a. Educational aspects of global surgery often fail to transfer knowledge effectively to host communities, with training sometimes irrelevant to local needs. Visiting trainees may face insufficient supervision and be unprepared for local challenges, impacting both their effectiveness and local training opportunities.
3. DOMAIN 3: Research, Monitoring, and Evaluation

   a. Global surgical research can marginalize LMIC researchers and fail to secure dual ethics reviews, limiting local benefits. Research may exploit vulnerable populations, and inadequate monitoring of outcomes hampers quality improvement.

4. DOMAIN 4: Engagement in Collaborations and Partnerships
   a. Sustainability in global surgical partnerships is often overlooked, with inadequate local involvement and problematic donation practices. Dependence on external aid can disrupt local systems and hinder long-term self-sufficiency.

These findings mirrored those detailed by Cunningham et al in 2019, who identified clinically care, education, research and advocacy as central pillars in establishing ethical global pediatric surgery partnerships [45]. These pillars are again reiterated by GICS foundational principles: infrastructure, service delivery, training and research [10]. These four metrics served as the foundations for the OReCS [10]. The OReCS document was a foundational document for GAPS and thus all of the domains identified are evaluated through this novel capacity assessment tool. The data generated from the GAPS tool can support an ethical framework for partnerships and help guide policy decisions and advocacy efforts aimed at improving pediatric surgical care globally. By providing concrete evidence of disparities and the impact of resource gaps, the tool empowers healthcare leaders to advocate for greater investment and reform in pediatric surgical care within an ethical framework.

Methodological Insights and Contributions

The development of the GAPS tool followed a structured, multi-phase methodology that adhered to established best practices in health measurement scale construction. This process was informed by existing methodologies for scale development as described by DeVellis et al. and Streiner et al. and entailed item generation, response scaling, item selection, iterative refinement, and item evaluation [34, 35, 85]. Each of these iterative stages contributed to the overall validity and reliability of the tool, essential for its effectiveness in evaluating pediatric surgical capacity in low-resource settings.

Devising the Items

The initial phase was dedicated to devising a comprehensive set of items that encapsulated the construct of pediatric surgical capacity in L-MICs. This step was rooted in an extensive systematic review of existing capacity assessment tools, relevant literature and theoretical frameworks [10, 20, 32, 34, 35, 38], to ensure that all aspects of pediatric surgical needs were comprehensively covered, as outlined in Manuscript 1 [28]. This foundational step was critical for establishing content validity by identifying relevant domains such as human resources, material resources, outcomes, accessibility, and training and creating an item pool that comprehensively reflected these domains.

Scaling Responses

The second phase involved developing response scales that captured the gradations of each item effectively[34, 35]. The GAPS tool employed ordinal scaling, featuring structured response options designed to distinguish between varying levels of institutional capacity. These response scales were specifically tailored to differentiate between the predefined designations of basic and advanced surgical care, aligning with the core construct of interest: pediatric surgical capacity. This phase emphasized clear response anchors to guide consistent interpretation by respondents, contributing to the tool's reliability and accuracy. This design choice allowed the tool to effectively stratify healthcare facilities by their ability to deliver basic or advanced pediatric surgical care.

GAPS Version 1, which comprised an initial set of 139 items distributed across five domains. The preliminary item set underwent a rigorous evaluation by a panel of experts, which included pediatric surgeons, global health specialists, and allied professionals. Pretesting the chosen scaling format through expert review and preliminary trials ensured that respondents could interpret and engage with the scale consistently. This expert consultation facilitated the refinement of item content, ensuring alignment with practical realities and needs in L-MICs. This led to the expansion and revision of the item pool, culminating in GAPS Version 2, with 168 items.

Selecting and Refining the Items

Following the generation and scaling of items, the next step was to select items based on their empirical performance. This involved pilot testing GAPS Version 2 across 65 institutions covering eight countries and analyzing item response rates and discriminatory power to ensure that only the most relevant items were retained for the final tool version.

Pilot testing of GAPS Version 2 was conducted between February 2017 and March 2018 in 65 institutions spanning low income, low-middle income, and upper-middle income countries.

The pilot included various types of institutions, including public, faith-based, and non-governmental hospitals, to capture a broad spectrum of pediatric surgical capacities. Data from the pilot were analyzed to evaluate item response rates and discriminatory power between basic surgical care and advanced surgical care facilities.

Empirical analysis of pilot data assessed item performance, focusing on their discriminative power to distinguish between basic and advanced care facilities. Statistical analyses focused on item performance, using measures such as Fisher's exact test to determine the ability of items to distinguish between different levels of care, bolstered by expert-defined "ideal" responses for each item. Items with high discriminatory power and response rates above 75% were retained, while those with poor performance or ambiguity were revised or removed. This empirical approach ensured that GAPS Version 3, comprising 64 items, reflected the tool's intended construct and provided robust differentiation between levels of care. Item Response Theory was then applied to ensure that each item aligned with the intended latent trait and contributed meaningfully to the scale's overall reliability[34, 35].

The outcomes and accessibilities section in the piloted GAPS Version 2 garnered a limited response rate, a phenomenon attributable to the longstanding challenges familiar to research in the developing world. Challenges surrounding data collection are a well-known barrier to research in L-MICs [52, 77]. Many L-MICs may lack the infrastructure or resources to consistently track and report outcomes, making it challenging to prove quality improvement with any intervention [86]. Furthermore, many institutions in L-MICs may not have established mechanisms for reporting adverse outcomes or monitoring patient access, leading to underreporting [87]. Outcome reporting in HIC is often accompanied with burdensome data collection due to the multitude of data variables in data collection tools such as American College of Surgeons' National Surgical Quality Improvement Program (ACS NSQIP) with over 130 patient and operative variables. Dimick et al showed that for five core general surgery operations procedure-specific hospital quality measures can be adequately risk-adjusted by limiting the number of variables to five (ASA class, functional status, congestive heart failure, dialysis, and bleeding disorder) [88]. Following this, Ullrich et al reported on a minimum dataset needed for adequate risk adjustment to predict perioperative mortality in pediatric surgical procedures in LMICs. They showed that three objective variables (Diagnosis, Procedure, and District) generated an area under the receiver operated curve (AUROC) of 0.92. An excellent result considering the minimum data required and that an AUROC value of 1.0 indicates the model can completely differentiate between two groups. However, this study was conducted in a single center and the model has yet to be externally validated. Though not available during the development of GAPS as described in this thesis, such an abbreviated dataset may be useful in determining outcomes data in the future.

Evaluating the items

Item Response Theory (IRT), specifically the Graded Response Model (GRM), was chosen to assess the reliability of the GAPS tool due to its detailed item-level analysis and ability to measure item discrimination and difficulty effectively. IRT focuses on item-level characteristics, allowing for a more nuanced understanding of each item's contribution to the overall scale. This level of analysis is critical for a multi-faceted tool like GAPS, which assesses various domains. The use of the GRM within the IRT framework provided insights into the discrimination and difficulty parameters of each item. Discrimination indicates how well an item can differentiate between institutions at different levels of capacity. Items with high discrimination values were retained as they effectively distinguished between basic and advanced pediatric surgical care. Difficulty parameters allowed understanding of what point on the capacity spectrum an item was most informative, ensuring that items targeted the appropriate levels of the construct for the settings being assessed. IRT allowed for the addition or removal of items without impacting the scale's overall reliability and validity. This feature was valuable for GAPS, as it may need to evolve over time to adapt to new insights or changes in global pediatric surgery practices. With IRT, any modifications to the tool can be seamlessly incorporated while maintaining consistency in its measurement properties.

Content Validity

Content validity was achieved during the phases of item development and expert consultation. As described in Manuscript 2, GAPS Version 1 was rigorously scrutinized by a panel of six authors to ensure its comprehensibility, clarity, relevance, and alignment with the unique challenges faced in LMICs. Thereafter, GAPS Version 1 underwent further content validation by a 37-person panel of experts in pediatric surgery, representing a diverse assortment of healthcare settings. They provided feedback on the relevance, utility, and clarity of each data point, as well as the pertinence of its inclusion in the tool. Additionally, participants were asked to identify standard responses for all categories based on the level of the healthcare facility.

Construct and Criterion Validity

Construct validity was achieved during the phases of item development and expert consultation through discriminative validation, comparing the responses of institutions categorized as basic surgical care (1st - and 2nd -level hospitals) and advanced surgical care (3rd-level hospitals and national children's hospitals) as described in Manuscript 2. Items were evaluated for their alignment with predefined "ideal" responses per level of care established by the expert panel. Responses for each question were categorized into basic surgical care and advanced surgical care. These responses were then evaluated against an objective assessment of the hospital's classification to verify whether they aligned with the hospital's official governmental designation.

Criterion validity was explored in Manuscript 3 through analyses that compared GAPS scores to external health indicators. The relationship between GAPS scores and variables such as human development index, under-five mortality rate, and neonatal mortality rate was assessed. This phase reinforced the tool's predictive validity, demonstrating that higher GAPS scores correlated with improved health outcomes, affirming the tool's capability to act as a reliable metric for guiding strategic interventions in global pediatric surgical care.

CHALLENGES AND LIMITATIONS

A notable limitation of the GAPS tool is that it does not directly assess surgical output, such as the number of surgeries performed, what those surgeries are, surgical outcomes, or postoperative recovery metrics. While the tool focuses on evaluating resources, accessibility, and institutional capacity to perform pediatric surgery, it encompasses few measures that reflect the effectiveness or success of surgical procedures themselves. GAPS Version 1 comprised 22 items related to outcomes, while GAPS Version 2 expanded to include 27 outcome-related items. Of these, GAPS Version 3 retained only three items: (1) What is the 30-day follow-up rate after major procedures?; (2) How often is the airway and ventilation continuously monitored while providing anesthesia? (3) How often is adequacy of ventilation monitored with capnography?. All outcomes related to morbidity, mortality, or specific procedures were excluded after the pilot study due to a lack of responses. This omission indicates that, although GAPS provides valuable insights into a center's capacity to deliver surgical care, it does not offer a full assessment of performance quality or patient outcomes. Integrating such metrics would require a rigorous outcome monitoring system, which most centers in L-MICs do not currently have in place[52, 87, 89, 90]. This lack of established outcome tracking infrastructure poses a significant challenge to incorporating robust surgical output measures and highlights an area for future development and capacity-building in L-MICs to support comprehensive evaluations.

While the pilot study confirmed the tool's utility in varied settings, certain limitations were noted, including the inability to assess test-retest reliability due to logistical constraints in multi-country deployment. To maximize the number of institutions and countries included in the pilot study, it was not feasible to administer GAPS to the same respondents (test-retest reliability) or to different respondents within the same institution over time (inter-rater reliability). These limitations, acknowledged in the Manuscript 2, highlight the need for further studies to explore the temporal stability of the tool, an example of this is test-retest reliability. Test-retest reliability refers to the consistency of a measure or test over time. It assesses the extent to which the same results are obtained when the same test is administered to the same subjects under similar conditions on two or more separate occasions. This type of reliability indicates the stability and dependability of the instrument or method being tested. Assessment of test-retest reliability for GAPS would have provided valuable insight into the temporal stability of the tool. Moreover, the pilot study did not assess inter-rater reliability. Inter-rater reliability is a measure used to assess the degree of agreement or consistency between different raters or observers evaluating the same phenomenon. It indicates how consistently different individuals produce similar ratings or assessments when examining the same subject under comparable conditions. Both test-retest reliability and inter-rater reliability assessments were not feasible during the pilot study due to practical limitations. The pilot was conducted in clinically demanding settings where resource constraints, time limitations, and the variability of participants' availability posed significant challenges. Test-retest reliability was hindered by the logistical difficulty of repeated evaluations within the same centers under similar conditions. Similarly, inter-rater reliability, which requires multiple raters independently scoring the same items to ensure scoring consistency, was not achievable due to the limited availability of participants.

In the early phases of GAPS, while expert consultations were essential for content validation, the process inherently carries a degree of subjectivity, which may introduce bias in item selection and relevance assessments. Although the expert panel included a range of specialists from various healthcare settings, potential bias could arise from the composition of the panel, which may not fully represent all perspectives or challenges encountered in different L-MICs. The pilot phase took place in eight countries with 63% being in the DRC and 70% in low-income countries. Though the pilot included diverse settings with significant variability in resources, the proportion of responses from a single country could have potentially influenced item performance on statistical analyses and carry an impact on the tool's generalizability. Understandably, the collection of outcome data in L-MIC is fraught with difficulty [38, 52, 63, 87, 89, 91-96]. Moreover, data collection relied on self-reported information provided by representatives from the institutions which carries inherent risk of response bias. In the development of GAPS many institutions lacked robust monitoring systems for surgical outcomes, leading to limited response rates for certain sections of the tool. Moreover, due to poor outcomes data within countries and institutions, during criterion validity external health indicators such as human development index, under-five mortality rate, and neonatal mortality rate were used.

An important limitation of the present study was the use of hospital levels as the primary outcome for differentiating between questions included in the GAPS tool in Phase II (Manuscript 2). The criticism that this methodological approach is inherently subject to circular logic is a valid one; the classification of hospital levels is typically determined by governmental or administrative organizations using criteria that are neither standardized nor consistent across different countries or health systems. Though there have been efforts made to normalize the definitions of first-level, second-level, third-level, and national children's hospital through the DCP3 and OReCS documents, these definitions have not yet been widely implemented[22, 24, 25, 31, 32, 48]. Consequently, hospital-level designations by country may differ widely in their definitions, scope, and clinical capabilities from what is described in the DCP3 or OReCS documents. Notwithstanding, these administrative classifications do not necessarily correspond directly to patient outcomes or actual clinical performance, thus limiting their relevance as valid indicators of institutional surgical capacity. Despite recognizing these critical limitations, hospital-level classification was selected as the outcome metric primarily due to the absence of other consistently available metrics for children's surgical capacity across participating institutions. Alternative potential outcome measures, such as postoperative morbidity, mortality rates, or complication rates, were explored but proved infeasible due to inconsistent availability, unreliable documentation, and lack of standardized reporting practices among the surveyed hospitals. Patient-level outcomes or case complexity assessments could have theoretically provided more clinically relevant insights but would have introduced considerable variability and been logistically impractical given the scope and resources of the study. Consequently, using predefined governmental or administrative hospital levels, despite their inherent flaws, was the most pragmatic option available to facilitate standardized comparisons across a diverse set of international institutions. It provided the only realistically implementable surrogate measure capable of supporting initial validation and discrimination of items within the GAPS tool. This choice thus represented a practical compromise, prioritizing feasibility and cross-context applicability over ideal, but unachievable, precision. Future studies should aim to develop, or leverage validated clinical bellwether procedures and universally accepted pediatric surgical outcome metrics, which could serve as more objective and clinically relevant standards to meaningfully assess and compare surgical capacity across varied low-resource settings.

Another notable limitation in the development of GAPS (Manuscript 2) was retaining the question, "*The presence of pediatric general surgeons/ urologist*", despite it not demonstrating significance in IRT analysis. At the time of the development of GAPS, defined bellwether procedures for children's surgery had not yet been investigated. Since then, potential bellwether procedures have emerged—including laparotomy, repair of tracheoesophageal fistula, gastroschisis management, anorectal malformation repair, congenital diaphragmatic hernia repair, omphalocele management, intestinal atresia repair, Hirschsprung's pull-through, appendectomy, inguinal hernia repair, cleft lip repair, shunt insertion for hydrocephalus, myelomeningocele closure, trauma craniotomy, external ventricular drain insertion, and open fracture treatment[63, 97-104]. These procedures, among others, may represent more clinically relevant indicators of surgical capacity and outcomes than the mere presence of specialist providers. Indeed, healthcare professionals in low-resource settings frequently task-share, possess varied types and levels of training, and may lack formal credentials commonly expected in high-resource environments[101, 105, 106]. Thus, tracking how many complex pediatric procedures, such as esophageal atresia repairs, are successfully performed might more accurately reflect institutional surgical capacity than the presence of a pediatric surgeon or urologist alone[43, 54]. Unfortunately, we could not integrate such bellwethers into the GAPS tool because they had not been identified when this study commenced in 2016, nor have they since undergone comprehensive validation for children's surgery. Additionally, reliable data collection mechanisms to quantify procedure volumes and outcomes across low-resource settings remain insufficient[87, 101]. Incorporating such procedures into GAPS would significantly enhance the tool's ability to directly assess surgical output and patient outcomes, marking an essential step toward a more outcome-driven evaluation framework. Although beyond the scope of this study, validating and incorporating bellwether procedures into GAPS represents an important and promising direction for future research.

Beyond the methodological limitations, maintaining the relevance of the GAPS tool as pediatric surgical practices and global health standards evolve will require periodic updates. Future plans to ensure this include developing a free application accessible to stakeholders. This app would not only enhance the tool's usability and accessibility but would also enable continuous data collection from its deployment in various centers. The aggregated data would be used to regularly update the GAPS score and refine the tool, ensuring it remains current and effective in assessing pediatric surgical capacity.

CHAPTER 7: SUMMARY

SUMMARY

The first article, *GAPS Phase I: Systematic Review of Capacity Assessment Tools in Pediatric Surgery*, provided an overview of existing tools for evaluating pediatric surgical capacity and highlighted their limitations. The review identified a significant gap in comprehensive tools designed specifically for assessing pediatric surgical capabilities in L-MICs. This article set the foundation for developing the GAPS tool by emphasizing the need for a targeted, reliable, and validated instrument.

The second article, *GAPS Phase II: Development and Pilot Results of the Global Assessment in Pediatric Surgery*, outlined the creation and pilot testing of the GAPS tool. The study included item development, expert consultation, and pilot testing across 65 institutions in various L-MICs. The results confirmed the 64-item tool's ability to discriminate effectively between basic and advanced surgical care levels, establishing its construct validity and highlighting the importance of reliable, context-specific assessments for pediatric surgical capacity.

The third article, *GAPS Phase III: Incorporation of Capacity-Based Weighting in the Global Assessment for Pediatric Surgery*, focused on refining the GAPS tool by integrating a capacity-based weighting system. This enhancement allowed for better representation of each domain within the tool and improved its correlation with key health indicators such as the human development index and mortality rates. The study underscored the tool's predictive capacity, demonstrating its value in informing partnerships and resource allocation in LMICs.

The significance of the GAPS tool lies in its targeted approach to addressing the global challenge of assessing pediatric surgical capacity in LMICs. By providing a validated, comprehensive instrument, GAPS offers stakeholders in global health an evidence-based method to identify gaps in pediatric surgical care, promote effective partnerships, and guide strategic interventions to improve surgical outcomes for children in resource-limited settings.

This thesis contributes both theoretically and practically to the field of global surgery by developing an evidence-based tool that fills a critical gap in pediatric surgical assessment. Conceptually, it advances the understanding of how pediatric surgical capacity can be conceptualized and measured across diverse healthcare settings. The thesis incorporates robust methodologies such as systematic reviews, expert panel feedback, and modern statistical approaches like Item Response Theory, enhancing the rigor and comprehensiveness of the assessment tool. Practically, the GAPS tool provides healthcare systems, policymakers, and global health organizations with a reliable means of evaluating pediatric surgical capacity, supporting targeted improvements and fostering international collaborations aimed at strengthening surgical care in L-MICs.

Future Directions

Currently, the GAPS tool is already demonstrating its practical utility in a significant international effort led by a Swiss team at Lucerne University. This team is leveraging the GAPS tool to assess pediatric surgical capacity in over 20 institutions across Burkina Faso. This large-scale application underscores the tool's adaptability and value in real-world settings, providing critical insights into healthcare capacity and guiding strategic interventions. The outcomes of this ongoing project will offer further validation of the GAPS tool and contribute to its continuous improvement, setting a precedent for similar initiatives in other regions. This example highlights how the GAPS tool can serve as an effective instrument for assessing and enhancing surgical care capacity in L-MICs.

Future directions for the GAPS tool include creating a free, user-friendly application to enhance accessibility and utilization worldwide. Such an application would enable multiple participants within the same institution to complete the GAPS tool, allowing for assessments of test-retest reliability and inter-rater reliability. This functionality would not only support the tool's continuous refinement but also allow institutions to monitor their progress over time, especially during and after partnerships with HICs or other L-MICs. Additionally, the application would provide participants with visual and numeric representations of the overall GAPS score and its five domains, as outlined in Manuscript 3, making the data more interpretable and actionable. Ultimately, the goal is for the GAPS tool to become a cornerstone in global surgical assessments, driving strategic improvements and fostering data-driven partnerships that enhance pediatric surgical care across the world.

REFERENCES

1. Meara, J.G., et al., *Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development.* Int J Obstet Anesth, 2016. **25**: p. 75-8.
2. Fowler, Z., et al., *Academic Output in Global Surgery after the Lancet Commission on Global Surgery: A Scoping Review.* World J Surg, 2022. **46**(10): p. 2317-2325.
3. Data., W.H.O.p.b.O.W.i. *World map region definitions*. Available from: https://ourworldindata.org/world-region-map-definitions.
4. Data, W.P.P.p.b.O.W.i., *Age Structure*.
5. 2023), U.N.I.-a.G.f.C.M.E. *Child and Infant Mortality*. Available from: https://ourworldindata.org/child-mortality?insight=youth-mortality-rates-vary-widely-around-the-world#key-insights.
6. Bickler, S.W., M.L. Telfer, and B. Sanno-Duanda, *Need for paediatric surgery care in an urban area of The Gambia.* Trop Doct, 2003. **33**(2): p. 91-4.
7. Smith, E.R., et al., *Waiting Too Long: The Contribution of Delayed Surgical Access to Pediatric Disease Burden in Somaliland.* World J Surg, 2020. **44**(3): p. 656-664.
8. Smith, E.R., et al., *Is Global Pediatric Surgery a Good Investment?* World J Surg, 2019. **43**(6): p. 1450-1455.
9. Woods, A., C. Shofner, and B. Hodge, *International pediatric surgery partnerships in sub-Saharan Africa: a scoping literature review.* Glob Health Action, 2022. **15**(1): p. 2111780.
10. Global Initiative for Children's, S., *Global Initiative for Children's Surgery: A Model of Global Collaboration to Advance the Surgical Care of Children.* World J Surg, 2019. **43**(6): p. 1416-1425.
11. Butler, M.W., et al., *The Global Paediatric Surgery Network: a model of subspecialty collaboration within global surgery.* World J Surg, 2015. **39**(2): p. 335-42.
12. Greive-Price, T., H. Mistry, and R. Baird, *North-South surgical training partnerships: a systematic review.* Can J Surg, 2020. **63**(6): p. E551-E561.
13. Ekenze, S.O., O.O. Onumaegbu, and O.E. Nwankwo, *The current status of international partnerships for child surgery in sub-Saharan Africa.* Int Surg, 2014. **99**(5): p. 616-22.
14. Krakauer, K.N., et al., *Evaluating Global Surgery Partnerships From Low and Middle Income Country Perspectives.* J Surg Res, 2024. **296**: p. 681-688.
15. Jedrzejko, N., et al., *A systematic review of global surgery partnerships and a proposed framework for sustainability.* Can J Surg, 2021. **64**(3): p. E280-E288.
16. Riviello, R., et al., *Role of collaborative academic partnerships in surgical training, education, and provision.* World J Surg, 2010. **34**(3): p. 459-65.
17. Grimes, C.E., et al., *Guidelines for surgeons on establishing projects in low-income countries.* World J Surg, 2013. **37**(6): p. 1203-7.
18. Shrime, M.G., A. Sleemi, and T.D. Ravilla, *Charitable platforms in global surgery: a systematic review of their effectiveness, cost-effectiveness, sustainability, and role training.* World J Surg, 2015. **39**(1): p. 10-20.
19. Roche, S.D., P. Ketheeswaran, and V.J. Wirtz, *International short-term medical missions: a systematic review of recommended practices.* Int J Public Health, 2017. **62**(1): p. 31-42.
20. Goodman, L.F., et al., *The Global Initiative for Children's Surgery: Optimal Resources for Improving Care.* Eur J Pediatr Surg, 2018. **28**(1): p. 51-59.

21. Surgery, G.I.f.C.s. *Global Initiative for Children's Surgery*. 2022 [cited 2024; About US]. Available from: https://www.globalchildrenssurgery.org/about-us/
.
22. Oldham, K.T., *Optimal resources for children's surgical care.* J Pediatr Surg, 2014. **49**(5): p. 667-77.
23. Seyi-Olajide, J.O., et al., *Inclusion of Children's Surgery in National Surgical Plans and Child Health Programmes: the need and roadmap from Global Initiative for Children's Surgery.* Pediatr Surg Int, 2021. **37**(5): p. 529-537.
24. Global Initiative for Children's, S., *Guidelines for Different Levels of Care*, in *OPTIMAL RESOURCES FOR CHILDREN'S SURGICAL CARE*. 2019: GICS website. p. 50.
25. Global Initiative for Children's, S., *Supplies, Equipment and Infrastructure*, in *OPTIMAL RESOURCES FOR CHILDREN'S SURGICAL CARE*. 2019: GICS website. p. 31.
26. Global Initiative for Children's, S., *Guidelines for Different Levels of Care*, in *OPTIMAL RESOURCES FOR CHILDREN'S SURGICAL CARE*. 2019: GICS. p. 4.
27. Yousef, Y., et al., *GAPS phase II: development and pilot results of the global assessment in pediatric surgery, an evidence-based pediatric surgical capacity assessment tool for low-resource settings.* Pediatr Surg Int, 2024. **40**(1): p. 158.
28. Yousef, Y., et al., *A systematic review of capacity assessment tools in pediatric surgery: Global Assessment in Pediatric Surgery (GAPS) Phase I.* J Pediatr Surg, 2019. **54**(4): p. 831-837.
29. Yousef, Y., et al., *Risk stratification in pediatric perforated appendicitis: Prospective correlation with outcomes and resource utilization.* J Pediatr Surg, 2018. **53**(2): p. 250-255.
30. Butler, M., et al., *Guidelines and checklists for short-term missions in global pediatric surgery: Recommendations from the American Academy of Pediatrics Delivery of Surgical Care Global Health Subcommittee, American Pediatric Surgical Association Global Pediatric Surgery Committee, Society for Pediatric Anesthesia Committee on International Education and Service, and American Pediatric Surgical Nurses Association, Inc. Global Health Special Interest Group.* J Pediatr Surg, 2018. **53**(4): p. 828-836.
31. . Essential Surgery: Disease Control Priorities, Third Edition (Volume 1), ed. H.T. Debas, et al. 2015, Washington (DC).
32. Global Initiative for Children's, S., *Optimal Resources for Children's Surgical Care: Executive Summary.* World J Surg, 2019. **43**(4): p. 978-980.
33. Bickler, S.N., et al., *Global Burden of Surgical Conditions*, in *Essential Surgery: Disease Control Priorities, Third Edition (Volume 1)*, H.T. Debas, et al., Editors. 2015: Washington (DC).
34. DeVellis, R.F. and C.T. Thorpe, *Scale development : theory and applications*. Fifth edition. ed. 2022, Thousand Oaks, California: SAGE Publications, Inc. xvii, 298 pages.
35. L. Streiner, D., G. R. Norman, and J. Cairney, *Health measurement scales*. 6th edition. ed. 2023, New York: Oxford University Press. pages cm.
36. Cronbach, L.J. and P.E. Meehl, *Construct validity in psychological tests.* Psychol Bull, 1955. **52**(4): p. 281-302.

37. Mock, C., et al., *Global Surgery: Effective Involvement of US Academic Surgery: Report of the American Surgical Association Working Group on Global Surgery.* Ann Surg, 2018. **268**(4): p. 557-563.
38. Binda, C.J., et al., *Defining a Framework and Evaluation Metrics for Sustainable Global Surgical Partnerships: A Modified Delphi Study.* Ann Surg, 2024. **279**(3): p. 549-553.
39. Naluyimbazi, R., S. Nimanya, and P. Kisa, *Anatomy and lessons of partnerships in global pediatric surgery.* Semin Pediatr Surg, 2023. **32**(6): p. 151353.
40. Debas, H., et al., *Academic Partnerships in Global Surgery: An Overview American Surgical Association Working Group on Academic Global Surgery.* Ann Surg, 2020. **271**(3): p. 460-469.
41. Guilfoyle, R., A.D. Morzycki, and A. Saleh, *What makes global healthcare partnerships successful? A systematic review.* Glob Public Health, 2022. **17**(5): p. 662-671.
42. Naidoo, G., et al., *A roadmap for starting, growing and sustaining a comprehensive pediatric surgery service in a low resource area.* Semin Pediatr Surg, 2023. **32**(6): p. 151350.
43. Banhidy, F.P. and N.F. Banhidy, *The Role and Duty of Global Surgery in Increasing Sustainability and Improving Patient Care in Low and Middle-Income Countries.* Cureus, 2022. **14**(10): p. e30023.
44. Grant, C.L., et al., *Ethical considerations in global surgery: a scoping review.* BMJ Glob Health, 2020. **5**(4): p. e002319.
45. Cunningham, A.J., et al., *Ethics in Global Pediatric Surgery: Existing Dilemmas and Emerging Challenges.* World J Surg, 2019. **43**(6): p. 1466-1473.
46. Nuss, S., et al., *Adaptation and validation of the Children's Surgical Assessment Tool for Rwandan district hospitals.* Glob Health Action, 2024. **17**(1): p. 2297870.
47. Lin, Y., et al., *Identifying essential components of surgical care delivery through quality improvement: An updated surgical assessment tool.* Int J Surg, 2020. **82**: p. 103-107.
48. Task Force for Children's Surgical, C., *Optimal resources for children's surgical care in the United States.* J Am Coll Surg, 2014. **218**(3): p. 479-87, 487 e1-4.
49. Yousef, Y., et al., *Delayed access to care and unmet burden of pediatric surgical disease in resource-constrained African countries.* J Pediatr Surg, 2019. **54**(4): p. 845-853.
50. Wimmer, S., et al., *Assessing the inclusion of children's surgical care in National Surgical, Obstetric and Anaesthesia Plans: a policy content analysis.* BMJ Open, 2023. **13**(4): p. e051248.
51. Vigliotti, V.S., et al., *Modeling the Scale-up of Surgical Services for Children with Surgically Treatable Congenital Conditions in Somaliland.* World J Surg, 2022. **46**(10): p. 2489-2497.
52. Ullrich, S.J., et al., *Pediatric surgical quality improvement in low- and middle-income countries: What data to collect?* Surgery, 2022. **171**(4): p. 1067-1072.
53. Truche, P., et al., *Defining Surgical Workforce Density Targets to Meet Child and Neonatal Mortality Rate Targets in the Age of the Sustainable Development Goals: A Global Cross-Sectional Study.* World J Surg, 2022. **46**(9): p. 2262-2269.
54. Stewart, B.T., et al., *Exploring the Relationship Between Surgical Capacity and Output in Ghana: Current Capacity Assessments May Not Tell the Whole Story.* World J Surg, 2018. **42**(10): p. 3065-3074.

55. Gajewski, J., et al., *Critical shortage of capacity to deliver safe paediatric surgery in sub-Saharan Africa: evidence from 67 hospitals in Malawi, Zambia, and Tanzania.* Front Pediatr, 2023. **11**: p. 1189676.
56. Yousef, Y., et al., *GAPS Phase III: incorporation of capacity based weighting in the global assessment for pediatric surgery.* Pediatr Surg Int, 2024. **40**(1): p. 292.
57. Y, Y., et al., *Development and Pilot Results of the Global Assessment in Pediatric Surgery (GAPS): An Evidence-Based Pediatric Surgical Capacity Assessment Tool for Low-Resource Settings: Phase II.* World Journal of Surgery, 2023.
58. *World Bank Country and Lending Groups*, in *Data Help Desk*, T.W.B. Group, Editor. 2018.
59. *World Bank Country and Lending Groups*. World Bank Group.
60. *Human Development Index (HDI)*, in *Explore HDI*, H.D. Reports, Editor. 2023, United Nations Development Programme.
61. Butler, M.W., *Developing pediatric surgery in low- and middle-income countries: An evaluation of contemporary education and care delivery models.* Semin Pediatr Surg, 2016. **25**(1): p. 43-50.
62. Okoye, M.T., et al., *Can Economic Performance Predict Pediatric Surgical Capacity in Sub-Saharan Africa?* World J Surg, 2016. **40**(6): p. 1336-43.
63. PaedSurg Africa Research, C., *Paediatric surgical outcomes in sub-Saharan Africa: a multicentre, international, prospective cohort study.* BMJ Glob Health, 2021. **6**(9).
64. Philipo, G.S., et al., *Lessons from developing, implementing and sustaining a participatory partnership for children's surgical care in Tanzania.* BMJ Glob Health, 2020. **5**(3): p. e002118.
65. Abdul-Mumin, A., et al., *Burden of Neonatal Surgical Conditions in Northern Ghana.* World J Surg, 2020. **44**(1): p. 3-11.
66. Concepcion, T.L., et al., *Barriers to Surgical Care Among Children in Somaliland: An Application of the Three Delays Framework.* World J Surg, 2020. **44**(6): p. 1712-1718.
67. Ozgediz, D. and D. Poenaru, *The burden of pediatric surgical conditions in low and middle income countries: a call to action.* J Pediatr Surg, 2012. **47**(12): p. 2305-11.
68. Vissoci, J.R.N., et al., *Disparities in surgical care for children across Brazil: Use of geospatial analysis.* PLoS One, 2019. **14**(8): p. e0220959.
69. Data, C.a.I.M.-p.b.O.W.i. *Child and Infant Mortality*. Available from: https://ourworldindata.org/child-mortality?insight=most-children-die-from-preventable-causes#key-insights.
70. *Institute of Health Metrics and Evaluation*. GBD compare [cited 2024; Available from: https://vizhub.healthdata.org/gbd-compare/
71. Ullrich, S.J., et al., *Burden and Outcomes of Neonatal Surgery in Uganda: Results of a Five-Year Prospective Study.* J Surg Res, 2020. **246**: p. 93-99.
72. Kim, N.E., et al., *Cost-effectiveness of neonatal surgery for congenital anomalies in low-income and middle-income countries: a systematic review protocol.* BMJ Paediatr Open, 2020. **4**(1): p. e000755.
73. Banu, T., et al., *Surgically Correctable Congenital Anomalies: Reducing Morbidity and Mortality in the First 8000 Days of Life.* World J Surg, 2023. **47**(12): p. 3408-3418.
74. Gosselin, R., D. Ozgediz, and D. Poenaru, *A square peg in a round hole? Challenges with DALY-based "burden of disease" calculations in surgery and a call for alternative metrics.* World J Surg, 2013. **37**(11): p. 2507-11.

75. Poenaru, D., *The burden of pediatric surgical disease in low-resource settings: Discovering it, measuring it, and addressing it.* J Pediatr Surg, 2016. **51**(2): p. 216-20.
76. Osen, H., et al., *Validation of the World Health Organization tool for situational analysis to assess emergency and essential surgical care at district hospitals in Ghana.* World J Surg, 2011. **35**(3): p. 500-4.
77. Greenberg, S.L., et al., *An investment in knowledge: Research in global pediatric surgery for the 21st century.* Semin Pediatr Surg, 2016. **25**(1): p. 51-60.
78. Peltzer, K., *Injury and social determinants among in-school adolescents in six African countries.* Inj Prev, 2008. **14**(6): p. 381-8.
79. Saxton, A.T., et al., *Economic Analysis of Children's Surgical Care in Low- and Middle-Income Countries: A Systematic Review and Analysis.* PLoS One, 2016. **11**(10): p. e0165480.
80. Butler, E.K., et al., *Epidemiology of pediatric surgical needs in low-income countries.* PLoS One, 2017. **12**(3): p. e0170968.
81. Price, R., E. Makasa, and M. Hollands, *World Health Assembly Resolution WHA68.15: "Strengthening Emergency and Essential Surgical Care and Anesthesia as a Component of Universal Health Coverage"-Addressing the Public Health Gaps Arising from Lack of Safe, Affordable and Accessible Surgical and Anesthetic Services.* World J Surg, 2015. **39**(9): p. 2115-25.
82. Lippa, L., et al., *Clinical Capacity Building Through Partnerships: Boots on the Ground in Global Neurosurgery.* Neurosurgery, 2024. **95**(4): p. 728-739.
83. Lin, H., et al., *Clinical spine care partnerships between high-income countries and low-and-middle-income countries: A scoping review.* PLoS One, 2023. **18**(10): p. e0287355.
84. Nataraja, R.M., et al., *Overview of a novel paediatric surgical simulation-based medical education programme in Myanmar.* ANZ J Surg, 2020. **90**(10): p. 1925-1932.
85. Streiner, D.L., *Starting at the beginning: an introduction to coefficient alpha and internal consistency.* J Pers Assess, 2003. **80**(1): p. 99-103.
86. Riviello, R. and J.W. Scott, *Closing the data gaps for surgical care delivery in LMICs.* Lancet Glob Health, 2016. **4**(3): p. e138-9.
87. Ng-Kamstra, J.S., et al., *Use and definitions of perioperative mortality rates in low-income and middle-income countries: a systematic review.* Lancet, 2015. **385 Suppl 2**: p. S29.
88. Dimick, J.B., et al., *Risk adjustment for comparing hospital quality with surgery: how many variables are needed?* J Am Coll Surg, 2010. **210**(4): p. 503-8.
89. Knight, S.R., et al., *Systematic review of the use of big data to improve surgery in low- and middle-income countries.* Br J Surg, 2019. **106**(2): p. e62-e72.
90. Collaborators, G.B.D.C.M., *Global, regional, national, and selected subnational levels of stillbirths, neonatal, infant, and under-5 mortality, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015.* Lancet, 2016. **388**(10053): p. 1725-1774.
91. Anderson, G.A., et al., *Data Improvement Through Simplification: Implications for Low-Resource Settings.* World J Surg, 2018. **42**(9): p. 2725-2731.
92. Fayers, P.M. and D. Machin, *Quality of life : the assessment, analysis, and reporting of patient-reported outcomes*. Third edition. ed. 2016, Chichester, West Sussex, UK ; Hoboken, NJ: John Wiley & Sons Inc. p.
93. Hamad, D., et al., *Defining the critical pediatric surgical workforce density for improving surgical outcomes: a global study.* J Pediatr Surg, 2020. **55**(3): p. 493-512.

94. Ibrahim, G.M., D.W. Cadotte, and M. Bernstein, *A framework for the monitoring and evaluation of international surgical initiatives in low- and middle-income countries.* PLoS One, 2015. **10**(3): p. e0120368.
95. Vicente-Crespo, M., et al., *Institutionalizing research capacity strengthening in LMICs: A systematic review and meta-synthesis.* AAS Open Res, 2020. **3**: p. 43.
96. Wright, N.J. and C. Global PaedSurg Research, *Management and outcomes of gastrointestinal congenital anomalies in low, middle and high income countries: protocol for a multicentre, international, prospective cohort study.* BMJ Open, 2019. **9**(8): p. e030452.
97. Vanderburg, R., et al., *Age at Primary Cleft Lip Repair: A Potential Bellwether Indicator for Pediatric Surgery.* Plast Reconstr Surg Glob Open, 2021. **9**(6): p. e3657.
98. O'Neill, K.M., et al., *Bellwether Procedures for Monitoring and Planning Essential Surgical Care in Low- and Middle-Income Countries: Caesarean Delivery, Laparotomy, and Treatment of Open Fractures.* World J Surg, 2016. **40**(11): p. 2611-2619.
99. Poenaru, D. and J.O. Seyi-Olajide, *Developing Metrics to Define Progress in Children's Surgery.* World J Surg, 2019. **43**(6): p. 1456-1465.
100. Ford, K., et al., *Gastroschisis: Bellwether for neonatal surgery capacity in low resource settings?* J Pediatr Surg, 2016. **51**(8): p. 1262-7.
101. Qazi, S.H., et al., *Global perspective of paediatric surgery in low and middle income countries.* J Pak Med Assoc, 2019. **69(Suppl 1)**(1): p. S108-S111.
102. Global PaedSurg Research, C., *Mortality from gastrointestinal congenital anomalies at 264 hospitals in 74 low-income, middle-income, and high-income countries: a multicentre, international, prospective cohort study.* Lancet, 2021. **398**(10297): p. 325-339.
103. Dewan, M.C., et al., *Pediatric neurosurgical bellwether procedures for infrastructure capacity building in hospitals and healthcare systems worldwide.* Childs Nerv Syst, 2018. **34**(10): p. 1837-1846.
104. Alslaim, H.S., et al., *Tracheoesophageal fistula in the developing world: are we ready for thoracoscopic repair?* Pediatr Surg Int, 2020. **36**(5): p. 649-654.
105. Mock, C.N., et al., *Essential surgery: key messages from Disease Control Priorities, 3rd edition.* Lancet, 2015. **385**(9983): p. 2209-19.
106. Falk, R., et al., *Surgical Task-Sharing to Non-specialist Physicians in Low-Resource Settings Globally: A Systematic Review of the Literature.* World J Surg, 2020. **44**(5): p. 1368-1386.