










COMMENTARY

Creating a data commons: The International Soft Tissue SaRcoma ConsorTium (INSTRuCT)

Kirk D. Wyatt¹  | Suzi Birz² | Douglas S. Hawkins³  | Veronique Minard-Colin⁴  |
David A. Rodeberg⁵ | Monika Sparber-Sauer^{6,7}   | Gianni Bisogno⁸  |
Ewa Koscielniak^{6,7}  | Gian Luca De Salvo⁹  | Martin Ebinger¹⁰  |
Johannes H. M. Merks¹¹ | Suzanne L. Wolden¹²  | Wei Xue¹³ |
Samuel L. Volchenbom² | for the INternational Soft Tissue SaRcoma ConsorTium
(INSTRuCT)¹

¹Division of Pediatric Hematology/Oncology, Roger Maris Cancer Center, Sanford Health, Fargo, North Dakota, USA

²Department of Pediatrics, University of Chicago, Chicago, Illinois, USA

³Division of Hematology/Oncology, Department of Pediatrics, Seattle Children's Hospital, University of Washington School of Medicine, Seattle, Washington, USA

⁴Department of Paediatric and Adolescent Oncology, Gustave-Roussy, Villejuif, France

⁵Division of Pediatric Surgery, University of Kentucky, Lexington, Kentucky, USA

⁶Klinikum der Landeshauptstadt Stuttgart, Olgahospital, Zentrum für Kinder-, Jugend -und Frauenmedizin, Pädiatrie 5 (Pädiatrische Onkologie, Hämatologie, Immunologie) Stuttgart Cancer Center, Stuttgart, Germany

⁷Medical Faculty, University of Tübingen, Tübingen, Germany

⁸Hematology Oncology Division, Department of Women's and Children's Health, University Hospital of Padova, Padua, Italy

⁹Clinical Research Unit, Istituto Oncologico Veneto IOV IRCCS, Padua, Italy

¹⁰Department Pediatric Hematology/Oncology, Children's University Hospital, Eberhard Karls University of Tuebingen, Tübingen, Germany

¹¹Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands

¹²Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, New York, USA

¹³Department of Biostatistics, Children's Oncology Group Statistics and Data Center, University of Florida, Gainesville, Florida, USA

Correspondence

Samuel L. Volchenbom, 900 E. 57th St,
Chicago, IL 60637, USA.
Email: slv@uchicago.edu

Abstract

In this article, we will discuss the genesis, evolution, and progress of the International Soft Tissue SaRcoma ConsorTium (INSTRuCT), which aims to foster international research and collaboration focused on pediatric soft tissue sarcoma. We will begin by highlighting the current state of clinical research for pediatric soft tissue sarcomas, including rhabdomyosarcoma and non-rhabdomyosarcoma soft tissue sarcoma. We will then explore challenges and research priorities, describe the development of INSTRuCT, and discuss how the consortium aims to address key research priorities.

Abbreviations: DUA, data use agreement; INRG, International Neuroblastoma Risk Group; INSTRuCT, International Soft Tissue SaRcoma ConsorTium; MOU, memorandum of understanding; NRSTS, non-rhabdomyosarcoma soft tissue sarcoma; OS, overall survival; PCDC, Pediatric Cancer Data Commons; RMS, rhabdomyosarcoma; STS, soft tissue sarcoma.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Pediatric Blood & Cancer* published by Wiley Periodicals LLC.

KEYWORDS

data aggregation, pediatrics, sarcoma

1 | BACKGROUND

1.1 | Soft tissue sarcoma

Soft tissue sarcoma (STS) comprises 7% of pediatric malignancies, with 40% of those representing rhabdomyosarcoma (RMS) and the remainder being classified as non-RMS STS (NRSTS). While 5-year overall survival (OS) for embryonal RMS is 66%–68%, outcomes for patients with alveolar RMS are less favorable, with a 5-year OS of 39%–50%.^{1,2}

Compared to RMS, the study of NRSTS presents unique challenges. While more than half of pediatric STS are classified as NRSTS, this diverse group includes over 75 biologically unique sarcomas.³ Due to the rarity of each subtype of NRSTS, it has historically been studied out of necessity in “basket” trials where the varied diagnoses are treated uniformly, without an ability to take into account the diversity and unique biological behavior of each of these subtypes.⁴ Recent discoveries about the molecular underpinnings of STS subtypes have uncovered the molecular diversity within NRSTS, highlighting the biological distinctness of each of these tumor types.

Research has traditionally been conducted by cooperative groups, and discoveries have been limited by the paucity of patients with rare disease subtypes. Due to the rarity of these subtypes, international collaborative relationships have developed to pool data on rare subtypes.^{5–8} Following this rich tradition of international collaboration, there was a desire to create a more comprehensive and reusable dataset and formally organized consortium for the study of pediatric STS. Pediatric STS experts engaged the Pediatric Cancer Data Commons (PCDC) team to assist in forming a consortium—and ultimately a data commons⁹—and developed the INternational Soft Tissue SaRcoma ConsorTium (INSTRuCT; Figure 1).

2 | INTRODUCING INSTRuCT

2.1 | Beginnings

Under the auspices of the PCDC, INSTRuCT was established in 2017 to foster international research and collaboration focused on pediatric STS. INSTRuCT’s membership is derived from North American and European cooperative groups (Children’s Oncology Group [COG], European paediatric Soft tissue sarcoma Study Group [EpSSG], and Cooperative Weichteilsarkom Studiengruppe [CWS]) and includes data from prior studies sponsored by the International Society of Paediatric Oncology (SIOP) Malignant Mesenchymal Tumour Committee and Italian Association of Pediatric Hematology and Oncology (AIEOP) Soft Tissue Sarcoma Committee.

The foundational meetings took place as add-ons to established international research conferences. Six in-person meetings took place between October 2017 and October 2019 before transitioning to virtual meetings due to the COVID-19 pandemic. Each meeting was structured to include a plenary as well as breakout sessions focused on disease type (e.g., RMS vs. NRSTS) or clinical discipline. Initial clinical discipline work groups included pathology/biology and surgery, and later expanded to include radiation oncology and imaging. The early meetings established the structure of INSTRuCT and set overarching priorities, whereas subsequent meetings led to more granular discussions that established discipline-specific research questions and led to the development and approval of a standardized RMS data dictionary.

2.2 | Establishing priorities

One founding charge was collection and aggregation of high-quality international clinical trial datasets into a data commons, which would increase access to cohorts of rare disease subtypes. Developing an international consensus for risk stratification in RMS was identified as the initial use for the data commons. Such an undertaking had recently been conducted by the International Neuroblastoma Risk Group (INRG), which similarly convened an international consortium of pediatric cancer researchers and combined international datasets in an effort to revise and standardize risk stratification for children with neuroblastoma.¹⁰ A member of the INRG Task Force was present at the founding INSTRuCT meeting to share background and insights learned in the process. Another priority was to focus on existing data from completed studies, rather than attempting to add new data to existing datasets.

Key clinical questions relating to RMS risk stratification included re-examination of primary tumor site designation as “favorable” or “unfavorable.” During the initial meetings, clinical trials from which data could be contributed were identified, and data elements to be contributed were selected so that work on a standardized data dictionary could begin. Priorities that emerged during subsequent meetings

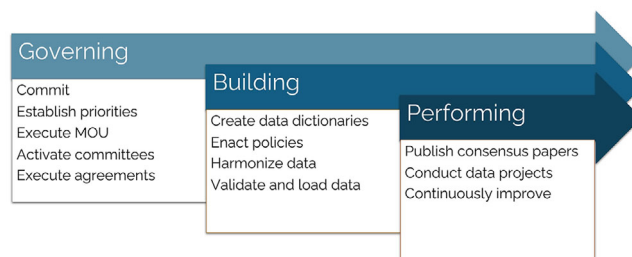


FIGURE 1 Summary of formation of INternational Soft Tissue SaRcoma ConsorTium (INSTRuCT)

included identification and assessment of genetic markers, developing methods to link clinical data to genomic data, and improving outcomes in patients with relapsed and metastatic disease. A final aim was to develop international consensus on management of RMS and NRSTS, both to inform the conduct of international clinical trials but also to inform clinicians of “best practices.”

2.3 | Clinical consensus building

One early task was to identify and standardize data elements to be included in a data dictionary to facilitate harmonization of datasets. The harmonization process highlighted similarities and differences between clinical trials and cooperative groups, including approaches to diagnosis and treatment. Data elements that were collected on case report forms were inventoried and analyzed, including how they were defined and collected. It quickly became apparent that consensus among international groups could facilitate retrospective harmonization of datasets from past clinical trials and—more importantly—could facilitate advances in clinical care, the conduct of future collaborative trials, and combination of new datasets. This prompted the development of four “consensus opinion” statements from the collaborative groups as the first publications to emerge from the consortium.^{11–14} These consensus opinion papers incorporated review of the literature and clinical trials treatment protocols along with expert opinion to provide specific consensus recommendations to guide clinical care and the conduct of future clinical trials.

3 | GOVERNANCE

3.1 | Governance structure

One early task was to establish a governance structure and define processes to be followed by the consortium. In particular, the specific roles and responsibilities of individuals and groups needed to be defined, and the relationships between the consortium, data contributors, and service providers needed to be detailed. Finally, policies were needed to outline the process for ingesting, requesting, and accessing data, in addition to dissemination of data.

3.2 | Memorandum of understanding

One foundational document is the memorandum of understanding (MOU). Parties to the MOU include cooperative groups and institutions with data to contribute. In the case of INSTRuCT, using the INRG as a model, the cooperative groups formed the consortium on a foundation of trust, built on a long history of collaboration. The cooperative groups signed an MOU and created an Executive Committee.

3.3 | Executive Committee

Responsibilities of the Executive Committee described in the MOU included strategic planning; coordinating, appointing, and changing the data commons manager; amending the MOU as necessary; approving and managing membership; reviewing and approving requests to access data within the data commons; reviewing and approving contributions of data to the data commons; and approving of grant or funding applications submitted on behalf of, or which rely upon, the consortium.

The INSTRuCT Executive Committee is composed of a representative designated by each cooperative group to serve as co-chairs, two other representatives of each cooperative group (to be selected to ensure a balanced and complete representation of clinical disciplines and geography), two statisticians appointed by mutual agreement of the cooperative groups, and one representative of the PCDC, which serves as the data commons service provider. A consortium manager—with both program management and science expertise—plans, organizes, and facilitates meetings; handles project requests; and acts as a liaison between the consortium and the PCDC. The consortium manager also coordinates individual work groups and tracks action items to ensure goals of the consortium are met and coordination across the consortium.

The cooperative groups determined that the members of the Executive Committee will make decisions and handle disputes by mutual, unanimous agreement and consensus. These responsibilities and the decision-making framework bolstered trust among cooperative groups willing to bring their data into the PCDC, as they provided reassurance that the contributed data will be used responsibly and only for projects approved by the cooperative groups through their representation on the Executive Committee, which reviews project requests.

3.4 | Work groups

The clinical discipline work groups are composed of members from each of the cooperative groups, with one member agreeing to serve as lead. To ensure communication with INSTRuCT leadership, Executive Committee members agreed to serve as a liaison to each work group. The Executive Committee drafted a charge to each work group. The Executive Committee opted to stagger the launch of the groups. NRSTS, pathology/biology, and surgery all convened for the first time in March 2018, followed by imaging in March 2020, and radiation oncology in October 2020.

In March 2021, the Executive Committee completed a review of the work group membership and progress. With 3 years of experience as guidance, the guidelines for work group composition were confirmed and term limits for the work group leads were adopted. For membership, the process used in 2018, whereby work group members are selected by the cooperative groups, was confirmed. As with the original arrangement, work group leads are selected by the work group members, and it was determined that the leads will be asked to serve 3 years, with the opportunity to extend another 3 years if they are willing and if the work group so decides, or the work group can name a new lead.

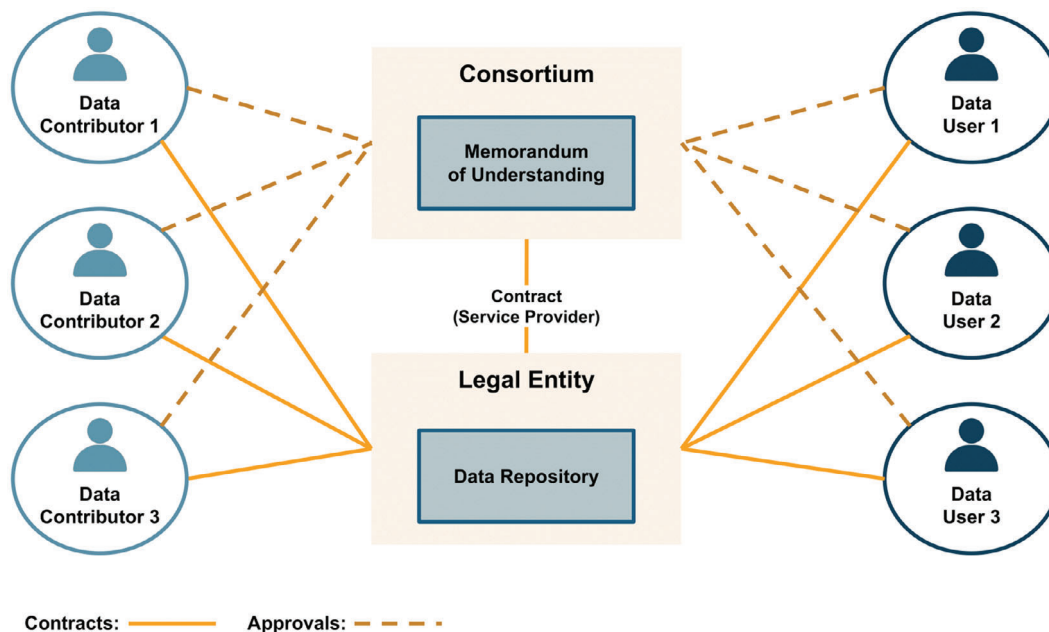


FIGURE 2 Relationship between data contributors, consortia, service provider, and data users. Reproduced with author permission (original figure published under Creative Commons Attribution Non-Commercial No Derivatives 4.0 License)⁹

The statistical work group consists of statisticians, data managers, clinicians, and PCDC data and technical team representatives. The first charge was the creation of the harmonized RMS data dictionary. Additional responsibilities included ensuring the quality and ongoing growth of the data dictionary by working with the PCDC team to create quality control checks and identify data elements to be added to the data dictionary. Critical to the success of the commons is the partnership between the PCDC team and the cooperative group members with detailed understanding of the disease, the data, and the research questions that will be supported by the commons. This model has been adopted and improved in subsequent disease-oriented groups that have engaged with the PCDC, with the key elements of the multidisciplinary approach to this work remaining as a central principle.

3.5 | Data contributor agreement

To meet regulatory requirements, the PCDC (through the University of Chicago) entered into data contributor agreements (DCAs) with each cooperative group or institution that intended to transfer data to the data commons. Of note, the University of Chicago, as the data commons service provider, enters into this agreement directly with the data contributor. INSTRuCT is not a legal entity and is neither directly contributing data nor directly receiving the data and is not a party to the agreement (Figure 2). The DCA provides the mechanism for the data contributor to outline what data are being transferred to the PCDC and any restrictions. The DCA also assures the contributor that the PCDC will release the data only to authorized users.

3.6 | Data access and project request process

Once data have been combined into a data commons, policies and procedures are necessary to govern access to the data. Early efforts led to a cohort explorer tool that is freely available to any user who registers to use it. The cohort explorer tool was designed as a means to identify whether suitable cohorts of patient data were available to proceed with specific research projects to answer clinical questions.

Two mechanisms by which a request for access to line-level data can be initiated were devised: project requests can be initiated from within an existing INSTRuCT work group or by an independent investigator not as part of a work group.

3.6.1 | INSTRuCT work group-initiated project requests

When forming the INSTRuCT clinical discipline work groups, one of the charges to each work group was to provide “the forum for international discussion and consensus building, with the goal of publishing these expert opinions for broad dissemination and use by pediatric oncologists, surgeons, radiation oncologists, radiologists, and pathologists worldwide.”¹⁵ Another charge was to identify projects in which the cooperative groups could jointly engage. Because work groups are composed of STS clinical experts and researchers from across the globe, work groups are an ideal avenue for hypothesis generation, refinement, and testing.

For project requests that arise from a work group, a process was devised to ensure appropriate representation from the cooperative

groups, in keeping with the aim of INSTRuCT to be balanced across the contributors and the cooperative groups to retain their internal processes for naming investigators to projects. Project ideas are regularly reported to the INSTRuCT Executive Committee and all INSTRuCT participants. The representatives to the work groups consult with their respective cooperative group leads as the teams are formed.

3.6.2 | Investigator-initiated project requests

For projects that arise outside of INSTRuCT work groups, a process was needed to identify overlapping projects, ensure project feasibility, support diverse representation on the study team, provide guidance, regulate access to the data, and ensure the data are handled responsibly. The INSTRuCT Executive Committee developed a project request form (Supporting Material), which any investigator worldwide can complete and submit for review. This request is circulated to all Executive Committee members. Following a discussion at an Executive Committee meeting, the consensus decision is provided to the requesting investigator. The decision will include comments, requested or suggested changes to the project, and suggestions of consortium members with discipline-specific expertise who can collaborate on the project to ensure appropriate representation from each cooperative group contributing data for the project. It was felt that incorporation of INSTRuCT members with an intimate understanding of the clinical topic and research studies will greatly enhance the research conducted with the data. A statistician with experience analyzing the dataset may also be nominated to facilitate data analysis.

3.7 | Publication policy

The INSTRuCT Executive Committee quickly recognized the importance of a publication policy. After reviewing the INRG publication policy, the committee had a robust discussion to identify the key elements of the policy, inclusive of project team involvement, authorship, acknowledgements, and review process. Goals for authorship policies focused on ensuring appropriate representation from each contributing group and specialty (including statistical specialists), in addition to assignment of authorship and co-authorship. INSTRuCT has adopted the International Committee of Medical Journal Editors (ICMJE) authorship criteria for determining co-authors, as it is the de facto standard used within the medical literature.¹⁶ To support the training goals of INSTRuCT, full professors are asked to consider appointing an early-career investigator as the senior author, or co-senior authorship with a young investigator.

As with the project request process, the publication policy addresses projects undertaken by the INSTRuCT work groups (and therefore sponsored by the cooperative groups) and projects initiated by investigators (who may be members of one or more of the cooperative groups). The processes are the same for work group- and investigator-initiated projects, with the exception of author selection. For projects proposed by investigators, the INSTRuCT Executive

Committee takes an active role in ensuring a complete and balanced project team that can address the proposed project hypothesis and use of the cooperative group data in INSTRuCT.

3.8 | Data use agreement

To meet regulatory requirements, the PCDC (through the University of Chicago) enters into data use agreements (DUAs) with each investigator who will receive data pursuant to an approved project (Figure 2). As is the case with DCAs, the University of Chicago, as the data commons service provider, enters into this agreement directly with the data user and their institution. INSTRuCT is not a legal entity and is not directly providing data and therefore it is not a party to DUAs. The DUA provides the mechanism for the PCDC to detail the terms of use of the data being transferred from the PCDC. These agreements include any specific requirements from the DCA under which the data being provided were contributed.

4 | REMOVING PATIENT IDENTIFIERS

To achieve compliance with patient privacy regulations, patient identifiers are removed from the dataset in accordance with the Health Insurance Portability and Accountability Act (HIPAA) "Safe Harbor" method.¹⁷ All steps are performed by data contributors, and identifiers are never transmitted to the PCDC. All dates in source data are replaced by the patient's age (in days) at the time of the observation. The year of diagnosis is retained to facilitate contextualization of the standard of care and treatments used at the time. Patients are assigned pseudonymized identification numbers by an honest broker, but linkage between this honest broker number and patient identifiers is never shared with the PCDC. The inclusion of the honest broker identifiers for research participants from the European Union meant that the data were considered "pseudonymized" according to the General Data Protection Regulation (GDPR). In these instances, the data contributor agreements include the required GDPR clauses for pseudonymized data. To adhere to those commitments, for approved projects that will include data protected by the GDPR clauses in the data contributor agreement, the DUA includes the required GDPR clauses.

5 | CREATING A HARMONIZED DATA DICTIONARY

With the initial structure of INSTRuCT in place, efforts shifted to developing standardized data dictionaries. Two data dictionaries were generated: one for RMS and another for NRSTS. The data dictionaries served to define all possible data elements that could be contributed to the common dataset, including permissible values, in a common language. Common data elements, such as patient age at diagnosis and primary tumor site, were represented identically in the two data dictionaries so that data across diseases were harmonized whenever

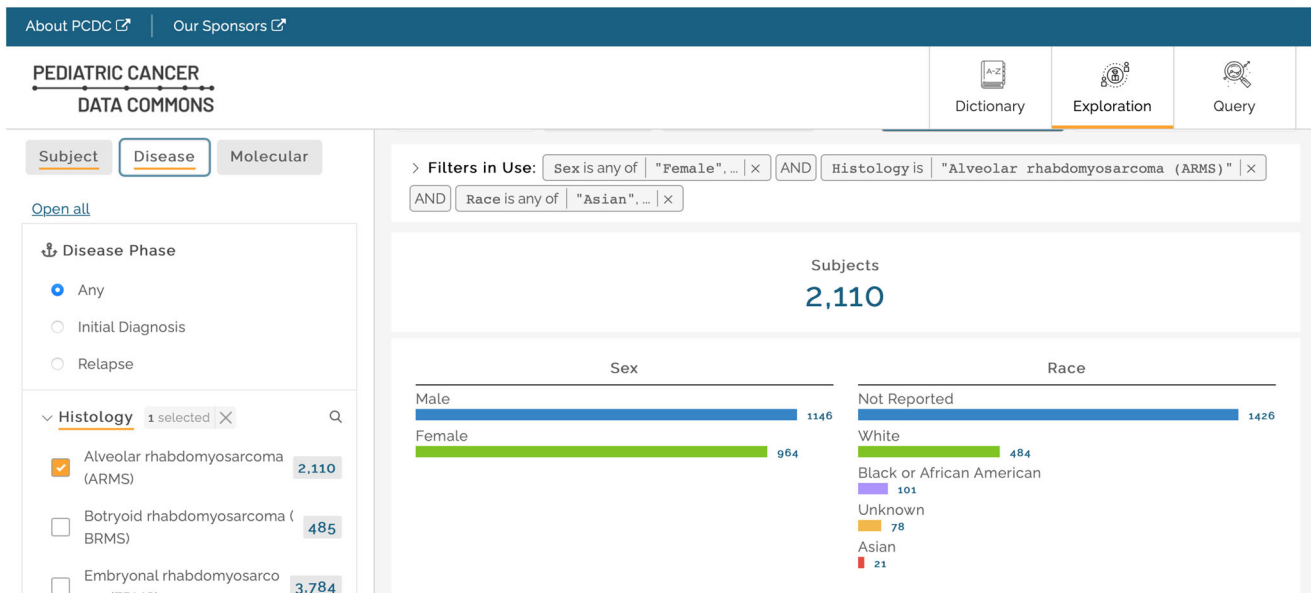


FIGURE 3 New Pediatric Cancer Data Commons (PCDC) cohort explorer

possible. Throughout the process, data elements within the data dictionary were mapped to the NCI thesaurus as a means to adopt a standardized terminology.¹⁸

6 | DEVELOPING A COHORT EXPLORER

To facilitate assessment of available data and cohorts and assess feasibility of proposed studies, a cohort explorer was developed. A similar tool had already been developed for the INRG data commons, and the tool was ported for use with the INSTRuCT data commons. Various levels of access control, on a spectrum from full access without registration to access restricted only to approved users, were discussed. It was decided that the cohort discovery tool would be freely available to all users who register to use it, while line-level data access would be restricted to those with approved project requests. Recently, the PCDC has launched a new front-end data exploration platform that facilitates cohort exploration (Figure 3) and will pilot additional visualization tools.

7 | EARLY CHALLENGES

As a consortium committed to making decisions by mutual, unanimous agreement and consensus, many topics required significant discussion before arriving at a consensus. Among early topics, access to cohort discovery in the PCDC data portal and access to submit a project proposal garnered the most discussion. Discussions focused on whether access to pooled data should be limited to only individuals of the cooperative groups that contributed data or to all. The decision—which is now the PCDC standard—is to allow anyone to create an account and perform cohort discovery. Anyone seeking line-level access to data may

submit a project proposal, which the Executive Committee will review and then determine whether access to line-level data will be granted.

Harmonizing data elements also generated detailed discussions. With the general principle to not sacrifice any data through binning or grouping, the groups set out to bring together clinical trials data elements with different value sets—for instance, ensuring details such as tumor site “cheek” collected in one trial are not lost when the value of “face” was collected in other trials. These were laborious deliberations requiring input from clinicians and data standard experts.

With data contributors located across the globe, ensuring full compliance with local and regional regulatory requirements posed a significant challenge. In some cases, additional regulatory “gates” needed to be lifted through modifications of standard data contributor agreements to meet requirements. These requirements included creating a study protocol for an ethics committee, entering into a multilingual study agreement, and providing supplementary privacy documentation. The planned INSTRuCT primary analyses and the first research projects were delayed while awaiting execution of all agreements to receive the data.

Finally, the topic that continues to generate the most discussion is that of authorship. A publication policy addressing project review, project team membership, and acknowledgements was adopted early on after much discussion, and the policy continues to be re-examined and updated, including requirements for acknowledging members of cooperative groups that provided the data included in the research.

8 | PROGRESS

As we have outlined, remarkable progress has been made in the past 4 years. INSTRuCT collaborated closely with the PCDC to establish a uniform data dictionary, which harmonized the data fields used across

cooperative groups and clinical trials (<https://commons.cri.uchicago.edu/instruct/>). With a common clinical language established, five data contributors have pooled over 4600 patient cases into the data commons. Up until the time of writing this report, nine manuscripts have been published^{11–15,19–22} and nine projects are in progress.

9 | GROWING WITH THE PCDC

Along with INRG, INSTRuCT was an early member of the PCDC. As the data service provider for disease-oriented consortia, PCDC seeks to create a singular data commons, which serves as a repository for clinical pediatric cancer data. PCDC also seeks, where available, to link clinical data to existing datasets in other repositories, such as genomic data. As with its role in the development of INSTRuCT, the PCDC facilitates the development of new disease-focused consortia by assisting with establishment of governance processes and entering into MOUs and DCAs. To date, the PCDC has engaged with 10 disease-based groups. Eight of these have signed memoranda of understanding and seven have established a data dictionary.

10 | FUTURE DIRECTIONS

We are confident that the establishment of INSTRuCT and contribution of clinical pediatric STS data to the PCDC will transform pediatric cancer research. With more open access to pediatric cancer data, we expect the number of published clinical research studies on pediatric STS to increase, and we anticipate an increase in the number of studies linking data from multiple sources, including genomic data, to further our understanding of pediatric STS.

In addition to providing a powerful tool to disease consortia and established researchers, we hope the PCDC will engage young investigators. While the enriched dataset will allow for analyses that were previously performed at the trial level to be performed across a larger set of patients, it will also allow us to ask new questions that we were unable to ask at the trial level. The existence of disease-based data for multiple pediatric cancers—all formatted according to a standard data dictionary—will facilitate cross-discipline research projects, such as studies on late effects of treatments that are common across diseases.

Although the PCDC aims to capture a rich and complete dataset from clinical trials data, many children with cancer are not enrolled on a clinical trial.²³ A European study²⁴ reported that approximately half of children and adolescents with STS enrolled on a clinical study, though enrollment rates climbed to 74% when the subset of patients with RMS was considered. Reasons for nonparticipation include being seen at a site that does not offer clinical trials participation, parental refusal to participate, or no clinical trial being open at the time of diagnosis.

Moreover, among patients who are enrolled in clinical trials, the corpus of data on those patients is limited to fields included on case report forms designed to support the study aims. The data fields must necessarily be constrained due to the time and effort required to collect and validate each data element. As we look for new ways to

expand the scope of data shared within the data commons, methods to link to external data repositories and to include data on patients not included within clinical trials are necessary. The PCDC has already demonstrated the ability to link to external genomic data by leveraging application programming interfaces (APIs) and data standards. In the future, the widespread use and support of data standards, such as Health Level 7 (HL7) Fast Health Interoperability Resources (FHIR) may allow us to add additional data elements from local medical records and to include patients who did not participate in clinical trials in bulk using automated processes.^{25,26} Challenges include the need for methods to disambiguate patients and validate bulk data, as well as legal, regulatory, and privacy concerns. The ability to collect real-world data may be particularly critical for advancing our understanding of rare cancers, including the many subtypes of NRSTS, for which the conduct of subtype-specific clinical trials is challenging or impracticable. Only by pooling real-world treatment and outcome data will sufficient statistical power be achieved to assess efficacy of treatments for many of these rare pediatric cancers.

11 | CONCLUSION

Over the past 4 years, INSTRuCT has brought together an international community with diverse representation to advance progress on pediatric STS. As never before, any researcher in the world can now access clinical pediatric cancer data on thousands of patients to conduct novel research studies. While this accomplishment is momentous, perhaps the largest and most lasting impact will be unification of an international community toward a common goal while fostering international collaboration and maintenance of clinical data standards to ensure the sustainability of data commons.

ACKNOWLEDGMENTS

INSTRuCT and the PCDC are supported in part by Cancer Research Foundation, Children's Research Foundation, Comer Development Board, KickCancer, King Baudouin Foundation, Rally Foundation for Childhood Cancer Research, Seattle Children's Foundation from Kat's Crew Guild through the Sarcoma Research Fund, St Baldrick's Foundation, and The Andrew McDonough B+ Foundation. This work is made possible through the efforts of COG, CWS, EpSSG, SIOP Malignant Mesenchymal Tumour Committee, and AIEOP Italian Soft Tissue Sarcoma Committee.

CONFLICT OF INTEREST

Monika Sparber-Sauer reports participation in advisory boards related to hemophilia treatments with Sobi and Roche and will receive honorarium from Bayer for work on infantile fibrosarcoma. The remaining authors declare no conflict of interest.

ORCID

Kirk D. Wyatt  <https://orcid.org/0000-0002-3856-2055>

Douglas S. Hawkins  <https://orcid.org/0000-0003-3602-1375>

Veronique Minard-Colin  <https://orcid.org/0000-0002-0296-5207>

Monika Sparber-Sauer  <https://orcid.org/0000-0001-9551-2399>
 Gianni Bisogno  <https://orcid.org/0000-0003-4462-5523>
 Ewa Koscielniak  <https://orcid.org/0000-0002-1519-7569>
 Gian Luca De Salvo  <https://orcid.org/0000-0001-7451-8476>
 Martin Ebinger  <https://orcid.org/0000-0002-4229-8058>
 Suzanne L. Wolden  <https://orcid.org/0000-0002-5234-4134>

REFERENCES

- Amer KM, Thomson JE, Congiusta D, et al. Epidemiology, incidence, and survival of rhabdomyosarcoma subtypes: SEER and ICES database analysis. *J Orthop Res*. 2019;37(10):2226-2230.
- Ries LAG, Smith MA, Gurney JG, et al. *Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975–1995*. NIH Pub. No. 99-4649. National Cancer Institute; 1999.
- Fletcher CDM. *WHO Classification of Tumours of Soft Tissue and Bone*. 4th ed. IARC Press; 2013.
- Milgrom SA, Million L, Mandeville H, Safwat A, Ermoian RP, Terezakis S. Non-rhabdomyosarcoma soft-tissue sarcoma. *Pediatr Blood Cancer*. 2021;68(2):e28279.
- Merks JH, De Salvo GL, Bergeron C, et al. Parameningeal rhabdomyosarcoma in pediatric age: results of a pooled analysis from North American and European cooperative groups. *Ann Oncol*. 2014;25(1):231-236.
- Oberlin O, Rey A, Lyden E, et al. Prognostic factors in metastatic rhabdomyosarcomas: results of a pooled analysis from United States and European cooperative groups. *J Clin Oncol*. 2008;26(14):2384-2389.
- Raney B, Anderson J, Jenney M, et al. Late effects in 164 patients with rhabdomyosarcoma of the bladder/prostate region: a report from the international workshop. *J Urol*. 2006;176(5):2190-2195.
- Oberlin O, Rey A, Anderson J, et al. Treatment of orbital rhabdomyosarcoma: survival and late effects of treatment – results of an international workshop. *J Clin Oncol*. 2001;19(1):197-204.
- Plana A, Furner B, Palese M, et al. Pediatric Cancer Data Commons: federating and democratizing data for childhood cancer research. *JCO Clin Cancer Inform*. 2021;5:1034-1043.
- Cohn SL, Pearson AD, London WB, et al. The international neuroblastoma risk group (INRG) classification system: an INRG task force report. *J Clin Oncol*. 2009;27(2):289-297.
- Rogers TN, Seitz G, Fuchs J, et al. Surgical management of paratesticular rhabdomyosarcoma: a consensus opinion from the Children's Oncology Group, European paediatric Soft tissue sarcoma Study Group, and the Cooperative Weichteilsarkom Studiengruppe. *Pediatr Blood Cancer*. 2021;68(4):e28938.
- Rudzinski ER, Kelsey A, Vokuhl C, et al. Pathology of childhood rhabdomyosarcoma: a consensus opinion document from the Children's Oncology Group, European Paediatric Soft Tissue Sarcoma Study Group, and the Cooperative Weichteilsarkom Studiengruppe. *Pediatr Blood Cancer*. 2021;68(3):e28798.
- Morris CD, Tunn PU, Rodeberg DA, et al. Surgical management of extremity rhabdomyosarcoma: a consensus opinion from the Children's Oncology Group, the European Pediatric Soft-Tissue Sarcoma Study Group, and the Cooperative Weichteilsarkom Studiengruppe. *Pediatr Blood Cancer*. 2020:e28608. <https://doi.org/10.1002/pbc.28608>
- Lautz TB, Martelli H, Fuchs J, et al. Local treatment of rhabdomyosarcoma of the female genital tract: expert consensus from the Children's Oncology Group, the European Soft-Tissue Sarcoma Group, and the Cooperative Weichteilsarkom Studiengruppe. *Pediatr Blood Cancer*. 2020:e28601. <https://doi.org/10.1002/pbc.28601>
- Hawkins DS, Bisogno G, Koscielniak E. Introducing INSTRuCT: an international effort to promote cooperation and data sharing. *Pediatr Blood Cancer*. 2020:e28701. <https://doi.org/10.1002/pbc.28701>
- Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. ICMJE; 2021. Accessed February 2, 2022. <http://www.icmje.org/icmje-recommendations.pdf>
- US Department of Health & Human Services. *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*. US Department of Health & Human Services; Accessed June 3, 2022. <https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html#safeharborguidance>
- de Coronado S, Haber MW, Sioutos N, Tuttle MS, Wright LW. NCI thesaurus: using science-based terminology to integrate cancer research results. *Stud Health Technol Inform*. 2004;107(1):33-37.
- Ferrari A, Orbach D, Sparber-Sauer M, et al. The treatment approach to pediatric non-rhabdomyosarcoma soft tissue sarcomas: a critical review from the International Soft Tissue Sarcoma Consortium. *Eur J Cancer*. 2022;169:10-19.
- Casey DL, Mandeville H, Bradley JA, et al. Local control of parameningeal rhabdomyosarcoma: an expert consensus guideline from the International Soft Tissue Sarcoma Consortium (INSTRuCT). *Pediatr Blood Cancer*. 2022;69(7):e29751.
- Ferrari A, Spunt SL, Sparber-Sauer M, et al. Controversies and challenges in the management of paediatric non-rhabdomyosarcoma soft tissue sarcomas. *Lancet Child Adolesc Health*. 2022;6(4):221-223.
- Hettmer S, Linardic CM, Kelsey A, et al. Molecular testing of rhabdomyosarcoma in clinical trials to improve risk stratification and outcome: a consensus view from European paediatric Soft tissue sarcoma Study Group, Children's Oncology Group and Cooperative Weichteilsarkom-Studiengruppe. *Eur J Cancer*. 2022;172:367-386.
- Parsons HM, Penn DC, Li Q, et al. Increased clinical trial enrollment among adolescent and young adult cancer patients between 2006 and 2012–2013 in the United States. *Pediatr Blood Cancer*. 2019;66(1):e27426.
- Ferrari A, Trama A, De Paoli A, et al. Access to clinical trials for adolescents with soft tissue sarcomas: enrollment in European pediatric Soft tissue sarcoma Study Group (EpSSG) protocols. *Pediatr Blood Cancer*. 2017;64(6):e26348.
- Jones J, Gottlieb D, Mandel JC, et al. A landscape survey of planned SMART/HL7 bulk FHIR data access API implementations and tools. *J Am Med Inform Assoc*. 2021;28(6):1284-1287.
- Mandl KD, Gottlieb D, Mandel JC, et al. Push button population health: the SMART/HL7 FHIR bulk data access application programming interface. *NPJ Digit Med*. 2020;3(1):151.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wyatt KD, Birz S, Hawkins DS, et al. Creating a data commons: The International Soft Tissue Sarcoma Consortium (INSTRuCT). *Pediatr Blood Cancer*. 2022;69:e29924. <https://doi.org/10.1002/pbc.29924>