

ARC Proposal – Phase Two

I. Mission Impact and Innovation:

The purpose of our proposal is to bring together a multi-disciplinary group of artificial intelligence (AI) researchers at CCHMC to build a self-sustaining image-focused AI core, establish streamlined multi-modal data harmonization pipelines, and carry out meaningful well-defined pilot projects (**Figure 1**). To accomplish these aims we will establish/consolidate infrastructure (e.g., computing resources, personnel) to support enterprise-wide image-based AI research and clinical translation as well as develop and clinically translate four meaningful AI tools via multi-modal pilot projects. We also will organize AI educational workshops for CCHMC clinical and research faculty/staff to accelerate AI adoption. At the end of three years, we will deliver the following: clinical implementation of multiple state-of-the-art AI algorithms that align with POPT pillars, an image-based AI core that will be self-sustaining and available to all of CCHMC for research and clinical purposes, and the necessary infrastructure to enable institutional AI researchers.

Medical images are helping fuel the advancement and adoption of AI in healthcare. This proposed investment in computer and data science will help establish CCHMC as a leader in pediatric AI innovation, discovery, and translation. CCHMC has one of the largest repositories of clinical digital pediatric imaging data (e.g., radiologic, cardiac, histologic, endoscopic, etc.) and clinical data (e.g., laboratory, “-omics” data) in the world. Yet, despite this wealth, we have been limited in our use of historical data for large-scale discovery and tend to remain in our individual silos, using only data with which we are familiar and that are easily accessible. Specifically, there is a lack of interoperability that inhibits integrating orthogonally important imaging and clinical data. This lack of data integration is a key institutional deficiency. Furthermore, we are resource-constrained in terms of computing resources for AI and the number of personnel available to carry out AI clinical and research projects. This proposal affords us the opportunity to invest in AI resources (including hardware and people), accelerate discovery and clinical translation, and improve patient care. Using AI methods, we believe the rate of discovery and clinical translation can be accelerated beyond the pace that is typical using conventional research methods.

By assembling a diverse team of passionate and expert clinicians and researchers, and upon hiring personnel with specialized skills for the AI core (**Appendix 1**), we will develop and implement image-based AI tools to tackle four unique unmet clinical needs (**Appendices 2-5**). These projects include: 1) creation of an automated bone age interpretation tool and modern-day bone age atlas; 2) automated LDA (lines, drains, and airways) detection and EMR reconciliation; 3) automated liver histologic fibrosis staging with multiple outputs (e.g., METAVIR, Ishak, % area collagen); and 4) automated CT/MRI organ segmentation (e.g., lungs, airway, heart, liver, spleen, fat and muscle). The creation and implementation of these algorithms will provide valuable knowledge and experience that will help future AI efforts, generate preliminary data showing our capabilities in the field, and establish pipelines that will enable innumerable future clinical and research AI projects.

II. Approach:

This ARC will fund four well-defined, multi-disciplinary image-based AI research projects and enable the creation of a self-sustaining AI core to support CCHMC researchers. This AI core along with the assembled ARC team will complete and clinically translate the four meaningful image-based AI projects over a three year period. All projects will also be aligned with the one or more pillars of POPT (Care, Cure, Community, and/or Culture) as well as with CCHMC’s Digital Transformation efforts. We will work closely with the Information Services and Digital Transformation leadership to ensure that the algorithms created can ultimately be implemented and positively impact patient care. The four projects detailed below will use CCHMC data for all stages of algorithm development, although outside datasets may be solicited to promote generalization and minimize bias, particularly if we want to deploy an algorithm beyond CCHMC or pursue commercialization. We expect the four projects below to be successful based on the expertise of the individuals involved, availability of data, knowledge

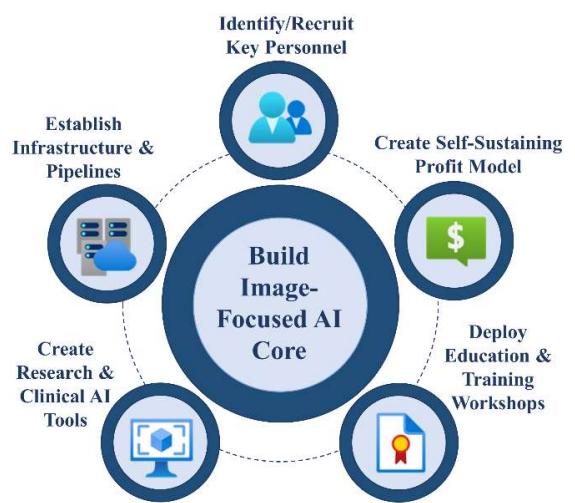


Figure 1.

of required methods, etc. These projects involve multiple types of images and disciplines, address unmet clinical and research needs, and will provide valuable preliminary data and serve as the foundation for future works.

Individual Projects

Project #1: Creation of a Modernized Bone Age Interpretation Tool using Artificial Intelligence (Appendix 2)

Radiographic bone age assessment is used in numerous settings to guide medical (e.g., endocrine) and surgical (e.g., orthopedic) treatment. The currently used reference standard for determining skeletal maturity was published by Greulich and Pyle nearly 70 years ago using a primarily White cohort from northern Ohio. The rate of skeletal maturation has accelerated over the past several decades, partly due to earlier onset of puberty. The objective of this project is to create an automated bone age interpretation tool that will base bone age on a library of more than 20,000 trauma hand radiographs acquired at CCHMC. This volume of data will allow for both sex and race/ethnicity-specific bone age assessment. Such an algorithm aligns with POPT (Care, Community) and should considerably improve bone age estimation accuracy while decreasing inter-radiologist variability.

Using Deep Learning (DL) architecture, an AI algorithm for automated assessment of bone age will be created. Our very large radiographic dataset will be divided into training, validation, and testing cohorts. The output of this algorithm, estimated skeletal age in months, will automatically be imported into the radiology report along with other important clinical data, such as actual age in months and demographics. We anticipate no problems in completing this task as we have sufficient input data and Dr. Rama Ayyala, a co-champion of this project, has developed related algorithms prior to her arrival at CCHMC (e.g., PMC6113150). This project will be considered successful when all bone age exams performed at CCHMC receive an automated interpretation that prepopulates the imaging report with a mean error within \pm 6 months. Such an algorithm has considerable potential for commercialization, as most pediatric radiology and endocrinology practices in the U.S. and throughout the world currently use the outdated Greulich and Pyle reference standard.

Project #2: Automated Radiographic LDA (lines, drains, and airways) Detection (Appendix 3)

Hundreds of chest radiographs are acquired weekly at CCHMC to evaluate the presence and locations of vascular catheters (lines), drains (e.g., chest tubes), and airways (e.g., endotracheal tubes). Evaluation of “LDAs” is important as malpositioned devices can have major clinical implications (and have been the cause of several CCHMC serious safety events). Furthermore, identifying these devices is a labor-intensive task for radiologists, with devices commonly unrecognized due to distracting objects located outside of the patient. The objective of this project is to create an object detection algorithm that will identify the presence and locations of LDAs based on thousands of CCHMC radiographs obtained in our various intensive care units. This project aligns with POPT (Care) and promotes both standardization and patient safety.

Using DL architecture, an AI algorithm for automated assessment of LDAs will be created. Our large dataset will be divided into training, validation, and testing cohorts. Patients of all ages will be included in these datasets to promote generalizability. The output of this algorithm will be an annotated image in PACS (or One View) that identifies the locations of LDAs. This annotated image will be available to both radiologists and clinical services. With additional work, it may also be possible to import the algorithm’s results directly into imaging reports as well as evaluate agreement with the “Lines, Drains and Airways” section in Epic. As some patients may have up to 5-10 different types of LDAs, we may need to focus initially on devices that are more likely to cause patient harm if malpositioned. This project will be considered successful when all chest radiographs performed on CCHMC inpatients undergo algorithmic assessment for LDAs. While simple algorithms exist in this space (e.g., endotracheal tube tip identification), no such algorithm exists to detect the spectrum of LDAs, particularly in children. Such an algorithm has considerable potential for commercialization.

Project #3: Automated Liver Histologic Fibrosis Staging with Multiple Outputs (Appendix 4)

Histologic assessment of core biopsy specimens is the current reference standard for detecting and quantifying the presence and severity of liver fibrosis. Fibrosis staging may use one of several semi-quantitative systems (e.g., METAVIR, Ishak) and is subject to imperfect inter-pathologist agreement ($k=0.3-0.7$). This level of agreement can adversely impact patient care and deleteriously impact research studies using histology as a reference standard. The objective of this project is to create an algorithm that will perform automated histologic fibrosis staging of digitally scanned slides, with outputs including commonly used semi-quantitative fibrosis stages and percent area collagen. This project aligns with POPT (Care, Cure) and promotes standardization.

Using DL architecture, an AI algorithm for automated assessment of histologic liver fibrosis will be created. Our dataset will include pediatric and adult CCHMC patients with a variety of causes of acute and chronic liver

disease in order to promote generalizability; we also expect to have additional adult histologic data available from the outside institutions through a co-existent R01 (PI: He, R01EB030582). The reference standard will be the consensus interpretations of at least 3 hepatopathologists. The output will be a report that includes multiple measures of liver fibrosis, including Ishak and METAVIR staging as well as percent area collagen. The main potential limitation is the time needed for human annotation of liver histologic specimens and associated inter-pathologist agreement. This project will be considered successful when all clinical and research liver biopsy specimens can undergo automated fibrosis staging, with the results exportable to clinical pathology reports. Such an algorithm has clinical and research applications and potential for commercialization.

Project #4: Automated CT / MRI organ segmentation (e.g., lungs, heart, liver, fat, and muscle) (Appendix 5)

Automated segmentation of organs and tissues from cross-sectional CT and MRI images has a variety of clinical and research applications. Examples include clinical reporting of organ volumes (e.g., liver, heart, lung), evaluating suitability for organ transplantation/partial resection, segmentation of muscle and fat for evaluation of sarcopenia and fat depots, and extraction of organs for further radiomic/DL AI assessments (e.g., measuring liver fibrosis without invasive biopsy, determining percent of lung affected by air trapping/emphysema). The objective of this project is to create AI algorithms that allow accurate segmentation of numerous organs of clinical or research interest, a typically laborious task for humans. This project is aligned with POPT (Care, Cure).

Using DL architecture, a series of AI algorithms will be created that allow automated segmentation of a variety of organs and tissues. The creation of such algorithms is feasible due to our large amounts of institutional CT and MRI data and local AI expertise in this area. The reference standards will be manually created “masks” drawn by human image analysts. The output will be accurate segmentations of organs and tissues that allow further quantitative and AI evaluations. We anticipate no major problems in completing this project, assuming we have adequate access to image analysts to create the requisite organ/tissue “masks”. This project will be considered successful when accurate segmentations of multiple organs and tissues are translated into clinical workflows and available to be used by CCHMC researchers. These algorithms have considerable potential for commercialization, in part because they will have been trained, validated, and tested on pediatric-rich datasets.

III. Timeline:

During months 0-12, we will establish our image-focused AI core, including hiring necessary personnel and acquiring necessary computing resources. Institutional review board (IRB) approvals will be obtained during this period. During months 7-24, necessary input data will be obtained and annotated, and AI algorithms will be trained, validated, and tested; algorithms also will be evaluated for bias/fairness. During months 25-36, we will focus on the clinical implementation of our four algorithms, working with IS and imaging/pathology IT. Our efforts also will be presented at key meetings and published in peer-reviewed journals. We anticipate the AI core will be self-sustaining at ~36-48 months, primarily via support from ARC members and their labs as well as faculty/staff researchers across CCHMC. Minimal space will be needed for incremental personnel and hardware.

IV. Plan:

A detailed Business Plan is presented in **Appendix 6**. The proposed AI core will broadly impact CCHMC by bringing state-of-the-art image-based AI to the masses (basic science, translational, and clinical researchers from all areas of CCHMC) to answer key research and clinical questions. The proposed AI core ultimately will become financially sustainable through external grant funding/fee-for-service activities, intellectual property development and commercialization, external collaborations with industry partners, and by adding value to healthcare system through improved outcomes and efficiencies. We anticipate ARC team members will submit 1-3 NIH R01 or equivalent grants in years 2 and 3 that will employ the services of this AI core beyond year 3.

V. Annual Milestones:

Year 1: Successful image-based AI core creation (including acquisition of necessary hardware and personnel); IRB approvals in place for proposed projects; AI algorithm input data secured.

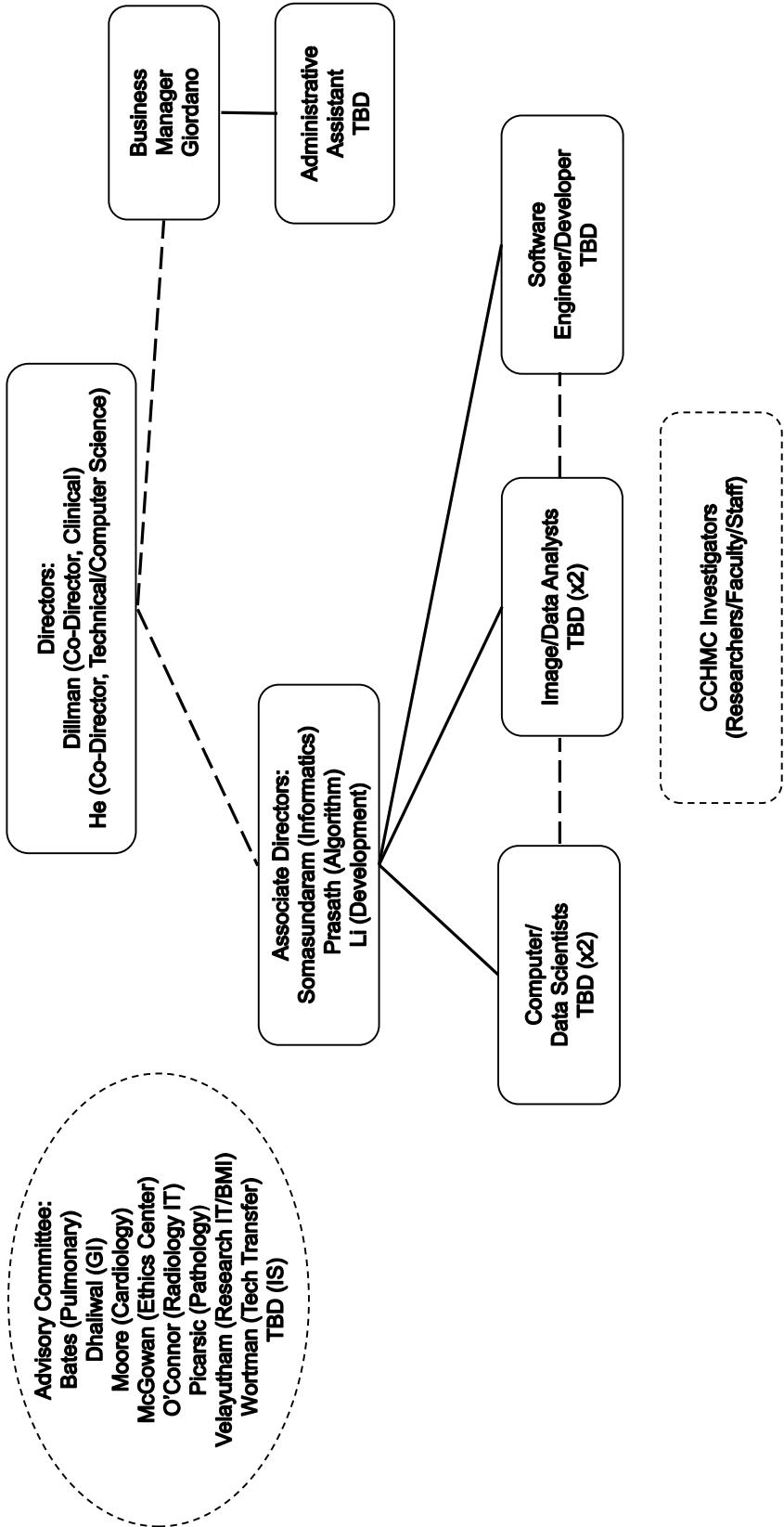
Year 2: Training, validation, and testing of proposed AI algorithms; creation of an AI education series (seminars and workshops) for CCHMC faculty/staff; 1-3 NIH grants submitted (based on ARC data and plan to use AI core).

Year 3: Clinical implementation of AI algorithms, working with IS/enterprise imaging; initiate commercialization dialog with potential industry partners; 1-3 NIH grants submitted (based on ARC data and plan to use AI core).

VI. Detailed Budget (CCRF budget forms)

A detailed budget is presented in **Appendix 6**, along with the Business Plan.

Appendix #1 Image-Based AI Core Structure



Appendix #2

Project Title: Creation of a Modernized Bone Age Interpretation Tool using Artificial Intelligence

Co-Champions: Rama S. Ayyala, MD; Elan Somasundaram, PhD

Objective: Bone age assessment is used in a variety of clinical settings to guide patient treatment. The current standard of reference for interpretation is an atlas created in the 1950s by Greulich and Pyle (GP) using left hand radiographs from a cohort of Caucasian children in northern Ohio. This reference is outdated given many factors that affect growth and maturation in children have changed. The objective of this project is to create a modernized bone age atlas and interpretation tool using an artificial intelligence (AI) algorithm developed using Deep Learning (DL) algorithms and trained on pediatric trauma hand radiographs. The goal of this project is to create a tool that is applicable to the current pediatric local population, taking into account a variety of demographic factors such as age, sex, race, and ethnicity as well as changes in skeletal maturation rate that have occurred over the past 70 years.

Background and Significance/Impact: Bone age radiographs are utilized in clinical practice for the diagnostic workup of a variety of pediatric conditions, including numerous endocrine, metabolic, and growth disorders. The standard of care reference for bone age interpretation currently is the GP *Radiographic Atlas of Skeletal Development of the Hand and Wrist*, which was created in the 1950s based on left hand radiographs of white children in the suburbs of northern Ohio. Since the 1950s, the overall growth and maturation of children has changed. Puberty onset in girls has been trending earlier in the United States and other developed countries (PMID 32040143); in the United States boys have also been documented to have earlier onset of puberty (PMID 23085608). This alteration in puberty timing is thought to be related to rising rates of obesity in the population and exposure to endocrine hormone disruptors in the environment (PMID 18245510). The creation of a modernized bone age reference utilizing a current pediatric patient population would take these changes into account and potentially provide improved clinical care. A modernized bone age reference will also take into account differences in race, ethnicity, and potentially geographical location, which is lacking in the current reference standard.

Not only would this tool be more applicable to today's pediatric population and in theory help clinicians make more informed medical and surgical decisions, it would also streamline workflow for those interpreting these examinations, who are typically pediatric radiologists and endocrinologists. The GP method currently entails using a hardcopy book containing pictures of normal radiographs for various ages of children and having clinicians perform a direct visual comparison of the patient's left hand image and radiographic standard within the atlas. This can be cumbersome and time-consuming, especially with variability of the quality of the radiograph pictures based on condition of the book. This contributes to inter and intrareader variability (PMID: 30069585, 29141916, 25581985), which can impact initial accuracy as well as subsequent evaluations, potentially making a difference in clinical treatment. Creating an automated interpretation tool that can be clinically implemented utilizing a modernized bone age atlas can improve workflow by lessening the time associated with the bone age interpretation as well as eliminating inter-reader variability.

Innovation: Previous AI algorithms for bone age interpretation have utilized bone age (hand) radiographs with GP interpretations by radiologists as the ground truth. The project we propose is novel in that the AI algorithm will be trained on pediatric trauma hand radiographs using the patient's chronological age as the ground truth. In addition, the algorithm can be trained on subsets of trauma hand radiographs for patients of various races and ethnicities to create more accurate interpretations in these cohorts, thus eliminating a form of inaccuracy and bias that is present in the current system. A preliminary algorithm created by a group led by Dr. Ayyala has shown promising results and will be utilized as the foundation to build an optimal algorithm for this project (<https://pubs.rsna.org/doi/10.1148/ryai.2020190198>).

Approach: This will be a multidisciplinary project involving Radiology, Endocrinology, Information Science and Radiology IT. A dataset of more than 20,000 pediatric trauma hand radiographs obtained over the past 10 years at CCHMC has been identified. This data will be stripped of the protected health information (while documenting age, sex, race, and ethnicity from electronic medical records) and processed into a format that can be utilized for algorithm development. This large dataset will be divided into cohorts for algorithm training, validation, and testing. The chronological age of the patients will be used as the ground truth, and the algorithm also will take into account sex, race and ethnicity of the patients. Benchmark convolutional neural network (CNN) architectures pre-trained for the image-net classification problem will be re-trained and evaluated for bone-age prediction using the validation set and the optimal CNN architecture will be selected. The CNN architecture will also be customized to account for the patient demographic factors. The final algorithm

will be tested on a cohort of previously unseen trauma hand radiographs to establish final diagnostic accuracy. Once the accuracy of the algorithm is determined and deemed acceptable (~6 month inter-rater variability), the tool will be integrated into CCHMC clinical workflow so that every bone age radiograph receives an automated interpretation prior to radiologist viewing.

In addition, the bone age assessment algorithm created will be tested on a large cohort of bone age radiographs performed over the past 10 years at CCHMC and compared to the human radiologist interpretations based on the GP method. This will show how this algorithm varies from the GP interpretations, and therefore help delineate changes in maturation in the United State pediatric population over the past 70 years. After this, independent datasets from various institutions from different geographical locations, nationally and globally, will be obtained to evaluate generalizability of the algorithm against various demographic factors.

Finally, representative hand radiographs from children of various age, sex, race, and ethnicity groups will be used to build an electronic virtual bone age atlas that is based on our modern patient population.

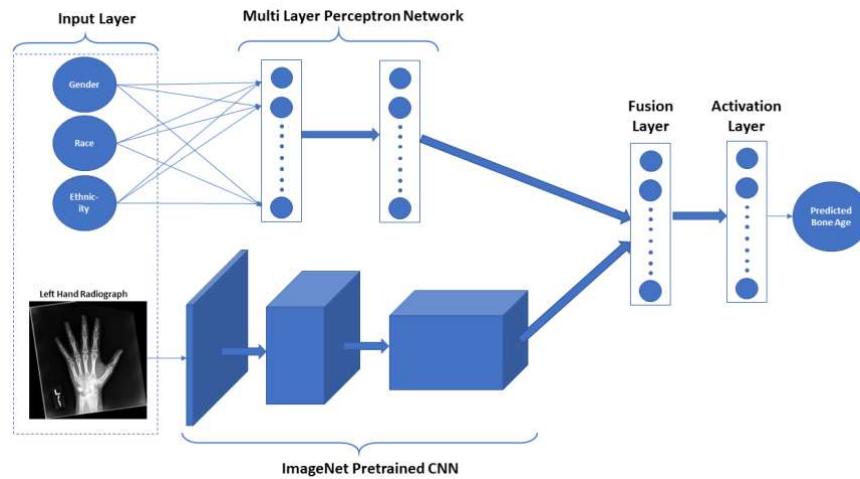


Figure 1: Deep Learning algorithm framework for bone-age prediction using left hand radiographs and demographic data.

Clinical Implementation Strategy: The final algorithm will be packaged in micro-services based containers that seamlessly integrate with existing clinical data pipelines. The containers will then be deployed onto clinical-grade servers designed to support multiple AI algorithms. To facilitate ease of clinical integration, the algorithm will automatically generate an industry standard DICOM Structured Report object containing the associated patient and exam metadata, along with the AI calculated bone age measurements. The resulting report object can then be automatically exported using the DICOM communications protocol to a variety of clinical and information systems. The targeted systems include Voice Dictation systems for integration of AI bone age calculation data into the diagnostic report as well as the Picture Archiving and Communication System (PACS) and Enterprise Archive for long term clinical use and storage. The intent of the clinical integration is that the AI calculated bone age measurements are automatically pre-populated into the exam's diagnostic report prior to interpretation by a Radiologist.

Commercialization Potential: The current standard reference tool is the GP hard copy atlas, which has limitations as described above. This algorithm will be a natural successor that potentially will improve accuracy in the modern pediatric population, providing better medical care and clinical decision support, as well as improving workflow by increasing speed of interpretation. Therefore, this tool can be commercially used by pediatric radiologists and endocrinologists in a variety of settings given the technological portability and use of standardized DICOM structured report output.

Alignment with CCHMC Mission and POPT/Digital Transformation: This project targets the Care and Community aspects of the CCHMC/POPT mission. Bone age radiograph interpretations play an integral role in the clinical care of patients with a wide array of medical and surgical disorders. Accurate interpretation taking into account known differences among race and ethnicity is important to direct appropriate care. This project will transform the current standard reference, which is outdated and based on a homogenous population cohort, to be applicable to the modern, diverse pediatric population. The creation of an automated tool that can be commercialized will allow a wide range of radiologists and clinicians with variable resources, whether it is at a tertiary level academic medical center or in a small community practice, to be able to promptly obtain accurate bone age interpretations.

Appendix 3

Project Title: Automated Radiographic Lines, Drains, and Airways (LDAs) Detection

Co-Champions: Alexander J. Towbin, MD; Maya Dewan, MD; Lili He, PhD

The placement of Lines, Drains, and Airways (LDAs) is an essential part of the critical care delivered in the intensive care and postoperative units.¹ Patients with an LDA are often imaged to determine the presence and location of their lines (e.g., vascular catheters), drains (e.g., chest tubes, mediastinal drains), and/or airways (e.g., endotracheal tubes, tracheostomy tubes) as malpositioned LDA devices can have major clinical implications. The **objective** of this multidisciplinary project is to create a multistep algorithm that will identify the presence and location of clinically important LDAs based CCHMC radiographs obtained in our various intensive care units (ICUs), compare the identified lines to the current LDA record in Epic, determine if each LDA is concordant with what is reported in Epic, and generate a draft radiology report (**Figure 1**). Specifically, we will create a novel AI algorithm using deep learning architecture for automated assessment of LDAs based on cutting edge image object detection and segmentation techniques. The algorithm will take single view chest X-rays (CXR) images as input and automatically output annotated images with the delineation of LDAs. The proposed AI algorithm will be developed and tested using ~1,000 radiographic portable inpatient CXRs.

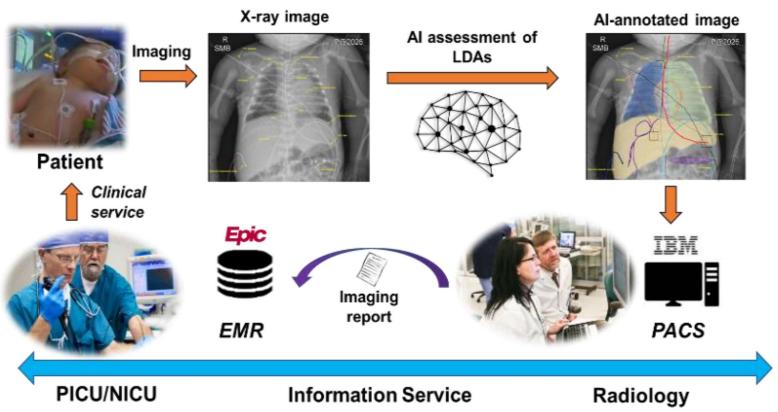


Figure 1. A conceptual overview of the LDA detection project.

A. Significance/Impact: The significance of proposed project is high, because identifying radiographic LDAs is paramount for patient safety and is a labor-intensive task for radiologists, with devices commonly unrecognized due to distracting objects located outside of the patient.²⁻⁶ Hundreds of chest radiographs are acquired weekly at CCHMC to assess LDAs, requiring a tremendous work effort from the Department of Radiology. The official LDA reconciliation process currently only occurs following admission or transfer of a patient to one of the three ICUs at CCHMC. If not for the work effort, the formal LDA reconciliation process would ideally occur during all inpatient CXR interpretation as malpositioned LDAs can cause complications such as airway compromise, vascular arrhythmias, vascular damage, cardiac perforation, bowel perforation, aspiration, and death. Currently, more than 23,000 single view CXRs are performed each year at CCHMC. While timely and accurate identification of LDAs typically occurs, these devices have been associated with numerous serious safety events, serious harm, and many other safety events (such as unplanned extubations) over the past 5 years.

B. Innovation: This project will provide a very desirable tool to the CCHMC community, allowing automated and timely assessment of LDAs. The proposed project is innovative in terms of the following aspects:

B.1. Novel deep learning model to evaluate radiographic LDAs: Literature exists on assessment of LDAs using segmentation methods,^{2,6} but none of existing approaches utilize the integration of object segmentation and detection strategy to delineate the LDAs devices, locate the tips of clinically important LDAs, and compare the list to the LDA record in the electronic medical record.

B.2. Novel full spectrum detection of LDAs: While simple LDAs assessment algorithms exist (e.g., to identify the location of ETT tips),^{3,9} no such algorithm exists to detect the full spectrum of LDAs, particularly in a pediatric population, and none are currently in routine clinical use.

C. Approach: A conceptual overview of the proposed AI algorithm LDAnet is shown in **Figure 2**.

C.1. Scientific premise. Deep learning has been applied to detect/segment the LDAs,^{2-4,6,9} but the prior methods are still not ready for clinical routine usage, particularly in children. Recent breakthroughs in both object segmentation and detection models enable a fine-grained recognition of challenging subtle characteristics of images.^{10,11} This will boost the performance of LDA evaluation model, ready for daily clinical routines.

C.2. Scientific rigor. Our research design closely follows the guidance for imaging research on artificial intelligence provided by the journal *Radiology* Editorial Board¹² to achieve robust, fair, and reproducible results:

1) To eliminate the main confounding factor (i.e., CXR imaging device bias), we will conduct data harmonization; 2) Consideration of biological variables: Sex is not a confounder, because this project is to assess radiographic LDAs devices. In contrast, age may be a confounder. The types of LDAs may vary for patients in different age groups (e.g., neonates, infants, or older children). If necessary, we will design and train different customized models for patients in different age groups.

C.3. Ground truth data annotation. We will conduct data annotation to create ground truth for model training. Specifically, we will manually delineate region map, object map (i.e., LDAs), and bounding box of LDAs tips. This will be performed by experienced data analysts, guided by CCHMC radiologists and ICU physicians.

C.4. LDAnet model design and training. The proposed LDAnet includes three components: object detection (orange panel), region localization (green panel), and LDA assessment (blue panel) (**Figure 2**). Object detection component will delineate LDAs devices; region localization component will outline the regions of interest; and then LDAs assessment component will utilize both region and object maps to detect the tips of LDAs and output the final annotated CXR image. We will design customized residual U-Net models^{11,13} for both region localization and object detection components. The LDAs assessment component will be an object detection model to locate the tips of LDAs with small bounding boxes. We will apply the state-of-the-art Mask R-CNN model for LDAs' tips detection.¹⁰ The proposed LDAnet will have two outputs 1) an automated annotated image with delineated regions of interested, identified LDAs, and bounding boxes to highlight the tips of LDAs, and 2) structured text identifying the line and the anatomic location of its tip. The text output will be used to compare to the real-time LDA record in Epic and determine if each identified line is concordant or discordant with what is documented. The image-based output will be sent as a DICOM secondary capture image to the PACS while the text output will be automatically incorporated into the draft radiology report using DICOM SR.

C.5. LDAnet model evaluation plan. We will train and validate the model using nested k-fold cross-validation on development set with 800 samples. Then, we will test the model on unseen test dataset with 200 samples. Model performance will be evaluated with Dice similarity score, balanced accuracy, mean squared distance error.

C.6. Potential issue and alternative solution. Due to the intensity variability, one single LDAs device may be erroneously detected as two or more separated segments. In this case, we will apply a probabilistic Hough line transform algorithm¹⁵ to merge the significant nearby contours to generate a smoothly curved LDAs trajectory.

C.7. Expected outcomes and endpoints. The expected outcome of this project is an independently validated LDAs assessment AI algorithm. This project will be considered successful when all chest radiographs performed on CCHMC inpatients undergo algorithmic assessment for LDAs that are more likely to cause patient harm if misidentified or malpositioned.

D. Research team and environment. To accomplish this multi-disciplinary project, we will apply our diverse team's expertise in pediatric critical care, radiology, computer science, and biostatistics to this study.

E. Clinical Implementation Strategy: The algorithm will be triggered after the completion of each inpatient CXR. The clinical workflow would remain unchanged for all providers. Each inpatient CXR would automatically be assessed. The output text file would be compared to the patient's LDA report using a real-time connection to Epic. Besides the Epic connection and the hardware the LDAnet algorithm will run on, all components of this workflow currently exist and are in routine clinical use at CCHMC.

F. Commercialization Potential: Automated LDAs assessment would be a welcomed additional safety strategy for all centers who provide care for ICU patients, both pediatric and adult.

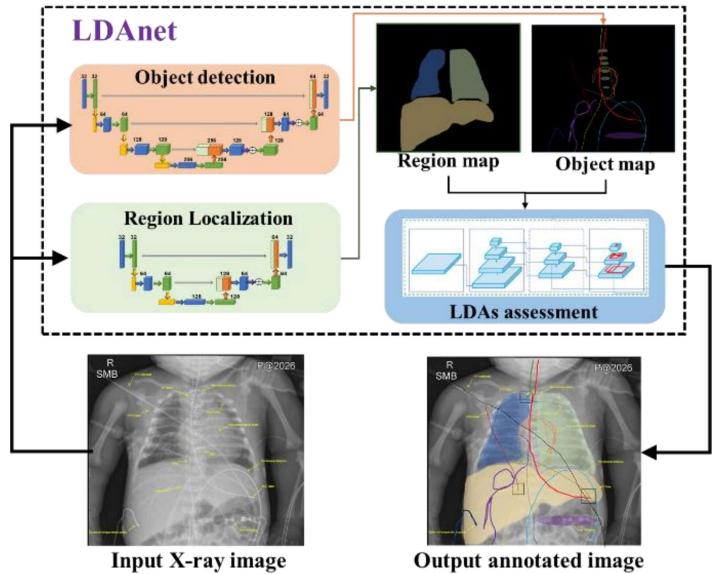


Figure 2. The proposed AI algorithm (LDAnet) to automatically identify the presence and locations of clinically important LDAs.

G. Alignment with CCHMC Mission and POPT/Digital Transformation: This project supports and strengthens the CARE POPT pillar at CCHMC. By improving the speed, accuracy, and consistency of identification and reporting of LDAs within children, we will improve medical care for our sickest patients.

References

1. Brooks AJ, Mahoney PF, Rowlands B. ABC of tubes, drains, lines and frames: John Wiley & Sons; 2009.
2. Lee H, Mansouri M, Tajmir S, Lev MH, Do S. A Deep-Learning System for Fully-Automated Peripherally Inserted Central Catheter (PICC) Tip Detection. *J Digit Imaging* 2018;31:393-402. PMID: 28983851; PMCID: PMC6113157.
3. Lakhani P. Deep Convolutional Neural Networks for Endotracheal Tube Position and X-ray Image Classification: Challenges and Opportunities. *J Digit Imaging* 2017;30:460-468. PMID: 28600640; PMCID: PMC5537094.
4. Lakhani P, Flanders A, Gorniak R. Endotracheal Tube Position Assessment on Chest Radiographs Using Deep Learning. 2021;3:e200026.
5. Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* 2017;284:574-582. PMID: 28436741.
6. Subramanian V, Wang H, Wu JT, Wong KCL, Sharma A, Syeda-Mahmood T. Automated Detection and Type Classification of Central Venous Catheters in Chest X-Rays. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*; 2019; Cham. Springer International Publishing. p 522-530.
7. Khasawneh FA, Smalligan RD. Guidewire-Related Complications during Central Venous Catheter Placement: A Case Report and Review of the Literature. *Case Rep Crit Care* 2011;2011:287261. PMID: 24826318; PMCID: PMC4010052.
8. Srivastav R, Yadav V, Sharma D, Yadav V. Loss of guide wire: a lesson learnt review of literature. *J Surg Tech Case Rep* 2013;5:78-81. PMID: 24741424; PMCID: PMC3977329.
9. Singh V, Danda V, Gorniak R, Flanders A, Lakhani P. Assessment of Critical Feeding Tube Malpositions on Radiographs Using Deep Learning. *J Digit Imaging* 2019;32:651-655. PMID: 31073816; PMCID: PMC6646608.
10. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. 2017. p 2961-2969.
11. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention* 2015:234-241.
12. Bluemke DA, Moy L, Bredella MA, Ertl-Wagner BB, Fowler KJ, Goh VJ, Halpern EF, Hess CP, Schiebler ML, Weiss CR. Assessing Radiology Research on Artificial Intelligence: A Brief Guide for Authors, Reviewers, and Readers-From the Radiology Editorial Board. *Radiology* 2019;192515. PMID: 31891322.
13. Li H, Chen M, Wang J, Illapani VSP, Parikh NA, He L. Automatic Segmentation of Diffuse White Matter Abnormality on T2-weighted Brain MRI Using Deep Learning in Very Preterm Infants. 2021;0:e200166.
14. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 2012.
15. Kiryati N, Eldar Y, Bruckstein AMJPr. A probabilistic Hough transform. 1991;24:303-316.

Appendix #4

Project title: Automated Detection and Scoring of Liver Fibrosis from Histopathology Digital Slides for Standardized Clinical Reporting

Co-Champions: Jennifer Picarsic MD; Surya Prasath, PhD

Objective: The prevalence of liver disease in children is increasing, with non-alcoholic fatty liver disease (NAFLD), the most common cause of chronic liver disease, affecting 1 in 10 children. The objective of this study is to develop and validate artificial intelligence (AI) driven algorithms for automated scoring of histologic liver fibrosis based on validated systems such as Metavir, Ishak, and percent area fibrosis using digitally scanned slides. We aim to develop a histopathologic image-based AI tool that considers the digital whole slide image variations across a large cohort of patients and automates the output of various validated scoring systems for liver fibrosis.

Background and Significance/Impact: There are a multitude of liver diseases which present in childhood (biliary atresia in infancy, autoimmune liver disease (e.g., primary sclerosing cholangitis), and most commonly NAFLD) that warrant longitudinal assessment for liver fibrosis, an important predictor of morbidity and mortality and one of the major indicators for determining hepatic functional status and predicting outcomes of liver diseases and liver allograft dysfunction (PMID: [30465125](#); [30335697](#); [32536769](#)). The sequelae of complications include liver transplantation (or re-transplantation) and hepatocellular carcinoma in late childhood/young adults (PMID: [32536769](#); [30959975](#)). Thus, recognizing and developing therapies to arrest or reverse hepatic fibrosis are becoming critical in many liver diseases including childhood diseases such as biliary atresia and familial cholestasis (PMID: [18026579](#)).

Histologic assessment of core needle biopsy specimens is the current reference standard for detecting and quantifying the presence and severity (stage) of liver fibrosis (PMID: [16531536](#)). The importance of this is evident from the many classification systems that have been proposed to assess liver fibrosis starting with the Ishak 6 stage system originally proposed for hepatitis. Since then, we have the Metavir system, the Scheuer system, and more recently systems designed for assessing fibrosis in circulatory disorders, especially Fontan livers, nonalcoholic steatohepatitis (NAFLD staging), and transplant allografts (PMID: [20101752](#); [15915461](#); [16531536](#)).

Histologic liver fibrosis staging uses one of various semi-quantitative systems (e.g., Metavir, Ishak) and is subject to less than perfect inter-pathologist agreement ($k=0.3-0.7$) (PMID: [10395036](#); [16531536](#)). This level of agreement has the potential to impact patient care and can adversely impact research studies using histology as the reference standard or ground truth (PMID: [16531536](#)). The objective of this project is to create an algorithm that will perform automated histologic staging of digitally scanned liver biopsy slides, with outputs including more commonly used semi-quantitative stages (e.g., Metavir, Ishak) as well as percent area fibrosis for standardized clinical reporting.

No such system truly exists to uniformly standardize liver fibrosis reporting across all diseases and between pathologists, but digital imaging with AI-based quantification of liver fibrosis is one technique that may allow for this by providing automated and reproducible results (PMID: [16531536](#)). This proposal gives us the chance to develop a novel algorithm with multiple different scoring systems as outputs, including total percent fibrosis using digital techniques on pediatric liver biopsy specimens. This work also may serve as the foundation for future studies to develop a single International scoring system for liver fibrosis.

Innovation: Despite the existence of various AI models for assessing pathologic images (PMID: [30603216](#)), no such prior AI techniques exist for the standardized evaluation of liver fibrosis in the pediatric setting. Using deep learning (DL) models, namely convolutional neural networks (CNNs), this project will strive to develop an optimal AI model that can provide automated assessment of histologic liver fibrosis. The developed AI based frameworks will be tested on histopathological imaging data from one of the world's largest pediatric cohort of chronic liver disease patients. Furthermore, we will test our final algorithm in adult patients to evaluate generalizability.

Approach: This multi-disciplinary project involves pediatric pathologists, gastroenterologists, and biomedical informaticians along with Information Services (IS) and Information Technology (IT) experts for clinical implementation. A dataset of more than 300 whole slide histologic images (WSIs) from at least the last three years at CCHMC with associated anonymized clinical metadata will be utilized. Our approach will include a careful evaluation of the image data to account for potential confounders and will apply automated preprocessing, including deconvolution, stain normalization to handle various imaging characteristics across histology variations, and adjusting for size/length of the core biopsy. We will leverage deep CNN models pre-trained on the large-scale cancer genome atlas (TCGA) by utilizing *transfer learning* along with adaptive learning on our in-house pathologists curated data for automated detection of liver fibrosis and staging (**Figure 1**). We will test various blocks such as the convolutional, inception, and ResNet along with *finetuning* of the top layers and test against multiple pathologists curated scores (including a group consensus score) of WSIs. The optimal model will be chosen by using hyperparameter optimization on the number of layers and convolutional blocks to achieve optimal accuracy in scores. The final systematic AI-based automated scoring will be compared with ground truth to standardize the process and develop a high level of concordance. A prior approach by Dr. Prasath's lab has shown promising results assessing brain glioblastoma multiforme histopathologic image data (PMID: [30603216](#)) and is currently being expanded for predicting disease course in pediatric ulcerative colitis (UC) with similar histopathology WSI and associated clinical data. These algorithms will be utilized as a foundation to build an optimal CNN model for scoring fibrosis in all patients as well as in Fontan and allograft livers. There is also a strong potential to expand to other liver biopsy scoring systems (e.g., NAFLD/NASH).

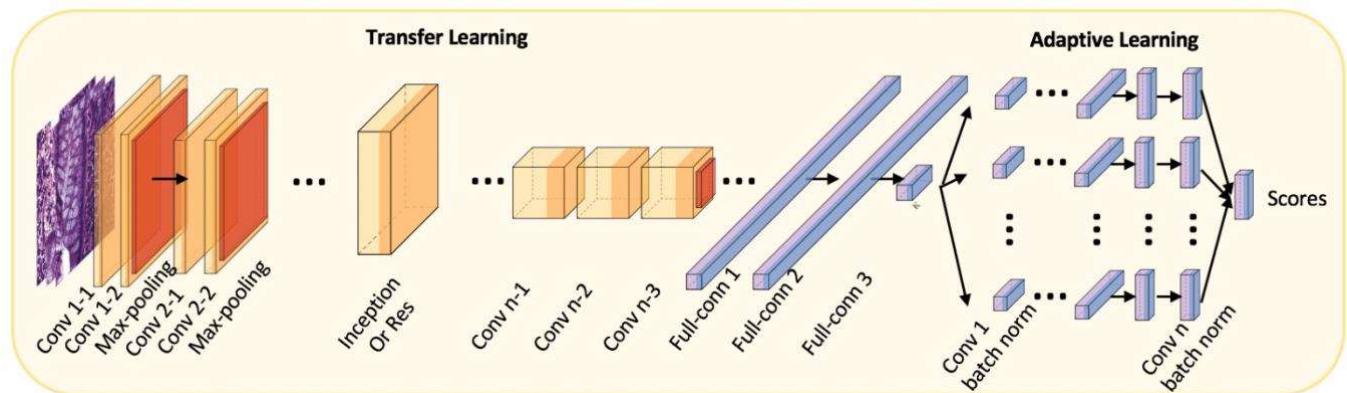


Figure 1: Deep Learning algorithm framework for liver fibrosis scoring using digital histopathologic image data.

Clinical Implementation Strategy: Once validated, this system will lend itself to use in daily clinical practice as an objective and standardized tool to supplement the current subjective assessment by individual pathologists reading slides, improving reproducibility between and within pathologists as well as reproducibility over time. The standardized clinical reporting of the various scores will be incorporated into clinical pathology core workflows. This tool ultimately may develop into a more standardized system of scoring liver fibrosis that could be used across institutions. Furthermore, this algorithm can be made available to CCHMC researchers that use histology as the reference standard for liver fibrosis.

Commercialization Potential: Once developed and adequately tested, there is the ability to market this product as a unique tool that improves accuracy and standardization of liver histologic fibrosis assessment. We anticipate interest from both children's and adult hospitals as well as industry, where pathologists deal with the similar dilemma of subjective scoring systems for liver fibrosis evaluation. Such an algorithm also can serve as ground truth for drug companies investigating anti-fibrotic therapies and for the development of novel liver fibrosis biomarkers, such as noninvasive imaging methods.

Alignment with CCHMC Mission and POPT/Digital Transformation: This project aligns with both the Care and Cure pillars of POPT as well as our Digital Transformation efforts. Specifically, we will transform traditional glass slides that have remained unchanged for decades with regards to technology into high-resolution digital images, that can then be evaluated using state-of-the art AI-based methods for both clinical and research applications. We expect these applications to improve patient care, enhance clinical workflows, and enable research.

Appendix #5

Project Title: Automated CT / MRI organ segmentation

Co-Champions: Alister Bates, PhD; Samuel Brady MS PhD; Ryan Moore, MD

Objective: Automated segmentation of organs and tissues from cross-sectional CT and MRI images.

Background and Significance/Impact: Using Deep Learning architecture, a series of AI algorithms will be created that allow automated segmentation of various organs and tissues. The creation of such algorithms using convolutional neural networks is feasible due to our large amount of institutional cross-sectional CT and MRI data and AI expertise in this area. The purpose of this project is to allow accurate segmentation of numerous organs of clinical/research interest, a typically laborious task when performed by humans. Examples include clinical reporting of organ volumes (e.g., liver, spleen, heart, lung), evaluating appropriateness for organ transplantation, segmentation of muscle and fat depots for evaluation of sarcopenia and visceral vs. subcutaneous fat, and segmentation of organs and lesions for radiomic and deep learning predictions (e.g., amount of liver fibrosis without needle biopsy, percent of lung affected by BPD/emphysema).

Innovation: Three major areas of organ/tissue segmentation will be explored: 1) lung and airways, 2) cardiac, and 3) abdominal organ/tissue (e.g., liver, spleen, kidneys, pancreas, fat, and muscle). First, segmentation of the upper and central airways allows more detailed analysis of airway diseases than by invasive endoscopy or imaging alone, namely:

detailed geometric analysis of segmented tracheas can predict the increase in breathing effort due to airway diseases and diagnose tracheomalacia in neonates as accurately as bronchoscopy^{1,2} (**Figure 1**). Many airway diseases involve dynamic collapse, requiring segmentation at several time points during the breathing cycle to incorporate airway motion into geometric and airflow models. Airway segmentation is an extremely time-consuming process, particularly using MRI, due to the many complex structures surrounding the airway such as the epiglottis and arytenoids. Similarly, lung segmentation at different lung volumes is extremely useful because it can be used to demonstrate regional ventilation, thereby determining the local lung function. Segmenting the lung at different points in the respiratory cycle can also be used to calculate basic clinical parameters such as tidal volume or classifying how much of the lung is abnormally emphysematous or fibrotic. Additionally, semantic lung segmentation algorithms can characterize lung regions, such as cysts, “ground glass opacities”, or bronchial wall thickening, all of which is currently identified by a radiologist, but is time consuming and can be difficult to quantify the degree of a particular lung abnormality. Automatic segmentation of these regions allows for user-invariant quantification of the percentage of lung tissue in these regions. Therefore, automated segmentation of the upper airway and lungs significantly improves the feasibility of using these techniques in large number of subjects, allowing research studies of larger populations and ultimately translation into the clinic. A current investigation is underway involving lung segmentation to use total lung volume for lung transplantation size matching.

Cardiac segmentation has several clinical benefits. First, patients with complex cardiac anatomy often undergo extensive 3D modeling for surgical planning³. This requires labor-intensive, detailed segmentation of all anatomic structures (**Figure 2**). Segmented 3D heart models can be 3D printed or transferred into robust virtual and augmented reality surgical simulation environments for detailed virtual surgical planning. Due to the highly interactive nature of these newer 3D modalities, the segmented surface models have become the predominant digital representation requested by surgeons (versus classic non-segmented, volume rendered models). Automated segmentation using deep learning convolutional neural networks is a necessary next step in cardiac segmentation to improve

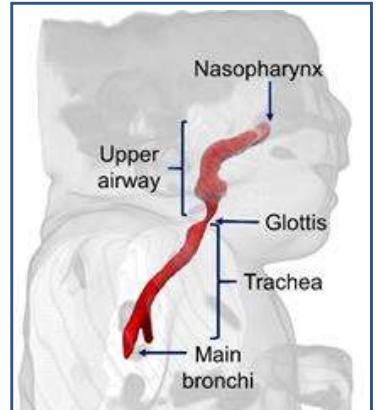


Figure 1. MRI at 40 weeks. Severe bronchopulmonary dysplasia and tracheomalacia. Preliminary model takes 1 min for airway auto segmentation.

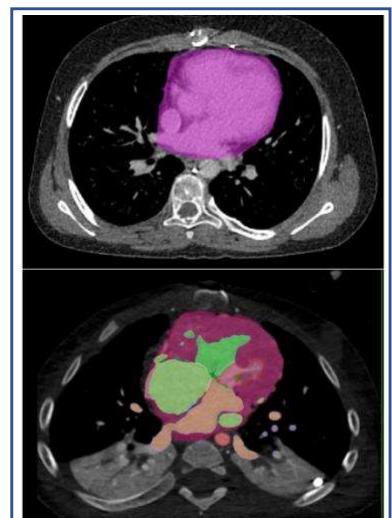


Figure 2. Examples of cardiac segmentation. Total cardiac volume in a Fontan (top) and detailed congenital heart segmentation for surgical planning (bottom).

throughput of 3D models for virtual surgical planning of complex congenital heart defects. Second, auto segmentation for pediatric heart transplantation donor pool expansion, using image-based total cardiac volume *in lieu* of the unreliable weight-based method as a surrogate of cardiac size. Concern for donor size mismatch is the second most common reason for refusal of donor hearts (~40% of all refusals)⁴⁻⁵. Image segmentation of chest computed tomography (CT) can provide a reliable measure of total cardiac volume by partitioning the heart from other thoracic structures (**Figure 2**)⁵⁻⁶. It has been hypothesized that heart segmentation to generate 3D total cardiac volumes (3D-TCV) may be the most accurate predictor of recipient-donor organ matching and could lead to expansion of the standard donor pool for transplant candidates⁷⁻⁸. However, accurate heart segmentation is limited by the need for extensive 3D image training and labor-intensive manual, and semi-automatic, processes. At this time, real-time heart segmentation of donors and recipients for matching purposes is not feasible due to the rapid response needed to accept donor offers. Our team has generated over 200 supervised segmentations of CT datasets to calculate 3D-TCV. We plan to develop a deep learning convolutional neural network to automate 3D-TCV. This will enable us to increase the number of segmented CT scans more efficiently in our “virtual donor pool”. Additionally, we are setting up a “federated learning” pipeline with multiple transplant centers so that we can further develop and train our TCV algorithm, thereby improving generalizability. The hope is that with a robust virtual donor pool, any pediatric heart transplant candidate can be rapidly matched to a donor to improve organ allocation and decrease waitlist mortality.

Third, automated abdominal multi-organ segmentation has, historically, been based on atlas registration and statistical fusion, which methods have proven to be difficult when implemented in a pediatric population due the wide range of body habitus; organ shapes, sizes, locations, and textures; and differing level of visceral fat. Deep learning networks have the advantage of discretizing the image and identifying organ boundaries, thus, not relying on model registration and fitting. Advances in abdominal-organ segmentation provides basic organ volume, shape, and location information and more critical morphologic data leading to the diagnostic prediction power for disease such as: inflammation, steatosis, fibrosis, and acute and chronic disease severity. Moreover, the role of multi-organ segmentation will provide a more holistic view of the image such as how a radiologist reviewed whole image leads to disease prediction and diagnosis. Computer-based radiomic feature identification will be enhanced when incorporating multiple organ feature analysis such as the use of liver and spleen volume, along with liver and spleen texture features, and liver tissue elasticity when predicting hepatic stiffness and passive hepatic congestion⁹. Beyond, the abdominal visceral organs, segmentation and quantification of skeletal muscle mass (**Figure 3**) is a general marker of overall pediatric growth and health and decreases (sarcopenia) have been linked to outcomes for various disease states, including surgical and non-surgical outcomes¹⁰.

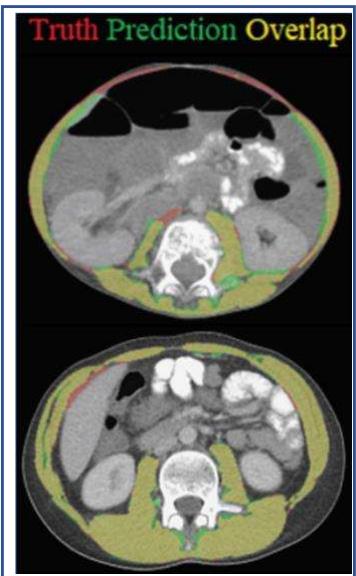


Figure 3. Preliminary data shows examples of skeletal muscle segmentation (top) 80% agreement, (bottom) 94% agreement. [ref 10]

Approach: AI-based segmentation is made feasible by the quality of the data used to train the algorithm. The initial task for this project will be to gather cross-sectional data from institutional repositories (e.g., the Department of Radiology Picture and Archiving Communication System). The data will be curated and identified to ensure heterogeneity of the data: using a variety of imaging types (i.e., different scanner makes/models—this is made possible due to the local data acquired on CCHMC scanners and the reference data acquired elsewhere but imported into our clinical image repositories), patient disease states (i.e., normal and abnormal—studies will be characterized based on diagnostic reports made available through keyword searches within the reporting databases), and a heterogeneous patient population (i.e., varying patient demographics—patient data will include variation by gender, body size, ethnicity, etc. as captured from patient EMR system). The reference standard will be manually created “masks” drawn by human data analysts. The output will be accurate segmentations of organs and tissues that allow further quantitative and AI evaluations. The definition of accuracy will be established by comparing AI-based segmentation models to inter- and intra-human segmentation accuracy. We anticipate no major problems in completing this project, assuming we have adequate access to image analysts to create the requisite organ/tissue “masks”.

Clinical Implementation Strategy: Current segmentation approaches used daily in the clinic are mostly all manual. Physicians identify patients that would benefit from additional information from organ/tissue segmentation. Typically, an imaging specialist will manually perform the segmentation. The resulting segmentation will then be used to provide quantitative measurements (e.g., liver volume in ml) or used in models or algorithms to supply the physician with additional clinical information for diagnosis; the entire process can take upwards of hours depending on the clinical workload for the imaging specialist. Automated segmentation can be implemented clinically as 1) a simplification to the current process that allows the current practice to be sped up by removing the manual segmentation step, or 2) allow for automated segmentation on all imaging studies wherein a report would be provided to the reporting physician; the physician would incorporate the AI-report into their final diagnostic analysis as needed. We anticipate extracted organs and tissues also will be used as part of clinically deployed deep learning models that diagnose disease and predict therapy response or outcomes.

Commercialization Potential: These algorithms have considerable potential for commercialization and wide acceptance into general practice, in part because they will have been trained, validated, and tested on pediatric and young adult datasets. These models are expected to have numerous clinical and research applications, and they will likely be critical preliminary data for future CCHMC grant submissions.

Alignment with CCHMC Mission and POPT/Digital Transformation: This project aligns with the Care and Cure pillars of POPT as well as our Digital Transformation efforts.

References

- ¹Hysinger EB et al. Chest 2020 157(3):595-602
- ²Gunatilaka CC et al. annals of the American thoracic Society 2020 17(10):1247-1256
- ³Moore RA et al. *Birth Defects Res.* 2018;110(13):1082-1090.
- ⁴Riggs KW et al. *J Thorac Cardiovasc Surg.* 2019 Jul 2. pii: S0022-5223(19)31329-7.
- ⁵Plasencia JD et al. *J Heart Lung Transplant.* 2017;36(4):S165.
- ⁶Szugye NA et al. *J Heart Lung Transplant.* 2019;38(3):331-333.
- ⁷Plasencia JD et al. *Pediatr Transplant.* 2018;22(8):1-11.
- ⁸Szugye NA et al. *J Heart Lung Transplant.* 2021 Feb;40(2):158-165.
- ⁹He L et al. AJR 2019 213(3):592-601
- ¹⁰Castiglione J et al. Radiology:AI 2021 3(2):e200130