

Expert-Guided Multi-Modal AI for Complex Pediatric Disease Management

Section 1: Use-Case & Problem Statement

Children with complex medical conditions, such as chronic diseases present unique treatment challenges. These children often arrive at specialized centers after receiving fragmented treatments across multiple institutions. Boston Children's Hospital (BCH), is ranked as #1 U.S. pediatric hospital for 9 years in a row, and is the world's largest pediatric research center. BCH frequently manages very complex pediatric diseases. A prime example is pediatric Inflammatory Bowel Disease (IBD), a condition marked by chronic inflammation of the digestive system. At BCH's IBD Center, over 2,000 pediatric patients receive ongoing care, with roughly 150 new diagnoses annually. Nationally, IBD affects nearly 3 million Americans, 10% of whom are children. Pediatric IBD poses distinct challenges due to its variable symptoms and complex clinical data, which demand time-intensive analysis. Inadequate long-term management often leads to surgeries, which are rarely curative. Repeated procedures risk complications, such as impaired nutrient absorption if an excessive amount of intestine is removed. Beyond physical effects, IBD disrupts growth, elevates cancer risks, and triggers psychosocial struggles like school absences, social isolation, and mental health disorders. Many challenges persist into adulthood, impacting careers, relationships, and quality of life, necessitating lifelong care. The financial burden is substantial - treating a child with IBD from diagnosis to adulthood costs \$500,000 to \$2.5 million. Families often endure lasting economic strain, which underscores the need for early intervention.

The development of a sophisticated clinical AI system capable of analyzing complex IBD data can address the gaps in managing this difficult chronic disease at a level that is currently not possible to meet due to the existing hospital resources being stretched by the disease's complexity. AI approaches that aim to extract key clinical information more efficiently, compute the well established disease severity scores that are otherwise time consuming to compute, and deliver risk assessments to optimize treatment decisions - could revolutionize treatment decision-making.

While large language models (LLMs) hold promise for extracting insights from electronic health records (EHRs), current AI systems face critical limitations in medical contexts: (1) they lack specialized domain knowledge for intricate pediatric cases, (2) struggle to integrate multi-modal data (e.g., imaging results, labs, clinical notes), and (3) often fail to provide transparent reasoning for their suggestions. Furthermore, most healthcare AI tools prioritize predictive analytics over the nuanced, evidence-based reasoning required for managing chronic pediatric conditions like IBD. To address these challenges, we propose to develop an expert-guided, multi-modal AI system that leverages three key assets: (1) BCH's unparalleled pediatric expertise, (2) access to one of the world's largest pediatric IBD datasets, and (3) advanced AI methodologies developed within Massachusetts' collaborative biomedical ecosystem from clinical AI experts at the highest level.

This proposed project will specifically target Crohn's disease, a subtype of IBD, with a focus on automating the Pediatric Crohn's Disease Activity Index (PCDAI)—the gold-standard metric for assessing disease severity in children. The PCDAI synthesizes symptoms, physical exam findings, lab results, and growth metrics to categorize disease states as remission (<10), mild (10–27.5), or moderate-severe (>30). However, calculating this severity score manually requires laborious aggregation of data from disparate sources, making it impractical for routine clinical use. By automating PCDAI tracking, our system aims to enable real-time severity monitoring, earlier intervention during flare-ups, and reduced hospitalizations while mitigating long-term complications. This innovation could shift pediatric Crohn's care from reactive to proactive management, improving outcomes. In addition to PCDAI, our system will also extract important disease severity metrics from imaging, endoscopy and histopathology data.

Additionally, this project will leverage AI to analyze patients' five-year clinical histories (1) to identify past medication failures (e.g., due to inefficacy, side effects) and (2) distinguish flare-ups between those caused by active inflammation vs confounders like infections. This directly complements the automated PCDAI system by providing clinicians with longitudinal insights into disease progression to optimize treatment. By synthesizing fragmented siloed data from EHRs, the AI system could flag historical trends that inform personalized care plans. Such granular analysis is currently unfeasible in routine practice due to the time-intensive nature of manually reviewing years of complex records.

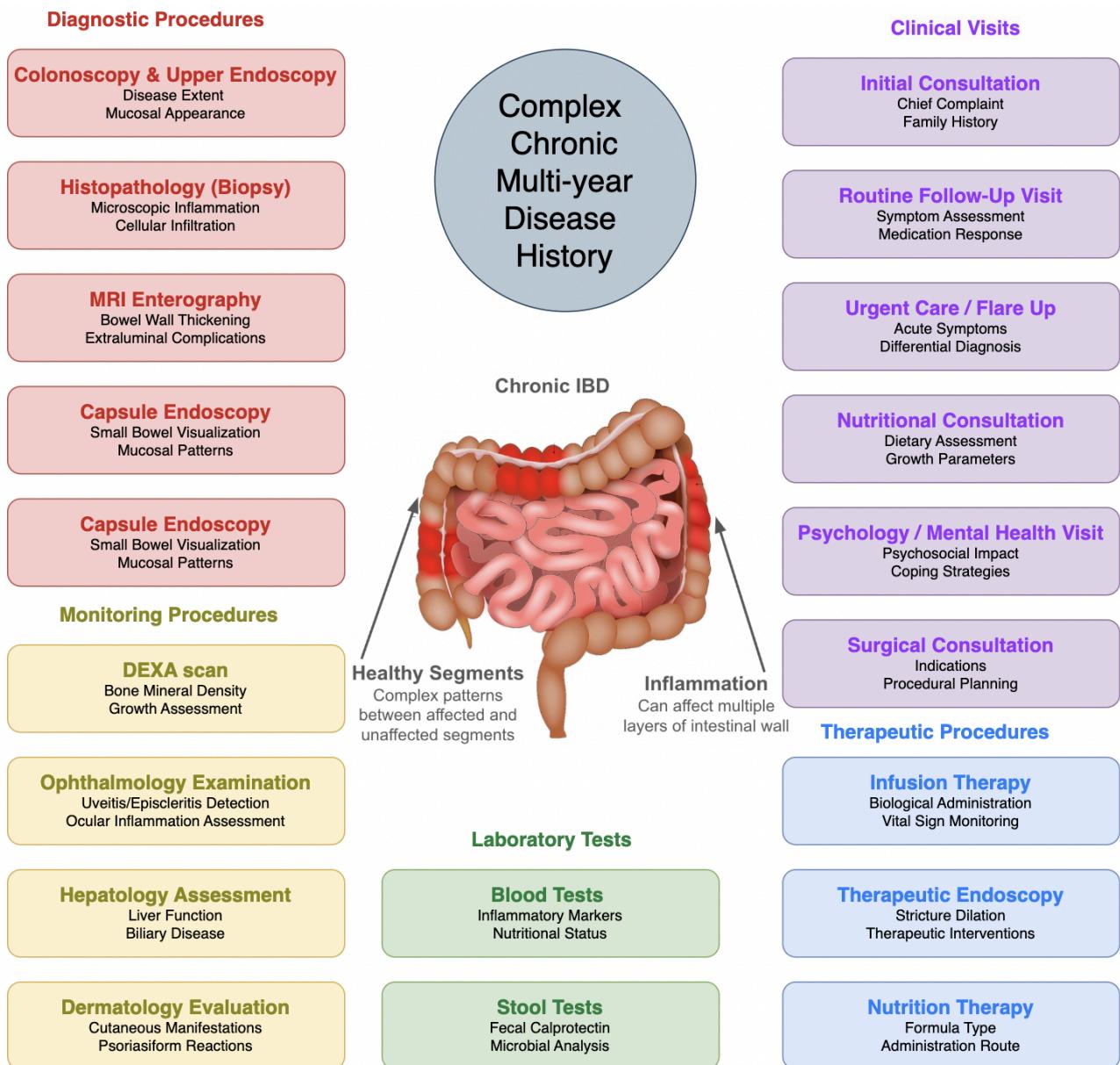


Figure 1: The Fragmented Data Landscape in Pediatric IBD Care. This figure illustrates the complex, fragmented nature of pediatric IBD management, where clinicians must manually synthesize data from numerous sources: diagnostic procedures, clinical visits, monitoring tests, laboratory results, and therapeutic interventions across years of patient history. This fragmentation makes comprehensive disease tracking, medication effectiveness assessment, and flare-up differentiation extremely difficult in routine clinical settings, often resulting in delayed interventions, suboptimal treatment decisions, and reactive rather than proactive disease management.

Section 2: Proposed Solution & Model Overview

We propose to build an expert-guided multi-modal AI system that is able to extract, analyse and score complex markers of Crohn's disease from multiple fragmented and siloed clinical data across the patient's record to generate actionable, explainable insights for clinicians, including clinical notes, radiology reports, and lab tests. The project is structured around three aims designed to bridge gaps in (1) data integration, (2) model training, and (3) clinical validation. First, we will create a domain-specific clinical data extraction framework using LLMs that are enhanced with clinician-guided steering mechanisms and advanced AI reasoning tools. This will enable us to accurately extract relevant data points from clinical notes, radiology reports and laboratory tests.

These LLMs will be trained to identify and synthesize critical Crohn's disease markers (e.g., treatment histories, flare triggers, lab trends) from unstructured EHR narratives. To ensure clinical relevance, we will employ reasoning architectures inspired by frameworks like ReACT (Reasoning and Action), which are similar to tools such as openAI's o3 and DeepResearch, which will iteratively refine outputs using expert feedback loops. Next, the extracted data will then be used to fine-tune models for IBD-specific reasoning, enabling them to replicate the nuanced decision-making patterns of BCH's specialists. To train robust, compliant AI models, we have partnered with Boston-based Amazon AI team to collaborate with them on their Bedrock and SageMaker platforms. Amazon will provide secure, HIPAA-compliant infrastructure—including high-performance GPU clusters and specialized tooling—to fine-tune state-of-the-art open-source LLMs at scale. This collaboration addresses critical gaps in BCH's internal capacity, as existing hospital systems lack the computational resources and AI engineering expertise required for large-scale model training.

Third, we will validate the AI system in real-world settings by collaborating with BCH's gastroenterology and radiology departments, which are key in frequent IBD visits. The model will be tested on its ability to (1) automate PCDAI scoring across all active Crohn's cases, together with additional scoring based on imaging, endoscopy/colonoscopy and histopathology results, (2) identify patterns in medication failures, and (3) distinguish IBD-related flare-ups from confounding factors like infections. Successful implementation could streamline care workflows, reduce diagnostic delays, and minimize unnecessary interventions—directly addressing the systemic inefficiencies outlined in earlier sections. In the next section we present a detailed technical overview of all three aims.

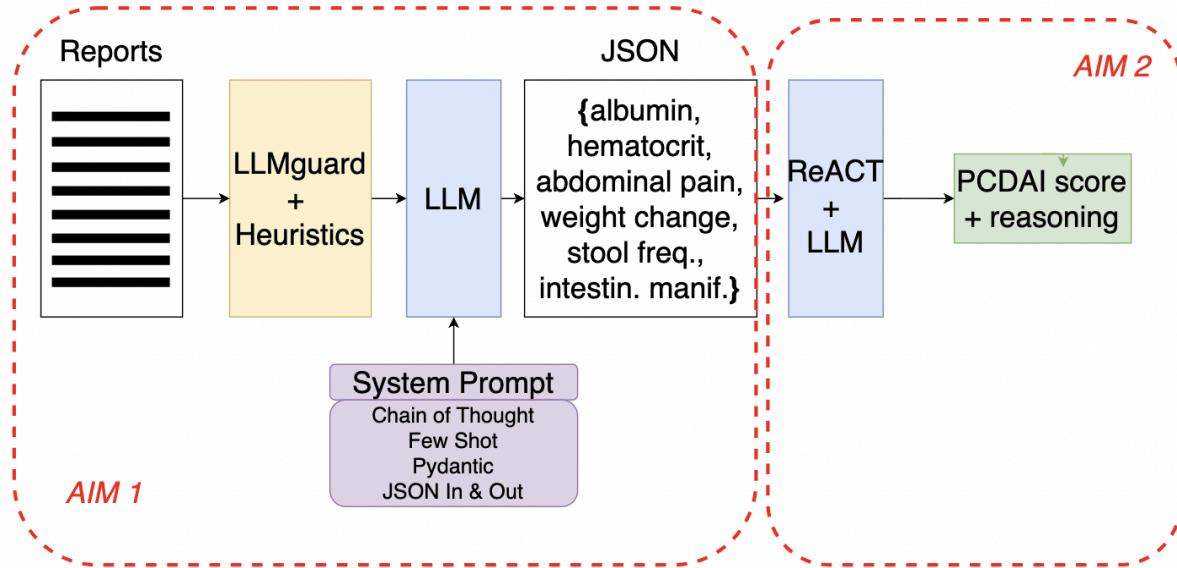


Figure 2. We illustrate our proposed expert-guided fine-tuned multi-modal AI system that will comprehensively assess pediatric Crohn's disease across patient populations. In Aim 1, we will build a comprehensive data extraction pipeline to standardize fragmented Crohn's disease data into structured formats aligned with medical taxonomies. In Aim 2, we will fine-tune a ReAct-style LLM (Reasoning → Action → Observation) to mirror pediatric gastroenterologists thought process during disease assessment and treatment planning, integrating multimodal data (notes, labs, imaging) and leveraging clinician expertise through iterative feedback loops during training.

Aim 1: Develop and Validate a Multi-Modal, Automated, Large-Scale Clinical Data Extraction

Framework: We will build a comprehensive data extraction pipeline to standardize fragmented Crohn's disease data into structured formats aligned with medical taxonomies. The framework will extract and harmonize: 1) **Unstructured clinical notes:** Using named entity recognition (NER) to identify key terms (e.g., symptoms, treatment histories); 2) **Radiology reports:** Codifying findings (e.g., wall thickening, fistulae) using RadLex, a standardized radiology lexicon; 3) **Endoscopy and histopathology reports:** Capturing mucosal inflammation patterns and disease extent. 4) **Laboratory results:** Extracting IBD-specific biomarkers (CRP, ESR, fecal calprotectin) with strict unit validation.

The framework will incorporate multiple safeguards to ensure reliability and clinical accuracy: 1) Domain-constrained Chain-of-Thought reasoning to mimic clinician logic, 2) IBD-specific few-shot examples to improve context-aware extraction, 3) Microsoft Guidance framework to enforce data integrity (e.g., valid lab ranges, measurement units). This structured dataset will form the foundation for training AI models that replicate expert clinical reasoning.

Aim 2: Fine-tune an Expert-Guided Clinical Reasoning Model with ReAct-style architecture

We will fine-tune a ReAct-style LLM (Reasoning → Action → Observation) to mirror the diagnostic logic of BCH's pediatric gastroenterologists. The model will integrate multi-modal data (notes, labs, imaging) to simulate clinician workflows and leverage clinician expertise through iterative feedback loops during training.

Technical Implementation: Base Model will be Llama 3.3 70B (8-bit quantized), chosen for its balance of performance and computational efficiency. HIPAA-compliant Amazon Bedrock cluster with 8x NVIDIA H100 GPUs will be used to enable scalable training unavailable in the hospital.

Training Pipeline: The training pipeline will include 3 steps: 1) **Domain-adaptive pre-training:** On de-identified IBD patient notes to build disease-specific knowledge. 2) **Instruction fine-tuning:** Using ReAct-style datasets to align outputs with clinical reasoning patterns. 3) **Reinforcement learning with human feedback (RLHF):** Refining outputs based on clinician preferences (e.g., prioritizing safety over novelty). This approach ensures transparency by replicating the structured reasoning clinicians use in practice.

Aim 3: Automated Population-Level IBD Scoring and Monitoring

We will deploy our expert-guided fine-tuned multi-modal AI system to comprehensively assess pediatric Crohn's disease across BCH's patient population. The system will produce: 1) **PCDAI scoring:** Generating real-time severity indices (remission/mild/moderate-severe) from integrated data, that are practically infeasible to gather in normal clinical setting due to time limitations; 2) **track medication effectiveness,** identifying failures and underlying causes; 3) **Flare-up differentiation:** Distinguishing true flare-ups due to inflammation from mimics (e.g., bacterial infections) using historical and biomarker trends. All these metrics are currently very difficult to track due to time-intensive nature and complex data integration requirements.

Validation: Metrics will include accuracy against expert consensus from two pediatric radiologists and two gastroenterologists (ICC >0.85), flare detection capability (sensitivity >0.85, specificity >0.90), and longitudinal precision.

Impact: By enabling continuous, system-wide monitoring, this tool will transform IBD care by preventing flare-ups, optimize medication selection, reduce hospitalizations, and improve outcomes for pediatric patients. It will also reduce delayed interventions, prevent unnecessary surgeries, and optimize therapies—directly addressing the care inefficiencies and costs outlined earlier.

Section 3: Data Generation & Curation

Data Sources: De-identified clinical records from Boston Children's Hospital (BCH) pediatric patients with confirmed Crohn's disease diagnoses, validated through clinical, endoscopic, radiologic, and histologic criteria. Patients with ulcerative colitis (UC) or unclassified IBD will be excluded.

Demographic Representation:

Age Range: 2–21 years, reflecting the pediatric population treated at BCH.

Gender: Balanced distribution aligned with current clinical demographics.

Ethnic Diversity: Reflect the racial and ethnic composition of BCH's IBD patient population, adhering to hospital protocols for equitable data aggregation.

Data Scope: Retrospective data spanning 10 years (2014–2024), extracted from BCH's EHRs.

Clinical documentation will primarily consist of gastroenterology visit notes (new consultations, routine follow-ups, urgent care encounters). These notes will be extracted from standardized EPIC templates that capture essential disease parameters including laboratory values (e.g., CRP, fecal calprotectin), medication histories, physical examination findings (e.g., growth metrics, abdominal tenderness), and treatment responses. Each record typically contains EHRs from 8-12 annual visits, providing a very large breadth of longitudinal data on disease progression and treatment outcomes. Additional less frequent clinical documents would include procedure notes (endoscopy, surgery), pathology reports (biopsy results, histologic activity indices), and interdisciplinary consultations (e.g., nutrition, physical therapy, psychology).

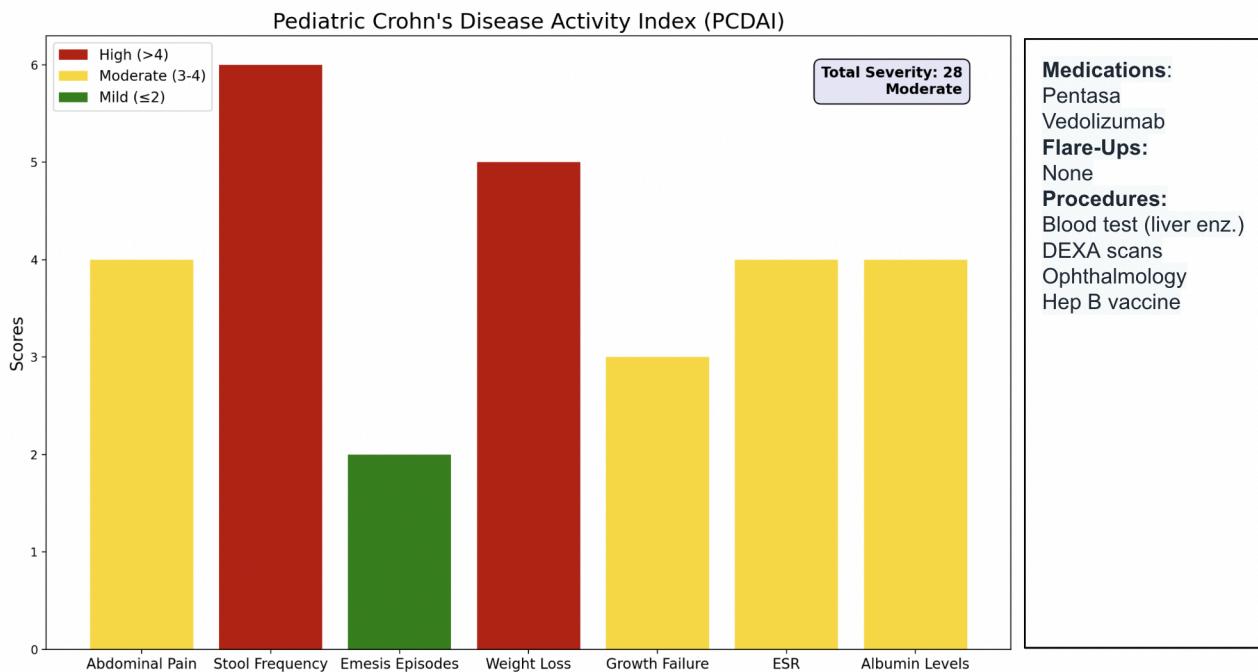


Figure 3. In Aim 3, we will deploy our expert-guided fine-tuned multi-modal AI system to assess pediatric Crohn's disease across BCH's patient population. The system will produce: 1) disease severity scoring including PCDAI and imaging based metrics, 2) track medication effectiveness from medical history, and 3) distinguish flare ups origin if present (e.g., bacterial infections). Here we show preliminary evaluation of a patient with moderate PCDAI score using a LLM with advanced reasoning capabilities. The patient is receiving Vedolizumab and Pentasa medications to reduce gut inflammation; and is referred to: DEXA scans as IBD patients show increase risk of osteoporosis, we well as ophthalmology exam and liver enzymes test as eyesight degradation and reduced liver function are common from immunomodulatory therapies. The AI is able to understand these complex therapies and extract the relevant markers from multiple clinical notes in the patient record history.

Data Extraction

Privacy & Anonymization:

To comply with HIPAA requirements, protected health information (PHI) will be scrubbed from the dataset through a multi-phase process. First, a rule-based named entity recognition (NER) system will remove explicit identifiers such as patient names, dates, medical record numbers, and demographic details. Next, the LLMguard software framework will perform secondary validation to detect subtler identifiers (e.g., rare abbreviations, contextual clues). Finally, a 70B-parameter LLM will scan all records for residual PHI patterns, followed by manual validation: 200 randomly selected notes will undergo human review to confirm anonymization efficacy. This layered approach minimizes re-identification risks while preserving clinical relevance for model training.

Medical Terminology Standardization: To ensure consistency and reliability in AI model training, clinical notes and radiology reports will undergo terminology standardization using the Unified Medical Language System (UMLS). This process will align variable or ambiguous terms, including mapping IBD-specific diagnoses and medications to standardized codes such as ICD-10 and RxNorm where applicable. To address institutional variations, a custom dictionary of BCH-specific abbreviations and clinician shorthand (e.g., "flares" vs. "exacerbations") will resolve local jargon. Clinical notes will also be parsed into standardized sections (e.g., Chief Complaint, Physical Exam, Assessment/Plan) using automated header recognition, ensuring uniform organization for downstream analysis. Additionally, all temporal references (e.g., "3 months post-diagnosis") will be converted to relative timestamps (e.g., "T+90 days") to enable longitudinal tracking of disease progression. This standardization mitigates ambiguities in raw clinical text and supports transparent, explainable AI reasoning in subsequent phases.

Training data preparation

Controlled Extraction Pipeline: To ensure clinically relevant and reliable data inputs, we will implement a structured extraction pipeline as described in Aim 1:

1. *Clinician-Driven Criteria Design:* Collaborating with BCH specialists, we will define a comprehensive taxonomy of indicators critical to IBD management (e.g., symptom severity, treatment response, complications). This taxonomy will reflect real-world clinical reasoning by observing and documenting how physicians assess patients during exams.
2. *Guided Keyword Extraction:* Locally deployed LLM, Llama 3.3 (70B), constrained by the Microsoft Guidance Framework, will systematically extract keywords from processed clinical data. This framework enforces strict rules to ensure consistency in identifying predefined indicator classes (e.g., lab trends, imaging findings).
3. *Expert Validation & Refinement:* Extracted keywords will be reviewed and refined by pediatric gastroenterologists and radiologists to ensure accuracy. Validated keywords will then be converted into instruction-tuning datasets (question-answer pairs). This data will be used to finetune the LLM model to replicate specialist decision-making.

Training Set Creation: The dataset will be partitioned into 80% training, 10% validation, and 10% testing sets to maximize generalizability. Stratification will ensure balanced representation across disease severity scores, age groups, and clinical subtypes (e.g., stricturing vs. penetrating Crohn's). K-fold cross-validation with k=5 will be used to evaluate model performance across diverse patient subgroups, reducing overfitting risks and ensuring reliability in real-world settings.

Training Protocol

Training will occur on Amazon Bedrock's enterprise-grade HIPAA-compliant infrastructure with a dedicated node equipped with 8x NVIDIA H100 GPUs. We will fine-tune Llama 3.3 70B (8-bit quantized), a state-of-the-art open-source LLM, with Low-Rank Adaptation (LoRA) to adjust model weights efficiently while maintaining computational feasibility.

Training will consist of three distinct stages for our Llama 3.3 70B (8-bit quantized) base:

1. Domain-Adaptive Pre-training: This phase focuses on familiarizing the model with IBD-specific terminology and clinical patterns using over 160,000 de-identified encounters, including gastroenterology notes, radiology reports, lab results, and endoscopy findings. The goal is to

enhance the model's ability to recognize clinical concepts (e.g. differentiating strictures from fistulas) and interpret longitudinal disease trends.

2. ReACT Framework and Instruction Finetuning: Using our curated IBD specific dataset, the model will undergo fine-tuning with Amazon SageMaker's parameter-efficient tools to optimize accuracy in scoring metrics like PCDAI and reasoning for disease exacerbations. Next, we will integrate the ReAct (Reasoning-Action-Observation) framework, training the model to replicate clinician workflows: analyzing clinical presentations (e.g., abdominal pain paired with elevated CRP), retrieving relevant historical data (e.g., prior medication responses), and synthesizing findings to assess disease activity or treatment failure. This phase leverages 200 expert-annotated IBD cases with detailed reasoning chains mirroring specialist decision-making.

3. Reinforcement Learning with Human Feedback (RLHF): The final training phase will incorporate clinician preference data across 200 paired comparisons to align model outputs with clinician priorities when scoring PCDAI indices, flare up diagnosis and medication assessments. Amazon SageMaker's RLHF toolkit will facilitate iterative refinement, while providing comprehensive tracking of model performance improvements.

Bias Mitigation & Validation

To ensure equitable performance, we will implement balanced sampling during training, ensuring equal representation of gender, ethnicity, and age groups (2–21 years). Post-hoc calibration will adjust decision thresholds for demographic subgroups (e.g., optimizing sensitivity across ethnicities). A comprehensive audit will evaluate accuracy and false-positive rates across patient subgroups, addressing disparities to ensure reliability in real-world application.

Section 4: Testing & Validation

Aim 1: Controlled extraction evaluation

Our data extraction will employ a three-stage validation. The initial extraction stage will utilize our foundational LLM with the Microsoft Guidance framework as mentioned in the previous section. Each extraction will be auto validated against existing structured EHR fields where available. The second stage will implement expert validation through random sampling of 10% of extracted records using a specially built validation interface, a web UI tool, for the manual review by GI specialists. Disagreements will be marked, creating a feedback loop to refine extraction rules for the foundational LLM system. Final quality assurance and evaluation will set minimum confidence thresholds, maintain audit trails for all extraction decisions in JSON like files, and calculate inter-rater reliability metrics between the automated system and expert reviewers. We will target extraction accuracy of >90% for critical IBD parameters (PCDAI components, disease location, treatment history) and >85% for secondary clinical indicators with Cohen's kappa coefficient >0.80 for inter-annotator reliability.

Aim 2: Model Training Evaluation

Training validation will follow a three-stage clinical pipeline. Stage one will assess extraction accuracy on 200 patient records against manual annotations by BCH gastroenterologists. Key parameters include disease classification, phenotype identification, and severity marker extraction. Disease phenotyping accuracy will be assessed through precision, recall, and F1 scores for disease location based on endoscopic and radiologic findings. Performance thresholds are set at >90% for primary classification and >85% for disease extent extraction. Stage two will comprise expert review by three pediatric gastroenterologists evaluating 60 complex cases with atypical presentations. The validation will focus on extraction accuracy of PCDAI scores, growth metrics, laboratory values, and treatment response indicators. Error analysis will quantify false positive/negative rates using a weighted confusion matrix. Error categories will be stratified by disease phenotype and severity. Discordant cases will undergo root cause analysis with classification into predefined categories: terminology misinterpretation, context misalignment, or missing critical indicators.

Aim 3: Core clinical metrics evaluation

Aim 3 focuses on three key clinical metrics where IBD disease management is to be improved:

PCDAI Score Validation: We will validate automated PCDAI scoring against expert-derived scores using intraclass correlation coefficient (ICC) with target threshold >0.85 . Three pediatric gastroenterologists will independently score 200 cases, establishing ground truth PCDAI values. Bland-Altman analysis will assess agreement across different severity ranges. Subcomponent accuracy for individual PCDAI elements (abdominal pain, stool frequency, well-being) will be measured separately.

Medication Effectiveness Tracking: System performance for identifying medication failures will be assessed using a curated test set of 60 treatment courses with documented outcomes. Evaluation metrics include precision and recall for detecting ineffective therapies (target >0.80), time-to-detection compared to clinical identification, and accuracy in attributing failure mechanisms (primary non-response vs. secondary loss of response).

Flare-up Detection and Classification: Validation will include targeted testing on 60 confirmed flare-up cases (30 infection-related, 30 disease progression). Performance targets include sensitivity >0.85 and specificity >0.90 for distinguishing infection-triggered flares from disease progression.

Longitudinal Performance Assessment

We will separately evaluate the system using 40 longitudinal patient records spanning 3+ years. Metrics will include consistency of PCDAI trend detection, agreement with clinician-identified disease trajectories, and stability of predictions across encounter types. Test-retest reliability will be measured on artificially fragmented clinical data to simulate incomplete documentation. This step is necessary to evaluate LLMs ability to assess disease progression.

Usability and Workflow Assessment

Our core goal is to present the system for clinicians to use. Clinician satisfaction will be evaluated using the System Usability Scale (SUS). Time-efficiency comparison between automated and manual PCDAI calculation will be quantified across 40 test cases. Information presentation effectiveness will be assessed through structured interviews with clinical users for flare up detection and medication effectiveness tracking.

Pilot Clinical Validation We will conduct a limited prospective evaluation ($n=40$) comparing standard care to AI-augmented assessment. Primary metrics will include correlation between automated and clinician-derived PCDAI scores, time-to-detection of disease flares, and accuracy of medication response prediction. A panel of three specialists will provide a blinded assessment of system outputs.

Section 5: Partnerships & Team

Sila Kurugol, PhD: is the director of the Quantitative Intelligent Imaging (QUIN) Research Laboratory At Boston Children's Hospital's Radiology Department, and an Assistant Professor of Radiology at Harvard Medical School. She received her BS, MS and PhD degrees in electrical and computer engineering. Over the last 20 years, her research focused on development of machine learning techniques for medical data analysis. She has extensive experience in using AI algorithms for automatically extracting imaging markers for improved management of Crohn's disease. She has been awarded several grants including an NIH R01 grant with direct focus on developing new imaging markers in IBD and Crohn's disease and evaluating them using PCDAI and other disease severity indices.

Serge Vasylechko, PhD: is a senior postdoctoral research fellow at Boston Children's Hospital and Harvard Medical School. He has acquired substantial knowledge on the AI applications for medical data during his PhD at Imperial College London. During the latter parts of his PhD, he had developed unsupervised deep learning methods for image reconstruction as well as frameworks that would combine multi-contrast sources. In the last 5 years he has been working with Dr. Kurugol to develop machine learning techniques for quantitative medical imaging. His recent work

includes using vision language models and large language models for extracting information from radiology reports to train an object detection algorithm for disease detection and localization. He has an extended track record of publications in top AI medical journals.

Jodie Ouahed, MD: is a pediatric gastroenterologist involved in the clinical care of many complex patients with focus on those with Inflammatory Bowel Diseases (IBD) and immune dysregulations. Dr. Ouahed is involved in several translational research projects focused on the genetics underlying early onset of IBD. Dr. Ouahed serves as an expert for the Department of Gastroenterology for Boston Children's Hospital Precision Medicine Service.

Lauren Collen, MD: Dr. Collen's clinical and research interests lie in the care of patients with IBD. She has special interest in understanding and treating patients with very early onset IBD (VEOIBD), defined as disease onset at less than 6 years of age, and those with disease that has been refractory to conventional therapies. Her research interests are in understanding the underlying pathways that drive disease in her IBD patients with the ultimate goal of advancing options for personalized medicine.

Andy Tsai, MD, PhD: is an Associate Professor of Radiology at Harvard Medical School and a board-certified pediatric radiologist. He has PhD in Electrical and Medical Engineering from MIT; and MD degree from Harvard Medical School. He has been a staff radiologist at BCH since 2010. He has collaborated with Dr. Kurugol to develop AI algorithms for medical data including generative AI models for image synthesis to improve computer aided detection and longitudinal data generation from cross sectional images using AI methods.

Lina Lu, MD: Dr. Lu is a pediatric radiologist with a focus on the clinical care of pediatric patients, with research interests primarily centered on the radiologic diagnosis and evaluation of abdominal diseases. Her prior research focused on utilizing CT to assess tumor response in patients with metastatic renal cell carcinoma undergoing treatment as well as evaluating the diagnostic performance of the Liver Reporting & Data System (LI-RADS) and the criteria for Organ Procurement and Transplantation Network (OPTN) for the detection of hepatocellular carcinoma (HCC) with MRI. Recent work includes aiding the optimization of diffusion-weighted MRI (DWI-MRI) to evaluate pediatric liver tumors as well as inflammatory bowel disorders, such as Crohn's disease.

Amazon Partnership:

To train robust, compliant AI models, we have partnered with Amazon AI team via their Boston-based Bedrock and SageMaker group. Amazon will provide secure, HIPAA-compliant infrastructure—including high-performance GPU clusters and specialized tooling—to fine-tune state-of-the-art open-source LLMs at scale. This collaboration addresses critical gaps in BCH's internal capacity, as existing hospital systems lack the computational resources and AI engineering expertise required for large-scale model training.

Section 6: Impact

Impact Potential and Strategic Alignment

Our expert-guided multimodal AI system addresses a critical societal challenge: improving care for children with chronic inflammatory bowel disease. By automating disease severity tracking and enhancing treatment decision-making, we aim to serve 2,000+ pediatric IBD patients at Boston Children's Hospital and set the course for improved management of this disease for the 300,000 children nationwide with this condition. The societal impact is substantial: IBD costs families \$500,000-\$2.5 million per child from diagnosis to adulthood, creating lasting economic strain. Our solution could significantly reduce this burden while minimizing the physical and psychosocial effects including growth disruption, cancer risks, school absences, and mental health disorders that often persist into adulthood affecting careers and relationships. This directly aligns with Massachusetts' healthcare priority sector, and positions the Commonwealth's Strategic AI Task Force goal of overcoming tangible societal problems in MA with AI. Within three years, we project tangible measurable outcomes as we target reduction in time-to-optimal-treatment and fewer complications requiring surgery.

Collaboration and Economic Development

This project creates a Massachusetts-centered collaboration between BCH (the world's largest pediatric research center) and the region's AI expertise by a partnership with Amazon AI team. The solution has significant commercial potential for expansion to other pediatric centers and IBD care.

Clinical Impact

The AI system will directly enhance clinical decision-making by automating extraction of more than 10 IBD-specific data points from disparate sources (laboratory tests, histopathology results, radiology studies and GI visits), reducing PCDAI calculation time from 20+ minutes to under 30 seconds. This will enable real-time disease severity monitoring for all 2,000+ BCH IBD patients and reduce treatment optimization delays from weeks to days. The system will identify medication failures and distinguish true inflammatory flares from infections with improved accuracy, preventing unnecessary escalation of immunosuppressive therapies.

Healthcare System Benefits

Implementation will deliver measurable system efficiencies. Thanks to the advanced extraction capability, the hospital will be able to track at a granular level the treatment process of this complex disease - for things such as medication failures, cause of flare ups and complex markers that make up the disease severity progression within the PCDAI index. The standardized extraction framework will reduce research data collection time, accelerating multi-center studies and biomarker discovery while establishing a replicable model for AI implementation across Massachusetts healthcare institutions.

Ethical and Responsible AI

Our HIPAA-compliant approach and strong partnership with Amazon create a template for ethical AI implementation in sensitive healthcare contexts that can be scaled to other centers around MA and beyond. Amazon partnership is critical for demonstration of responsible innovation in pediatric care while maintaining the substantial compute capacity that's necessary to drive these kinds of multi-modal AI systems in a clinical context. Amazon has expressed direct interest to maintain these trained models and use them for serving to the hospital in an ethical and responsible manner. This effort can help pioneer such industry collaborations for many institutions in MA.

Extending the Impact to Other Complex Diseases

While our proposal focuses specifically on pediatric IBD to demonstrate clear, measurable impact within the grant timeframe, the modular nature of our approach enables seamless extension to other complex chronic conditions requiring multi-modal data aggregation. Our framework readily adapts to cardio-genetics for congenital heart disease and rare pediatric liver cancers like hepatoblastoma—conditions similarly challenged by fragmented data and complex treatment decisions. Our support letters demonstrate strong interest from BCH's Cardiology Department, and our strong prior work with the oncology team on Liver Tumors provide natural pathways to amplify this investment's impact across Massachusetts' pediatric specialty care ecosystem.

Section 7: Project Plan

This project is planned for a 12-month period structured across four distinct phases with specific technical and clinical milestones. The project timeline incorporates critical dependencies between infrastructure development, model training, clinical validation, and deployment preparation.

Phase 1: Infrastructure Development (Months 1-3):

- Complete IRB approval and establish data sharing agreements with clinical departments
- Implement secure data architecture for EHR processing with HIPAA compliance
- Complete clinical note extraction pipeline with >90% accuracy on test dataset

Milestone Deliverable: Validated extraction pipeline capable of processing clinical documentation at scale

Phase 2: Model Development (Months 4-6)

- Complete curation of training data with expert annotation from GI and Rad departments
- Implement ReACT framework architecture and initial domain-specific fine-tuning
- Complete initial model training with validation on 100 retrospective cases

Milestone Deliverable: Working prototype system demonstrating baseline clinical reasoning capabilities for PCDAI scoring

Phase 3: Clinical Validation (Months 7-9)

- Complete comprehensive validation protocol with pediatric gastroenterologists
- Perform systematic performance testing across 200 defined clinical cases

Milestone Deliverable: Validated model meeting pre-specified performance benchmarks (>0.85 ICC for PCDAI scoring)

Phase 4: System Finalization (Months 10-12)

- Complete final validation with prospective cohort of 40 patients
- Generate documents reporting findings, prepare papers for publication
- Generate plans and roadmap for future deployment

Milestone Deliverable: System with comprehensive validation

Project Outcomes

1. **High-Quality Datasets:** Annotated training corpus of 2,000+ pediatric IBD clinical patients with expert-validated PCDAI scores
2. **Validated AI Model:** Expert-guided multi-modal AI system for automated PCDAI calculation with a target accuracy metric of >0.85 ICC agreement with clinical experts
3. **Performance Benchmarks:** Comprehensive documentation reporting model accuracy, reliability, and clinical utility metrics
4. **Compliance Documentation:** Determine ethical framework and privacy protection protocols for clinical AI implementation
5. **Deployment Plan:** Detailed implementation future roadmap for integration into BCH clinical workflows with collaborative partnership agreements

Reporting Schedule and Match Funding We will submit quarterly progress reports and invoices and a final comprehensive report upon completion to MassTech. The 25% match funding is awarded in-kind from BCH through effort of all clinicians and engineers and from the Amazon team through credits and in-kind technical support to help with training and validation of LLMs.

Section 8: Budget Narrative

Machine Learning Engineer (12 Calendar Months, Serge Vasylechko): will serve as the technical architect and primary implementer of the AI system, bringing expertise in large language models, medical imaging and medical data processing. They will be responsible for three core aspects of the project: (1) developing the multi-modal clinical data integration framework, (2) implementing the ReAct-style architecture for clinical reasoning, and (3) managing the model fine-tuning pipeline. Key responsibilities include designing and implementing the IBD-specific information extraction system and developing specialized fine-tuning approaches for the large language model. The engineer will work closely with clinical specialists to ensure technical implementations align with medical requirements and maintain high standards of accuracy and explainability, meeting with them weekly.

Senior Machine Learning Scientist (3 Calendar Months, Sila Kurugol): will provide critical oversight and strategic direction to ensure project feasibility and clinical integration. Their role involves translating complex clinical requirements into actionable engineering tasks and ensuring the technical architecture aligns with medical workflow needs. They will (1) review technical approaches, (2) evaluate implementation strategies, and (3) facilitate communication between technical and clinical team members. Their expertise is essential for assessing the scalability and

feasibility of proposed technical solutions and ensuring that the developed models meet both performance standards and clinical utility requirements. The multidisciplinary team will meet weekly and the technical team will meet twice every week to discuss progress and milestones.

Gastroenterologists (0.6 Calendar Months each, Drs. Jodie Ouahed and Lauren Collen): Two IBD specialist gastroenterologists will bring extensive expertise in pediatric IBD management. Their involvement is essential for multiple critical aspects: (1) defining and validating IBD-specific clinical parameters for LLM instruction tuning (2) curating appropriate patient cohorts, and (3) validating the AI system's clinical performance from diverse data sources including clinical notes, laboratory results, and treatment outcomes. The IBD team will collectively evaluate the AI model's scoring accuracy, assess the clinical reasoning behind score assignments, and validate the explainability of the model's decision-making process. Their expertise is crucial in ensuring the system maintains the high standards of clinical care.

Radiologists (0.36 Calendar Months each, Drs. Andy Tsai and Lina Lu): Two pediatric radiologists will contribute a quarter day per week, focusing on the imaging aspects of IBD assessment. Their role is crucial for: (1) standardizing the interpretation of MRI studies, (2) validating the automated extraction of imaging findings, and (3) ensuring the accuracy of image-text correlations in the multi-modal system. They will be instrumental in developing the ground truth for imaging biomarkers and validating the system's ability to integrate imaging findings with clinical data. Their expertise is essential for maintaining radiological accuracy and ensuring the system's imaging assessments align with clinical standards. The dual radiologist approach ensures robust validation of the system's imaging interpretation capabilities and provides comprehensive coverage of varying presentation patterns in pediatric IBD.

Computing Infrastructure and Resource (\$32,800 plus \$3000 credits from Amazon):

The project's computational backbone will be centered around a cluster of 8x H100 GPUs optimized for large-scale model training and fine-tuning using p5.48xlarge, which costs about \$113/hour. For 300 hours of computing, it will cost \$33,900. Storage for 2TB will cost around \$1900, totaling \$35,800. This infrastructure will support the intensive computational requirements for processing complex medical data and training specialized clinical AI models. The computing resources will be allocated across three main areas: (1) large-scale data preparation and standardization, including processing of clinical notes, imaging data, and laboratory results; (2) model development and fine-tuning, particularly for the adaptation of large language models to specialized pediatric use case; and (3) validation and testing environments for ensuring robust performance before clinical deployment. The budget includes storage, high performance computing and necessary technical support for optimizing model training processes and maintaining computational efficiency throughout the project lifecycle.

Attachment A
Application Cover Sheet

Name of Respondent The Children Hospital Corporation d/b/a Boston Children's Hospital			
Mailing Address 300 Longwood Avenue	City/Town Boston	State MA	Zip Code 02115-5724
Telephone 617-919-2729	Fax	Web Address www.childrenshospital.org	
Primary Contact for Clarification Sila Kurugol		Primary Contact E-mail Address sila.kurugol@childrens.harvard.edu	
Authorized Signatory Matthew Riley Foster		Authorized Signatory E-mail Address OSP@childrens.harvard.edu	
Legal Status/Jurisdiction (e.g., a Massachusetts Corporation, LLC, LLP, etc.) A Massachusetts Corporation with 501(c)3 status		Respondent's UEI No.: Z1L9F1MM1RY3	
		Respondent's EIN No.: 1042774441A1	

Attachment B
Massachusetts Technology Collaborative
Authorized Respondent's Signature and Acceptance Form

The undersigned is a duly authorized representative of the Respondent listed below. The Respondent has read and understands the NOFO requirements. The Respondent acknowledges that all of the terms and conditions of the NOFO are mandatory. By executing this Authorized Respondent's Signature and Acceptance Form, Respondent certifies that they (1) are in compliance with the terms, conditions and specifications contained in this NOFO, (2) acknowledges and understands the procedures for handling materials submitted to the Mass Tech Collaborative as set forth above, (3) agrees to be bound by those procedures, and (4) agrees that the Mass Tech Collaborative shall not be liable under any circumstances for the disclosure of any materials submitted to the Mass Tech Collaborative pursuant to this NOFO or upon the Respondent's selection.

The Respondent understands that, if selected by the Mass Tech Collaborative, the Respondent and Mass Tech Collaborative will execute an Agreement specifying the mutual requirements of participation. The undersigned has either (*please check one*):

- specified exceptions and counter-proposals to the terms and conditions of the [Capital Matching Grant Agreement](#); or
 agrees to the terms and conditions set forth therein;

The undersigned acknowledges and agrees that the failure to submit exceptions and counter-proposals with this response shall be deemed a waiver, and the Agreement shall not be subject to further negotiation.

Respondent agrees that the entire bid response will remain valid for sixty (60) days from receipt by the Mass Tech Collaborative.

I certify that Respondent is in compliance with all corporate filing requirements and State tax laws.

I further certify that the statements made in this response to the NOFO, including all attachments and exhibits, are true and correct to the best of my knowledge.

Respondent: The Children's Hospital Corporation d.b.a Boston Children's Hospital
(Printed Name of Respondent)

By: Matthew Riley Foster Digitally signed by Matthew Riley
Foster Date: 2025.02.27 16:30:12 -05'00'
(Signature of Authorized Representative)

Name: Matthew Riley Foster

Title: Senior Grant Officer

Date: 02/27/2025

Request for exceptions to the terms/negotiation for Massachusetts Technology Collaborative Capital Matchin Agreement

- Whole agreement
 - We request the ability negotiate the agreement's terms and conditions, should our application be selected for funding. Below are a few requests we can identify at this time, but it is not an exhaustive list, accordingly, we request the ability to negotiate if selected for funding
- Section 6. Indemnification and Liability
 - We request this flow in both directions:
 - Each Party shall indemnify, defend and hold harmless the other party and its successors and assigns, and all of its officers directors, lenders, shareholders, beneficial owners, trustees, partners, affiliates, agents and employees from and against any and all claims, suits, actions, judgments, demands, losses, costs, attorney's fees, expenses, damages and liability to the extent caused by, resulting from, or arising out of the intentional acts, negligent acts, errors, omissions, or allegations thereof, of the indemnifying Party, its employees, agents or representatives in the performance of work the SOW under the this Agreement.
- Section 11 Publicity
 - We request that this clause also flow in both directions. Further, as an academic institution performing fundamental research we request that there be no restrictions on our investigator's ability to publish their results
- Attachment 1 Statement of Work
 - Section 6. Ownership
 - We cannot grant MTC ownership of our deliverables, we request the ability to negotiate this. This is a grant not a service Agreement. We would request the opportunity to negotiate this further to negotiate a license
 - We cannot agree to our Background IP being disclosed without our permission nor can we agree to MassTech using our Background IP for MassTech's own purposes.
 - Accordingly, we request, "Participant understands all Participant Property provided under this Agreement is subject to disclosure as set forth above in Section 11, Public Records" be struck from the agreement.

**Massachusetts Technology Collaborative
Research and Development
Standard Budget and Invoice Template
Invoiced Costs**

Budgeting				
Applicant Information				
Applicant:		Sila Kurugol, PhD		
Budget Period:		12 months		
Title of Proposed Project		Expert-Guided Multi-Modal AI for Complex Pediatric Disease Management		
		MTC Funding Requested:	\$200,013.66	
Please note green cells contain formulas		Total Project Cost:	\$250,041.86	
Cost Elements		Total Project	Match Funds	MTC Funding
CAPITAL COST ELEMENTS				
Capital costs should directly contribute to the creation of long-term digital infrastructure assets, such as datasets, AI models, or deployment systems, etc. Only costs directly tied to these activities are eligible for reimbursement by MassTech.				
I. Capital Labor		Annual Salary	Percent Effort	
name/title				
Lauren Collen, MD		\$234,364.00	5%	\$11,718.20 \$2,343.64 \$9,374.56 0
Jodie Ouahed, MD		\$169,451.00	5%	\$8,472.55 \$1,694.51 \$6,778.04 0
Lina Lu, MD		\$426,400.00	3%	\$12,792.00 \$4,093.44 \$8,698.56 0
Andy Tsai, MD, PhD		\$448,056.00	3%	\$13,441.68 \$4,368.55 \$9,073.13 0
Sila Kurugol, PhD		\$163,000.00	25%	\$40,750.00 \$8,150.00 \$32,600.00 0
Serge Vasylechko, PhD		\$77,627.00	100%	\$77,627.00 \$15,525.40 \$62,101.60 0
Total Labor				\$164,801.43 \$36,175.54 \$128,625.89 0
II. Direct Labor Fringe Cost		Rate (%):	30%	\$49,440.43 \$10,852.66 \$38,587.77 \$0.00
III. Capital Subcontractors/Consultants		Firm name/title of consultant/type of services to be provided		
				\$0.00
				0
				0
				0
Total Capital Subcontractors/Consultants			\$0.00	\$0.00
IV. Capital Direct Materials				0
				0
				0
				0
				0
				0
Total Direct Materials			\$0.00	\$0.00
V. Other Capital Costs (list by type)				
Cloud computing (AWS) computing and storage			\$35,800.00	\$3,000.00 \$32,800.00 \$0.00
				0
				0
Total Other Capital Costs			\$35,800.00	\$3,000.00
VI. Capital Indirect Costs		Rate (%):		\$0.00 \$0.00 \$0.00 \$0.00
TOTAL Capital Project Expenditures, less Indirect Costs			\$250,041.86	\$50,028.20
TOTAL Capital Project Expenditures			\$250,041.86	\$50,028.20
Other Project Expenses				
These are not capitalizable under GAAP and are not reimbursable by MassTech. However, they can be covered by the applicant's required match funding (cash or in-kind). These costs include general and administrative tasks necessary for project execution but not directly tied to asset creation.				
I. Direct Labor		Hours	Rate	
name/title				
			\$0.00	\$0.00
			\$0.00	\$0.00
			\$0.00	\$0.00
			\$0.00	\$0.00
Total Direct Labor			\$0.00	\$0.00
II. Direct Labor Fringe Cost		Rate (%):		\$0.00 \$0.00
III. Subcontractors/Consultants		Firm name/title of consultant/type of services to be provided		
			\$0.00	\$0.00
			\$0.00	\$0.00
			\$0.00	\$0.00
			\$0.00	\$0.00
Total Subcontractors/Consultants			\$0.00	\$0.00
III. Direct Materials				
			\$0.00	\$0.00
			\$0.00	\$0.00
Total Direct Materials			\$0.00	\$0.00
IV. Travel				
V. Other Direct Costs (list by type)				
			\$0.00	\$0.00
			\$0.00	\$0.00
			\$0.00	\$0.00
Total Other Direct Costs			\$0.00	\$0.00
VI. General & Administrative Expense/Overhead		Rate (%):		\$0.00 \$0.00
Total Other Project Costs			\$0.00	\$0.00
Total PROJECT COSTS			\$250,041.86	\$50,028.20
Cost Share			25.01%	74.99%
Contractual Minimum Cost Share Met (Yes/No)			1	1
Additional detail in support of the budget should be provided on the Budget Assumptions tab				
Upon invoicing, additional detail in support of the match expenditures should be provided on the Supporting Schedule				



Boston Children's Hospital

Where the world comes for answers

Scott B. Snapper, MD PhD

Chief, Division Gastroenterology, Hepatology and Nutrition
Wolpolw Family Chair and Director, IBD Center
300 Longwood Avenue, Enders 664, Boston, MA 02115
scott.snapper@childrens.harvard.edu
617-919-4973 (office); 617-905-8867 (mobile); 617-730-0498 (FAX)

BRIGHAM HEALTH



BRIGHAM AND WOMEN'S HOSPITAL

Physician, Crohn's and Colitis Center
Division of Gastroenterology, Hepatology and Endoscopy



HARVARD MEDICAL SCHOOL TEACHING HOSPITAL

Egan Family Foundation Professor of Pediatrics
in the Field of Transitional Medicine
Professor of Medicine
Harvard Medical School

February 27, 2025

Dear Members of the Review Committee,

I am writing to express my strong support for the proposed expert-guided, multi-modal AI system for pediatric Inflammatory Bowel Disease (IBD) management. As Chief of the Division of Gastroenterology at Boston Children's Hospital and Director of our IBD Center, I see firsthand the challenges clinicians face in managing complex IBD cases. Our center cares for over 2,000 active IBD patients, with about 150 new diagnoses each year. Each case requires synthesizing diverse clinical data, including disease characteristics, treatment responses, and long-term outcomes, which is both time-consuming and complex.

Pediatric IBD poses unique challenges across medical, adherence, psychosocial, and economic domains. Medically, treatment regimens are complex, and optimizing dosing and monitoring disease activity can be difficult. Adherence is a major issue, with rates of nonadherence ranging from 50% to 88%, often due to developmental and psychosocial factors. The disease also significantly impacts patients' emotional well-being, quality of life, and transition to adult care. Economically, pediatric IBD is more costly to treat than adult IBD due to disease severity and the need for early, aggressive therapies. Balancing clinical effectiveness with cost considerations adds another layer of complexity. Addressing these challenges requires a multidisciplinary, patient-centered approach.

The AI system proposed under this project addresses a critical gap in IBD care by creating a framework to systematically analyze complex patient data and help standardize treatment decisions. I am particularly impressed by the project's focus on practical clinical implementation. The involvement of our IBD specialists, including Drs. Ouahed and Collen, and pediatric radiologists, Drs. Tsai and Lu, ensures the system will be grounded in real-world clinical needs. Their expertise in early-onset and treatment-resistant IBD will be invaluable in developing robust, clinically relevant models.

By leveraging advanced data analysis, this AI system has the potential to transform IBD care. It aims to provide personalized treatment plans, automate time-intensive disease severity assessments, and predict individual risks for complications, enabling earlier interventions. By analyzing EHR data from thousands of patients, the system could uncover patterns in disease progression, stratify patients into subgroups for targeted therapies, and streamline clinical workflows. Ultimately, this approach promises more precise, efficient, and patient-centered care, improving long-term outcomes and quality of life for those living with IBD.

This project aligns perfectly with our mission to advance personalized treatment approaches. I strongly support this proposal and believe it represents a significant step forward in using AI to improve IBD care.

Sincerely,



Scott B. Snapper, MD PHD
Chief, Division of Gastroenterology, Hepatology and Nutrition



Andy Tsai, MD, PhD

Associate Professor of Radiology, Harvard Medical School
300 Longwood Avenue, Boston, MA 02115
Email: andy.tsai@childrens.harvard.edu

Feb 24, 2025

Dear Members of the Review Committee,

I am writing to express my strong support for the proposed expert-guided multi-modal AI system for pediatric inflammatory bowel disease (IBD) management, submitted by Dr. Kurugol and her team to the Massachusetts AI Models Innovation Challenge.

IBD affects 1.6-3 million Americans, with approximately 10% of cases occurring in children and adolescents, making it a significant pediatric health concern. Half of these pediatric patients are believed to have Crohn's disease, a chronic and debilitating condition. Without precise diagnostic tools or imaging technologies capable of accurately assessing disease activity, severity, and treatment response, IBD often progresses over time. This frequently leads to surgical bowel resection, a procedure required by up to 75% of Crohn's patients. Unfortunately, surgery is rarely curative and can result in complications such as short bowel syndrome.

The impact of IBD extends far beyond physical symptoms. Pediatric patients often experience poor growth, anemia, infertility, and an increased risk of colon cancer later in life. The disease also disrupts their education, social development, and future employment. Many children with IBD miss significant school days, struggle to participate in sports and social activities, and face higher rates of depression, anxiety, and low self-esteem. These challenges often persist into adulthood, affecting their ability to maintain jobs and personal relationships.

Given the significant impact of IBD on patients, families, and society, I am particularly inspired by the project led by Dr. Kurugol. This initiative focuses on creating an AI system that integrates complex electronic health record (EHR) data from various sources to enhance IBD treatment decisions. The team plans to automate the extraction of clinically relevant information using large language models (LLMs) combined with expert input. They intend to fine-tune these LLMs to improve their ability to interpret diverse EHR data, including clinical notes, radiology reports, endoscopy and colonoscopy findings, and relevant lab results. With this refined AI model, the team aims to compute time-intensive disease severity scores, such as the Pediatric Crohn's Disease Activity Index (PCDAI), which helps classify the severity of Crohn's disease in pediatric patients. The data extracted from thousands of patients at multiple time points could be used to develop AI models that personalize treatment recommendations, estimate individual risk scores, and cluster patients into homogeneous groups based on disease markers. This clustering could help identify common genotypes or phenotypes, potentially offering new insights for disease management strategies.

By combining advanced AI techniques with the clinical expertise of Boston Children's Hospital, this project addresses a critical healthcare need and aligns with the Innovation Challenge's goals of fostering scientific discovery and responsible AI development.

As a staff radiologist at Boston Children's Hospital, I am especially impressed by the project's integrated approach to analyzing medical imaging alongside clinical data. The collaboration between the Quantitative Intelligent Imaging Lab (led by Dr. Kurugol), IBD specialists and the pediatric radiologists creates a strong multi-disciplinary team uniquely positioned to tackle this challenge. The involvement of two clinician-scientists from our nationally renowned IBD Center and two expert pediatric radiologists ensures that the AI system will be rigorously developed and evaluated.

Dr. Kurugol, the Principal Investigator, and Dr. Serge Vasylechko, a senior postdoctoral fellow and machine learning scientist, bring exceptional expertise in medical image analysis and AI algorithm development. Together, this team has the potential to create transformative tools that improve clinical management and outcomes for young patients with IBD.

The project's emphasis on explainable AI techniques and rigorous validation underscores its commitment to develop trustworthy, clinically deployable systems. Success in this endeavor would not only benefit pediatric IBD patients but also serve as a model for responsible AI implementation across Massachusetts' healthcare system.

I wholeheartedly endorse this proposal and believe it merits serious consideration for funding through the Mass Tech Collaborative AI Models Innovation Challenge. Please do not hesitate to contact me if you require additional information or have any questions.

Sincerely,

A handwritten signature in black ink, appearing to read "Andy Tsai".

Andy Tsai, MD, PhD
Associate Professor of Radiology | Harvard Medical School

Hi Serge & Sila,

It was great to meet you both. We at Amazon are very excited about the opportunity to be your partner and support you in this project with Amazon's Bedrock & Sagemaker platforms. Your proposed project on using LLMs to Crohn's disease management is an important and impactful project.

We have provided a budgetary estimate to provide you with the resources for inference and finetuning of a Llama 3 like model or similar, and on a cluster of 8x H100 or H200 GPUs or similar, together with HIPAA-eligible storage (with dual responsibility) and storage capacity of ~2Tb.

The budget estimate is around ~\$35,000 and we will be providing you with \$3,000 AWS Credits to support this project as demonstration of our partnership.

Best,

~Jim
508-944-6626

Jim Ray
Strategic Account Leader | Healthcare
Mobile: (508) 944-6626 | jimray@amazon.com
[AWS for Health](#)
[AWS Healthcare Marketplace](#)





HARVARD MEDICAL SCHOOL
TEACHING HOSPITAL

February 27, 2025

Dear Review Committee,

I would like to give my support for the proposed application titled “Expert-Guided Multi-Modal AI for Complex Pediatric Disease Management”, submitted by Dr. Kurugol and her team, inc. Dr Vasylechko, to the Massachusetts AI Models Innovation Challenge.

As Chief Innovation Officer at Boston Children's Hospital and the leader of the key AI initiatives across our institution, I am excited by this proposal's potential to transform pediatric IBD care. This project represents precisely the type of impactful clinically-grounded AI solution that our hospital aims to develop as part of our broader digital health strategy.

The burden of IBD on pediatric patients, their families, and our healthcare system is significant. The current approaches to IBD management are hindered by fragmented data and time-intensive analysis of this complex disease. The proposed AI system offers a compelling solution by automating the extraction and integration of clinically relevant information from diverse source.

Boston Children's Hospital is uniquely positioned to lead this initiative given our world-class status in pediatric expertise and unparalleled access to pediatric IBD data. The collaboration between the Quantitative Intelligent Imaging Lab, IBD specialists, pediatric radiologists, and Amazon AI team, creates a truly multidisciplinary team. At Boston Children's Hospital, our Innovation and Digital Health Accelerator has catalyzed numerous startups and collaborations with technology leaders including Google, Amazon, Apple, Microsoft, and OpenAI. We have witnessed firsthand how thoughtfully designed AI systems can improve clinical workflows and patient outcomes. This proposal represents an opportunity to extend this success to pediatric IBD care.

Sincerely,

John Brownstein, PhD
SVP and Chief Innovation Officer
Robert and Dana Smith Family Professor of Pediatrics, Boston Children's Hospital
Professor of Pediatrics and Biomedical Informatics, Harvard Medical School



Boston Children's Hospital

Where the world comes for answers



HARVARD MEDICAL SCHOOL
TEACHING HOSPITAL

Joan LaRovere, MD, MSc, MBA

Associate Chief Medical Officer, Boston Children's Hospital
Assistant Professor of Pediatrics, Harvard Medical School
Harvard Medical School
300 Longwood Avenue, Boston 02115
Email: joan.larovere@cardio.chboston.org

Feb 27, 2025

Dear Sir/Madam,

I am writing to express my strong support for the "Expert-Guided Multi-Modal AI for Complex Pediatric Disease Management" proposal by Dr. Kurugol and Dr. Vasylechko.

As the Associate Chief Medical Officer for Transformation at Boston Children's Hospital, I see tremendous value in this multi-modal AI approach for improvement of pediatric IBD care. The system's ability to extract and analyze complex disease markers from fragmented data sources addresses a critical clinical challenge.

Beyond IBD, I'm particularly excited about future applications to complex cardiac conditions. In my cardiovascular critical care work, I observe similar challenges with fragmented data that impede optimal treatment for congenital heart disease. The proposed AI framework could be adapted to transform management of many similar cardiac conditions.

Boston Children's Hospital's pediatric expertise, combined with Amazon's technical capabilities, provides an ideal environment for this innovative work. I fully support this proposal and its potential to advance care across multiple complex pediatric disease states.

Sincerely,

Joan LaRovere, MD, MSc, MBA

Associate Chief Medical Officer, Boston Children's Hospital
Assistant Professor of Pediatrics, Harvard Medical School



Bernardo Bizzo, MD, PhD

Senior Director, Mass General Brigham AI
Assistant Professor of Radiology, Harvard Medical School
Associate Chief Science Officer, ACR Data Science Institute
399 Revolution Dr., Sommerville, MA
Email: bbizzo@mgh.harvard.edu

Feb 28, 2025

Dear Committee,

I would like to express my strong support for the grant application with a title "Expert-Guided Multi-Modal AI for Complex Pediatric Disease Management"- a proposal submitted by Dr. Kurugol and Dr. Vasylechko.

As the Senior Clinical Director of the Mass General Brigham AI and a diagnostic radiologist, I recognize the significant potential of this proposal. The multi-modal approach to extracting and analyzing complex disease markers from disparate data sources addresses a critical need in pediatric care.

The proposal's emphasis on integrating radiological findings in Inflammatory Bowel Disease (IBD) — like wall thickening measurements, strictures, and fistulae—is particularly compelling. These imaging markers are essential for comprehensive IBD assessment but are often challenging to systematically track and analyze across patient populations.

Having led the development and validation of numerous clinical AI applications, I appreciate the rigorous methodology outlined in this proposal. This project has a potential to be translated into multiple other areas beyond IBD if successful.

The collaboration with Amazon provides the necessary computational infrastructure for this project, while BCH's extensive pediatric IBD dataset offers an unparalleled foundation for model training and validation.

I believe this work not only has tremendous potential to transform IBD care but could also establish a blueprint for AI-assisted management of other complex conditions requiring multi-modal data integration.

Sincerely,

Bernardo Bizzo, MD, PhD

Senior Director, Mass General Brigham AI
Assistant Professor of Radiology, Harvard Medical School



Marc Succi, MD

Founder and Executive Director, MESH Incubator

Associate Chair of Innovation and Commercialization, Mass General Brigham

Attending Radiologist, Emergency Dept, Mass General Brigham

Assistant Professor, Harvard Medical School

Email: msucci@mgh.harvard.edu

Feb 27, 2025

To Whom It May Concern:

I strongly support the grant application "Expert-Guided Multi-Modal AI for Complex Pediatric Disease Management" by Dr. Kurugol and Dr. Vasylechko.

Pediatric IBD patients in acute distress require rapid assessment of complex imaging findings alongside fragmented clinical histories. PCDAI score calculation in emergency settings would transform pediatric IBD patient management, enabling targeted interventions and improving clinical outcomes through real-time disease severity assessment.

The proposed AI system addresses this critical gap by integrating radiological findings (bowel wall thickening, strictures, fistulae) with clinical data. This innovation could significantly reduce emergency department admissions for IBD flares through earlier intervention and more precise disease monitoring.

From an innovation standpoint, automated PCDAI scoring may represent a market-ready solution to a persistent clinical challenge. By making comprehensive IBD assessment feasible at scale, this technology could transform emergency care pathways while generating valuable population-level insights.

Sincerely,

A handwritten signature in black ink, appearing to read "Marc Succi, MD".

Marc Succi, MD

BIOGRAPHICAL SKETCH

NAME: Sila Kurugol

eRA COMMONS USER NAME (credential, e.g., agency login): s_kurugol

POSITION TITLE: Assistant Professor in Radiology, Boston Children's Hospital and Harvard Medical School

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE	MM/YY	FIELD OF STUDY
Middle East Technical University, Ankara, Turkey	B.S.	05/03	Electrical and Electronics Engineering
Bilkent University, Ankara, Turkey	M.S.	09/06	Electrical and Electronics Engineering
Northeastern University, Boston, MA	Ph.D	08/11	Electrical and Computer Engineering
Brigham and Women's Hospital, Harvard Medical School	Postdoctoral	09/11-02/14	Medical Imaging/ Machine Learning
Boston Children's Hospital and Harvard Medical School	Postdoctoral	02/14-06/16	Medical Imaging

A. Personal Statement

My research has primarily focused on the development of advanced methods for quantitative imaging, image analysis, and machine learning to generate reliable markers from medical images, with a strong emphasis on translating these innovations into clinical practice. I hold a PhD focused on pioneering machine learning techniques for image analysis, and I have 14 years of research experience after PhD in leading and contributing to projects aimed at enhancing image acquisition and analysis through computational methods and machine learning techniques.

I have substantial experience as the Principal Investigator (PI) for several successful career development and foundation grants, as well as prior NIH grants (R01 and R21s). Our research findings have been recognized with prestigious awards, including the 2018 Caffey Award for Best Scientific Paper from the Society of Pediatric Radiology and several accolades from International Society of Magnetic Resonance in Medicine (ISMRM) Meetings between 2019-2024, including several summa Cum Laude Awards. Our work was also highlighted at conferences in 2020 and 2023 and our deep learning-based image segmentation work was honored as the best paper at an IEEE conference. Our research has been published in esteemed journals such as Medical Image Analysis, IEEE Transactions on Medical Imaging, and Magnetic Resonance in Medicine. In addition to my research, I contribute to several NIH review committees and play roles in the technical organization committees of various machine learning and imaging conferences, including MIDL and ISMRM. I also have been invited to

give several educational, scientific and key notes talks and to organize educational sessions and workshops for ISMRM, MICCAI and other prestigious societies in my field.

In summary, my experience from my current and previous projects involving medical image computing, machine learning and mathematical modeling is invaluable and highly significant project. Supported by an expert, interdisciplinary team, I believe we will also quickly translate the proposed novel AI techniques into clinical practice.

Ongoing and recently completed projects that I would like to highlight include:

Machine learning for fast motion compensated quantitative abdominal DCE-MRI

NIH R21EB029627-01

09/15/2020-09/14/2024 (NCE)

This application aims to develop improved retrospective motion correction methods for mitigating pediatric subject motion in dynamic contrast enhanced (DCE) MRI. The successful completion of these aims will also provide new, clinically important abdominal imaging capabilities, with real-time, motion-compensated image reconstruction and reliable real-time parameter estimation.

Role: Principal Investigator

Robust quantitative MR imaging markers of response to therapy in Crohn's Disease

NIH R01 DK125561

04/01/2021-12/31/2025

This project is aimed at developing and evaluating noninvasive, contrast and radiation-free quantitative imaging markers for assessing disease activity and for monitoring response to therapy in Crohn's disease. We aim to reduce the imaging time with estimated x4 acceleration with an accelerated image acquisition and a new, advanced parameter estimation technique. Another important goal is to develop and broadly disseminate open source software that will enable the standardized evaluation of quantitative markers for other diseases presently evaluated with DW-MRI that would benefit from the advanced diagnostic and assessment capabilities of the proposed technique.

Role: Principal Investigator

Relevant citations:

1. S. Wu, **S. Kurugol**, A. Tsai. Improving the Radiographic Image Analysis of the Classic Metaphyseal Lesion via Conditional Diffusion Models. Medical Image Analysis. 2024 Jul 25:103284.
2. S. Vasylechko, O. Afacan, **S. Kurugol**, Self-Supervised Denoising Diffusion Probabilistic Models for Abdominal DW-MRI, Proceedings of MICCAI Workshop on Computational Diffusion MRI (CDMRI), 2023, Oct. 8.
3. S.D. Vasylechko, S.K. Warfield, O. Afacan, **S. Kurugol**. Self-supervised IVIM DWI parameter estimation with a physics based forward model. Magnetic Resonance in Medicine. 2021. PMCID: PMC8627432
4. S.D. Vasylechko, S.K. Warfield, **S. Kurugol**, O. Afacan Improved myelin water fraction mapping with deep neural networks using synthetically generated 3D data, Medical Image Analysis, 2023. *contributed equally as last authors. Wu, S. **, Kurugol, S. and Tsai, A., 2024. Improving the radiographic image analysis of the classic metaphyseal lesion via conditional diffusion models. Medical Image Analysis, 103284. IF: 10.7.
5. Wu, S., **Kurugol**, S. Kleinman, P.K., Ecklund, K., Walters, M., Connolly, S.A. and Tsai, A., 2024. Deep generative model of the distal tibial classic metaphyseal lesion: radiologists' recognition and diagnostic performance of synthetic images. Radiology Advances, 1(2), July.
6. Vasylechko, S.D., Tsai, A., Afacan, O. Kurugol, S., 2025. Self-Supervised Denoising Diffusion Probabilistic Models for Abdominal DW-MRI. Magnetic Resonance in Medicine. In Press. IF: 3.74.

B. Positions and Honors

Positions and Employment

- 2021- Assistant Prof. of Radiology at Boston Children's Hospital and Harvard Medical School, Boston, MA
2016-2021 Instructor of Radiology at Boston Children's Hospital and Harvard Medical School, Boston, MA.
2014-2016 Research Fellow at Boston Children's Hospital and Harvard Medical School, Boston, MA.
2011-2014 Research Fellow at Brigham and Women's Hospital Radiology Department, Boston, MA.
2006-2011 Research and Teaching Assistant at Northeastern University, Boston, MA.

2003-2005 Research and Teaching Assistant at Bilkent University, Ankara, Turkey.

Other Experience and Professional Memberships

Reviewer, IEEE Transactions on Medical Imaging, Transaction on Image Processing, Magnetic Resonance Imaging in Medicine (MRM), Medical Physics, IEEE International Symposium on Biomedical Imaging (ISBI), Information Processing in Computer Assisted Interventions, Journal of Magnetic Resonance Imaging, IEEE Signal Processing Letters, Magnetic Resonance Materials in Physics, Medicine and Biology, International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI), International Society for Magnetic Resonance in Medicine Conference (ISMRM), Crohns and Colitis Foundation UK.

Member, International Society of Magnetic Resonance in Medicine (ISMRM), the Institute of Electrical and Electronics Engineers (IEEE), Society of Medical Image Computing and Computer Assisted Interventions (MICCAI), Society of Pediatric Radiology (SPR), AGA.

Area Chair, International Conference on Medical Imaging with Deep Learning (MIDL) 2019, 2022.

Program Committee Member, Annual ISMRM Conference 2022

Study Section Member, National Institute of Health (NIH) Grants 2020-2021

Reviewer, NIH Grants 2020-2023

Honors

2021	IEEE Int. Symposium on Computer-Based Medical Systems (CBMS) Best paper award winner
2019-2020	Recipient of ISMRM Summa Cum Laude Merit Award
2018	Society of Pediatric Radiology Caffey Award for Best Scientific Paper: Feed and Wrap MRU
2015	Recipient of ISMRM Magna Cum Laude Merit Award
2015	ISMRM Diffusion Study Group Award, Selected Among Best Papers on Diffusion MRI in ISMRM
2015	Award and invitation to present in ISMRM Workshop on <i>Diffusion-Weighted Imaging Outside the Brain: From Physics to Clinical Practice</i>
2015	ISMRM travel award
2012	ISBI travel award
2010	GPSA at Northeastern University travel award,
2010	CDSP Research Workshop Recognition Award
2008	ISBI travel award (for top student papers submitted to ISBI'08)
2006-2011	Awarded stipended graduate research assistantship by Northeastern University
2003-2005	Awarded with scholarship for graduate study, Bilkent University.
1999	Ranked 83rd among 1.5 million high school seniors in college entrance exam

C. Contribution to Science (Peer Reviewed Publications)

1. Quantitative MRI markers: Diffusion-weighted MRI has been increasingly used for the detection and characterization of abdominal abnormalities in liver, spleen and bowel. Challenges in abdominal DW-MRI include low SNR and respiratory, cardiac and peristalsis motion. One of my major contributions to the field has been development of novel computational models for diffusion weighted MRI of abdomen that increased the robustness and reproducibility of parameter estimation. These methods will potentially improve the clinical utility of quantitative DW-MRI parameters for assessment of Inflammation and fibrosis in Crohn's disease of the bowel. Recently we also developed new denoising models for accelerating the DW-MRI image acquisition and also improving the accuracy of quantitative markers.

- a. Y. Lamash, **S. Kurugol**, J. Perez-Rossello, M. Callahan, A. Bousvaros, M., Freiman, S.K. Warfield. Curved Planar Reformatting and CNN-Based Segmentation of the Small Bowel for Visualization and Quantitative Assessment of Pediatric Crohn's Disease from MRI. *Journal of Magnetic Resonance Imaging*. DOI:10.1002/jmri.26330. 2018. PMCID: PMC7205020
- b. **S. Kurugol**, M. Freiman, L. Domachevsky, O. Afacan, J.M. Perez-Rossello, M.J. Callahan S.K. Warfield, "Spatially-Constrained Probability Distribution Model of Incoherent Motion (SPIM) for Abdominal Diffusion-weighted MRI", *Medical Image Analysis*, doi:10.1016/j.media.2016.03.009, 2016. PMCID: PMC4903917
- c. S. Vasylechko, **S. Kurugol***, O. Afacan* Improved myelin water fraction mapping with deep neural networks using synthetically generated 3D data, *Medical Image Analysis*, 2023 *contributed equally.

- d. S. Vasylechko, O. Afacan, **S. Kurugol**, Self-Supervised Denoising Diffusion Probabilistic Models for Abdominal DW-MRI, Proceedings of CDMRI: MICCAI Workshop on Computational Diffusion MRI, 2023, Oct. 8.

2. Automated image analysis of CT and radiography images: Earlier work involved development of a segmentation algorithm for 3D/4D segmentation of structures in thoracic CT scans for radiation therapy planning. This image segmentation algorithm incorporated local deformations into a global parametric shape model to be used within a 3D level set algorithm to improve accuracy of esophagus segmentation in thoracic CT scans. Another project was quantification of cardiovascular disease phenotypes using image-based markers, which required automated image analysis of large cohorts of volumetric CT images. I developed automated aorta segmentation, vessel morphology quantification and mural calcification detection software that was applied on a large cohort of CT images from 2000 smokers to study associations between aortic morphology changes (like stiffening, unwrapping and dilation), calcification and cardiovascular disease events in smokers. Recently we worked on automated classification of radiography images of fracture related to child abuse from radiography images. We developed mask conditional generative models for synthesizing realistic images to improve accuracy and generalizability of our automated fracture detection approach.

- a. **S. Kurugol**, CE Come, A.A. Diaz, J.C. Ross, G.L. Kinney, J.L. Black-Shinn, JE Hokanson, MJ Budoff, G.R. Washko, R. San Jose Estepar, "Automated quantitative 3D analysis of aorta size, morphology, and mural calcification distributions." *Medical Physics*. Sep 2015, 42(9), 5467-78. PMCID: PMC4552704
- b. **S. Kurugol**, E. Bas, D. Erdogmus, J. Dy, G. Sharp, D. H. Brooks, "Centerline extraction with principal curve tracing to improve 3D level set esophagus segmentation in CT images", Proc. of the IEEE Eng. in Medicine and Biology Society (EMBC), August 2011. PMCID: PMC3349355
- c. **S. Kurugol**, N. Ozay, J. Dy, G. Sharp, D. H. Brooks, "Locally Deformable Shape Model to Improve 3D Level Set based Esophagus Segmentation", Proc. of International Conf. on Pattern Recognition (ICPR'10), August 2010. PMCID: PMC3127393
- d. S. Wu, **S. Kurugol**, A. Tsai. Masked Conditional Diffusion Models for Image Analysis with Application to Radiographic Diagnosis of Infant Abuse. Proceedings of DALI: The 3rd MICCAI Workshop on Data Augmentation, Labeling, and Imperfections, 2023, Oct. 12.

3. Machine learning in medical imaging: I have been leading several projects on development of deep learning models and active learning and semi-supervised learning techniques for medical image analysis problems. We have achieved outstanding image segmentation within seconds; and our methods provide accurate quantitative markers of disease to improve clinical decisions.

- a. J. Sourati, A. Gholipour, J.G. Dy, X Tomas-Fernandez, **S. Kurugol**, S.K. Warfield. Intelligent Labeling Based on Fisher Information for Medical Image Segmentation Using Deep Learning. *IEEE transactions on medical imaging*. 2019. PMCID: PMC7179938
- b. A. Mortazi, N. Khosravan, D.A. Torigian, **S. Kurugol**, U. Bagci, (2019). Weakly Supervised Segmentation by A Deep Geodesic Prior. *Lecture Notes in Computer Science (MLMI)*. Springer, 2019.
- c. M. Haghghi, S. K. Warfield, and **S. Kurugol**, "Automatic Renal Segmentation in DCE-MRI using Convolutional Neural Networks," in Proc. of IEEE Int. Symp. on Biomedical Imaging (ISBI) 2018. PMCID: PMC6248325
- d. A. Koçanaoğulları, C. Ariyurek, O. Afacan and **S. Kurugol**. Learning the Regularization in DCE-MR Image Reconstruction for Functional Imaging of Kidneys. *IEEE Access*, vol. 10, pp. 4102-4111, 2022, doi: 10.1109/ACCESS.2021.3139854. PMCID: PMC9348606

4. Motion and distortion correction techniques: We developed a 3D motion tracking and correction method for motion-robust diffusion weighted MRI (DW-MRI) of kidneys. We also developed a motion-compensated DW-MRI technique for quantitative imaging of Crohn's disease. These motion-robust imaging techniques provide the clinicians with accurate and reproducible quantitative imaging markers that will allow an objective assessment of disease, potentially leading to optimized clinical decisions.

- a. C. Ariyurek, A. Kocanaoğulları, O. Afacan, **S. Kurugol**. Motion-Compensated Image Reconstruction for Improved Kidney Function Assessment Using Dynamic Contrast-Enhanced MRI. *NMR in Biomedicine*, 2024.
- b. **S. Kurugol**, B. Marami, O. Afacan, S.K. Warfield, A. Gholipour, "Motion-Robust Spatially Constrained Parameter Estimation in Renal Diffusion-Weighted MRI by 3D Motion Tracking and

Correction of Sequential Slices”, In: *Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment*. RAMBO 2017. Lecture Notes in Computer Science, vol 10555. Springer. PMCID: PMC5810407

- c. **S. Kurugol**, M. Freiman, L. Domachevsky, O. Afacan, J.M. Perez-Rossello, M.J. Callahan S.K. Warfield, “Motion-robust parameter estimation in abdominal diffusion-weighted MRI by simultaneous image registration and model estimation”, *Medical Image Analysis*, 39, 124-132, 2017. PMCID: PMC5514879
- d. **S. Kurugol**, M. Freiman, L. Domachevsky, O. Afacan, J.M. Perez-Rossello, M.J. Callahan S.K. Warfield, “Motion compensated abdominal diffusion weighted MRI by simultaneous image registration and model estimation (SIR-ME)” In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (pp. 501-509). Springer, Cham. 2016. PMCID: PMC4636124

5. Motion-robust accelerated DCE-MRI: One of my recent work entails development of novel motion-robust MR imaging techniques and deep learning tools for non-sedated, radiation-free imaging of abdomen. We developed motion-robust, high spatiotemporal resolution dynamic contrast-enhanced MRI (DCE-MRI) technique and new automated analysis software to enable the robust and reproducible assessment of quantitative markers of perfusion and permeability in children. As a non-invasive, motion-robust, radiation-free imaging technique that can depict anatomical structures at much higher resolution, DCE-MRI offers, for the first time, diagnostic information that is unattainable with previous techniques. This technique will substantially improve clinical decision-making in various diseases of abdomen including kidneys, bowel and liver.

- a. J. Coll-Font, O. Afacan, J.S. Chow, S.K. Warfield, and **S. Kurugol**, “Modeling Dynamic Radial Contrast Enhanced MRI with Linear Time Invariant Systems for Motion Correction in Quantitative Assessment of Kidney Function, *Medical Image Analysis*, 2020. PMCID: PMC7735437
- b. J. Coll-Font, O. Afacan, J.S. Chow, R.S. Lee, A. Stemmer, S.K. Warfield, and **S. Kurugol**, “Bulk motion-compensated DCE-MRI for functional imaging of kidneys in newborns, *Journal of Magnetic Resonance Imaging*, 2019. PMCID: PMC7293568
- c. **S. Kurugol**, C.M. Seager, H. Thaker, J. Coll-Font, O. Afacan, R.C. Nichols, S.K. Warfield, R.S. Lee, and J. Chow, Feed and Wrap Magnetic Resonance Urography Provides Anatomic and Functional Imaging in Infants without Anesthesia. *Journal of Pediatric Urology*, 2019. PMID: 31889687
- d. **S. Kurugol**, O. Afacan, R. S. Lee, C. Seager, M. A. Ferguson, D. Stein, R. Nichols, M. Dugan, A. Stemmer, S.K. Warfield, J. S. Chow, “Prospective pediatric study comparing glomerular filtration rate estimates based on motion robust dynamic contrast enhanced MRI and serum creatinine (eGFR) to ^{99m}Tc DTPA,” *Journal of Pediatric Radiology*, 2020. PMCID: PMC7153988

- **Complete List of Published Work in My Bibliography:**

<https://www.ncbi.nlm.nih.gov/myncbi/sila.kurugol.1/bibliography/public/>

BIOGRAPHICAL SKETCH

NAME: Serge Vasylechko

eRA COMMONS USER NAME (credential, e.g., agency login): vasylechko

POSITION TITLE: Postdoctoral Fellow in Radiology, Boston Children's Hospital and Harvard Medical School

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE	MM/YY	FIELD OF STUDY
Imperial College of Science, Technology and Medicine, London, United Kingdom	M.Eng.	07/11	Biomedical and Electrical Engineering
Imperial College of Science, Technology and Medicine, London, United Kingdom	Ph.D.	08/20	Computer Science
Harvard Medical School / Boston Children's Hospital	Postdoctoral	06/20 - ongoing	Medical Imaging

A. Personal Statement

Dr. Vasylechko is a senior postdoctoral research fellow at Boston Children's Hospital and Harvard Medical School. He has acquired substantial knowledge on the AI applications for medical data during his PhD at Imperial College London. During the latter parts of his PhD, he had developed unsupervised deep learning methods for image reconstruction as well as frameworks that would combine multi-contrast sources. In the last 5 years he has been working with Dr. Kurugol to develop machine learning techniques for quantitative medical imaging. His recent work includes using vision language models and large language models for extracting information from radiology reports to train an object detection algorithm for disease detection and localization. He has an extended track record of publications in top AI medical journals.

B. Positions and Honors

Positions and Employment

2012-2014 Research and Teaching Assistant at Imperial College London
2014-2018 Managing Director at Blocks Wearables Limited (R&D in Bio-Electronics Engineering)
2019-2020 Research Assistant in Radiology at Boston Children's Hospital and Harvard Medical School
2020- Postdoctoral Fellow in Radiology at Boston Children's Hospital and Harvard Medical School

C. Contribution to Science (Peer Reviewed Publications)

1. Using generative priors for image reconstruction and data denoising: Quantitative diffusion weighted MRI in the abdomen provides important markers of disease, however significant limitations exist for its accurate computation. One such limitation is the low signal-to-noise ratio, particularly at high diffusion b-values. To address this, multiple diffusion directional images can be collected at each b-value and geometrically averaged, which invariably leads to longer scan time, blurring due to motion and other artifacts. We investigated a novel parameter estimation technique based on self supervised diffusion denoising probabilistic model that can effectively denoise diffusion weighted images and work on single diffusion image. The ssDDPM demonstrated superior performance over comparison methods in terms of image

quality, lesion conspicuity, and apparent diffusion map accuracy. It received higher scores in radiologist assessments and showed better lesion discrimination in ROC analysis.

(Under Review) **MRM 2024** Vasylechko, S., Tsai, A., Afacan, O., and Kurugol, S., 2024. Self Supervised Denoising Diffusion Probabilistic Models for Abdominal DW-MRI. *Magnetic Resonance in Medicine*.

MICCAI CDMRI 2023 Vasylechko, S., Afacan, O., and Kurugol, S., 2023. Self-Supervised Denoising Diffusion Probabilistic Models for Abdominal DW-MRI. In International Workshop on Computational Diffusion MRI, pp. 80-91. Cham: Springer Nature Switzerland

2. A generative model for synthesis of 3D datasets for quantitative MRI parameter mapping of myelin water fraction: Training robust neural networks for estimation of quantitative MRI parameters requires large amounts of data. Conventional approaches to tackling data scarcity use spatial augmentations, which may not capture a broad range of possible variations when only a very small initial dataset is available. Furthermore, conventional non linear least squares (NNLS) based methods are highly sensitive to noise, which means that high quality ground truth MWF parameters are not available for supervised training. We propose to leverage the biophysical model that describes how the MRI signals arise from the underlying tissue parameters to synthetically generate a wide variety of high quality data of the corresponding signals and corresponding parameters for training any CNN based architecture. We demonstrated that our synthetically trained neural network provides superior accuracy over conventional NNLS based methods under the constraints of naturally occurring noise as well as on synthetic low SNR images.

ISMRM 2021 **Vasylechko, S.D.**, Warfield, S.K., Kurugol, S. and Afacan, O., 2021. Synthesizing large scale datasets for training deep neural networks in quantitative mapping of myelin water fraction. In Proceedings of Int. Soc. of Magnetic Resonance Imaging.

MIDL 2022 **Vasylechko, S.D.**, Warfield, S.K., Kurugol, S. and Afacan, O., 2022. SynthMap: a generative model for synthesis of 3D datasets for quantitative MRI parameter mapping of myelin water fraction. International Conference on Medical Imaging with Deep Learning. PMLR.

MEDIA 2023 **Vasylechko, S.D.**, Warfield, S.K., Kurugol, S. and Afacan, O., 2023. Improved myelin water fraction mapping with deep neural networks using synthetically generated 3D data. *Medical Image Analysis*.

3. Self-supervised IVIM DWI parameter estimation with a physics based forward model: The goal of this study was to assess the robustness and repeatability of intravoxel incoherent motion model (IVIM) parameter estimation for the diffusion weighted MRI in the abdominal organs under the constraints of noisy diffusion signal using a novel neural network training method. The method is based on the principle of a physics guided self-supervised neural network that does not require supervision for training. Such approach is beneficial in conditions where the reference methods are not available, or are not robust enough to provide good supervision. This work is targeting evaluations towards accelerated IVIM DWI scanning which exhibit low SNR.

MRM 2022 **Vasylechko, S.D.**, Warfield, S.K., Afacan, O. and Kurugol, S., 2022. Self-supervised IVIM DWI parameter estimation with a physics based forward model. *Magnetic Resonance in Medicine*, 87(2), pp.904-914. PMCID: PMC8627432

ISMRM 2021 **Vasylechko, S.D.**, Warfield, S.K., Afacan, O. and Kurugol, S., 2021. Self-supervised IVIM DWI parameter estimation with a physics based forward model. In Proceedings of Int. Soc. of Magnetic Resonance Imaging.

4. Machine learning in medical imaging: Ultra fast MRI acquisition, such as single shot 2D echo planar imaging, is routinely used for acquisition of 2D slices of the moving fetal anatomy, which is free from motion

artefacts. Further processing is required to view this data as an anatomical 3D volume, which bears significant clinical benefit. Existing 3D reconstruction techniques require high resolution 2D images that have densely sampled the anatomy of interest from multiple orthogonal views, which is often difficult to achieve. The goal of this study was to investigate reconstruction of 3D brain images where acquisition of multiple orthogonal slice stacks is prohibitive due to high sensitivity to motion or tight timing constraints. To achieve this, we built a conditional generative adversarial network that can synthesize a prior for a chosen target contrast to serve as a seed for Slice to Volume Reconstruction (SVR) algorithms. As a result, individually scattered slices from a sparsely sampled image space could be combined with the synthetic volume of the same contrast to reconstruct a fully sampled 3D image. In the second part of the project, we built a fully automatic segmentation and reconstruction framework that can be used to propagate segmentation labels from a densely reconstructed high resolution prior to low resolution 2D slices in a different MRI contrast. The project's publications are ongoing and the first part has been published in ISMRM 2020.

ISMRM 2020 **Vasylechko, S.**, Hughes, E., Allsop, J., Fox, M., Rueckert, D., and Hajnal, J. V., 2020. Fetal and neonatal whole brain T2* mapping at 3T. In Proceedings of Int. Soc. of Magnetic Resonance Imaging.

5. **Motion robust acquisition of quantitative T2* maps in the developing brain:** Brain maturation involves complex structural and physiological changes that give rise to varying levels of T2* relaxation time. Knowledge of T2* values can thus help with evaluation of pathology by establishing its normative values. T2* values are a valuable biomarker for myelin microstructure and iron concentration, as well as an important tool towards achievement of optimal fMRI contrast. Significant challenge arises in neonatal and fetal MR imaging due to unpredictable and uncontrollable patient motion. In this study, S.Vasylechko had investigated approaches that would permit fast motion robust acquisition of fetal and neonatal quantitative T2* maps. The study had concluded with a novel method for ultra-fast imaging of the developing brain and the follow up clinical investigation had concluded that fetal T2* values are significantly higher than those previously reported in pre-term neonates, and decline with a consistent trend across gestational age. The investigations of this work had contributed towards the development of gradient echo based sequences for the large multi-university multi-year project of the Developing Human Connectome, funded by the European Research Council.

MRM 2015 **Vasylechko, S.**, Malamateniou, C., Nunes, R. G., Fox, M., Allsop, J., Rutherford, M., Rueckert, D., and Hajnal, J. V., 2015. T2* relaxometry of fetal brain at 1.5 Tesla using a motion tolerant method. In Magnetic Resonance in Medicine 73.5, pp. 1795–1802.

ISMRM 2013 **Vasylechko, S.**, Malamateniou, C., Nunes, R. G., Fox, M., Allsop, J., Rutherford, M., Rueckert, D., and Hajnal, J. V., 2013. T2* Measurement of Fetal Brain Using a Motion Tolerant Method. In Proceedings of Int. Soc. of Magnetic Resonance Imaging.

NeuroImage 2016 Ferrazzi, G., Nunes, R.G., Arichi, T., Gaspar, A.S., Barone, G., Allievi, A., **Vasylechko, S.**, Abaei, M., Hughes, E., Rueckert, D. and Price, A.N., 2016. An exploration of task based fMRI in neonates using echo-shifting to allow acquisition at longer TE without loss of temporal efficiency. In Neuroimage, 127, pp.298-306.

6. **Development of standardized engineering tools in Synthetic Biology:** To accelerate the field of synthetic genomics and parts-based synthetic biology there exists a need to develop well-characterized biological parts suitable for precise engineering on a chassis. In this project, we presented a design strategy that would build parts of new organisms on a base of standardized promoters in yeast to help to better characterize uncertainty and variability in gene expression. A highly constitutive *Saccharomyces cerevisiae* promoter, PFY1p, was identified by bioinformatic approaches, characterised in vivo and diversified at its core sequence to create a 36-member promoter library. The ability to diversify a promoter at its core sequences and then independently target Transcription Activator-Like Orthogonal Repressors (TALORs) showed great promise toward the design and construction of future synthetic gene networks that encode

complex "multi-wire" logic functions. The project was initiated and led by S. Vasylechko as part of his Masters thesis, and extended thereafter by B. Blount with further experiments.

PLoS ONE 2012 Blount, B.A., Weenink, T., **Vasylechko S.** and Ellis, T., 2012. Rational Diversification of a Promoter Providing Fine-Tuned Expression and Orthogonal Regulation for Synthetic Biology. In PLoS one, 7(3), p.e33279.

D. Research Support

Ongoing Research Support

Machine learning for fast motion compensated quantitative abdominal DCE-MRI

NIH R21EB029627-01

09/15/2020

This application aims to develop improved methods for mitigating pediatric subject motion in dynamic contrast enhanced (DCE) MRI. The successful completion of these aims will also provide new, clinically important abdominal imaging capabilities, with real-time, motion-compensated image reconstruction and reliable real-time parameter estimation.

Role: Research fellow

2019 SPR Multi-institutional Pilot Award

7/1/2019- 12/31/2021(NCE)

The goal of this one-year study is to evaluate a new non-sedated DCE-MRI based GFR measurement technique in three children hospitals located in USA.

Role: Research fellow