Pediatric & Lifespan

Data Science Conference

April 29–30, 2026

Artificial Intelligence

AI Agents

Clinical Care

Social Needs

Patient Voices

CHOC Research

GO BEYOND

Letter From *the* CEO

2026 marks a pivotal moment at Rady Children's Health as we continue into the second year of our parent organization's merger. It is also a year in which Artificial Intelligence (AI) and Agentic AI are poised to reshape society and accelerate a new era of medical and research innovation. These technologies are transforming how we ask research questions, generate evidence, and translate discoveries into improved clinical care. Yet even as innovation advances rapidly, our guiding principle remains constant: it starts with one child. Every study, dataset, and insight originates from a single patient whose experience inspires curiosity and drives meaningful contributions to scientific literature.

In this same spirit, we expect this conference to address the needs of patients in pediatrics and across the lifespan in ways that create direct impact for families and communities. Accordingly, this year's meeting focuses on how research in AI and Agentic AI may enhance clinical care, address social needs, and elevate the lived experiences of patients and families. Together, we will explore how intelligent systems can support clinicians, strengthen care pathways, and expand our understanding of the factors that shape health both inside and beyond the hospital.

This commitment to impact is reflected not only in our research and discussions but also in our collective action. Through collaboration across multiple teams and organizations, volunteers will pack 10,000 shelf-stable meals for distribution throughout Orange County and within Rady Children's clinics and hospitals. This effort demonstrates that the energy and engagement of this conference extend beyond innovation dialogue to supporting patients and families in real and immediate ways.

As you engage in presentations and conversations, remember that behind every research question is a child who challenges us to turn ideas into answers and hope into reality.

Kimberly Chavalas Cripe
Co-President & Chief Executive Officer
Rady Children's Health

Letter From the CSO

Welcome to the third annual Pediatric and Lifespan Data Science Conference, hosted by the CHOC Research Institute, now part of Rady Children's Health in Orange County. Each year, this event brings together clinicians, scientists, engineers, and community partners who share a deep commitment to improving the health and wellbeing of people from birth through adulthood.

Artificial intelligence continues to advance at a remarkable pace and is reshaping what is possible in medicine and in society. As these technologies evolve, we seek an understanding of how they can be used to improve the lives of children and the adults they will become. Just as a sequence of words defines the function and power of large language models, the sequence of events in our lives defines our function, health, and success. Time is a fundamental currency of AI, and time is the fundamental driving force of human growth, development, and progress. Therefore we are now presented with a tremendous opportunity to leverage this new technology to redefine healthcare throughout the lifespan, starting with the youngest patients in our care, through the adults that will follow.

To do this, we rely on rigorous research and responsible experimentation to ensure safe, effective, and equitable decision-making. Patients, families, and communities must remain at the center of every application. The future of healthcare will be shaped not only by the capabilities of these new tools but also by how well we integrate them into clinical workflows, decision-making, and real-world care. This progress depends on high quality evidence and strong national and global partnerships.

This year's conference highlights several important areas of research innovation. Generative AI is changing how we interpret and act on unstructured clinical text, and AI can now be used to interpret and act upon the data that underly all evidence-based healthcare. Vision language models are creating new opportunities in radiology and other imaging focused specialties by linking what clinicians see with what machine learning systems can quantify. The rapid rise of AI agents prompts us to study how human-in-the-loop intelligent systems can support the cognitive and creative work of physicians and care teams. In all of these areas, multimodal approaches utilizing data from multiple real-world sources are becoming increasingly important and are driving the need for more robust and insightful algorithms.

Our goal for this gathering has always been real world impact. This is aligned with our bedside-to -bedside research philosophy, in which insights begin at the bedside and return to the bedside with measurable impact. In that spirit, we invite you to take part in conversations that support learning, collaboration, and progress in both clinical care and the social factors that shape health across the lifespan.

Thank you for being part of this vibrant and dedicated community.

Terence Sanger, MD, PhD

Vice President, Chief Scientific Officer, Rady Children's Health Orange County
Professor of Electrical Engineering & Computer Science, *UCI*
Dept. Pediatrics Vice Chair for Research, *UCI School of Medicine*
Child Neurology & Movement Disorders, *Rady Children's Health Orange County*

Contents.

KEYNOTES

Annie Qu, PhD

Professor
Department of Statistics and Applied Probability
University of California, Santa Barbara

Dr. Annie Qu is a distinguished statistician and leading researcher in the fields of statistical learning, machine learning, and data science. She joined the Department of Statistics and Applied Probability at the University of California, Santa Barbara (UCSB) in 2025, bringing a profound record of scholarship and academic leadership.

Dr. Qu received her Ph.D. in Statistics from the Pennsylvania State University in 1998, and her research focuses on developing rigorous statistical methods for analyzing large-scale, unstructured, and heterogeneous data — including work in text analysis, recommender systems, network data, tensor imaging, and high-dimensional inference. Her innovations have wide-ranging impacts across biomedical studies, genomic research, public health, and the social sciences.

Before joining UCSB, Dr. Qu held prestigious appointments including Chancellor's Professor at the University of California, Irvine, and Data Science Founder Professor of Statistics and Director of the Illinois Statistics Office at the University of Illinois at Urbana-Champaign. A recipient of numerous honors such as the NSF CAREER Award, she is a Fellow of the Institute of Mathematical Statistics, the American Statistical Association, and the American Association for the Advancement of Science.

Dr. Qu's current research spans cutting-edge topics such as data integration for heterogeneous sources, causal inference, active learning, deep learning from a statistical perspective, and privacy-aware data analysis. She also serves as Co-Editor of the Journal of the American Statistical Association (Theory and Methods).

Katherine P. Andriole, PhD, FACR, FSIIM, FSPIE

Associate Dean for Health Artificial Intelligence Strategy and Innovation, UCLA David Geffen School of Medicine
Director, UCLA Center for AI and SMART Health
Professor, Radiological Sciences, UCLA David Geffen School of Medicine

Dr. Katherine P. Andriole is a distinguished leader in medical imaging informatics and clinical data science with more than three decades of experience at the nexus of engineering, medicine, and artificial intelligence. She is the newly appointed inaugural Associate Dean for Health Artificial Intelligence Strategy and Innovation at the David Geffen School of Medicine and Director for the UCLA Center for AI and SMART Health.

Dr. Andriole earned her degrees in Biomedical Engineering from Duke University and in Electrical Engineering and Medicine at Yale University, and completed postdoctoral fellowships in radiology at both UCLA and UCSF, where she played a pivotal role in designing and deploying early Picture Archiving and Communication Systems (PACS) before they became commercial standards in healthcare.

Her research spans medical informatics, image processing, data analytics, and machine learning - always focused on clinically relevant impact and improving patient care. A passionate educator and mentor, Dr. Andriole has developed formal courses, directed fellowships in biomedical imaging and informatics, and guided more than 100 trainees toward leadership in the field.

As a committed contributor to professional societies, she holds leadership roles with the Radiological Society of North America, Society for Imaging Informatics in Medicine, American College of Radiology Data Science Institute, and others. Furthermore, Dr. Andriole has been recognized with top honors including the RSNA Gold Medal, RSNA Honored Educator Award, SIIM Gold Medal, and the RADxx Trailblazer Award for her pioneering work and advocacy for women in imaging informatics.

Rick Whitted, MBA

President and Chief Executive Officer
U.S. Hunger

Rick Whitted is the President & Chief Executive Officer of U.S. Hunger, a national nonprofit operating at the intersection of food security, healthcare, and community engagement. Since becoming CEO in March 2020, Rick has led a strategic transformation of the organization—shifting from a primarily volunteer-driven hunger relief model to a market-aligned social impact enterprise built for long-term sustainability and systems integration. Under his leadership, U.S. Hunger has grown more than fivefold over six years through a disciplined focus on contract-based partnerships, scalable infrastructure, and measurable impact. Central to this evolution is a framework Rick calls "Market-Aligned Social Impact," grounded in a simple question: How does our mission solve a marketplace problem? This approach has enabled sustainable revenue while expanding mission delivery.

Rick repositioned the Hunger Project model from volunteer-centric events to a nationally scalable corporate engagement platform. By evolving the Million Meal Pack Event and building repeatable enterprise partnerships, the portfolio has expanded more than eightfold since 2020, strengthening both financial sustainability and national reach. In parallel, Rick has overseen the development of Full Cart as a healthcare-aligned food access infrastructure supporting Medicaid and vulnerable populations through dignity-centered engagement, integrated screening, referral coordination, and actionable data insights. The model preserves the leadership role of community-based organizations while enabling healthcare partners to advance health equity and system accountability.

He has also built a data and technology ecosystem that supports compliance, operational execution, partner reporting, and outcomes measurement across social determinants of health initiatives—positioning U.S. Hunger as a scalable, healthcare-integrated partner. Prior to joining U.S. Hunger, Rick spent nearly three decades in banking, specializing in small business, government, and nonprofit financial services. He brings that financial and operational discipline to mission-driven leadership. Rick holds a B.A. from Stetson University and an MBA from Nova Southeastern University.

He is an author and speaker on leadership and organizational culture and serves as an Amazon Web Services Social Enterprise Accelerator Teaching Fellow, RISE Fellow, and RISE Healthy Communities Conference Co-Chair.

Kara Goss, MD

Associate Professor of Medicine & Pediatrics
Pulmonary & Critical Care Medicine
UT Southwestern

Dr. Goss is an R01-funded physician scientist in the Division of Pulmonary and Critical Care Medicine at UT Southwestern Medical Center.

Trained in Internal Medicine, Pediatrics, and Adult Pulmonary and Critical Care including pulmonary hypertension and right ventricular failure, she has developed unique lifespan expertise in the long-term cardiopulmonary effects of extreme preterm birth.Her work utilizes state-of-the-art imaging and physiologic assessments to study this emerging high-risk population.

Clinically, she specializes in care of adults born preterm as well as management of a full range of pulmonary vascular diseases through the UT Southwestern Pulmonary Hypertension Clinical Program.

Data and Social Impact: 10,000 Meals for Health

Since the inaugural Pediatric and Lifespan Data Science Conference, the Research Institute has challenged participants to move beyond discussion and collaborate toward measurable outcomes. This approach aligns with the Institute's bedside-to-beside focus. In support of this mission, the conference steering committee has emphasized the importance of centering clinical, academic, and community dialogue during the conference on research innovation and initiatives that generate tangible, real-world impact.

The success of last year's conference, together with this sustained commitment to action, inspired Rick Whitted, CEO of U.S. Hunger, to support the packing of 10,000 Red Lentil Jambalaya meals as part of this year's conference. This vegan, gluten-free, shelf-stable meal has a 12-month shelf life, is simple to prepare, and is clearly labeled. This meal-packing initiative reflects the application of conference discussions on social needs and community engagement through a concrete community action. These meals will be distributed through Rady Children's Health clinics and facilities, as well as through community partners across Orange County.

This event was made possible through collaboration between U.S. Hunger and Rady Children's Health in Orange County, including its Department of Population Health, the Research Institute, and CHOC's Community Action Team.

Join us or stop by to show your support on April 29, 2026, at 5:00 PM in the auditorium next to the main conference stage.

Hackathon at the 2026 Pediatric and Lifespan Data Science Conference

The Hackathon at the 2026 Pediatric and Lifespan Data Science Conference is a hands-on, application-focused AI experience addressing real clinical challenges across pediatric, adult, and lifespan medicine. It brings together students, clinicians, researchers, engineers, and innovators in a collaborative environment where participants learn from one another while building solutions to meaningful, real-world problems. The atmosphere is welcoming and exploratory, encouraging teams to experiment, ask questions, and develop ideas together.

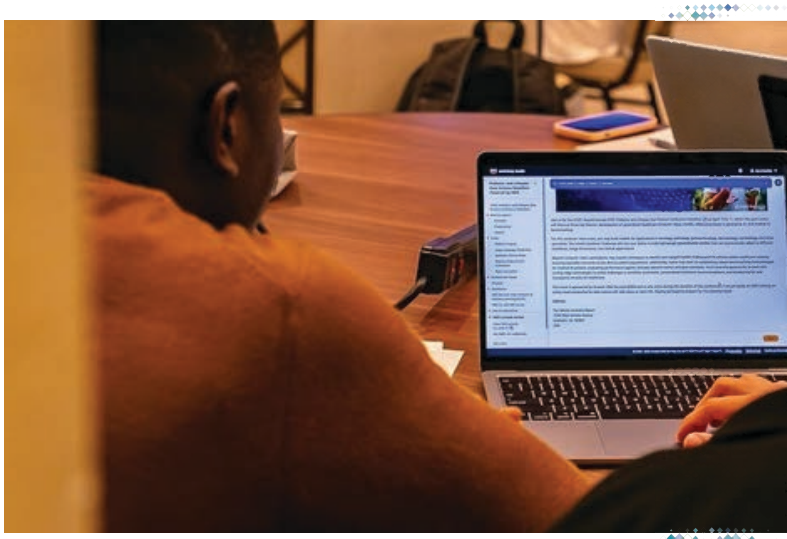
This year's Hackathon participants will develop the Care to Home Companion, a web-based or mobile application that helps caregivers stay on track with recovery plans, managing medications, follow-up appointments, and identifying warning signs, while connecting them to essential community resources. The application will focus on two core features to help caregivers stay on track with recovery plans when they leave the hospital.

- **Discharge Companion:** A web and mobile tool designed to translate hospital discharge plans into clear, actionable guidance for caregivers, helping families understand and manage care after leaving the hospital.
- **Social Needs Connectivity Application:** A platform focused on helping families identify and access community-based social resources, including food assistance, transportation, housing support, and other essential services.

These tracks give participants hands-on experience building practical AI-enabled applications, organized by Rady Children's Health in Orange County, formerly CHOC, powered by cloud-based tools from Amazon Web Services, and supported by experts from Slalom. Rather than emphasizing traditional data science tasks or curated datasets, the Hackathon focuses on rapid prototyping and exploring how AI-driven systems can be thoughtfully designed to support patients, families, and communities. Teams may choose to develop traditional applications or explore agentic AI approaches, depending on their interests.

Throughout the event, teams collaborate to create tools, prototypes, and AI-driven concepts that explore new approaches to improving healthcare delivery and experience. Mentors and technical experts from Rady Children's Health in Orange County, Amazon Web Services, and Slalom will be available throughout the Hackathon to provide guidance and support. In addition, Dr. Leo Anthony Celi will conduct a 50-minute AI educational workshop on the Health AI Systems Thinking for Community (HASTC) framework that guides participants through evaluating the ethical, legal, social, and real-world implications of health AI models prior to deployment.

The Hackathon offers more than the opportunity to build a prototype. Participants gain exposure to modern AI techniques, develop an understanding of different AI-enabled system designs, and collaborate with peers who bring diverse skills and perspectives. The experience also provides valuable opportunities to connect with mentors, industry partners, and fellow participants who share a commitment to using technology for positive impact. Many attendees leave with new skills, new collaborators, and a clearer vision for how their ideas can contribute to future healthcare innovation. Designed to be both practical and inspiring, the Hackathon is consistently one of the most memorable components of the conference.



Acknowledgements from the Director

Thank you for helping make the 3rd Annual Pediatric and Lifespan Data Science Conference a meaningful and successful gathering. Our goal each year is to strengthen the experience for attendees and to increase the value this event provides to institutions and communities across the world. Your continued involvement and engagement make this possible. To help us keep improving, we welcome your feedback through the confidential survey or by emailing me directly at lehwerhemuepha@choc.org.

We extend sincere appreciation to our keynote speakers: Annie Qu, PhD, Katherine Andriole, PhD, Rick Whitted, and Kara Goss, MD. Their contributions covered a broad range of topics in clinical care, social needs, and the growing opportunities for algorithmic innovation in pediatric and lifespan research. We are also grateful to the moderators who guided the panel sessions with clarity and expertise. Thank you to Aldo Faisal, PhD, Steven Martel, MD, Leo Anthony Celi, MD, PhD, Sandra Schultz, MHA, CPXP, Charles Golden, DO, John Henderson, and Terence Sanger, MD, PhD.

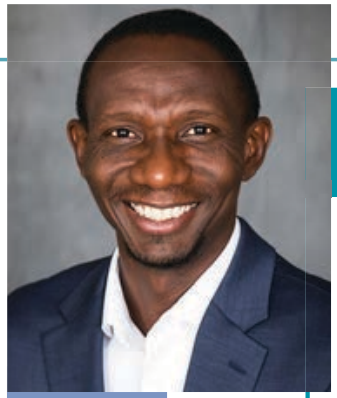
Our program was strengthened by individual presentations from Jake Ormond, PhD, Kamar Braish, MS, RHIA, Diana Ortiz, RN JD, Caroline Peika, MSc, Matthew Bainbridge, PhD, and Davide Ortolan, PhD. We also recognize all abstract podium and poster presenters, including the top submissions selected this year. Warm appreciation to Dr. Keith Feldman, PhD and the team of external Scientific Abstract Review Committee for the conference.

Our appreciation goes to US Hunger for collaborating with us on the 10,000 Meals for Health and Hope initiative. We are grateful for the support of Rick Whitted, Dr. Michael Weiss, DO, Erika Jewell, LCSW, ACM, the CHOC Community Action Team, and all volunteers who contributed to efforts addressing food insecurity and improving information about social needs across Orange County. In addition, we are grateful to Amazon Web Services (AWS) and Slalom for their close partnership and collaboration on the Hackathon addressing social needs and patient education.

Our deepest appreciation goes to our patient and family panel. Thank you to Lara MacLean, Grace Leroy Loge, and Alicia McCormick for sharing your time and insights this year and previously. We also thank the new Executive Planning Committee members for their direction and recommendations that improved this year's conference and will further elevate the 2027 conference.

Finally, heartfelt thanks to the entire planning committee, including the Research Institute and CME teams, for their dedication throughout the year.

Louis Ehwerhemuepha, PhD
Chair, Pediatric & Lifespan Data Science Conference
Director, Research Data Science (Computational Research), Rady Children's Health in Orange County
Adjunct Associate Professor of Pediatrics, University of California - Irvine



With Gratitude

Research Institute

Manager, Research Programs	Carol Davis-Dao, PhD
Supervisor, Advanced Research Data Services	Tatiana Moreno
Project Coordinator, Research Programs	Jorge Ramirez
Manager, Research Programs	Jo Rogers
Research Computational Scientist I	Aline Rohloff
Project Manager I	Dan Rucker
Manager, Research Programs	Keri Zabokrtsky

Continuing Medical Education

Director, Academic Affairs	Mary Hickcox
Manager, CME Programs & Events	Cristen Hemingway
CME Digital Events Project Manager	Maxine Manglicmot

Marketing

Manager, Service Line Marketing	Desireé DeLattre
Manager, Social Media	Diana Eichmann
Project Manager, Marketing	Daniel Holloway

Research Data Science at Rady Children's Health

The Research Computational and Data Science (RCS) team was established in 2021 by Terence Sanger, MD, PhD, and Phuong Dao, JD, to support data science and artificial intelligence research across Rady Children's Health in Orange County, formerly CHOC. From the outset, the team was designed to serve as a shared scientific resource for the hospital research community.

Under the leadership of Louis Ehwerhemuepha, PhD, the RCS team has advanced a long-term vision centered on building deep expertise in data-driven discovery. Core focus areas include classical and causal machine learning, natural language processing and generative AI, computer vision and advanced imaging, genomics, and reinforcement learning. This vision is reflected in the team's growing operational and clinical impact across the health system.

Several mission-driven units have been established, beginning with the Research Data Services unit, led by Tatiana Moreno, BS. The unit supports EHR-based research through data extraction and processing, cohort discovery, and retrospective analyses. Its scope spans structured clinical data, clinical notes, and biomedical imaging from the PACS system. The unit also enables access to multicenter resources such as Oracle EHR Real-World Data, with planned expansion to Epic's COSMOS database. Since its formation, the unit has supported more than 300 studies across most clinical specialties.

RCS has also established specialized programs in Multimodal Genomic Data Science and Advanced Imaging. The Multimodal Genomic Data Science unit, led by Ying-Chu Lo, PhD, provides bioinformatics support for bacterial, fungal, and viral research. The program integrates health system and public multi-omics data with EHR data using artificial intelligence methods such as multimodal fusion models to predict outcomes and identify relationships between molecular signals and clinical phenotypes.

The Advanced Imaging unit, led by Sumiko Abe, PhD, focuses on developing and applying neuroimaging and image-processing methods to advance research and clinical care in pediatric neurological disorders. Using multimodal approaches including MRI, diffusion tractography, CT, and functional imaging, the unit maps brain structure, connectivity, and function to guide clinical understanding and neuromodulation therapies such as deep brain stimulation. Improved visualization of biomedical images before complex brain surgeries has enabled adjustments to surgical plans, resulting in direct patient benefit.

Across all units, sustained interdisciplinary collaboration has resulted in a high volume of scholarly output. The team averages approximately one peer-reviewed publication per month across diverse clinical specialties and methodological approaches. These publications arise from jointly developed study protocols or from RCS-led investigations conducted in close collaboration with clinical partners. In parallel, RCS collaborates closely with other Rady Children's Health teams within the Information Systems Department (John Henderson, MBA, VP, Chief Information & Digital Officer) and Health Informatics (Steven Martel, MD, VP, Chief Health Information Officer, Orange County) to support the deployment of translational data science models and accelerate the progression of research discoveries into clinical practice. Additional internal partnerships include joint initiatives with the Medical Intelligence, Information, Investigation & Innovation (MI4) Institute, led by Anthony Chang, MD, MS, MBA, Chief Intelligence and Innovation Officer.

Beyond academic publication and clinical deployment, RCS expertise has also supported broader innovation activities. Separately from their roles within the organization, several team members have founded external AI startups, including Krv Labs and Lunara Intelligence. Krv Labs is building a platform that validates and monitors clinical AI models to identify failure modes before they affect patient care. Lunara Intelligence focuses on developing data infrastructures and artificial intelligence strategies for global social needs through a transatlantic consortium of partners.

To expand translational impact, RCS collaborates with Imperial College London, the University of California, Irvine, Chapman University's Schmid College of Science, John Snow Labs, Amazon Web Services, Slalom, and conference partners. A formal collaboration with Imperial College London, led by Aldo Faisal, PhD, focuses on learning optimal treatment policies for sepsis and supports the Nightingale AI initiative to develop large-scale multimodal foundation models for healthcare. Additional efforts include evaluating vision-language models for clinical care and developing semi-automated AI research tools using human-in-the-loop approaches.

This progress has been enabled by close collaboration with clinical teams, strong partnerships across the health system, and support from the Research Institute. At its core, the mission of RCS is to harness data science and artificial intelligence to improve pediatric health, advance patient-centered research, and build bridges to lifespan discovery that support a complete understanding of the patient journey.

Agenda Day One

Wednesday, April 29

Morning	Theme: AI & Agents for Clinical Care
7:00 - 8:00 AM	Check-in & Breakfast
8:00 - 8:05 AM	Welcome Note by Louis Ehwerhemuepha, PhD and Terence Sanger, MD, PhD
8:05 - 8:15 AM	Conference & Outcomes Updates by Louis Ehwerhemuepha, PhD
8:15 - 9:00 AM	Keynote Address | Representation Retrieval Learning for Heterogeneous Data Integration
9:00 - 9:50 AM	Panel Presentation | Agentic AI in Medicine: Current Applications, Promises & Challenges in Clinical Care
9:50 - 10:00 AM	Coffee Break
10:00 - 10:45 AM	Keynote Address | AI in Radiology: Experiences from Model Development to Clinical Implementation
10:45 - 11:45 AM	Panel Presentation | Transforming Clinical Workflows Using Assistive & Agentic Intelligence
11:45 - 12:00 PM	Presentation | Pediatric AI Clinician for Optimizing Sepsis Treatment: An Imperial College-CHOC Research Initiative
12:00 - 12:15 PM	Presentation | From Acuity to Action: Using Solventum AI with the Rothman Index to Improve Pediatric Documentation, Outcomes & Workflow
12:15 - 12:30 PM	Presentation | RCH Chat: Building Safe, Governed, GenAI for Pediatric Healthcare
12:30 - 1:30 PM	Lunch & Gourmet Coffee
Afternoon	**Theme: AI & Agents for Health-Related Social Needs**
1:30 - 2:00 PM	Keynote Address | From Lived Experience to Actionable Data: Advancing Whole-Child Care Through Social Risk Integration
2:00 - 2:45 PM	Panel Presentation | Social Needs Considerations in the Era of AI & Agentic Intelligence in Medicine
2:45 - 3:00 PM	Presentation | Bridging the Gap: Proactively Identifying Patients that Would Benefit from Genomic Sequencing Using AI
3:00 - 3:05 PM	Coffee Break
3:05 - 3:25 PM	Presentation | Nightingale AI - Foundation World Model for Health
3:25 - 3:40 PM	Presentation | An AI-Driven 3D RPE Digital Twin for Detecting Polarity Defects & Disease Signatures
3:40 - 4:00 PM	Abstract Podium Presentations Sponsored by Chapman University Schmid College of Science & Technology
4:00 - 5:00 PM	Abstract Poster Presentation
5:00 - 6:00 PM	10,000 Meals for Health & Hope

Agenda Day Two

Thursday, April 30

Morning	Theme: The Voice of Patients and Families, & Role of Industry in AI Innovation
7:00 - 8:00 AM	Breakfast Buffet
8:00 - 8:05 AM	Welcome Note by Coleen Cunningham, MD
8:05 - 8:55 AM	Keynote Address | The Lifelong Cardio-Pulmonary Footprint of Preterm Birth: Insights Across the Lifespan & Emerging Opportunities with AI
9:00 - 10:00 AM	Panel Presentation | Patient & Family Perspectives on Improving Care with AI & Data Science
10:00 - 10:15 AM	Coffee Break
10:15 - 11:10 AM	Panel Presentation | Industry Perspectives on Driving Clinical AI through Innovation & Partnerships
11:10 - 11:25 AM	Hackathon Results Overview by Tatiana Moreno, BS
11:25 - 11:55 AM	Panel Presentation | The Voice of the People: Open Forum for Collective Action in Pediatric & Lifespan Health
11:55 - 12:00 PM	Closing Notes by Louis Ehwerhemuepha, PhD

Winner of the Best in KLAS Public Cloud

AWS is honored to be rated again as the top performer for Public Cloud in the 2026 Best in KLAS:

Software & Services report
from @KLASResearch.

This award recognizes our continued commitment to customer obsession and the work our teams do every day to support our healthcare customers.

Learn more about why AWS was awarded this honor:

Panels & Presentations

Keynote Address | Representation Retrieval Learning for Heterogeneous Data Integration, with Annie Qu, PhD
April 29, 8:15 – 9:00 AM

In the era of big data, large-scale, multi-modal datasets are increasingly ubiquitous, offering unprecedented opportunities for predictive modeling and scientific discovery. However, these datasets often exhibit complex heterogeneity, such as covariate shift, posterior drift, and missing modalities which can hinder the accuracy of existing prediction algorithms. To address these challenges, we propose a novel Representation Retrieval (R2) framework, which integrates a representation learning module (the representer) with a sparsity-induced machine learning model (the learner). Moreover, we introduce the notion of "integrativeness" for representers, characterized by the effective data sources used in learning representers, and propose a Selective Integration Penalty (SIP) to explicitly improve the property. Theoretically, we demonstrate that the R2 framework relaxes the conventional full-sharing assumption in multi-task learning, allowing for partially shared structures, and that SIP can improve the convergence rate of the excess risk bound. Extensive simulation studies validate the empirical performance of our framework, and applications to two real-world datasets further confirm its superiority over existing approaches.

Panel | Agentic AI in Medicine: Current Applications, Promises, and Challenges in Clinical Care, Moderated by Aldo Faisal, PhD
April 29, 9:00 – 9:50 AM

This panel will explore the emerging role of agentic AI, which includes systems capable of autonomous decision-making and goal-directed behavior, in clinical medicine. Panelists will highlight current use cases in which agentic AI is supporting diagnostics, treatment planning, patient monitoring, and operational efficiency. With a focus on pediatric and lifespan health, this session invites researchers, clinicians, and technologists to engage in a forward-looking dialogue on the future of intelligent systems in healthcare.

Panelists:
Anthony Chang, MD, MBA, MPH, MS
Annie Qu, PhD
Christopher Boone, PhD, FACHE, FHIMSS
Robert B. Kelly, MD, FAAP

Keynote Address | AI in Radiology: Experiences from Model Development to Clinical Implementation, with Katherine P. Andriole, PhD, FACR, FSIIM, FSPIE
April 29, 10:00 – 10:45 AM

This keynote explores AI across the medical imaging pipeline, highlighting clinical examples, validation infrastructure, governance processes, and integration standards. It also outlines strategies for informatics and AI education, shares lessons learned from real-world deployment, and offers forward-looking insights on shaping the future of radiology across all patients of all ages.

Panels & Presentations

Panel | Transforming Clinical Workflows Using Assistive and Agentic Intelligence, Moderated by Steven Martel, MD, FAAP
April 29, 10:45 – 11:45 AM

This panel examines how assistive and agentic AI can enhance clinical workflows in responsible and scalable ways. Panelists will discuss real-world use cases, data science considerations, EHR and workflow integration, and the trust and governance models needed to support safe and accountable AI-enabled care across complex neonatal, adolescent, and adult clinical settings.

Panelists:
Terence Sanger, MD, PhD
Katherine P. Andriole, PhD, FACR, FSIIM, FSPIE
David Talby, PhD, MBA
Matthew Bainbridge, PhD

Presentation | Pediatric AI Clinician for Optimizing Sepsis Treatment: An Imperial College-CHOC Research Initiative, with Jake Ormond, PhD
April 29, 11:45 – 12:00 PM

This presentation introduces the Pediatric AI Clinician for optimizing sepsis treatment, an Imperial College-CHOC collaborative initiative extending the pioneering AI Clinician framework developed by Aldo Faisal and colleagues into pediatric medicine. The work applies reinforcement learning for critical care decision support to improve sepsis management for children.

Presentation | From Acuity to Action: Using Solventum AI with the Rothman Index to Improve Pediatric Documentation, Outcomes and Workflow, with Kamar Braish, MS, RHIA & Diana Ortiz, JD, RN
April 29, 12:00 – 12:15 PM

This short presentation highlights the use of the Rothman Index and related tools from Solventum to improve clinical documentation and reimbursement accuracy. Originating from a collaboration between two past conference attendees, the concept was rapidly prototyped and is now being integrated to enhance clinical care and operational efficiency.

Presentation | RCH Chat: Building Safe, Governed, GenAI for Pediatric Healthcare, with Caroline Peika, MSc
April 29, 12:15 – 12:30 PM

This presentation outlines Rady Children's Health's journey in developing RCH Chat, a safe and governed GenAI platform that provides HIPAA-compliant, secure access to generative AI for clinical, administrative, and clerical use. A system leader will highlight how enterprise-grade governance, guardrails, and workflow integration enable responsible AI adoption across one of the nation's largest pediatric health systems.

Keynote Address | From Lived Experience to Actionable Data: Advancing Whole-Child Care Through Social Risk Integration, with Rick Whitted, MBA
April 29, 1:30 – 2:00 PM

The session will explore how household-level lived experience can be translated into structured, interoperable data that meaningfully informs pediatric care, system design, and sustainable models of whole-child health.

Panels & Presentations

Panel | Social Needs Considerations in the Era of AI and Agentic Intelligence in Medicine, Moderated by Leo Anthony Celi, MD, MPH, MSc
April 29, 2:00 – 2:45 PM

This panel examines challenges the current AI revolution may pose to health-related social needs, as well as how emerging tools can transform data and insights to strengthen the social dimensions of health. Panelists will explore integrating social determinants into models, ensuring fairness and cultural sensitivity, and designing intelligent systems that support personalized, human-centered care across pediatric and lifespan populations.

Panelists:
Ezekiel Brooks
Lindsey Jarrett, PhD
Andrea Goubeaux, MSW, LSW
Katie Nguyen Kalvoda

Presentation | Bridging the Gap: Proactively Identifying Patients Who Would Benefit from Genomic Sequencing Using AI, with Matthew Bainbridge, PhD
April 29, 2:45 – 3:00 PM

This presentation explores how machine learning can be used to proactively identify patients most likely to benefit from genomic sequencing. By prioritizing candidates for testing, AI-driven approaches can help maximize the clinical impact of limited genomic resources, particularly in the diagnosis of complex and rare diseases.

Presentation | Nightingale AI – Foundation World Model for Health, with Aldo Faisal, PhD
April 29, 3:05 – 3:25 PM

Nightingale AI is a multimodal foundation world model for health, developed by Professor Aldo Faisal at Imperial College London. Trained on millions of clinical and scientific data sources, it integrates imaging, diagnostics, patient histories, and biomedical knowledge to generate actionable insights and power trustworthy, domain-specific AI tools for research and clinical care.

Presentation | An AI-Driven 3D RPE Digital Twin for Detecting Polarity Defects and Disease Signatures, with Davide Ortolan, PhD
April 29, 3:25 – 3:40 PM

This presentation introduces an AI-driven high-content image analysis platform using mask R-CNN and mathematical modeling to generate a quantitative 3D digital twin of the Retinal Pigment Epithelium (RPE) during polarization. The approach reveals non-stochastic cell-state transitions and organelle interactions, providing a powerful reference model for detecting intracellular defects in diseased RPE.

Panels & Presentations

Keynote Address | The Lifelong Cardio-Pulmonary Footprint of Preterm Birth: Insights Across the Lifespan and Emerging Opportunities with AI, with Kara Goss, MD
April 30, 8:05 – 8:55 AM

This keynote will address how preterm birth shapes cardio-pulmonary health from infancy through adulthood, synthesizing emerging evidence on long-term respiratory and cardiovascular risks. She will explore how multimodal data and AI can reveal subtle physiologic patterns, guide personalized prediction, and transform care for preterm-born individuals across the lifespan.

Panel | Patient and Family Perspectives on Improving Care with AI and Data Science, Moderated by Sandra Cortez-Schultz, MHA, CPXP & Charles V. Golden, DO, FAAP
April 30, 9:00 – 10:00 AM

Patients, families, and community members share their lived experiences and offer insights into how data science, AI, and agentic systems can better serve their needs.
This is a moment to center the voices that matter most and to ensure that the future of pediatric and lifespan healthcare is shaped by those it is meant to serve.

Panelists:
Lara MacLean
Grace LeRoy-Loge
Alicia McCormick, PhD
Peyton Averitt

Panel | Industry Perspectives on Driving Clinical AI Through Innovation and Partnerships, Moderated by John Henderson, MBA
April 30, 10:15 – 11:10 AM

This session invites our industry partners to share their perspectives on the evolving landscape of clinical AI and agentic systems. From cloud infrastructure and algorithm development to data platforms and funding models, industry plays a critical role in enabling the tools and technologies that drive innovation in healthcare.

Panelists:
James Hickman
Rod Tarrago, MD
Charles Friedman, PhD
Hide Harashima

Panel | The Voice of the People: Open Forum for Collective Action in Pediatric & Lifespan Health, Moderated by Terence Sanger, MD, PhD
April 30, 11:25 – 11:55 AM

This closing session brings together key stakeholders to synthesize the most pressing clinical challenges and opportunities identified throughout the conference. The discussion will focus on aligning priorities, defining concrete action items, and establishing next steps to be pursued following the conference.

Panelists:
Melanie Patterson, DNP, MHA, RN
Sameer S. Pathare, MD
Louis Ehwerhemuepha, PhD

Podium Presentations

Speaker	Title
Jeremy Wayland, MS	Nightingale: A Framework for Evaluating Robustness of Clinical ML Models
Sean Dornbush, DO, MBA	Differences in Outcomes at Pediatric vs. Adult Hospitals for 18–26-Year-Olds with Complex Chronic Conditions
Natanael Alpay, MS	Integrating Machine Learning for VAERS Vaccine Safety Signal Triage vs. Classical Disproportionality: A Pediatric and Lifespan Evaluation

Jeremy Wayland, M.S., is PhD student at the Technical University of Munich and Doctoral Researcher at Helmholtz Munich's Institute of AI for Health. He is a member of the AIDOS Lab, which focuses on research at the intersection of geometry, topology, and machine learning. He is cofounder and Chief Scientific Officer at Krv Analytics, a startup focused on improving AI evaluation frameworks to enable reliable, performant, and generalizable model deployments.

Dr. Sean Dornbush, DO, MBA, is an Internal Medicine Hospitalist and Pediatric Hospital Medicine Fellow at Rady's Children's Health in Orange County, Formerly CHOC, whose work focuses on outcomes for adolescents and young adults with complex chronic conditions across pediatric and adult hospital systems. His research leverages real-world data and health services methods to improve care transitions and the quality of inpatient care.

Natanael Alpay, M.S., is a first-generation American and a PhD student in Mathematics at the University of California, Irvine (M.S. Applied Mathematics). He holds four B.S. degrees from Chapman University in Mathematics, Physics, Computer Science, and French, and received the Ronald M. Huntington Outstanding Scholarship Award for top academic achievement.

Poster Presentation Index

#	Presenting Author	Poster Title
1	Moulayna Johnson Sidi Baba, MS	Digital Inequality in the Age of GenAI: How Rural versus Urban Parents Differ in AI Perceptions and Adoption
2	Bryan Trieu, BS	Addressing Structural Barriers to STEM Through Near-Peer Mentorship: Outcomes From the OPTIMUS-SPARK Program
3	Lois Sayrs, PhD	From Predictive to Prescriptive: A Socio-Technical Framework for Pediatric Digital Health Twins (DHTs), Complementary, and Prescriptive Validity
4	Anna Le	Comprehensive ASSESSMENT of t-cell exhaustion in allogeneic HSCT and car t-cell immunotherapy
5	Howard Huang, MS	Descriptive Analysis for Diabetes Patients in Pediatrics
6	Diamond Luong, BA	A 25-Year Retrospective Analysis of Prenatal Detection and Genetic Evaluation of Cleft Lip
7	Anna Pelc, MS, MBA, MPH	Pediatric Gene Therapies Authorizations and Costs in the U.S.
8	Wei Ling Amelia Lee, MBBS	Development of a Best-Practice Machine Learning Pipeline for Predicting Clinical Outcomes in Children in Low-Resource Settings
9	Wei Ling Amelia Lee, MBBS	Development of Interpretable Clinical Risk Scores for Predicting Paediatric Outcomes in Low-Resource Settings Using the AutoScore Framework
10	Judith Weng Zhu, BS	Predicting Long-Term Cardiac Effects in Preterm Born Individuals
11	Omer Odabas, MS	Anticoagulant Treatment Effects on Survival in Antiphospholipid Syndrome Using Targeted Learning and Electronic Health Record Data
12	Liora Mayats Alpay, MS	Heart Attack Incidence Rates and Post-Event Survival Across U.S. States, 2019–2024
13	Duncan Ofosu, PhD	SuperLearner Combined with LLM-Derived Features to Predict Severe Bleeding in Adults with Venous Thrombosis
14	Duncan Ofosu, PhD	Targeted Learning Analysis of Oral Anticoagulant and Risk of Thrombosis Recurrence in Adults

**Posters listed in submission order.*

Poster Presentation Index

#	Presenting Author	Poster Title
15	Eric Adams, MSc	Comparison of Treatment Effect of Stimulant and Non-Stimulant ADHD Medication on Incident Rates of Cardiovascular Disease: A Targeted Learning Study Based on MIMIC-IV Data
16	Kevin Truong, MD	The Great Note Bake Off: Comparison of Quality between Clinician Written Documentation and Generative Artificial Intelligence Chart Summarization Tools
17	Godfrey Ejiroghene Akpojotor, PhD	Smartphones Driven Data Science in Healthcare as Roadmap for Inclusivity in Emerging Economies
18	David Licerio, MS	Predicting ICU Readmissions Using Deep Learning and Natural Language Processing Techniques: A MIMIC IV Data Analysis
19	Reza Rezaei, MS	Trends in RSV, hMPV and Co-infections, 2019–2024
20	Adhithi Kanthan Lakshmikanthan, MS	Identifying potential adiposity-driven cardiovascular outcomes in males and females with supervised machine learning methods in public access databases
21	Mingfei Dong, PhD	AI-Enabled Interactive Dashboard for RSV Surveillance and Nirsevimab Uptake in Pediatric Populations
22	Mingfei Dong, PhD	From Structure to Function: Validating Automated fMRI Metrics of Pectus Severity Using Cardiopulmonary Performance and Dynamic Simulation
23	Damini Trehan	Longitudinal Patterns in Major Depressive Disorder: A Nationwide Claims Analysis (2019–2024)
24	Christian Andrade Herrera, BS	Longitudinal Cardiac Hemodynamics in a Hyperoxia-Induced Rat Model of Bronchopulmonary Dysplasia
25	Ijeoma Nnadozie, MS	Advancing Pediatric Precision Medicine Through Multimodal Imaging-Omics-Clinical Integration
26	Sierra Willens, MD	Automated Detection of Velopharyngeal Port Dynamics from Nasopharyngoscopy Videos using Deep Learning
27	Anuraag Tandon	Using Caregiver-Logged Behaviors to Predict Daily Needs in Minimally Verbal Autism
28	Leon Shpaner, MS	Reproducible Exploratory Data Analysis Frameworks for Pediatric and Lifespan Clinical Research

**Posters listed in submission order.*

Podium Presentation

Nightingale: A Framework for Evaluating Robustness of Clinical ML Models

Jeremy Wayland[1,2], Sidney Gathrid[2], Stuart Wayland[2]

[1]Rady Children's Health in Orange County, formerly CHOC
[2]Krv Labs

Background: Clinical machine learning models often show strong performance in validation but frequently fail at deployment due to data quality issues and population shifts. State-of-the-art evaluation techniques emphasize accuracy without systematically testing behavior under realistic stress conditions like sensor failures or physiologic variability.

Methods: Nightingale evaluates robustness via three metrics: (1) Resiliency—stressed-to-clean performance ratio measuring corruption tolerance, (2) Stability—prediction variance under jitter ($1/(1+\alpha \times \text{mean\_variance})$), (3) Generalizability—robustness to previously unseen patient populations. As a case study in sepsis prediction, we tested a LightGBM model based on Moor et al. (2021, MLHC) trained on PhysioNet 2019 data (Hospital A, n=1000) with seven vital signs. Stress-testing on Hospital B (n=1000) included MAP blackouts (60% missing), HR jitter (σ=1.0), and Age>60 subgroup analysis. Performance used PhysioNet Utility Score, which rewards timely alerts (6h before sepsis) while penalizing false alarms.

Results: The model achieved Resiliency=0.994 (near-perfect MAP outage tolerance), Stability=0.464 (moderate HR noise vulnerability), and Generalizability=0.470 (substantial Hospital B degradation—less than half the original utility). These results expose critical weaknesses in state-of-the-art sepsis prediction that standard validation misses. While the model tolerates missing data well, it shows marked sensitivity to measurement noise and population shifts. This pilot evaluation used reduced cohorts (n=1000 per site) to establish feasibility; the framework readily scales to full datasets on production infrastructure. Nightingale metrics demonstrate strong promise for systematic robustness assessment before clinical deployment, where physiologic variability and institutional differences remain critical challenges.

Podium Presentation

Differences in Outcomes at Pediatric vs. Adult Hospitals for 18–26-Year-Olds with Complex Chronic Conditions

Sean Dornbush, DO, MBA[1], Ricardo Aguilar, MS[1], Wendy Gray, PhD[1], Donald Phillips, MD, MPH[1], Alpesh Amin, MD, MBA[2]

[1]Rady Children's Health Orange in County, formerly CHOC
[2]University of California, Irvine

Background: Adult discharges from pediatric hospitals are increasing, with many patients having Complex Chronic Conditions (CCCs). These adults experience higher Pediatric Intensive Care Unit (ICU) utilization, longer length of stay (LOS), and greater costs than their pediatric counterparts. Outcome differences in readmission and LOS have been observed for specific CCCs at adult-only and mixed-service hospitals, but whether these and other differences extend to all CCC patients remains unclear. This study aims to assess the association between pediatric-only health systems (POH), adult-only health systems (AOH), and mixed health systems (MH) and clinical outcomes for patients aged 18–26 with CCCs.

Methods: We conducted a retrospective cohort study using Oracle EHR Real-World Database (OERWD) of observation and inpatient encounters for patients aged 18–26 with CCCs from January 1, 2023, to December 1, 2024, excluding childbirth-related admissions. Health systems were categorized as POH, MH, or AOH. Mixed-effects regression modeled associations between health category and outcomes, while adjusting for age, sex, race, ethnicity, primary insurance type, admission source, encounter type, CCC organ system, and severity as the encounter Medicare Severity Diagnosis Related Group cost-weight.

Results: A total of 34,746 encounters met inclusion criteria. The most common CCC was metabolic (36.48%). The distribution of encounters across health system categories was 14.6% at POH, 57.7% at MH, and 27.5% at AOH. Compared to POH, adjusted odds of 30-day ED return were higher at MH (OR 1.48, 95% CI: 1.13–1.94) and AOH (OR 1.76, 95% CI: 1.34–2.32). No statistically significant differences between health system categories were found for 30-day readmission (MH OR 1.04, 95% CI: 0.85–1.27; AOH OR 1.05, 95% CI: 0.85–1.29), ICU transfer (MH OR 1.91, 95% CI: 0.25–14.89; AOH OR 0.78, 95% CI: 0.10–5.97), and LOS (MH MR 0.95, 95% CI: 0.78–1.15; AOH MR 0.95, 95% CI: 0.79–1.15).

Podium Presentation

Integrating Machine Learning for VAERS Vaccine Safety Signal Triage vs. Classical Disproportionality: A Pediatric and Lifespan Evaluation

Natanael Alpay, MS[1]

[1]University of California, Irvine

Background:Classical vaccine safety signal detection ranks vaccine–adverse event (AE) pairs using disproportionality statistics such as PRR, ROR, IC, and EBGM-style shrinkage computed from spontaneous reporting systems. These rankings can be unstable under sparse counts and may mask age-dependent patterns that are important for pediatric and lifespan surveillance. We investigate whether machine learning (ML) built on the same count-table foundation, enriched with reporting context and temporal structure, can support more robust operational triage on VAERS.

Methods: We use U.S. VAERS data (CDC/FDA) from 2024–2025 (VAERSDATA, VAERSVAX, VAERSSYMPTOMS). Vaccine exposures are linked to MedDRA-coded symptom preferred terms and aggregated into quarterly vaccine–event panels. For each (quarter, vaccine, AE), we compute 2×2 report-count tables (a,b,c,d) and classical disproportionality metrics (PRR/ROR/IC and an EBGM-style shrinkage proxy). We then train regularized logistic regression and gradient-boosted trees using feature groups: stabilized counts, stacked classical scores, report-level context aggregates (e.g., seriousness indicators and concomitant vaccine counts), and stable temporal features (lags and smoothed trends). Evaluation uses time-respecting rolling windows (train on prior quarters, validate on the next quarter) and reports AUROC/AUPRC, precision@k, Brier score, and FDR-guarded operating points. We perform stratified analyses across age bins (0–1, 2–11, 12–17, 18–64, 65+, Unknown) and conduct ablations to assess the marginal value of temporal and context features.

Results: Across rolling windows, ML models provide calibrated, operationally useful rankings that support top-k triage under sparsity and reporting shifts. Age-stratified analyses show substantial heterogeneity in the highest-ranked vaccine–AE pairs across bins, and stability analyses quantify when rankings are consistent versus quarter-specific. Ablations indicate that temporal and report-context features can add value beyond classical scores, while emphasizing the need for stability-aware feature design in sparse settings.

Posters

Poster #1

Digital Inequality in the Age of GenAI: How Rural versus Urban Parents Differ in AI Perceptions and Adoption

Moulayna Johnson Sidi Baba, MS[1], Gillian R. Hayes, PhD[1], Candice Odgers, PhD, PhD[1], Kelli Dickerson, PhD[1], Aehong Min, PhD[1]

[1]University of California, Irvine

As Generative AI (GenAI) rapidly expands into youth spaces, particularly K-12 education, understanding parental perceptions of these technologies can help designers, researchers, and policymakers improve and promote equitable user experiences. We conducted a nationally representative survey of 2,826 parents in the United States and analyzed trends in GenAI adoption and perception by geography. Our results indicate significant differences between urban (N = 2,398) and rural (N = 428) parent populations in terms of perception, experience, and use of GenAI, including an urban adoption rate nearly double that of rural households. Our analysis indicates that it is a structural lack of perceived usefulness, and not engagement or perceived risk, that drive this adoption gap, and further, that even after controlling for factors like household income and educational attainment, there is still a significant geographic disparity in AI adoption. Multivariate modeling indicates that geography served as the strongest independent predictor of AI adoption in our sample (OR = 1.80), suggesting that the rural context may function as a unique barrier to technology appropriation. These results point toward a need for targeted interventions, moving beyond general awareness to community-embedded co-design and localized case studies that make AI utility salient within rural educational and economic environments.

Poster #2

Addressing Structural Barriers to STEM Through Near-Peer Mentorship: Outcomes From the OPTIMUS-SPARK Program

Bryan Trieu, BS[1], James Murphy, MD[1]

[1]University of California, San Diego, School of Medicine

Background: Youth from underserved communities face structural barriers to pursuing STEM careers, including limited mentorship, financial constraints, and reduced exposure to academic pathways. This study evaluates the impact of a novel medical student-led near-peer mentorship program designed to increase exposure to science and healthcare careers among underserved high school students in San Diego.

Methods: OPTIMUS (Outreach Program to Inspire Minority and Underrepresented Students) and SPARK (Summer Program to Accelerate Regenerative Medicine Knowledge) are UC San Diego–based summer research programs aimed at increasing diversity in STEM and clinical research. UC San Diego medical students serve as near-peer mentors, providing research training, college preparation, and exposure to STEM and healthcare careers. A mixed-methods survey was administered to OPTIMUS/SPARK participants (n=27) to assess career interests, scientific self-efficacy, college preparedness, barriers, and program satisfaction using Likert scales. Open-ended responses were analyzed using thematic analysis.

Results: Most participants (93%) reported aspirations to attend a four-year university. Interest in future careers was highest in healthcare (89%), followed by other STEM fields (63%) and research (30%). Despite high aspirations, participants reported ongoing concerns related to financial costs (63%) and time management (44%). Confidence in applying to college was moderate to high (59%), and all respondents (100%) reported that OPTIMUS/SPARK improved their perceived likelihood of college acceptance.

Conclusions: Near-peer mentorship programs may help mitigate early educational and professional development gaps by improving confidence, sense of belonging, and college preparedness. These findings highlight the potential of structured, equity-focused mentorship to support early STEM and healthcare pipeline development and contribute to efforts to promote equity and diversity in STEM fields.

Poster #3

From Predictive to Prescriptive: A Socio-Technical Framework for Pediatric Digital Health Twins (DHTs), Complementary, and Prescriptive Validity

Lois Sayers, PhD[1]

[1]Rady Children's Health in Orange County, formerly CHOC

Background: Has traditional pediatric data science reached an "empirical ceiling"? Current models are easily criticized for an adult-centric bias, overfitting, and a failure to account for the small-data paradox in rare diseases and complex developmental conditions. This white paper argues that improving predictive analytics without integrating the child's lived experience or the environmental constraints of the family, will continue to limit prediction unless we transition from descriptive analytics (what is happening) and predictive models (what may happen) to what I call prescriptive Hermeneutics (what should happen and how to facilitate it).

Methods: I synthesize post-empiricist critiques with 2026 AI capabilities to introduce the concept of a Socio-Technical Digital Health Twin (ST-DHT). This model moves beyond theories of knowledge undergirding first generation twinning, emulated trials, and algorithm-driven case matching to study the nature of a patient's reality.

Results: The ST-DHT model includes:

- Ecological Phenotyping: Incorporating family dynamics, caregiver stress, and Social Determinants of Health (SDOH) as modifiable variables.
- Bias-auditing: Employing tests of inherent bias throughout the modeling endeavor.
- Complementary Validity: A dual-validation protocol where statistical accuracy is triangulated with family-centered qualitative goals.
- Prescriptive Validity: Fitting the real world to the model by permitting the ST-DHT to serve as a normative map for proscribing a health trajectory.
- Positive-Looping Effect: Understanding how the model defines the child and how that definition requires change from the real world, specifically to secure social and environmental resources for the child.

Discussion: To illustrate these ideas, I provide a case example in Asthma.

Conclusion: The proposed the ST-DHT is a "whole child" approach to data science that is both mathematically rigorous and humanistically grounded.

Poster #4

Comprehensive ASSESSMENT of t-cell exhaustion in allogeneic HSCT and car t-cell immunotherapy

Anna Le[1], Adelyne Tran[2], Sofia Sato[3], Brian Rubio, BS[4], Olivia Marr[5], Rishikesh Chavan, MD[1]

[1]Children's Health in Orange County, formerly CHOC
[2]University of California, Los Angeles
[3]Yale University
[4]University of Southern California
[5]Chapman University

Background: T-cell exhaustion is a distinct differentiation trajectory caused by chronic antigen stimulation, characterized by progressive functional decline, sustained inhibitory receptor expression, and stable epigenetic changes. Exhaustion impacts immune outcomes following CAR T-cell therapy and Allogeneic stem cell transplantation (Allo-SCT). This review synthesizes exhaustion mechanisms, differentiation states, and functional consequences of T-cell exhaustion; identifies key markers and drivers; explores current measurement techniques; and future strategies to reduce relapse and improve therapeutic outcomes.

Methods: We conducted a literature review of studies from 2000 to 2025, focusing on experimental and clinical literature on T-cell exhaustion in Allo-SCT and CAR T-Cell therapy. Included studies examined exhaustion differentiation, molecular regulations, and profiling techniques. Studies focused on viral infection and autoimmune disease were excluded.

Results: T-cell exhaustion represents a distinct developmental pathway, progressing from precursor to terminally exhausted states driven by mechanistic drivers such as persistent antigen exposure, inadequate CD4+ help, and extrinsic modulators. In Allo-SCT, partial exhaustion may attenuate graft-versus-host disease (GvHD) but can compromise graft-versus-leukemia (GvL) activity; therefore, maintaining a balance to minimize relapse is crucial. In CAR T-cell therapy, exhaustion limits persistence and antitumor efficacy, contributing to relapse. Emerging approaches aim to detect and reverse exhaustion before clinical relapse.

Conclusion: T-cell exhaustion may serve as a predictive biomarker for relapse in CAR T-cell therapy and Allo-SCT, but standardized assessment methods are lacking. Developing a method to quantify antigen load alongside exhaustion markers may provide an early signal of relapse risk, allowing for a comprehensive monitoring of immune competence post-treatment.

Poster #5

Descriptive Analysis for Diabetes Patients in Pediatrics

Howard Huang, MS[1], Oliver Gu, MS[1]

[1]Chapman University

Background: Diabetes is one of the most common chronic metabolic disorders affecting both adult and pediatric populations, with growing concern about early-onset and youth-onset disease. Oracle Real-World Database provides an opportunity to characterize patient demographics, lab results, and medications at scale. Descriptive analyses of such data are important for understanding disease burden and informing future studies of treatment effectiveness.

Methods: We conducted a descriptive study of diabetes patients identified from a large, de-identified clinical database derived from the Oracle Real-World Database. Patients were identified using ICD-10 diagnosis codes for diabetes mellitus. Cohorts were constructed by integrating diagnosis records with laboratory measurements, medication prescriptions, and encounter data. Data processing and cohort construction were performed using distributed computing tools, including Python, R, PySpark and SQL, to manage large-scale longitudinal data. Descriptive statistics were used to summarize demographics, including age and sex, as well as lab results, such as hemoglobin A1c levels and estimated Glomerular Filtration Rate (GFR). Diabetic medication usage was also summarized.

Results: The study included a large and diverse population of pediatric patients with diabetes. Preliminary analyses showed variability in age distribution, laboratory monitoring frequency, and medication use among pediatric patients. Differences were also observed in patterns of healthcare utilization, including encounter frequency and follow-up intervals. These findings highlight heterogeneity in clinical management and monitoring in real-world setting.

Poster #6

A 25-Year Retrospective Analysis of Prenatal Detection and Genetic Evaluation of Cleft Lip

Diamond Luong, BA[1], Melissa D. Kanack, MD[1], Kristen Wigby, MD[1], Amanda A. Gosman, MD[1], Benjamin B. Massenburg, MD[1], Marilyn Jones, MD[1]

[1]University of California, San Diego

Background: Prenatal ultrasound has been standard in U.S. obstetric care since the 1970s, with the American College of Obstetricians and Gynecologists recommending at least one examination between 18 and 22 weeks of gestation. When a fetal structural anomaly is identified on ultrasound, diagnostic testing via chorionic villus sampling or amniocentesis is recommended for chromosomal evaluation. Understanding longitudinal shifts in prenatal diagnosis and genetic evaluation can provide insight into how evolving clinical technologies shape pediatric care trajectories. This study evaluated temporal changes in prenatal cleft lip detection, gestational age at diagnosis, and patterns of karyotypin over 25 years.

Methods: A retrospective review was conducted of 957 pediatric patients with unilateral or bilateral, complete or incomplete cleft lip evaluated at Rady Children's Hospital–San Diego from 2000 to 2024. Patients with and without associated anomalies were included. For each year, the proportion diagnosed prenatally and the mean gestational age at diagnosis were calculated and stratified by phenotype. Among prenatally diagnosed cases, annual proportions of patients who were karyotyped were assessed. Temporal associations were evaluated using linear regression and Pearson correlation ($p < 0.05$).

Results: Prenatal detection increased significantly over time for total cleft lip ($r = 0.937$, $p < .001$), complete cleft lip ($r = 0.937$, $p < .001$), and incomplete cleft lip ($r = 0.812$, $p < .001$), rising from 22.6% in 2000 to 100% in 2024. Karyotyping among prenatally diagnosed patients increased modestly from 57.1% to 76.2% ($r = 0.449$, $p = .025$). Mean gestational age at diagnosis decreased from 25.3 to 22.2 weeks ($r = -0.495$, $p = .012$).

Conclusions: Over the past 25 years, prenatal cleft lip diagnosis has become more frequent, earlier in gestation, and increasingly accompanied by genetic evaluation, paralleling advances in imaging and multidisciplinary prenatal care.

Poster #7

Pediatric Gene Therapies Authorizations and Costs in the U.S.

Anna Pelc, MS, MBA, MPH[1], Mana Rahimian, PharmD, MS[1], Adeola Bakare, MS[1], Larence Brown, PharmD, PhD[1], Marc Fleming, PhD[1], Vaishali Shukla, PhD[1], Rosa Rodriguez-Monguio, PhD[2], Enrique Seoane-Vazquez, PhD[1]

[1]Chapman University
[2]University of California, San Francisco

Background: Gene therapies represent an emerging treatment option for pediatric conditions. This study examined FDA authorizations and costs of gene therapies for pediatric indications.

Methods: Data on gene therapy authorizations were obtained from the FDA website. Treatment cost per gene therapy was defined as the wholesale acquisition cost at market entry (Micromedex Red Book). Prices were inflation-adjusted to 2025 US dollars using the consumer price index. Data were analyzed descriptively.

Results: The FDA authorized 27 gene therapies from 2015–2025, of which 13 (48.1%) had pediatric indications (2 pediatric-only and 11 adult/pediatric), and 14 (51.9%) were adult-only. Twelve of 13 (92.3%) therapies for pediatrics were single-use, and all had orphan designations. Pediatric-only or adult/pediatric indications were most common in the therapeutic classes: alimentary tract and metabolism (100%; n=3), musculo-skeletal system (100%; n=3), dermatologicals (100%; n=2), and blood and blood forming organs (50.0%; n=3).

The cost at market entry was available for all except etuvetidigene autotemcel. Onasemnogene abeparvovec and tisagenlecleucel have indications for multiple patient populations (pediatric, adult/pediatric, adult). Costs for the different indications were considered separately in the analysis. The median costs at market entry were $3,198,294 (IQR $830,397; n=3) for pediatric-only, $2,792,244 (IQR $1,613,748; n=10) for adult/pediatric, and $498,811 (IQR $208,997; n=15) for adult-only. Excluding antineoplastics, median costs at market entry were $3,198,294 (IQR $830,397; n=3) for pediatric-only, $2,997,859 (IQR $926,880; n=9) for adult/pediatric, and $2,525,155 (IQR $3,088,830; n=5) for adult-only.

Conclusions: Gene therapies with pediatric indications represent approximately half of FDA-authorized gene therapies, with most for genetic diseases. Median costs at market entry were higher for gene therapies for pediatrics compared to adult-only.

Poster #8

Development of a Best-Practice Machine Learning Pipeline for Predicting Clinical Outcomes in Children in Low-Resource Settings

Wei Ling Amelia Lee, MBBS[1], Sham Lal, PhD[1]

[1]London School of Hygiene and Tropical Medicine

Background: The application of machine learning (ML) to developing prediction models for clinical outcomes for paediatric patients in low resource settings remains an area for research opportunities. Using a secondary paediatric dataset from Mozambique, we present the design of a robust ML pipeline that applies best practices in ML model development to predict clinical outcomes in paediatric patients with general acute illness.

Methods: ML models were developed on a dataset derived from the Health and Demographic Surveillance System in Manhica District, Mozambique, Africa. The final sample size included a total of 26538 inpatient admissions over the period of 1996-2024. The clinical outcome of choice for this analysis was inpatient length of stay > 7 days. Random Forest (RF), XGBoost (XGB), and Multi-Layer Perceptron (MLP) models were developed in R. Adherence to best practices in predictive ML model development was guided by the Prediction model Risk of Bias Assessment Tool (PROBAST) framework. This included the application of nested cross- validation on the training data. This method is advocated to avoid data leakage and enable more accurate estimates of generalisation of performance. The tidymodels metapackage allowed pre-processing steps to be specified programmatically and applied consistently within each nested cross-validation resample. The data was highly imbalanced with a ratio of 9.5:1. Therefore, model performance was evaluated using discrimination and calibration metrics, with particular emphasis on the calibration slope and clinical decision curves.

Results: In the ML model development stage of the study, all 3 models (RF, XGB and MLP) have quite similar area under the Receiver Operator Curve (AUROC) performance, with XGBoost having the highest AUROC of 0.77 (95% CI: 0.74-0.79). Model choice in these scenarios would depend on consideration of performance with calibration, or net benefit from the decision curve analysis.

Poster #9

Development of Interpretable Clinical Risk Scores for Predicting Paediatric Outcomes in Low-Resource Settings Using the AutoScore Framework

Wei Ling Amelia Lee, MBBS[1], Sham Lal, PhD[1]

[1]London School of Hygiene and Tropical Medicine

Background: The field of predictive modelling has become increasingly dynamic with the integration of machine learning (ML) techniques, translating into a growing role in healthcare decision support. However, as ML models become more complex, there is a growing demand for transparent explanations of the ML models construction. This transparency is crucial for fostering human trust in model deployment. We demonstrate how this can be achieved using a contemporary R programming library, AutoScore. This was applied on a secondary dataset from Mozambique to develop models to predict clinical outcomes in paediatric patients with general acute illness from a low resource setting.

Methods: The dataset involves routine data collected by the Health and Demographic Surveillance System in Manhica District, Mozambique, Africa. A total of 26538 inpatient admissions over the period of 1996-2024 were analysed. Domain knowledge was applied to select 24 clinical and 19 sociodemographic potential candidate predictors. In this study of paediatric patients, the clinical outcome of choice was length of stay >7 days. The AutoScore_1.0.0 framework was implemented to produce a parsimonious clinical risk score.

Results: The AutoScore programme selected eight variables including clinical malnutrition, weight for age z-score, age (days), standardised heart rate and respiratory rate, temperature, number of household members educated and number of days with fever. The clinical score generated ranged from 0 to 92 and had a similar performance to that of the conservative ML models. Model performance on the test set scored an area under the Receiver Operator Curve (AUROC) of 0.75 (95% CI: 0.72-0.78). At a cut-off score of 29 which is the best score balancing sensitivity and specificity, sensitivity is 62.6% and specificity, 75.6%.

Poster #10

Predicting Long-Term Cardiac Effects in Preterm Born Individuals

Judith Weng Zhu, BS[1], Christian Andrade Herrera, BS[1], Kara Goss, MD[2], Naomi Chesler, PhD[1], Colleen Witzenburg, PhD[3], Pim Oomen, PhD[1]

[1]University of California, Irvine
[2]University of Texas, Southwestern
[3]University of Wisconsin, Madison

Survival of preterm birth, defined as a delivery before 37 weeks of gestation, is a relatively recent medical advance. Because the oldest preterm born individuals are just reaching their 40s, our understanding of lifelong complications remains limited. A well-known complication is bronchopulmonary dysplasia (BPD), a chronic lung disease that occurs following ventilator use and oxygen therapy. BPD includes pulmonary hypertension which may alter right ventricular geometry and function [1,2,3]. However, the long-term cardiac effects remain poorly understood. Computational modeling has emerged as a tool to address this gap, allowing researchers to simulate long-term trajectories. In this study, we developed a computational model to understand the mechanisms of cardiac growth under healthy and BPD conditions.

We used our previously published computational model to simulate cardiac mechanics and hemodynamic circulation [4]. We modelled healthy somatic cardiac growth from birth through adulthood by scaling cardiac and circulation parameters based on body weight [5,6]. We then simulated the lung injury in BPD by increasing the parameter representing pulmonary vascular resistance.

Our model demonstrates that BPD induces right ventricular hypertrophy via increased wall strain, driven by elevated preload secondary to heightened pulmonary vascular resistance. Conversely, our model shows that BPD lowers left ventricular volume compared to the healthy case.

While several preclinical studies have measured changes in the right ventricle following BPD, less is known about the left ventricle. Our simulations showed contradicting growth patterns under BPD, where the right ventricle grows larger while the left ventricle smaller compared to somatic growth. Future work will focus on testing mechanisms other than, or in combination with, pulmonary vascular resistance, such as pulmonary vascular compliance and cardiac contractility.

Poster #11

Anticoagulant Treatment Effects on Survival in Antiphospholipid Syndrome Using Targeted Learning and Electronic Health Record Data

Omer Odabas, MS[1], Cyril Rakovski, PhD[1]

[1]Chapman University

Background: Antiphospholipid syndrome (APS) is a chronic autoimmune condition associated with recurrent thrombosis and long-term morbidity. Lifespan implications are substantial, particularly for younger patients who require lifelong anticoagulation. Although warfarin remains standard therapy, direct oral anticoagulants (DOACs) are increasingly prescribed despite conflicting evidence regarding long-term outcomes. Robust real-world survival analyses are needed to inform treatment decisions across the lifespan.

Methods: We conducted a retrospective cohort study using de-identified Oracle Health electronic health record data. Adults with confirmed APS diagnosis and new initiation of warfarin or DOAC therapy were followed for up to 365 days. The primary outcome was all-cause mortality. To estimate causal survival effects while addressing confounding, we implemented targeted maximum likelihood estimation (TMLE) with Super Learner ensembles for treatment and outcome modeling. Risk differences and risk ratios were calculated, and heterogeneity of treatment effects was explored across age strata to assess lifespan-related variation.

Results: Among 9,012 eligible APS patients, adjusted one-year mortality risk was higher in patients receiving DOAC therapy compared with warfarin. The average treatment effect corresponded to an absolute increase in mortality risk of approximately 1.3 percentage points for DOAC therapy. Age-stratified analyses suggested greater relative harm among older patients, while treatment effects were attenuated in younger subgroups. Survival curves derived from targeted learning estimates demonstrated consistently higher estimated survival probabilities for warfarin-treated patients across follow-up.

Poster #12

Heart Attack Incidence Rates and Post-Event Survival Across U.S. States, 2019–2024

Liora Mayats Alpay, MS[1], Reza Rezaei, MS[1], Eric Adams, MSc[1], Cyril Rakovski, PhD[1]

[1]Chapman University

Background: Real-world claims from IQVIA insurance claims data can help us understand patterns heart attacks (acute myocardial infarction (AMI)) occurrence across demographic and geographic strata.

Objective: We estimate heart attack incidence and survival rates and identify global and stratum-specific (sex, age, socio economic status, locations) longitudinal trends.

Data & Methods: We analyzed de-identified longitudinal claims (2019–2024; 2025 partial. AMI events were defined by ICD-10 I21 or I22** (primary diagnosis). Incident cases counted a patient's first AMI across the study window. Denominators were the number of unique patients with any claim in each year and stratum. Rates were summarized as Number/Denom × 100,000 overall, by sex, age bands (<18, 18–34, 35–49, 50–64, 65–84, 85+), and by state (five ranked by average rate).

Results: Yearly incidence and survival rates showed a marked decline in 2020 decline followed by a peak in 2022.

Conclusions: Claims-based monitoring reveals heterogeneity in AMI burden across demographic groups and states and provides an actionable baseline for subsequent causal analyses (e.g., TMLE) on post-AMI outcomes and care-pathway improvements.

Poster #13

SuperLearner Combined with LLM–Derived Features to Predict Severe Bleeding in Adults with Venous Thrombosis

Duncan Ofosu, PhD[1], Eric Adams, MSc[1], Mubarak Iddrisu, MS[2], Parthiv Sheth, MD[3], Cyril Rakovski, PhD[1]

[1]Chapman University
[2]University of Massachusetts, Boston
[3]University of San Diego

Background: Accurate prediction of outcomes in patients with thrombosis is critical for guiding post-discharge management. While structured Electronic Health Records (EHR) are typically used in predictive modeling, valuable clinical context remains embedded in unstructured medical notes. We carried out the most complete predictive analysis of thrombosis patient outcomes, using an extensive list of covariates that combined structured data with important variables extracted from unstructured medical notes. We evaluated predictive performance of major bleeding at clinically relevant follow-up intervals.

Methods: We analyzed adult thrombosis cases from the MIMIC-IV database. Discharge summaries were processed using fine-tuned large language models (LLMs) to extract key variables, including family history of clotting disorders and classification of thrombosis as provoked or unprovoked. These features were combined with structured clinical variables and used to train a Super Learner ensemble comprising statistical, machine learning and deep learning models. Predictive performance was evaluated for the outcomes: major bleeding events within 3 and 6 months. The SuperLearner model performance was assessed using negative log-likelihood (NLL) and area under the receiver operating characteristic curve (AUC-ROC) on a held-out test data.

Results: The Super Learner ensemble that incorporated both structured and LLM-derived covariates consistently outperformed models based on structured data alone. Inclusion of note-level features reduced NLL by 0.2% to 1.2% across outcomes. Improvements were consistent across cross-validation folds and different learner combinations.

Poster #14

Targeted Learning Analysis of Oral Anticoagulant and Risk of Thrombosis Recurrence in Adults

Duncan Ofosu, PhD[1], Mubarak Iddrisu, MS[2], Eric Adams, MSc[1], Parthiv Sheth, MD[3], Cyril Rakovski, PhD[1]

[1]Chapman University
[2]University of Massachusetts, Boston
[3]University of San Diego

Background: Comparative effectiveness of oral anticoagulants remains central to thrombosis management. While randomized trials have compared Vitamin K Antagonists (VKA) and Factor Xa inhibitors, real-world treatment effects require robust adjustment for clinical confounding. Electronic Health Records (EHR) contain both structured variables and clinically relevant information embedded in discharge summaries, which are often not used in causal analyses.

Methods: We analyzed adult patients diagnosed with thrombosis using the MIMIC-IV database. Discharge summaries were processed using fine-tuned large language models (LLMs) to extract clinically meaningful covariates, including thrombosis classification (provoked versus unprovoked) and family history of clotting disorders. These note-derived features were combined with structured demographic, laboratory, and comorbidity variables.

The causal effect of VKA (primarily warfarin) versus Factor Xa inhibitors (apixaban, edoxaban, rivaroxaban) on thrombosis recurrence was estimated using Targeted Maximum Likelihood Estimation (TMLE). Super Learner ensembles were used to estimate both the outcome regression and treatment mechanism prior to the targeting step, yielding doubly robust and efficient effect estimates.

Results: After adjustment for both structured and LLM-extracted covariates, VKA treatment was associated with a higher absolute risk of thrombosis recurrence compared with Factor Xa inhibitors. The estimated risk increase was 3.2% (95% CI: 1.8%–4.6%) at the earlier follow-up(3-months) horizon and 3.5% (95% CI: 2.1%–5.0%) at 6-months after diagnosis.

Poster #15

Comparison of Treatment Effect of Stimulant and Non-Stimulant ADHD Medication on Incident Rates of Cardiovascular Disease: A Targeted Learning Study Based on MIMIC-IV Data

Eric Adams, MSc[1], Duncan Ofosu, PhD[1], Ehsan Yaghmaei, PhD[1], Liora Mayats Alpay, MS[1], Peter Chung, MD[2], Alexandru Vajiac[1], David Licerio, MS[1], Mubarak Iddrisu, MS[3], Cyril Rakovski, PhD[1]

[1]Chapman University
[2]University of California, Irvine
[3]University of Massachusetts, Boston

Background: Cardiovascular risk needs be accounted for when deciding the right ADHD medication, as different medication classes may pose different risks to patients with certain demographic and pre-existing conditions.

This study pulled patients with ADHD from the MIMIC-IV, a longitudinal EHR database. Individuals with cardiovascular conditions present at the time of initial ADHD diagnosis were excluded.

The final analytic cohort includes 591 patients, of whom 467 (79.0%) received stimulant therapy and 124 (21.0%) received non-stimulant therapy at baseline.
Targeted Learning is the optimal causal inference method that focuses directly on the target estimand of treatment contrast while using flexible data-adaptive approaches to adjust for confounding, reducing bias, and increasing robustness.

Methods: We estimated causal effect of baseline stimulant and non-stimulant therapy on 1-year incident cardiovascular events using TMLE, reporting average treatment effect as a risk difference (RD) (ATE = E[Y_{A=1}] – E[Y_{A=0}]).

We used Super Learner ensemble to predict both cardiovascular outcome regression and treatment mechanism (propensity score), and then applied TMLE's targeting step to refine prediction and obtain doubly robust, efficient effect estimates with 95% confidence intervals.

Results: TMLE risk difference was -0.007, suggesting stimulant were associated with 0.7% lower adjusted 1-year risk than non-stimulants. The 95% CI was [-0.04055224, 0.02724406], indicating no statistical difference at 0.05 level.

Conclusion: Overall, we found no statistically clear cardiovascular RD between the two medication classes after adjustment. Next, we will improve confounder adjustment by incorporating ADHD severity derived from discharge notes. We fine-tuned severity classifier using reinforcement learning (GRPO) and will merge severity labels into structured cohort. We will rerun TMLE to see whether adjusting for severity changes effect estimate and confidence interval.

Poster #16

The Great Note Bake Off: Comparison of Quality between Clinician Written Documentation and Generative Artificial Intelligence Chart Summarization Tools

Kevin Truong, MD[1], Roderick Eguilos, DO[1], Kenneth Leung, MD[1], Abraham Lin, MD[2], Nina Zhu, MD[2], Tamera Hatfield, MD[1]

[1]University of California, Irvine
[2]University of California, Los Angeles

Background: Generative AI tools embedded in the EHR can automate inpatient summarization1, with potential to reduce documentation burden and clinician burnout2. Independent evaluation of documentation quality, reliability, and safety remains limited.

Methods: We conducted a pilot study of 30 inpatient discharges. For each chart, five outputs were saved: clinician discharge summary (DCS), Draft Hospital Course Summary (DHCS), and three Inpatient Insights (II) views. Six physician reviewers scored all outputs using a validated PDSQI-9 instrument (9–11 domains including accuracy, synthesis, internal consistency, hallucination, and bias)3. Two reviewers evaluated each chart. Paired t-tests compared DCS and DHCS across domains. Inter-rater reliability was assessed using intraclass correlation coefficients (ICC) and weighted kappa. Structured safety flags were recorded.

Results: Composite PDQI scores were higher for AI outputs than clinician DCS (DCS 4.39 vs DHCS 4.82; II range 4.71–4.83). In paired analyses, DHCS outperformed DCS in thoroughness (mean difference –1.07, $p<0.001$), usefulness (–0.43, $p=0.003$), organization (–0.32, $p=0.028$), synthesis (–0.42, $p=0.009$), and overall composite score (–0.36, $p=0.0003$). Inter-rater reliability for composite scores was moderate for DCS (ICC(2,k)=0.566) and II Recent (ICC(2,k)=0.619), with lower stability due to score clustering. AI errors noted in the summarizations included hallucination (10%), incorrect diagnosis/treatment (23%), misleading timeline (20%), and clinical omission (17%), with no bias flags.

Discussion/Conclusion: In this pilot, AI-generated hospital course summaries demonstrated better perceived documentation quality compared to clinician-authored summaries. However, clinically meaningful errors and timeline distortions were present, reinforcing the need for physician oversight. Generative AI may augment inpatient documentation workflows but cannot yet function independently without structured review safeguards.

Poster #17

Smartphones Driven Data Science in Healthcare as Roadmap for Inclusivity in Emerging Economies

Godfrey Ejiroghene Akpojotor, PhD[1], Oghenerurho Akpojotor, BSc[1], Moses Imah, MSc[2]

[1]Delta State University, Abraka, Nigeria
[2]Paul Odili College of Medicine Margaret Lawrence University, Abuja, Nigeria

The use of smartphones in healthcare delivery has been key in the Doctors without Borders (DWB) success in emerging economies. However, its use in health care research has been limited due to its limited hardware resources. The implication is that the emerging economies are alienated from the growing use of Data Science in healthcare. Interestingly, our project to achieve inclusivity and diversity of the Python which is the best programming language for these disciplines has been remarkably successful as we have achieved AI-Powered and cloud Computational science and engineering among our students and non-student learners. Starting with adopting QPython in Android phones as mobile computational laboratory for science and engineering anywhere, anytime and anyhow, we extended the project to include Pydroid for its best educational offline IDE experience with excellent library support. However, to bypass local hardware constraints of both QPython and Pydroid, we escalated the project to include Google Colab. Therefore, the purpose of this talk is to explore how this proven roadmap for inclusivity and diversity has been extended to healthcare. In particular, we consider here the level of its adoption in pediatric by developing the full stack, from modeling to production deployment using data in this discipline. In this way we have created the inception of adoption of Data Science in health care in emerging economies using smartphones which are relatively low cost, easy to maintain and low power requirements.

Poster #18

Predicting ICU Readmissions Using Deep Learning and Natural Language Processing Techniques: A MIMIC IV Data Analysis

David Licerio, MS[1], Cyril Rakovski, PhD[1]

[1]Chapman University

Background: Hospital readmissions are costly and preventable. Unplanned re-hospitalizations in 2004 was $ 17.4 billion. 27% of children's hospital readmissions were found to be preventable. The objective of this paper is to predict readmission from all-causes in the ICU within 30 days.

Methods: Using the MIMIC IV database, a predictive model is developed to predict ICU readmission. The populations of ICU admissions is 73,000, with 50,920 unique patients. Python is used for the predictive modeling. Patients are selected to be part of the modeling if they are readmitted for all-causes within 30 days of ICU admission. The predictive modeling combines several methods of Natural Language Processing including Latent Dirichlet Allocation, Latent Semantic Analysis, Singular Value Decomposition, and word embeddings using transformer based models. These methods are used to generate the input data for the model. The model uses a combination of several predictive models to make the final prediction for each patient.

Results: In its current state, the model has achieved .81 AUC.

Conclusions: The methods show that combing multiple NLP techniques to extract information from text data, and combining with chart data, lab data and prescription data can be used to identify patients at high risk of readmission. The model is very sensitive in that it often marks patients who are not readmitted to be readmitted, which is one aspect that can be improved in the predictive modeling.

Poster #19

Trends in RSV, hMPV and Co-infections, 2019–2024

Reza Rezaei, MS[1], Sun Yang, PhD[1], Cyril Rakovski, PhD[1]

[1]Chapman University

Background: Human metapneumovirus (hMPV) and respiratory syncytial virus (RSV) are major causes of adverse outcomes such as bronchitis, pneumonia, hospitalization, and death, yet large-scale comparative epidemiologic data across demographic and geographic dimensions in the United States remain limited. We provide hMPV, RSV, and hMPV&RSV co-infections characterization of incidence and mortality rates, demographic patterns, and geographic variation over the 2019-2024 time period.

Methods: We conducted a retrospective cohort study using the IQVIA insurance claims database, analyzing approximately 8 million individuals annually between 2019 and 2024. The hMPV and RSV cases were identified using the appropriate ICD-10 codes. Incidence rates per 100,000 population were calculated by year, age group, sex, and geography at state and 3-digit ZIP code levels. Survival was assessed at 1, 6, and 12 months following diagnosis.

Results: We identified 10,385 hMPV cases, 143,621 RSV cases, and 360 co-infections. Annual incidence ranged from 11.7 to 32.1 per 100,000 for hMPV and from 106.8 to 425.9 per 100,000 for RSV. Both viruses showed marked declines in 2020, followed by pronounced post-pandemic rebounds, with RSV peaking in 2022 and hMPV in 2023. RSV incidence was highest among children aged 0 to 2 years, whereas hMPV demonstrated a bimodal age distribution with elevated burden in young children and adults aged 60 years and older. Males experienced higher incidence for both infections. hMPV patients exhibited substantially greater comorbidity burden, including cardiovascular disease, diabetes, and chronic kidney disease. Geographic heterogeneity was pronounced at both state and local levels. Twelve-month survival was 99.0% for RSV and 96.6% for hMPV.

Poster #20

Identifying potential adiposity-driven cardiovascular outcomes in males and females with supervised machine learning methods in public access databases

Adhithi Kanthan Lakshmikanthan, MS[1], Madeline Chu[1], Derek Bivona, PhD[2], Kenneth C. Bilchick, MD, PhD[2], Pim Oomen, PhD[1]

[1]University of California, Irvine
[2]University of Virginia

Background: Cardiovascular diseases and obesity are two epidemics in the world today that are closely linked. Obesity, and visceral adipose tissue specifically, has been identified as a crucial risk factor for heart failure and other cardiovascular diseases. However, combinatorial exploration of sex, menopause status, adiposity, and cardiovascular biomarkers remains insufficient but addressable with public datasets.

Methods: We use supervised machine learning techniques to identify cardiovascular features of interest in a controlled access public database, the Jackson Heart Study (JHS) with ~3k patients. The dataset was first cleaned (handling missing values and correlated features), and two targets were tested (1) a binary classifier target (CVD History), and (2) a regressor target using left ventricular mass index and E/A (metric for diastolic function). Linear and non-linear tree-based models find important features. We visualize feature influence and non-linearity with SHAP and PDP plots, and tree-based splitting order with a single decision tree.

Results: (1) For the binary classifier target, logistic regression (0.78 accuracy) where fat, menopause status, sex, and LDL were top features. Random forest (0.94 accuracy) identified top features to be adiponectin, leptin, ldl, and fat. However, though accounted for in the algorithm, the target class used was highly imbalanced. (2) For the regressor target, linear model (R^2=0.69) identified top features to be age, ventricular diameter, presence of hypertrophy and AFib (p-value < 0.05), while XGBoost (R^2=0.67) identified age, ventricular diameter, presence of hypertrophy and obesity category. SHAP identified adiponectin and fat (fat not a factor in females only model).

Next steps: (1) We will improve model prediction and performance to elucidate menopause status in females, and influence of obesity with increased age. (2) We will investigate the MESA database similarly.

Poster #21

AI-Enabled Interactive Dashboard for RSV Surveillance and Nirsevimab Uptake in Pediatric Populations

Mingfei Dong, PhD[1], Britanny Winckler, MD[1], Allison Yeh[2], Lois Sayers, PhD[1]

[1]Rady Children's Health Orange County, formerly CHOC
[2]University of California, Irvine

Background: Respiratory syncytial virus (RSV) remains a leading cause of hospitalization among infants and young children, with prevalence varying across neighborhoods and geoeconomic strata. Following the introduction of Nirsevimab for RSV prevention in 2023, accurate monitoring of spatiotemporal disease trends and prophylaxis uptake has been increasingly important. Traditional static reports limit clinicians' ability to explore data and frame analyses. We aimed to develop an interactive, AI-enabled analytical platform to support on-demand RSV surveillance and data-driven decision-making within our institution.

Methods: We built a secure, web-based dashboard using Python that integrates encounter-level RSV data and Nirsevimab administration records with population-adjusted denominators at the ZIP Code Tabulation Area (ZCTA) level. The platform enables visualization of RSV trends and uptake across time and geography. An embedded AI-powered chatbot interface translates clinicians' natural-language questions into predefined statistical workflows with corresponding visualization specifications. The system executes these analyses upon request and returns standardized outputs, including effect estimates, test statistics, and p-values. All analyses are conducted on de-identified synthetic datasets within a secure institutional environment to ensure privacy and reproducibility.

Results: The dashboard supports flexible exploration of RSV burden and prevention uptake patterns across time and geographic regions. The AI agent enables clinicians to generate visualizations and reproducible statistical comparisons without programming expertise, reducing reliance on ad hoc analytic requests and shortening turnaround times while maintaining methodological rigor.

Discussion: This project demonstrates a scalable framework for AI-assisted, on-demand geospatial surveillance in pediatric health care and a modular design extendable to other pediatric population health initiatives.

Poster #22

From Structure to Function: Validating Automated fMRI Metrics of Pectus Severity Using Cardiopulmonary Performance and Dynamic Simulation

Mingfei Dong, PhD[1], Zhuo Chen, PhD[1], Lois Sayers, PhD[1]

[1]Rady Children's Health Orange County, formerly CHOC

Background: Traditional pectus severity indices rely on static geometric measurements, which do not fully depict physiologic impairment—the primary driver of symptoms and treatment decisions. Using a richly characterized CHOC cohort with thoracic imaging and cardiopulmonary exercise testing (CPET), we evaluated whether automated imaging-derived metrics better reflect functional impact rather than anatomical appearance alone.

Methods: A novel automated imaging analysis pipeline, recently developed by collaborators at the University of Genova, was applied to extract deformity-related thoracic metrics from imaging data. Aside from validating against legacy geometry-based indices, we also assessed clinical relevance through: (1) association with CPET outcomes, and (2) patient-specific cardiopulmonary simulation modeling cardiac compression and pulmonary restriction under physiological stress. These analyses evaluated whether imaging-derived features reflect patient-specific exercise limitation and simulated functional compromise associated with pectus.

Results: Automated imaging metrics showed meaningful associations with CPET indicators of exercise limitation and with modeled reductions in cardiac and pulmonary performance.

Discussion: Validating imaging metrics against physiologic performance shifts severity assessment from appearance-based to impact-based characterization.
Future work will extend this framework through smartphone-enabled 3D scanning for at-home monitoring, enabling:

- Day-to-day longitudinal tracking of chest wall trajectory
- Evaluation of response to conservative treatments (e.g., bracing, physical therapy)
- Monitoring of post-surgical recovery dynamics
- Richer, trajectory-based characterization of pectus severity beyond single timepoint imaging

Linking longitudinal structural change with functional indicators may support more personalized assessment of disease progression and treatment effectiveness.

Poster #23

Longitudinal Patterns in Major Depressive Disorder: A Nationwide Claims Analysis (2019–2024)

Damini Trehan[1], Cyril Rakovski, PhD[1]

[1]Chapman University

Background: Major depressive disorder (MDD) remains a leading cause of morbidity across the lifespan, with increasing concern regarding earlier age of onset. Large-scale administrative claims data enable characterization of real-world diagnostic patterns across developmental stages.

Methods: We conducted a longitudinal analysis using a nationwide IQVIA administrative claims database (2019–2024). MDD was identified using ICD-10 codes F32 (single episode) and F33 (recurrent). Annual unique diagnosed patients were calculated and stratified by age group (0–11, 12–17, 18–25, 26–49, 50+), sex, and geographic region. Temporal trends were examined descriptively across developmental strata.

Results: The dataset included 4,756,454 records representing 2,591,352 unique individuals with MDD. Annual unique diagnosed patients increased from 671,902 in 2019 to 842,534 in 2024, representing a 25.4% rise. Adolescents (12–17) experienced a 16.8% increase over the study period. Females consistently accounted for approximately 69% of diagnosed individuals across all study years. Geographic variation was observed across states.

Conclusion: Population-scale claims data reveal sustained, age-dependent increases in diagnosed MDD from 2019 to 2024. These findings provide a descriptive foundation for subsequent analytic modeling and targeted intervention strategies across developmental stages.

Poster #24

Longitudinal Cardiac Hemodynamics in a Hyperoxia-Induced Rat Model of Bronchopulmonary Dysplasia

Christian Andrade Herrera, BS[1], Judith Weng Zhu, BS[1], Fernando Trinidad[1], Polaris Pham[1], Colleen Witzenburg, PhD[2], Kara Goss, MD[3], Naomi Chesler, PhD[1], Pim Oomen, PhD[1]

[1]University of California, Irvine
[2]University of Wisconsin, Madison
[3]University of Texas, Southwestern

Medical advances in neonatal health care have greatly improved in the last two decades, leading to the increased survival rate of extreme preterm born infants (babies born between 24 and 28 weeks of gestation). Unfortunately, this population has an 80% increased chance of developing Bronchopulmonary Dysplasia (BPD). BPD is characterized by the underdevelopment of the lung, including a reduced number of alveolar sacs and pulmonary microvasculature, which leads to impaired cardiopulmonary function. It significantly affects the pulmonary arteries and can lead to secondary conditions, including right ventricular dysfunction and failure. However, the longitudinal impact of BPD throughout the lifespan is unknown, and somatic growth occurs in tandem with this disease burden. Our goal for this study was to determine BPD-driven cardiac hemodynamic changes throughout the lifespan and the implications for cardiac growth and remodeling. We used a standard BPD hyperoxia-exposed rat model to mimic extreme preterm birth. Newborn rat pups were exposed to 85% O2 for 14 days postnatally and allowed to mature in room air conditions at 21% O2 until postnatal day (P)21. Hyperoxia-exposed (Hx) male and female pups, as well as normoxia-exposed controls (Nx), were anesthetized for echocardiography of the left ventricle (LV) and invasive pressure measurements in the right ventricle (RV) before being humanely euthanized at P21, P35, and P90, respectively. Preliminary results indicate altered left ventricular function at P21, followed by stabilization by P35, with no significant changes by P90. RV pressure is elevated at P21 and remains elevated at P90. The data collected quantifies hemodynamic and developmental changes with BPD throughout the lifespan and will drive the development of computational models of single organs, such as the heart and lungs, and of coupled models that use combined data to assess growth and remodeling in this subpopulation.

Poster #25

Advancing Pediatric Precision Medicine Through Multimodal Imaging-Omics-Clinical Integration

Ijeoma Nnadozie, MS[1], Ying-Chu Lo, PhD[2], Louis Ehwerhemuepha, PhD[2]

[1]Keck Graduate Institute
[2]Rady Children's Health Orange County, formerly CHOC

Integrating multiparametric clinical imaging data (MRI, CT, X-ray) with multi-omics profiles and clinical documentation may enhance risk stratification and support precision medicine. The project aims to develop a scalable, flexible framework that harmonizes clinical imaging, omics data, and structured and unstructured clinical records (EHR data and clinical notes) into a unified analytical platform. Our methodology implements a standardized pipeline of pediatric low-grade glioma (pLGG). To begin the project, we will prioritize integrating image and transcriptomics data. For multi-omics datasets, we will use standardized pipelines and in-house classification tools. For imaging analysis, multiparametric MRI scans, including structural and diffusion-weighted sequences, undergo rigorous quality control procedures. We will apply machine-learning (ML) approaches, including support vector machines with feature selection and nested cross-validation, to classify imaging-omics subgroups. Survival and risk analyses will be conducted using penalized Cox proportional hazards models to quantify the predictive value of radiomic, molecular, and clinical features. Crucially, we introduce a novel Deep Learning (DL) architecture to automate feature extraction directly from MRI volumes, eliminating the subjectivity of manual feature engineering. By integrating these DL-derived features with clinical variables and transcriptomic data, we surpass the predictive performance of traditional models. This multimodal integration provides a superior tool for predicting progression-free survival (PFS) and treatment response, particularly in high-risk patients who are not adequately classified by clinical variables alone. In this framework, we provide a robust and flexible analytical pipeline for integrating multi-omics, imaging, and clinical notes. This framework enables biomarker discovery, identification of biological subgroups, and precision medicine research across diverse pediatric diseases.

Poster #26

Automated Detection of Velopharyngeal Port Dynamics from Nasopharyngoscopy Videos using Deep Learning

Xinfei Miao, BS[1], **Sierra Willens, MD**[2], Peyman Kassani, PhD[1], Shivang Trivedi[2], Ja'Neil Humphrey, MD[3], Jaime Perry, PhD[4], Miles Pfaff, MD[2]

[1]Rady Children's Health in Orange County, formerly CHOC
[2]University of California, Irvine
[3]Loma Linda University
[4]East Carolina University

Background: Accurate prediction of velopharyngeal port (VPP) closure is critical in assessing velopharyngeal insufficiency (VPI) because it allows for correct identification of VPP closure failure during oralized speech. Manual review and analysis of VPP is time-consuming, relies on expert interpretation, and is susceptible to subjective error. In this study, we applied a sequence-level artificial intelligence (AI) approach to NP videos to predict VPP closure status. This approach may help reduce analytical burden in multimodal VPI research and support clinical assessment and workflows.

Methods: Twenty-five NP video clips consisting of 78,849 frames of 567 sequences were analyzed. NPs were segmented and identified as VPP full closure (closed) versus patent (open). After preprocessing, the clips were analyzed by Naso-Net, a custom temporally aware convolutional neural network (CNN) that automatically learns spatial features within each frame and temporal features through their progression over time between frames.

Results: Naso-Net achieved a sequence-level accuracy of 80.22% and an area under the receiver operating characteristic curve (AUC) of 82.36% for VPP prediction. The learned frame weights allowed the model to focus on critical frames and captured the temporal progression of velopharyngeal movement during speech. Compared to standard ResNet50 with 73.44%, the proposed Naso-Net gained nearly 7% improved accuracy due to the inclusion of the temporal progression.

Discussion/Conclusion: The reliable performance of Naso-Net demonstrates that automated sequence-aware analysis in the presence of the temporal-aware modelling can improve VPP patency prediction. It holds promise for an improved VPI diagnostic process and better patient outcomes. This sequence aware design addresses limitations in earlier modelling or acoustic only studies. Future direction includes training Naso-Net to predict VPI through multicenter validation.

Poster #27

Using Caregiver-Logged Behaviors to Predict Daily Needs in Minimally Verbal Autism

Anuraag Tandon[1], Gauri Arora[1], Armaan Bassi[1], Niranjan Bhagwat[1], Moulik Chatterjee[1], Melina Drazien[1]

[1]University of California, Los Angeles

Minimally verbal autism affects 30% of individuals on the autism spectrum, requiring caregivers to infer daily needs from indirect cues. Existing tools are invasive, threaten privacy, and lack continuous tracking. We hypothesize that systematic analysis of structured, timestamped caregiver logs of activities, behaviors, and vocalizations will reveal reproducible patterns predicting unmet needs before distress escalates.

We developed RoutineLog, a mobile app where caregivers log activities, vocalizations, and behaviors. Caregivers initially input the most frequent activities significant to their child, creating color-coded buttons that log behavior start and end times. These form longitudinal event timelines that preserve order, timing, and co-occurrence across the day. Simulated data is generated by extrapolating from the observed behavior frequencies of a single patient. In each independent trial, the frequencies will be randomly reassigned across the predefined behavior categories, with additional random error incorporated to reflect natural variability. Through contingency tables and chi-squared tests, we will train Higher-Order Markov Models and Sequential Transformers to predict upcoming needs from prior event sequences. The system is specifically designed to support personalized care, individualized alerts, and AI-driven caregiver guidance.

We expect dense caregiver-logged time series to reveal consistent links between behaviors and core needs, identifying reliable predictive signals and their lead times. Machine learning models trained on these sequences will generate personalized predictions of emerging needs or escalating distress, enabling proactive caregiving rather than reactive responses.

Using simulated caregiver-logged data to preliminarily test our model, we aim to develop a scalable, non-invasive tool that translates everyday observations into predictions, improving quality of life for minimally verbal individuals and their caregivers.

Poster #28

Reproducible Exploratory Data Analysis Frameworks for Pediatric and Lifespan Clinical Research

Leon Shpaner, MS[1], Oscar Gil, MS[2]

[1]University of California, Los Angeles, Los Angeles, CA
[2]University of California, Riverside, Riverside, CA

Background: Exploratory data analysis (EDA) is a critical step in clinical data science workflows, yet it is often conducted in ad hoc and poorly documented ways. In pediatric and lifespan research, where datasets frequently contain heterogeneous cohorts, missing values, and sensitive attributes, inconsistent exploratory analysis can lead to downstream modeling errors and reduced reproducibility.

Methods: We developed the EDA Toolkit, an open-source Python library designed to standardize and accelerate exploratory data analysis workflows. The framework integrates automated data profiling, visualization, statistical summaries, and preprocessing tools within a reproducible environment. The toolkit is designed to work alongside standard scientific Python libraries and supports transparent documentation of data transformations and analytic decisions.

Results: The toolkit enables analysts to rapidly generate structured exploratory reports, identify cohort imbalances, evaluate feature distributions, and document preprocessing steps in a consistent manner. By automating common exploratory tasks while preserving analyst control, the framework reduces variability across projects and facilitates collaboration among research teams.

Discussion: Reproducible EDA frameworks can improve transparency, efficiency, and methodological consistency in pediatric and lifespan data science projects. Tools that standardize early-stage analytic workflows help ensure that downstream modeling and inference are based on well-documented and interpretable data preparation processes.

Event Sponsors

Acadia Pharmaceuticals is a biopharmaceutical company focused on developing innovative therapies for neurological disorders and rare diseases with significant unmet need. The company is driven by a patient-centered approach, working closely with patients, caregivers, and healthcare providers to advance meaningful treatments. With a proven portfolio and a growing pipeline, Acadia continues to lead in bringing new therapies to underserved populations.

Our AWS mission is to enable access and delivery of person-centered healthcare, drive improved outcomes at a lower cost, and accelerate the digitization and utilization of healthcare data.

The Schmid College of Science and Technology at Chapman University guides students as they push the boundaries of knowledge to drive discovery through collaboration and innovation. Our undergraduate and graduate STEM students engage in interdisciplinary research through mentorship, advanced lab techniques, and remote fieldwork, providing real-world experience that helps them achieve their goals.

Chiesi USA, is a specialty pharmaceutical company focused on advancing innovative therapies in neonatology, rare diseases and critical care. With a strong commitment to improving patient outcomes, Chiesi partners with healthcare professionals to deliver clinically differentiated treatments supported by research, education, and real-world evidence. In the neonatal space, Chiesi is dedicated to supporting the most vulnerable patients through therapies that promote lung development and improve survival outcomes.

Community Action Partnership of Orange County (CAP OC) supports individuals and families facing economic hardship by addressing immediate needs such as food access, housing support, and utility assistance. By pairing critical relief with long-term programs including financial literacy, youth development, and home weatherization, CAP OC empowers communities to build pathways to lasting stability. Through strong partnerships with local stakeholders, CAP OC drives meaningful, sustainable change across Orange County.

Event Sponsors

Databricks is the Data and AI company. More than 20,000 organizations worldwide — including Intermountain, AT&T, Bayer, Block, Mastercard, Rivian, Unilever, and over 60% of the Fortune 500 — rely on Databricks to build and scale data and AI apps, analytics and agents. Headquartered in San Francisco with 30+ offices around the globe, Databricks offers a unified Data Intelligence Platform. This platform includes Agents, the ability to ask questions of the data for reporting, building applications, data ingestion, and a governance catalog—all on the Databricks Lakehouse.

Rubrik, the Security and AI company, operates at the intersection of data protection, cyber resilience and enterprise AI acceleration. The Rubrik Security Cloud platform is designed to deliver robust cyber resilience and recovery including identity resilience to ensure continuous business operations, all on top of secure metadata and data lake.

Slalom is a fiercely human business and technology consulting company that leads with outcomes to bring more value, in all ways, always. From strategy through delivery, our agile teams across 52 offices in 12 countries collaborate with clients to bring powerful customer experiences, innovative ways of working, and new products and services to life.

U.S. Hunger is a 501(c)(3) nonprofit that has activated more than 1,000,000 volunteers to distribute over 200,000 million meals across the globe. U.S. Hunger addresses hunger by engaging volunteers, educating on the root causes of food insecurity, and creating access to nutritious meals for families and individuals in need.