National Institutes of Health
Office of Strategic Coordination—The Common Fund

Precision Medicine with AI: Integrating Imaging with Multimodal Data (PRIMED-AI) Workshop

March 11-12, 2025

Revised on May 5, 2025

This meeting summary was prepared by Rose Li and Associates, Inc., under contract to the Office of Strategic Coordination (OSC) of the National Institutes of Health (NIH). This summary represents the opinions and perspectives of the workshop participants, which do not necessarily reflect the perspectives of the NIH, its staff, the Department of Health and Human Services (HHS), the federal government, or the goals or structure of the PRIMED-AI program. Contributions to this summary by the following individuals are gratefully acknowledged: Jessica Tennis, Bethany Stokes, Nancy Tuvesson.

Table of Contents

Acronym List

ACR	American College of Radiology
AI	artificial intelligence
fMRI	functional magnetic resonance imaging
MIDRC	Medical Imaging and Data Resource Center
NEI	National Eye Institute
NIBIB	National Institute of Biomedical Imaging and Bioengineering
NIH	National Institutes of Health
PRIMED-AI	Precision Medicine with AI: Integrating Imaging with Multimodal Data

Meeting Summary Day 1

Welcoming Remarks

Kaitlyn Browning, PhD, Office of Strategic Coordination, National Institutes of Health (NIH)

Dr. Browning welcomed participants to the Precision Medicine with AI: Integrating Imaging with Multimodal Data (PRIMED-AI) workshop held by NIH Common Fund. This 2-day workshop identified key opportunities, complexities, and challenges in the emerging space of AI for precision medicine. The meeting consisted of four sessions focused on (a) developing algorithms that leverage multimodal data to solve clinical needs, (b) accessing and preparing AI-ready imaging and multimodal data to enable interoperability, (c) validating and implementing clinical imaging and multimodal data, and (d) ethical considerations for using imaging based multimodal AI clinical decision support tools.

NIH Institute Directors Introduction

Michael Chiang, MD, National Eye Institute (NEI), NIH; Bruce Tromberg, PhD, National Institute of Biomedical Imaging and Bioengineering (NIBIB), NIH

Drs. Chiang and Tromberg highlighted the potential promise of multimodal AI models for precision medicine as well as barriers to implementing these AI models. Health data are inherently multimodal, with various data types interacting to influence patient health. Large-scale, multimodal health databases that include imaging data have created a pathway to multimodal health AI development. These AI models can provide a more complete picture of patient health, which can support the personalized medicine approaches that communities need. Multimodal AI models can also improve imaging data utility by enhancing contrast resolution, reducing hardware requirements, and providing new insights.

While AI models and related devices show promise in research, their broad implementation requires careful consideration of unresolved questions, such as their specific functions and clinical uses. Questions remain about how to best support and enable AI infrastructure(s), interoperability, data use, data sharing, and billing. Adding to the complexity of AI implementation, AI algorithms must be redeveloped and revalidated across various implementation sites. This workshop sought to address these barriers by bringing together enthusiastic experts across and beyond NIH.

Session 1: Developing Algorithms That Leverage Multimodal Data to Solve Clinical Needs

Introduction

Co-Chairs: Caroline Chung, MD, PhD, MD Anderson Cancer Center; Leanne Williams, PhD, Stanford University

Issues with Variation in Data Precision

Biomedical data vary in their precision, granularity, and timescales of measurement, which complicates the process of integrating multimodal datasets for biomedical AI model development. Variations in data precision and collection practices can also lead to technical challenges in interpreting AI model outputs. AI models point to connections in biomedical data that go beyond those that humans can identify, but some of these connections may be *technical* rather than *biological*. The co-chairs highlighted an AI model developed to determine whether lungs in images were healthy or infected with SARS-CoV-2. Because most training images depicted infected adult lungs and healthy children's lungs, the model conflated age with being infected with SARS-CoV-2, thereby failing to perform its task accurately. This example highlights the technical challenge of identifying which AI-identified connections are biological and should therefore be investigated further.

Co-chairs emphasized the importance of collecting metadata, or the contextual information surrounding data collection. Including metadata in the AI development process enables human interpreters to discern between biological AI-identified connections and those that are purely technical, or possibly irrelevant. Panelists agreed that rigorous data governance should include requirements for consistent metadata collection.

Precision Medicine for Early Detection and Heterogeneity

Many current diagnostic tools assume patients are representative of an average disease state. AI models can improve on this assumption and aid in moving toward precision medicine—which can enable early detection and heterogeneity stratification based on unique features of a patient within a larger population. The co-chairs described an AI model trained on functional magnetic resonance imaging (fMRI) and diagnostic data that learned to separate individuals with depression from healthy controls, enabling early detection of at-risk individuals and informed diagnostic decisions. A second example noted how an AI model learned to stratify four subtypes of depression based on fMRI and diagnostic data. Depression confers a large burden of illness yet remains difficult to understand physiologically without AI-identified insights (e.g., the aforementioned subtypes). These subtypes separated patients who responded better to behavioral interventions from those who responded better to medication intervention, highlighting AI models' potential utility in precision medicine.

Panel Discussion

Panelists: Conor Liston, MD, PhD, Cornell University; Faisel Mahmood, PhD, Harvard University; Ziad Obermeyer, MD, University of California, Berkeley School of Public Health; Cui Tao, PhD, Mayo Clinic; John Paul (JP) Yu, MD, PhD, University of Wisconsin

Prioritizing AI Functionality

Broadly, AI models can perform three types of functions. First, AI models can fuse orthogonal data modalities to predict treatment responses and other outcomes. Second, AI models can combine data across modalities to improve feature representation within those data. Third, AI models can predict data in one modality based on data from other modalities, which may be especially useful when a particular modality is difficult to measure or not readily available. Across these three functions, panelists agreed that the most useful AI models will be those that answer important clinical questions that are difficult for clinicians to solve and that AI models can solve readily. AI developers can maximize benefits to patient care by identifying questions that meet both criteria.

Need for High-Quality Benchmarks

Discrete, high-quality benchmarks can be used to test and assess AI model performance in a systematic, standardized way. Panelists emphasized the need for a wide range of high-quality benchmarks to evaluate AI models' performance (especially for foundation models and multimodal generative AI models). Panelists noted that current health care applications lack these benchmarks. Some benchmark evaluation mechanisms will ensure that models trained on one type of data can perform the desired analysis on a new data type. Dataset benchmarks that assess the quality of data collection methods and data integration methods will also help standardize and ensure the quality of multimodal AI models. Panelists also emphasized that these benchmark assessments must be implemented frequently and consistently to gauge AI model quality.

Addressing and Leveraging Noisy Data

Some evidence suggests that AI models trained on noisy data may be more accurate in analyzing low-quality, real world input data than models exclusively trained on high-quality data. Data with a wide range of qualities may also be useful for discovering novel data connections with AI, and generalized AI models that contrast across one or two modalities can improve pattern recognition for each of those modalities. However, high-quality data are ideal for making clinical decisions about a single patient and for highly targeted or single-purpose AI models. Governance that specifies metadata standards can improve data quality for clinical decision making and targeted AI models. Additionally, novel AI models may be able to improve the labeling of patient imaging if these models can be trained on longitudinal metadata collected on the imaged patients.

AI Connections and Human Understanding

AI models can often identify data correlations that human researchers cannot. One of the panelists highlighted a multimodal AI model that they and their team are developing that analyzes electronic health record (EHR) and electrocardiogram data for patients visiting the emergency room. They built the model to identify which patients may have had strokes and it has identified correlations beyond those that are classically associated with stroke. Their team observed that AI-identified correlations can be used to understand individualized treatment effects when the biology underlying the correlations is understood. Other panelists stated that AI-identified correlations may benefit patient care *before* the underlying biology is understood and highlighted the example of neuropsychiatry (i.e., a complex field that has benefited and progressed through AI advancements).

Team-Based Approach to Implementing Multimodal AI

Panelists agreed that implementation of AI in clinical care should include a team of medical specialty providers, computer and applied scientists, researchers, and technologists. Such a team-based approach would ensure that patients benefit from the comprehensive analysis of their multimodal health data and AI models' insights into that data. AI models could be implemented in a manner akin to a tumor board, wherein experts across a range of relevant disciplines collaborate to discuss potential treatment options. Panelists agreed that, to contribute to these teams, medical students should be trained on AI's basic operations and how to use AI models; the level of training a medical student needs will depend on their own goals and intended use for AI. For example, a biomedical researcher will likely need to know more about AI models than a primary care provider.

Questions and Answers

Democratizing Health Care Access with AI Models

AI models that predict data in one modality based on another data modality may be able to expand data collection access. For example, an AI model based on mobile phone images of the retina is being evaluated for its ability to predict critical heart failure. Using this and similar strategies, inexpensive tests can be used to predict the outcome of costly tests. AI models can leverage datasets that are more accessible, cheaper, easier, and faster to analyze, which enables providers to expedite and democratize patient care.

Task-Specific AI Models

Panelists cautioned that generative AI often hallucinates false information and is therefore very difficult to use effectively for deterministic and definitive testing purposes. Instead, panelists highlighted the benefits of task-specific AI models that can be evaluated for their ability to accomplish their task. If a model is task specific, a developer could successively add new modalities and thereby rigorously test how much benefit each new data modality confers to the model's performance.

Session 2: Accessing and Preparing AI-Ready Imaging and Multimodal Data to Enable Interoperability

Introduction

Co-Chairs: Curtis Langlotz, MD, PhD, Stanford University; Tanveer Syeda-Mahmood, PhD, IBM

Acquiring Large Volumes of Data

Acquiring the large volumes of nationally representative data needed to train AI models is a common challenge for developers. Health care data are especially difficult to gather in large quantities because they are more inherently multimodal and often occur in image-based more than text-based formats, which expands the dataset size. Although some vendors make health care data available, these data are often too expensive for most AI developers to procure. Synthetic data generation strategies can help expand training data. Other creative strategies like contrastive self-supervision, whereby an AI model is trained on both matched and mismatched reports about health imaging data, can reduce the volume of data needed to train AI models.

Data Curation

Data curation involves both deidentifying training, test, and validation data to maintain patient privacy, and labeling data to enable AI models to learn. Deidentification is a highly complicated process. At this time, AI models for deidentifying patient data can be freely obtained by academic researchers and are available for commercial use under nonexclusive licensing. Establishing appropriate data labels and then manually applying these labels to the large quantities of data needed to train AI models is not feasible, indicating a need for high-precision AI labeling tools.

Panel Discussion

Panelists: Maryellen Giger, PhD, University of Chicago; Benjamin Haibe-Kains, PhD, University of Toronto; Greg Sorenson, MD, DeepHealth, Inc.

Technical Data Sharing Solutions

Panelists affirmed the need for large volumes of health data to develop multimodal AI models and observed that *technical* and *process* solutions are both necessary to promote data sharing. Technical data sharing solutions include infrastructure expansion and development, data harmonization, and building algorithms that can interpret multiple data types and qualities. Many technical solutions for data sharing, such as tokenization to link multimodal health care data, are currently only available to well-resourced organizations because of their cost. The Medical Imaging and Data Resource Center (MIDRC) is an open-source data commons that provides harmonized, deidentified, searchable datasets to researchers at cost, helping investigators form relevant datasets.

Process Data Sharing Solutions

Process data sharing solutions are those that depend on effective data collection and curation regulations to ensure patient privacy when health care data are linked across modalities.

Current process solutions are more effective within institutions than across them. Cross-institutional process solutions will be essential to safe and broad data sharing. Panelists urged policymakers and data-collecting organizations to (a) create incentives for data sharing and (b) lower the costs of building federated datasets to facilitate robust AI development. Data-collecting organizations are more likely to trust platforms that are built by a broad community with established data sharing principles, such as MIDRC, to protect privacy, while they may not trust single academic or for-profit institutions in the same way.

Sustainability

NIH data sharing requirements promote sustainability of AI models by ensuring that AI models can be continuously assessed and improved with new datasets. A sustainable data ecosystem requires both centralized and distributed datasets to ensure that AI models are robust to a wide variety of data. To promote a sustainable data ecosystem, panelists called for investing in both low-resourced organizations to enable them to develop their data collection and curation infrastructure and high-resourced organizations to enable them to develop guidance on best practices for data sharing. MIDRC, as an example, supports a sustainable data ecosystem by maintaining a searchable dataset marketplace for use in AI development. MIDRC keeps its own datasets sustainable by sequestering 20% of the data it receives for in-house evaluations that seek to determine whether externally developed AI models are fit-for-purpose in the models' intended populations.

Questions and Answers

Costs of Data Ownership

Panelists raised concerns about private ownership of health care data. Although some panelists suggested that individual patients could be paid for their health care data, others observed that this payment model could create perverse incentives for seeking unnecessary care. Additionally, many health care AI development companies cannot afford to pay the numerous individuals who would need to contribute to a functioning health care dataset. Health care data are dramatically more expensive to collect than other personal data, such as internet use data purchased by online advertisers, increasing the cost of creating large health care datasets. Organizations that create large health care datasets incur costs to curate and harmonize the data to share them, and panelists suggested that funding agencies could create grants to offset these costs.

Benchmark Datasets

AI models must be continuously monitored and evaluated for temporal and geographic data drift to ensure that they accurately account for population-level data correlations as both the models and the U.S. population evolve over time. Conducting these evaluations requires a benchmark dataset that can be used as a testing input. These benchmark datasets must evolve over time, both to reflect changes in the population and to avoid learning the test, a phenomenon by which AI models can accurately assess benchmark datasets but not novel input data. Therefore, benchmark datasets must be stratified and assessed for their population representativeness. One panelist noted that his team recently conducted an independent assessment of 10 commercial AI models and found that none of the AI models was as effective on a benchmark dataset as their developers presumed, underscoring the importance of independent AI model assessment.

Data Curation and Harmonization

Whereas *training* data should vary in quality to ensure AI models that are robust to data quality fluctuations, *test* data should be high-quality to effectively evaluate the model. A combination of real and synthetic training data can promote AI model robustness. Rather than collecting only small, high-quality datasets with highly systematic collection protocols, AI developers should endeavor to collect and harmonize as much training data as possible to generate unforeseen discoveries and ensure the most robust AI models. This method requires detailed knowledge of the quality and provenance of the data. To accurately assess data quality, AI developers must perform effective data harmonization and curation, including ensuring retention of critical metadata.

Meeting Summary Day 2

Introduction to Day 2

Karlie Sharma, PhD, National Center for Advancing Translational Science, NIH

Dr. Sharma reviewed the PRIMED-AI workshop's first two sessions, highlighting from Session 1 the need for both data and metadata to ensure meaningful insights from AI-identified correlations across datasets that vary in granularity. From Session 2, she highlighted data sharing considerations related to data governance, privacy, and benchmarking as well as the financial resources, infrastructure, and inclination needed to share datasets.

Session 3: Validating and Implementing Clinical Imaging and Multimodal AI

Introduction

Co-Chairs: Karandeep Singh, MD, University of California, San Diego; Pamela Woodard, MD, Washington University in Saint Louis

Data and Algorithmic Monitoring Ecosystem

Successful implementation of AI models in clinical care settings requires careful consideration of a range of health care data and algorithmic variables. Considering how data and algorithmic monitoring will be conducted, including monitoring frequency, designation of responsible party (e.g., developer, vendor, user, governance groups, federal agencies), and methodology, is also critically important. The American College of Radiology (ACR) Recognized Center for Healthcare-AI is an example of a governance group that assesses AI and grants seals of recognition to centers using AI. In contrast, the Digital Medicine Society approves products rather than centers.

Development of AI Models

Governance groups institute reporting guidelines for each step of the AI model development process. However, the process often results in AI models that have issues with fit-for-purpose, validation, implementation, adoption, and generalizability. Panelists noted that a major factor in AI model governance is that AI models that are well-researched are not widely implemented, while AI models that are widely implemented are not well-researched. The co-chairs presented examples of AI models at either end of this spectrum and observed that, in addition to being well-researched during development, AI models must be monitored where they are implemented because AI model performance depends greatly on the context of deployment. However, the best practices for monitoring models remain an open question.

Panel Discussion

Panelists: Tessa Cook, MD, PhD, University of Pennsylvania; Woojin Kim, MD, ACR Data Science Institute; Matthew McDermott, PhD, Columbia University; Shannon McWeeney, PhD, Oregon Health and Science University

Challenges with Validation and Transportability

Validating each AI model deployment requires additional levels of both data and transparency in the training and validation data. Further, validating each AI model deployment requires significant infrastructure in clinical settings and may seriously impact clinical workflows, both of which are expensive.

AI models effective at one site are not always effective across sites (*i.e.,* an issue of transportability). While to date validation studies have typically been single-site, retrospective studies, panelists suggested that multi-site, prospective studies would be more effective in determining both validity and transportability. Panelists observed that effective tools to measure AI model transportability have not yet been developed. Complicating multi-site evaluations further, some developers deploy 'model recipes'—series of data transformations applied to datasets—rather than models, making it challenging to trace errors across different implementations

Run Charts and Metrics to Measure AI Implementation

Panelists agreed that end users and developers should share responsibility for validating AI models. Many clinicians are already overburdened and cannot take responsibility for checking that AI models retain validity over time. Panelists proposed that AI run charts, automated systems that alert clinicians when errors arise with AI models, could be implemented alongside AI models to ease the burden of monitoring AI model validity. Industry groups and societies can also share responsibility for AI model validation. For example, ACR's Assess-AI registry continuously compares AI model output and radiology reports to evaluate concordance, which can help end users determine when to request software updates from vendors.

Panelists also discussed the metrics by which AI model implementation should be judged. Although the goal of introducing a new tool into the clinic is clinical utility, clinical utility cannot be reduced to a number, complicating the ability to measure and therefore to value appropriately. Users of AI models should also consider the efficiency, clinical accuracy, workflow metrics, and return on investment of any implemented AI model. Although health data are inherently multimodal, panelists recommended that each modality added to an AI model should be tested against these metrics to determine individual contributions to the desired outcomes.

Variation Across Clinical Fields in AI Implementation

Some specific AI models, including models for pathology and screening for diabetic retinopathy, can run more autonomously than other models because these models have shown consistent accuracy. The IDx diabetic retinopathy screening model is the first AI model approved by the U.S. Food and Drug Administration to run autonomously. Even for these models, panelists cautioned that a human-in-the-loop testing approach is always necessary, although more rigorous testing is needed for newer models.

Questions and Answers

Next Steps for Promoting AI Validation

Panelists considered funding, infrastructure, and technical improvements that could foster AI validation. Because funding agencies seek innovative studies, academia or industry may be more appropriate for funding validation studies that are important but not often innovative.

Modernizing health care infrastructure across clinical sites could support validating AI models by enabling sites to produce consistent data for training and testing the models. Panelists also encouraged AI developers to consider how to include metadata in AI models, including by developing AI models that can extract data from EHR charts.

Patient Use of Generative AI Models

Panelists cautioned against patients using generative AI models for medical advice, both because generative AI models often hallucinate false information and because submitting personal data to these models risks privacy concerns. Patients often use generative AI models because they are more accessible than health care providers, showcasing overall health care access issues. Panelists highlighted the dangerous accessibility gap that could be established if some patients can access human therapists while others rely on AI model chatbot therapists. Because removing generative AI from patients is likely infeasible, one panelist also encouraged providers to direct patients to the best tools and advise them on their use.

Learning from Site-to-Site Variation

Differences in model performance across AI model implementation sites can illuminate site differences and potentially biological pathways. Some ongoing projects seek to investigate the potential for site-to-site variation. Data must be integrated precisely to ensure that AI model performance differences across sites are truly biological rather than technical.

Session 4: Ethical Considerations for the Use of Imaging-Based, Multimodal AI Clinical Decision Support Tools

Introduction

Chair: Judy Gichoya, MD, MS, Emory University

AI-Identified Imaging Biomarkers

AI models have demonstrated the potential to identify imaging biomarkers for a specific disease that humans are unable to detect. One AI model can understand features of breast imaging to produce breast cancer risk scores based on imaging data alone, which can be helpful for patients who are unable to provide family history information. Finding biomarkers with imaging holds promise to be more cost-effective than the genomics-based biomarkers that represent the current state of the art. AI-identified imaging biomarkers could also be used to identify subgroups of patients from a personalized medicine perspective, stratifying beyond traditional categories such as racial and ethnic groups to predict patients' response to treatment options. One study of breast cancer oncotypes identified 110 treatment pathways for patients with breast cancer diagnoses and highlighted the need to find optimum treatments for all patients.

Technological Progress Toward Combining Datasets

The chair outlined methods for combining datasets across modalities, noting that technological progress in dataset combination has led to increasingly precise stratification of patients' risk scores and disease states. For example, a multimodal model of the risk of endometrial cancer extracts patches of whole-slide images, merges information from those images with anatomical information, and assigns a risk score that correlates strongly with genomic risk scores and has been rigorously externally validated. Other technologies, such as automatic labeling at different size scales, can also increase the amount of data available for use in AI models. AI models are highly sensitive to misleading or missing labels on imaging data, underscoring the importance of accurate and thorough labeling.

Panel Discussion

Panelists: Paul Gross, Cerebral Palsy Research Network; Maia Hightower, MD, MPH, MBA, Equality AI; Xiaoqian Jiang, PhD, UTHealth

Privacy Risks of Multimodal Data

Combining data across modalities increases risks to patient privacy and complicates the processes necessary to protect that privacy. Imaging data are often more difficult to deidentify than other patient data because imaging includes key anatomical signatures of patients. Even when data are deidentified, combinations of rudimentary data quickly lower the barrier to reidentifying patients. Each health data modality implies its own set of risks, not only to patient privacy but also for imputing bias into the AI models they are used to train. These risks must be carefully considered and addressed to mitigate risks before combining data. To reduce risk, AI developers can use federated data streams, which allow generalized linear mixed effects AI models to learn from patterns in data. However, preparing data for use with other federated data requires a great deal of data curation and additional computational power compared to other AI models.

Ethics of Patient Consent

Panelists raised concerns about patients' ability to consent to data collection for AI model development. Because of AI's inherent continuous development, current consent practices cannot anticipate every potential use of patient data in AI models. Outside of the AI model context, many patients feel pressured to consent to data collection procedures, such as routine imaging, because alternatives to data collection leave patients without necessary medical advice. Use of AI models exacerbate these concerns by reducing the ability to protect patient privacy and obfuscating how patient data will be used. Panelists agreed that the benefits of discovery currently outweigh the risks to patient privacy but cautioned that reidentification may lead to harm to individuals, such as targeting by insurers, which may alter the risk-benefit analysis.

Governance

AI regulations are the minimum requirements for AI use, such as abiding by the Health Insurance Portability and Accountability Act, and, like many policies, are often developed slowly over time. AI models develop rapidly and therefore require more flexible and aspirational AI governance principles, which can guide the development and use of AI models. Implementing AI evaluations across every deployment site is an important aspect of AI governance that will help AI users ensure that AI models serve the needs of the local communities. However, many local healthcare sites that may implement AI models lack the necessary capacity to develop and implement rigorous, evolving evaluations, so panelists highlighted the need to identify responsible parties for these governance measures.

Questions and Answers

Implementing Reliable AI

Transparency throughout the AI development process is essential to the development of reliable models. Implementing high-quality benchmarking can improve transparency and help ensure that multimodal AI models are deployed safely. Implementing AI models in many small steps (i.e., incrementalism) rather than all at once may also help improve public trust in AI models by mitigating serious privacy and health risks due to mistakes made by AI models.

Addressing Uniformity and Duplication in Health Datasets

Due in part to differences in health care access, existing health datasets tend to skew toward a wealthier, sicker, and homogenous population than the general population. Skewed training, testing, and validation data are highly detrimental to assessing AI model performance. Synthetic data can help resolve this problem by expanding the currently underrepresented segments of the population more rapidly than structural or policy changes. Duplicate data in federated datasets can also lead to skewed datasets and overrepresentation of some characteristics and health conditions. Privacy-preserving data linkage can help remove duplicate data from various sites. Although data duplication is not specific to multimodal data, multimodal data can facilitate linking data across modalities with high confidence and therefore increase the chances of successfully identifying duplicate data for removal.

Public–Private Partnerships to Develop Ethical AI Models
Panelists agree that public–private partnerships are the ideal spaces for AI development to ensure ethical AI development. These partnerships represent a broad range of interests that may not benefit from typical private sector revenue models yet may enable the pursuit of innovative ideas that address major unmet needs.

Concluding Remarks

Michael Chiang, MD, NEI, NIH; Bruce Tromberg, PhD, NIBIB, NIH

Drs. Chiang and Tromberg thanked panelists, participants, and organizers and highlighted that all four sessions included discussions of the importance of data sharing for developing robust AI models. In 2023, NIH implemented a data sharing requirement for funded research, recognizing the importance of data sharing for accelerating scientific discovery. NIH aims to incentivize data sharing further with its Data Sharing Index Challenge, which seeks proposals for a metric, the Data Sharing Index or S-index, that measures how effectively researchers share data.

To stay up to date on new PRIMED-AI activities and announcements, join the listserv: https://go.nih.gov/Z6e8AlR.

Appendix A: Agenda

Day 1: March 11, 2025

11:00 – 11:15 am	**Welcoming Remarks** *Kaitlyn Browning, PhD, Office of Strategic Coordination, National Institutes of Health (NIH)*
11:15 – 11:30 am	**NIH Institute Director Introduction** *Michael Chiang, MD, National Eye Institute, NIH; Bruce Tromberg, PhD, National Institute of Biomedical Imaging and Bioengineering, NIH*
11:30 – 1:45 pm	**Session 1: Developing Algorithms that Leverage Multimodal Data to Solve Clinical Needs** *Moderator: Keyvan Farahani, PhD, National Heart, Lung, and Blood Institute, NIH* *Co-Chairs: Caroline Chung, MD, PhD, MD Anderson Cancer Center; Leanne Williams, PhD, Stanford University* *Panelists: Conor Liston, MD, PhD, Cornell University; Faisel Mahmood, PhD, Harvard University; Ziad Obermeyer, MD, University of California, Berkeley School of Public Health; Cui Tao, PhD, Mayo Clinic; John Paul (JP) Yu, MD, PhD, University of Wisconsin*
1:45 – 2:45 pm	**Lunch**
2:45 – 5:00 pm	**Session 2: Accessing and Preparing AI-Ready Imaging and Multimodal Data to Enable Interoperability** *Moderator: Steve Henle, PhD, National Eye Institute, NIH* *Co-Chairs: Curtis Langlotz, MD, PhD, Stanford University; Tanveer Syeda-Mahmood, PhD, IBM* *Panelists: Maryellen Giger, PhD, University of Chicago; Benjamin Haibe-Kains, PhD, University of Toronto; Greg Sorenson, MD, DeepHealth, Inc.*
5:00 pm	**Adjourn**

Day 2: March 12, 2025

11:00 – 11:15 am	**Introduction to Day 2** *Karlie Sharma, PhD, National Center for Advancing Translational Science, NIH*
11:15 – 1:30 pm	**Session 3: Validating and Implementing Clinical Imaging and Multimodal AI** *Moderator: Elizabeth Powell, PhD, National Institute on Alcohol Abuse and Alcoholism, NIH* *Co-Chairs: Karandeep Singh, MD, University of California, San Diego; Pamela Woodard, MD, Washington University in Saint Louis* *Panelists: Tessa Cook, MD, PhD, University of Pennsylvania; Woojin Kim, MD, American College of Radiology Data Science Institute; Matthew McDermott, PhD, Columbia University; Shannon McWeeney, PhD, Oregon Health and Science University*

1:30 – 2:30 pm	**Lunch**
2:30 – 4:45 pm	**Session 4: Ethical Considerations for the Use of Imaging-Based, Multimodal AI Clinical Decision Support Tools** *Moderator:* *Marcel Salive, MD, MPH, National Institute on Aging, NIH* *Chair:* *Judy Gichoya, MD, MS, Emory University* *Panelists:* *Paul Gross, Cerebral Palsy Research Network; Maia Hightower, MD, MPH, MBA, Equality AI; Xiaoqian Jiang, PhD, UTHealth*
4:45 – 5:00 pm	**Concluding Remarks** *Michael Chiang, MD, National Eye Institute, NIH; Bruce Tromberg, PhD, National Institute of Biomedical Imaging and Bioengineering, NIH*
5:00 pm	**Adjourn**