Research Idea: Early diagnosis of Autism Spectrum Disorder (ASD) and timely intervention is widely recognized as critical for achieving improved cognitive, behavioral and social outcomes.[1] Despite a growing list of suspected risk factors,[2]–[5] the etiology of Autism is still unclear. Even with increasingly widespread adoption of screening with standardized checklists at 18 and 24 months, the median age of diagnosis for ASD remains at over 4 years. Starting with a positive initial screen, a clinical diagnosis of ASD is a frustrating multi-step process spanning 3 months to 1 year, often delaying time-critical interventions.[2]–[4], [6] One obvious driver for these delays is the vast number of false positives encountered in the current initial screening. For example, the M-CHAT/F, the most widely used screener,[1], [7] produces over 85 false positives out of every 100 flagged for diagnostic evaluation, significantly inflating wait times[5], [8] especially in rural and underserved communities. Further, current screening tools are insensitive to language barriers and cultural issues. Because of these limitations, too offten primary care provides apply a wait and see approach until they have access to educational and psychology professionals at school age.[1], [9] The need for better screening tools is thus paramount.

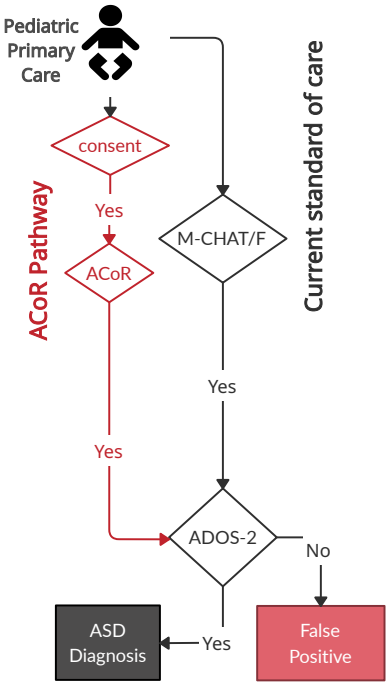

Fig. 1. Patient processing: Parallel pathways through M-CHAT/F and ACoR screens, and a flag in either triggers ADOS-2 evaluation.

In this study, we plan to prospectively validate the efficacy of our recently reported[10] machine-inferred **digital biomarkers** for autism, prospectively in a clinical study, carried out in a pediatric primary care practice at the University of Chicago. The **ASD Co-morbid Risk score (ACoR)** uses machine learning (ML)-based pattern discovery on diagnostic codes already present in individual patient files from past medical encounters, with no questionnaires and no new laboratory tests or blood-work. We plan to demonstrate stand-alone ACoR efficacy, and compare its effectiveness against existing tools such as the M-CHAT/F, with children between 16-26 months of age. Our rationale is informed by the extensively documented comorbidities of ASD, *e.g.*, with epilepsy,[11] gastrointestinal disorders,[12], [13] mental health disorders,[14] insomnia, decreased motor skills,[15] allergies including eczema,[12], [13] immunologic[16]–[18] and metabolic[12], [19], [20] disorders. While ASD presentation is highly variable, deep pattern recognition on the individual longitudinal history of diagnostic codes is expected to reveal uncharted associations enabling precise screening for at-risk patients.

Orthogonal to questionnaire based detection of behavioral signals, the proposed tool can reduce socio-economic, ethnic and demographic biases[21] to elicit more objective and stable results — with zero administrative burden on clinicians and parents. With a team comprising machine learning experts, pediatric clinicians and developmental experts with extensive history of research and service in the field (and past collaboration), we hope to demonstrate that ACoR can significantly improve outcomes by either substantially boosting sensitivity or slashing the current false positive rate. The study aims are:

Aim 1: Reduce false positives in current screening protocols. The current high false positive rate exacerbates post screen wait-time, increasing diagnostic delay. To evaluate the *hypothesis: ACoR reduces up to 50% of false positives*, we will track the cases in which M-CHAT/F triggers a flag, but ACoR does not. We will evaluate the positive predictive value (PPV) of ACoR, under high specificity conditions ($> 95\%$), and also evaluate if ACoR replicates high sensitivity observed in preliminary (retrospective) studies without losing specificity (See Fig. 2 for preliminary results, showing AUC close to 85% under 2 years, and promising early disambiguation between control and positive cohorts).

Aim 2. Evaluate the statistical relationship between ACoR and M/CHAT-F, and estimate a joint or conditional operational protocol. *Hypothesis: The uncertainties or errors in the two tests are are statistically independent.* Thus, we will evaluate our ability to boost performance by conditioning the sensitivity-specificity trade-offs on the M-CHAT/F score of individual patients.

Aim 3. Evaluate the effectiveness of ACoR in a demographically diverse population with a range of socio-economic confounders. *Hypothesis: A questionnaire-free approach has the potential to mitigate biases that arise from limitation of language, cultural barriers, and demographic diversity, e.g.* disproportionately failing to diagnose children with average to above-average intelligence in diverse populations,[22] and under-reporting of symptoms by parents or primary care-givers due to cultural differences.[23]

Aim 4. Characterize heterogeneity presentation by relating it to patterns in medical history, and predictive co-

morbidities. Heterogeneous presentation is a key barrier in the mechanistic understanding of ASD pathobiology. *Hypothesis: We can characterize distinct classes and/or hierarchies of co-morbidities by analyzing medical histories*. This will foster new insights into intrinsic classes of the underlying disease processes.

Thus, we plan to exploit observed co-morbidities in children who ultimately meet the criteria for ASD to develop a risk estimation pipeline, and predict future clinical diagnosis under 2 years of age. Orthogonal to checklists, we aim to reduce the median diagnostic age for ASD, by reducing the long post-screen wait times,[5] by significant boosts in positive predictive value, reduction in false positives, and increased sensitivities at little or no loss of specificity, and at no additional administrative burden or resource utilization.

To achieve the specific aims, recruited patients (with caregiver consent) will be administered M-CHAT/F and ACoR, and scheduled for the gold standard ADOS-2 (Fig. 1). For all assessments, basic demographic information, recruitment site, medications and diagnoses assigned by the current clinical treatment team, will be obtained from the parent/caregiver and medical record. The feasibility of proposed design has been validated with the same study team with a intramural pilot grant, with low sample size (Of 5 ACoR flags produced in this pilot, all were diagnosed with ASD by ADOS-2).

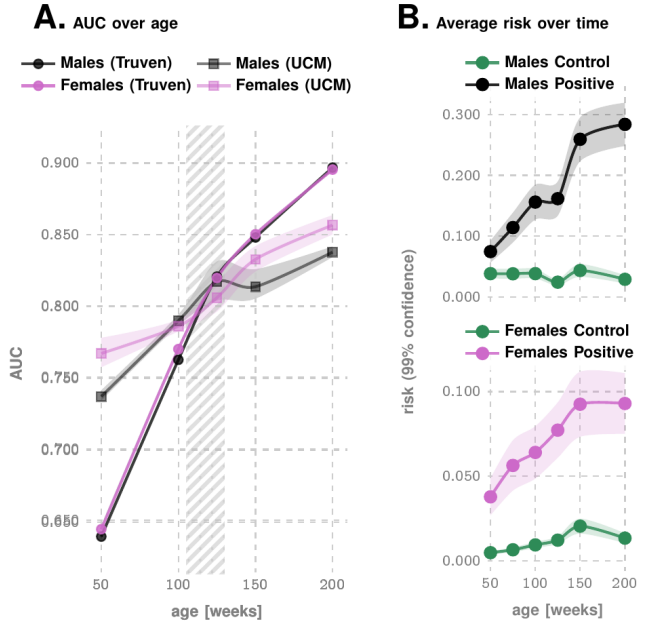

Fig. 2. ACoR performance in retrospective studies. Panel A illustrates AUC achieved as a function of patient age, for 2 independent datasets (Truven & UCM): we achieve $> 80\%$ AUC for either gender from shortly after 2 years. Panel B illustrates risk variation with time for the control and the positive cohorts.

Impact: A screening capability independent of existing tools, deployable as an automated module of a standard EHR system at the point of care, requiring no behavioral observations or new blood-work or laboratory tests has potential to transform ASD care.

ASD presentation has significant heterogeneity, with no simple comorbidity consistently signaling future diagnosis; our algorithms distill robust actionable signatures under such stochastic scenarios. Thus, ACoR opens the possibility of a new screening modality for neuropsychiatric diseases beyond ASD.

Innovation: Standardized questionnaires attempt to measure risk by direct observation of behavioral symptoms by untrained observers (parents). Hence the current screening tests are only as good as the ability of the questions to discern and disambiguate behavior in infants and toddlers on casual observation, and on the ability of parents and caregivers to interpret and answer the items without bias. This has led to possibility of under-diagnosis in minority communities as reflected by the lower apparent prevalence among African-American and Hispanic children. Also, children with average or higher-than-average cognitive abilities seem to have been under-diagnosed as reported is large scale population studies.[1] Borderline cases are typically problematic to screen for due to the possibility of subjective interpretations of the questionnaire. Responses to checklists are clearly confounded by a host of socio-economic (SES) variables, potential interpretive biases, and cultural differences. The heterogeneity of presentation also causes issues: a potential plurality of symptom classes makes it hard for clinicians to recognize borderline cases, or combine observed co-morbidities with scores from current screening tools.

Here we aim to validate ACoR, which operationalizes the fact that ASD has a wide range of co-morbidities.[1] ACoR can address the aforementioned challenges of ASD screening, by leveraging predictive signatures of elevated risk gleaned from past medical history of individual patients alone which are available at the point-of-care, and using no questionnaires, or additional laboratory tests.

Personnel: Our research team comprises Ishanu Chattopadhyay (IC, lead-PI, machine learning (ML)), Peter J. Smith and Michael Msall (Developmental pediatrics), & James Mitchell (pediatric primary care). PJS and MM have multi-decade experience in service and research in autism and developmental disorders, and have made substantial contributions to the field. MM has been the Section Chief of UChicago developmental pediatrics and is medical director of UChicago Early Intervention Outreach and co-director of the JP Kennedy Research Center on Intellectual and Developmental disabilities. JM is an experienced primary care pediatrician, and IC has a track record of designing high impact predictive analytics in medicine.