



Mapping infant neurodevelopmental precursors of mental disorders: How synthetic cohorts & computational approaches can be used to enhance prediction of early childhood psychopathology



Joan Luby^{a,*}, Norrina Allen^b, Ryne Estabrook^b, Daniel S. Pine^c, Cynthia Rogers^a, Sheila Krogh-Jespersen^b, Elizabeth S. Norton^d, Lauren Wakschlag^b

^a Washington University School of Medicine, 4444 Forest Park Avenue, St. Louis, MO, 63108, USA

^b Northwestern University Feinberg School of Medicine & Institute for Innovations in Developmental Sciences, 633 N. St Clair, 19th Floor, Chicago, IL, 60611, USA

^c National Institute of Mental Health (NIMH) Intramural Research Program, Building 15K, Room 110, MSC 2670, Bethesda, MD, 20814, USA

^d Northwestern University, Department of Communication Sciences and Disorders, 2240 Campus Drive, Evanston, IL, 60208, USA

ARTICLE INFO

Keywords:

Child characteristics
Developmental psychopathology
Risk prediction
Computational methods

ABSTRACT

Bridging advances in neurodevelopmental assessment and the established onset of common psychopathologies in early childhood with epidemiological data science and computational methods holds much promise for identifying risk for mental disorders as early as infancy. In particular, we propose the development of a mental health risk algorithm for the early detection of mental disorders with the potential for high public health impact that applies and adapts methods innovated in and successfully applied to early detection of cardiovascular risk. Specifically, we propose methods to advance risk prediction of early developmental psychopathology by creating synthetic cohorts that contain complete behavioral and neural data in the first years of life, as the basis for a robust and generalizable risk algorithm. The application of computational approaches within synthetic cohorts, an approach increasingly applied in psychiatry, may be particularly well suited to advancing risk prediction in early childhood mental health. We propose new research directions using these methods to generate an early childhood mental health risk calculator that could significantly advance early mental health risk detection to direct preventive intervention and/or need for more intensive assessment within a pragmatic framework for maximal clinical utility. The availability of such a tool in early childhood, a period of high neuroplasticity, holds promise to reduce the burden of mental disorder by identifying risk early in the clinical sequence and delivering prevention that targets the neurodevelopmental vulnerability phase.

1. Introduction

There has been an increasing awareness that many major mental disorders have developmental antecedents, some of which manifest as early neurodevelopmental vulnerability beginning in infancy (Casey, Oliveri, & Insel, 2014; Mittal & Wakschlag, 2017). Advances in developmentally sensitive assessment methods over the last two decades have detected such neurodevelopmental vulnerabilities for a number of common mental disorders including mood (depression) and anxiety disorders (Generalized Anxiety Disorder, Separation Anxiety Disorder), disruptive disorders and Attention Deficit Hyperactivity Disorders (ADHD) (Clauss & Blackford, 2012; Wakschlag et al., 2019). Further, the advent of developmentally appropriate diagnostic interviews and dimensional measures that assess for the age adjusted symptom

manifestations early in life have facilitated recognition of impairing forms of psychopathology as detailed above as early as age 2 (Egger & Angold, 2006; Wakschlag et al., 2005, 2019). The importance of the earliest possible identification and intervention in mental disorders is underscored by the relatively large effect sizes and cost effectiveness of interventions delivered earlier rather than later in life during periods of high plasticity (Campbell et al., 2014; Dawson, Ashman, & Carver, 2000). Based on this, there is an increasing imperative to identify malleable markers of risk at the earliest phase of the clinical sequence (i.e. high risk children) to prevent escalation of vulnerability to clinically impairing symptoms (Finlay-Jones et al., 2019). Translation of this knowledge to application requires pinpointing which vulnerable children are likely to go on to develop psychopathology and which are not to inform clinical-decision making and target resources efficiently

* Corresponding author. 4444 Forest Park Avenue, St. Louis, MO, 63108, USA.
E-mail address: lubyj@wustl.edu (J. Luby).

<https://doi.org/10.1016/j.brat.2019.103484>

Received 8 February 2019; Received in revised form 4 September 2019; Accepted 25 September 2019

Available online 26 September 2019

0005-7967/ © 2019 Published by Elsevier Ltd.

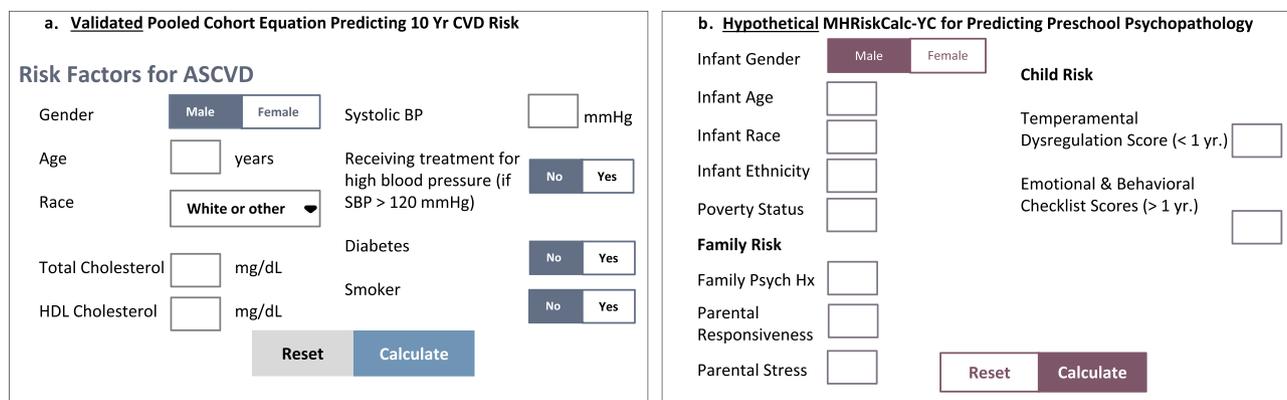


Fig. 1. Clinical risk calculators: Validated CVD pooled cohort equation (PCE) & hypothetical RiskCalc-MHYC for survey measures.

(Ozonoff, 2015). Currently, the lack of reliable methods for this determination is a major impediment to clinical application.

Distinction between normative extremes and markers of risk: sensitivity and specificity: There has been a burgeoning developmental literature identifying infant markers of risk as well as increasing emphasis on the identification of neural substrates of emergent psychopathology (Bilgin et al., 2018; Graham et al., 2016; Hay et al., 2014; Rogers et al., 2017). Autism Spectrum Disorder (ASD) represents one area where such data are actively shaping clinical guidelines using measures of typical and atypical social development, cognition, brain function, and behavior (Bosl, Tager-Flusberg, & Nelson, 2018; Ozonoff et al., 2014). However, in other areas of psychopathology and to a lesser extent in ASD, much more work is needed before these findings can meaningfully inform clinical decision making. We here focus on the common preventable psychopathologies of early childhood, i.e., internalizing and externalizing syndromes including depression, anxiety, ADHD and disruptive disorders. Achieving reliable identification of internalizing/externalizing risk, requires a big data approach to reliably demarcate the distinction between normative variation and transient developmental extremes from markers of clinical risk that predict high risk for impairing psychopathology. As neurodevelopmental vulnerability to psychopathology in infancy is a relatively new concept, there are not yet population-based estimates of natural course remittance (i.e., the percentage of young children who exhibit neurodevelopmental vulnerability but do not develop psychopathology), critical information that is established in other neurodevelopmental disorders. However, while neurodevelopmental vulnerability exponentially increases risk of psychopathology, many children who exhibit such vulnerability will not go on to develop mental health problems (Chronis-Tuscano et al., 2009; Ozonoff, 2015) (e.g. 60% of children with language delays do not go on to develop language disorders). Thus, a careful balance of sensitivity and specificity in this domain is necessary before such markers can be harnessed to inform the earliest risk detection. Deriving optimal “sensitivity-specificity” cut-points are particularly important in early childhood when developmental extremes that might be labelled as symptoms in older children are more likely to reflect transient and normative variation (Cole, Luby, & Sullivan, 2008; Wakschlag, Tolan, & Leventhal, 2010). We propose that this determination be based on high probabilistic risk of impairing psychopathology by preschool age (Wakschlag et al., 2019).

Risk calculators in early mental health: The most widely used Framingham risk calculator has transformed clinical care and prevention in cardiovascular disease and has accelerated precision medicine approaches. Its utility lies in its modern risk prediction approach that utilizes the most parsimonious, lowest burden and/or most readily available set of indicators that achieve sensitive and specific risk prediction (Pencina & D’Agostino Sr, 2012). The lack of validated risk calculators represents a major gap in translational science in mental health as advances in neurodevelopmental discovery and validation of

early childhood mental disorders has yet to be leveraged for practical application to clinical decision making. Risk calculators could greatly facilitate pragmatic public health efforts towards prevention and early intervention in common and preventable mental disorders by providing a widely accessible tool to guide clinical decision making about young children’s mental health. In stark contrast to physical disorders (Corbelli et al., 2014; Pencina & D’Agostino Sr, 2012), pragmatic risk calculation is nascent in psychiatry (Bernardini et al., 2017). For prevention and intervention targeting neurodevelopmental vulnerability to mental disorder, we lack a “science of when to worry” (Wakschlag et al., 2019). Such a tool could provide clinically-feasible empirically derived guidance for clinical decision making in the first 3 years of life, when these disorders have their roots (Shaw, 2013). A risk calculator could serve as the basis for tiered prevention. Even at preschool age, where the validity of many forms of psychopathology are well-established, entrenched “they’ll grow out of it” myths (Luby, 2012) continue to impede wide acceptance of the need for early identification and prevention. Thus, tools are now needed that “embrace rather than erase” developmental heterogeneity, harnessing it for clinical decision making and tailored prevention in the first years of life, an approach we have termed “Mental Health, Earlier” (Wakschlag et al., In Press).

Fig. 1 shows the Framingham calculator and a hypothetical infant mental health calculator optimized for clinical use using only survey measures. A key aspect of the risk prediction method that has proven so successful in Framingham is its emphasis on a pragmatic assessment approach (Glasgow, 2013). A key element of this is that its risk prediction models give preference to low intensity methods, requiring more intensive methods (such as observation, interviews and imaging) to demonstrate sufficient clinical added value to justify their added burden and cost (Lloyd-Jones, 2010; Pool, Ning, Wilkins, Lloyd-Jones, & Allen, 2018). This has been widely studied in cardiovascular disease but remains underexplored in mental health prediction.

Availability of developmentally appropriate surveys and clinical interviews: Developmentally-sensitive surveys and clinical interviews have identified valid clinical patterns that show predictive validity across the infant-toddler-preschool period (Biedzio & Wakschlag, 2018; Briggs-Gowan, Carter, Bosson-Heenan, Guyer, & Horwitz, 2006; Egger, H. L. & Angold, A., 2006; Gaffrey & Luby, 2012; Lorber, Del Vecchio, & Slep, 2015). More recently, developmentally-specified dimensional measurements designed to capture an ordered normal-abnormal spectrum have shown utility for mapping to neural correlates (Grabell et al., 2018; Wakschlag et al., 2018). Direct assessments including observations during evocative tasks also enable performance-based assessment of key processes (e.g., standardized clinical observation, visual attention) that complement parental perspective and enhances earlier detection of children at high risk as well as those at intermediate risk (Halperin & Marks, 2019; Miller, Iosif, Young, Hill, & Ozonoff, 2016; Wakschlag et al., 2005). The availability of valid multi-level biobehavioral methods suggest that a neurodevelopmental toolkit

for early clinical risk identification is within reach, although such algorithms have yet to be developed. Integrative approaches that weight unique information from these different sources including generation of multiple thresholds of risk that are optimized for clinical feasibility are a key next step (Bufferd, Dyson, Hernandez, & Wakschlag, 2016).

The role of neural and other biomarkers: There has been a proliferation of feasible and developmentally sensitive neurobehavioral measures in the last decade, providing the capacity for more intensive characterization of infants and young children. Such quantitative measures are particularly important in early childhood to address the difficulties of making inferences about the internal mental states of young children. The use of neural measures in infancy including non-invasive methods such as electroencephalography (EEG) and structural and functional magnetic resonance imaging (fMRI) have provided new clues to alterations in brain structure and function detected at birth and during the first year of life. Many of these measures have proven useful to identify very early markers of ASD and other forms of psychopathology. For example, resting state-functional MRI (rs-fMRI) is a tool that elucidates the functional connectivity of brain networks in the sleeping infant without requiring task engagement (Eggebrecht et al., 2017; Rogers et al., 2017; Smyser et al., 2016; Sylvester et al., 2017). Rogers and colleagues as well as others have used this technique in longitudinal studies to relate infant functional connectivity to subsequent measures of early childhood psychopathology. Variability in early childhood internalizing symptoms has been linked to variation in functional connectivity of the amygdala (Graham et al., 2018; Rogers et al., 2017) and cortical networks (Sylvester et al., 2017) as early as the neonatal period. Others have demonstrated that functional connectivity during later infancy can predict ASD symptoms (McKinnon et al., 2019) and diagnoses (Emerson et al., 2017) at age 2 years. Conversely from behavior to brain, early childhood depressive syndromes have been associated with alterations in patterns of cortical thinning in middle childhood (Luby et al., 2016) as well as concurrent alterations in hippocampal volume (Barch et al., in 2019). Consistent with this, we have also recently discovered associations between cortical thickness and early transdiagnostic indicator of INT/EXT risk, i.e., early irritability, a pattern replicated across two independent and diverse samples (Norton et al., Manuscript in preparation).

Along this line, EEG and event-related potentials (ERPs, EEG responses time-locked to stimuli) have also been useful and highly pragmatic tools to assess neural correlates of emotion in infants and young children. For example, resting EEG patterns in children as young as 3 months old have been found to predict ASD diagnosis with greater than 95% sensitivity/specificity, which is far greater than is currently possible with behavioral measures (Bosl et al., 2018). Resting EEG asymmetry has also been associated with mental health outcomes; children with low positive emotionality and high negative emotionality at age 3 subsequently developed reduced left hemisphere activity, a pattern associated with depression in adults (Goldstein et al., 2018). ERPs during reward or error detection tasks have proven useful for parsing heterogeneity in clinical outcomes for children with behaviorally homogenous risk profiles at preschool age. ERP responses during reward tasks differentiate acutely depressed from healthy preschool children (Barch et al., 2019; Belden et al., 2016) as well as predict response to treatment (Barch et al., 2019). ERPs also may serve to differentiate clinical heterogeneity, informing which young irritable children will subsequently develop internalizing versus externalizing psychopathology (Kessel et al., 2016). Although generally not included in standard neuroimaging consortia protocols, EEG has high potential clinical utility relative to MRI, with the advantage of being low cost and feasible in a clinic setting, despite the disadvantage of offering poor localization of the neural signal. Evidence from other non-invasive methods such as eye-tracking and cardiac orienting response obtained via biosensor, also hold promise for detecting abnormality in the first year of life more reliably than behavioral indicators and with less cost and time required (Finlay-Jones et al., 2019).

Limitations of neural markers: Despite intriguing progress in this domain, neural markers have not yet proven reliably useful in clinical decision making for the detection and diagnosis of risk states and clinical mental disorders. Importantly, infant MRI and EEG/ERP studies to date have largely included small sample sizes and the replicability of findings has not been assessed (McWeeny et al., Manuscript in preparation), much less jointly considered in terms of their relative contribution to clinical risk prediction. Further, the lack of specificity of many of these neural findings may prevent them from being pragmatically useful in clinical risk algorithms. However, pending further progress, neural markers are potentially very fruitful domains as some neural substrates may serve as screeners signaling the need for more intensive assessment. In other cases, neural markers may outperform behavioral measures in predicting relevant mental health outcomes in terms of greater accuracy or earlier assessment (Finlay-Jones et al., 2019).

Need for neurobehavioral norms in early childhood: The mapping of early emotional, behavioral, and neural patterns as basis for neurodevelopmental norms that draw clear distinctions between typical and atypical trajectories are needed as the first step in developing the proposed risk calculator. Ideally, this would take the form of a pediatric growth chart that would map expectable milestones of emotional and behavioral regulation. While a great deal of early neurodevelopmental data is now available, these have not yet been used to generate meaningful developmental norms for public health use. It is clear that this would require a big data approach to achieve the needed power and precision, including sufficient numbers of cases and to account for the complex individual differences that characterize this period of rapid development. However, given the complexity of brain-behavioral relationships, and the numerous and multi-faceted forms of data that are generated, it is less clear how to combine these big, multi-level data sets in a clinically informative way. Clinical utility requires not only reliable and valid differentiation of salient behavioral patterns but also feasibility, which requires parsimonious optimized algorithms. Notably, this contrasts with typical research designs where the most comprehensive and deep assessment methods are considered a strength. The key methodologic question for clinical application of a neurodevelopmental risk algorithm thus being, “when more is not better, what is enough?” (Pickett et al., 2009).

Utility of risk prediction in medicine: Risk prediction models have been successfully employed in other areas of medicine (Pencina & D'Agostino Sr, 2012). Whereas traditional predictive models are post-hoc (i.e. the outcome is known), prognostic risk prediction models statistically estimate the probability that a specific individual in the population will develop a condition in the future, based on a parsimonious set of risk indicators. These models are derived from large studies to generate reliable and generalizable estimates with two objectives: (1) assign each individual in the population a score based on some combination of risk that can be translated into a personalized estimate of probability of developing a disorder over a specific period of time and (2) assign a categorical risk classification (e.g., positive/negative) for clinical decision making about therapeutic indications. These models have the added advantage of parsimony and therefore clinical application because incremental predictive value is based on potential for reclassification and/or discrimination rather than simply statistical significance (Pencina & D'Agostino Sr, 2012). They are further optimized for clinical feasibility because intensity/expense of methods has a higher threshold for inclusion based on demonstrating incremental utility. Risk prediction algorithms have been successfully developed using epidemiologic data science for many physical diseases, including incorporation into standard clinical care for the prevention of CVD (Pencina & D'Agostino Sr, 2012). Indeed, the Framingham risk score and the Pooled Cohort Equation are currently the gold standard for CVD risk assessment, predicting 10 year coronary event risk from the “simple seven” parameters (Greenland, LaBree, Azen, Doherty, & Detrano, 2004).

Computational models to inform risk prediction in mental health: Of note, the question of how neural markers might inform clinical risk prediction and diagnostic processes has been difficult in psychiatry in general, much less applied to early childhood mental health. The rapid and often non-linear developmental changes in early life known across these domains further complicate such attempts. To begin to address this issue, there has been an emerging interest in the use of computational risk prediction models as a tool to combine complex data bridging clinical and neuroscience research. This approach, increasingly explored in adult psychiatry provides methods for integrating multiple, dense data streams and big data to achieve greater precision in diagnostic criteria, risk prediction and treatment response (Bernardini et al., 2017). This paradigm shift moves from emphasis on statistical significance of between group differences to an explicit focus on risk prediction, noting that such approaches often identify different indicators (Paulus, Huys, & Maia, 2016). While promising, to date these studies have focused on populations that are already far advanced along the clinical risk trajectory. A central tenet of the Mental Health, Earlier framework is the use of risk prediction models beginning in the first years of life or potentially in some cases even at birth. This is the crucial next frontier for realizing neurodevelopmentally-based prevention prior to the onset of mental disorders (see Fig. 1). In particular, mental disorders are a domain in which the earlier identification of risk is particularly potent because this capitalizes on periods of greater neuroplasticity for more powerful intervention effects.

The advent of computational psychiatry employing both data-driven (e.g. machine learning) and theory-driven (e.g., Bayesian model comparison) methods now enables identification of regularities in large neurodevelopmental datasets with application to understanding brain:behavior relationships, reflecting the larger imperative towards cumulative science as the wave of the future (Curran, 2009). Such findings may be particularly useful to inform the prediction of dimensional clinical phenotypes in heterogeneous populations (Ferrante et al., 2018). The capacity for computational approaches to map covariation of risk markers at brain, behavior and contextual levels as they unfold across development is of particular relevance to identifying infant predictors of mental health outcomes. *The novelty of this approach, if successful, is that it would enable for the first time identification of vulnerability to mental disorder beginning from birth.* This method is highly applicable to generating risk prediction models in mental disorders where a central determination is when neural data may have added value for clinical classification over and above less expensive and more feasible measures. Fig. 1 highlights this clinical risk calculation approach for established CVD methods and a theorized calculator for predicting probabilistic risk of preschool psychopathology.

The field of infant mental health is one that has the potential to be greatly catalyzed by such computational approaches. The stage is set for such work due to advances in neuroimaging and assessment techniques described above, coupled with challenges in decision-making approaches that balance false positives and false negatives, particularly salient in early childhood. The clinical promise of early neurodevelopmental discovery has also been impeded by diagnostic uncertainty based on concerns about false positives at this young age (e.g., “stigmatizing young children”) (Luby, 2012; Ozonoff, 2015). However, importantly, the cost of false negatives (e.g., the negative cascade that is initiated when early patterns of dysfunction go unchecked (Shaw & Gilliam, 2017) is also high. Further, the central feature of early intervention for the dysregulation that presages both internalizing and externalizing problems is promotion of self-regulation, an endeavor that benefits all children (Smith et al., In Press; Wakschlag et al., 2019). Moreover, the rapid changes in brain-behavior relations across development necessitate creation of flexible tools that can be adapted in ways that function similarly at different ages (Mittal & Wakschlag, 2017). As such, early childhood mental health is a domain in which computational tools that can account for variation in timing of the same construct might be uniquely useful.

2. Synthetic cohorts and Computational Models to advance risk prediction

Risk prediction models require large, prospective “pooled” cohorts in which “disease-free” participants are followed to capture disease onset over a set time (Pencina & D’Agostino Sr, 2012). In an ideal world we would have a very large and generalizable cohort with data that included all potential surveys, interviews, direct assessment and neural markers of interest starting at birth and spanning every age of interest to develop a mental health risk prediction tool. However this ideal dataset is not currently available nor is it feasible. Integrative data analyses (IDA) methods have been developed to analyze multiple separate datasets as if they were a single one. To initiate a cost effective and expedient approach to data integration for early childhood mental health research we can borrow methods innovated within epidemiologic data science to advance risk prediction in CVD.

As a result of method/design variance across studies, traditional pooling methods have extensive missing data, introducing imprecision and potential bias. This constrains the ability to capture normative variation, sensitive periods, and other nonlinear patterns of central importance to neurodevelopmental risk prediction (Pine, 2017). One integrative data analysis method for pooled data utilizes observed data from each cohort using statistical models that take into account missing data due to varying assessment points across cohorts (Allen, Ning, Jones, Zhao, & Siddique, 2017). However, because not all the common and unique elements are consistent across samples, a large amount of missing data precludes traditional pooled methods from estimating detailed longitudinal developmental patterns (Siddique et al., 2015). Our emphasis on the use of the synthetic cohort approach within the broader context of computational methods resting on pooling of multiple datasets, is the unique advantage the synthetic approach has for *maximizing* full information making it optimal for capturing nuanced developmental variation that might be lost if whole time points were routinely treated as missing. In particular, relative to other pooling approaches, the synthetic method enables a richer developmental characterization. It does this by enabling the analysis of the same individuals across the full developmental period of interest by imputing missing time points for each individual specifically. In contrast, traditional pooling methods rest on the assumption that individuals are interchangeable taking into account measured, but not unmeasured, differences between the individuals to inform developmental patterns.

To address these limitations, a novel epidemiologic method was developed by several of the co-authors (omitted blind review) for CVD risk prediction with minimal loss of precision (e.g., < 1%) relative to traditional pooling. This synthetic cohort approach uses multi-level, multiple imputation models to fill in missing data arising from variations in design (i.e. different exam ages and/or exam components) across pooled studies. The resultant synthetic cohort combines observed and imputed behavioral data at all timepoints to substantially reduce loss of information. CVD research has used a synthetic cohort approach when traditional pooled cohort methods were not feasible. In contrast, the synthetic cohort approach treats varied measures and time-points inherent in a pooled approach as a “missing data problem,” using multiple imputation to fill in gaps while taking into account individual and cohort-specific differences (Siddique et al., 2015). These models can be used to impute longitudinal data incorporating assessment/age (temporal effects/calendar year), individual (birth cohort, demographics), and cohort-specific variables to account for sampling and measurement differences. In this method, the missing data for each individual (e.g., the preschool data for infancy cohorts and vice versa) are treated as missing data and missing observations are multiply imputed. (Of note, at this time, imputation is proposed for behavioral data as methods for imaging imputation are not fully developed.) A major advantage to this approach for neurodevelopmental studies is that it uses full information for imputation. Thus, if various studies collected data on the same construct but at different timepoints, imputation will

be based on data from all children who provide data on that construct from other studies and from other pertinent data on the cohort missing the specific timepoint from prior and following timepoints. This adds the richness that is so important to developmental modeling. Of note, although missingness due to cross-cohort design differences (e.g. measure selection differences) is distinct from missingness due to individual participant differences within a cohort (e.g., attrition, incomplete data response), synthetic methods treat this as data missing (relatively) at random (MAR) based on the likelihood that it is ignorable for making inferences (Brincks et al., 2018; Curran & Hussong, 2009).

The resulting “synthetic cohort” is then composed of a combination of observed and imputed data at every time point for every measure across all cohorts (Siddique et al., 2015). This synthetic cohort approach has been successfully employed for modeling of emergent CVD risk as young as 8 years and improved risk estimation of morbidity and mortality at earlier ages with minor trade-offs in precision (e.g., < 1% relative to traditional pooling (Allen et al., 2014, 2017). The increased precision yielded by the synthetic cohort approach is of particular importance for early life mental health risk determination as modeling in early development must account for extensive normative variation, and the potential for sensitive periods and other non-linear factors in longitudinal patterning. The generation of such a large synthetic cohort also offers the opportunity to incorporate cutting edge neural, specifically fMRI, and behavioral markers.

Thus, we propose that synthetic cohorts can offer sample sizes with sufficient statistical power to enable transformative, high impact advances in infant mental health. Such cohorts might also offer the necessary precision to detect developmental trends that could not be achieved from a single cohort alone. In particular, the generation of a large synthetic cohort is necessary to achieve more precise mapping of the typical:atypical brain-behavior trajectories from birth to age 3 in the context of a complex psychosocial environment known to critically impact development (Barch, Belden, Tillman, Whalen, & Luby, 2018; Bick & Nelson, 2016). Frequent assessments are particularly critical during this rapid and steep early developmental trajectory with multiple internal and external factors of influence. The application of such a process is particularly critical in early childhood, when detection of non-linear growth requires dense sampling.³⁵ Data harmonization and synthesis of unique brain-function and developmentally-defined behavioral phenotypes beginning in infancy is necessary to open new avenues for more powerful exploration of the earliest precursors of psychopathology using computational methods. Using this approach by harmonizing existing and ongoing data collection, it is possible to develop in a reproducible multi-level clinical algorithm for the earliest detection of risk for mental disorders that can be disseminated as an online calculator. Such an innovation is necessary to accelerate the pace of translation from discovery to clinical implementation in mental disorders.⁴⁴

3. Trade-offs in using the synthetic cohort approach

While the synthetic cohort approach specifically and the risk prediction modeling framework overall are very promising for accelerating the pace of translation from neurodevelopmental discovery to infant mental health risk calculation, the necessary harmonization is certainly not without its disadvantages and practical challenges. By virtue of synthetic and integrative modeling approaches’ emphasis on embracing the importance of between-study variability, this introduces substantial challenges to internal validity (Curran & Hussong, 2009). These have been discussed extensively elsewhere (Brincks et al., 2018; Curran & Hussong, 2009) with emphasis on treating them in combination for greatest effectiveness, and are briefly summarized here. Threats to internal validity are particularly salient since the integration of multiple diverse datasets in itself increases external validity relative to a single study (Curran & Hussong, 2009).

3.1. Power vs. precision

Synthesizing cohorts provides far more power than any single cohort could provide, which is of key importance since imaging costs constrain sample sizes. However, there is also loss of the precision that comes with the more in depth approach of any single study. One major element that speaks to feasibility of the synthetic approach is the fact that virtually all large neurodevelopmental studies include measures of core constructs requisite for infancy-based risk calculation (e.g., indicators of self-regulation, temperament, developmental functioning, parenting etc) (Smith et al., In Press). This typically means that virtually all salient datasets have common anchoring items drawn from varied scales (Curran & Hussong, 2009; Kolen & Brennan, 2014; Smith et al., In Press). But from this flows the inherent challenge in data harmonization such that heterogeneous measures are used to assess the same construct. While most studies will measure similar key constructs, methods will differ. One approach to this is z-score transformation. However, the trade-off is that this precludes generating an absolute risk threshold from any particular measure for an individual (because the z-score only provides information in regard to relative place on a continuum rather than a calibrated score for that instrument). IRT methods have also been used to bridge items common across different methods with tests of dimensionality of the latent construct and concomitant scale construction (Brincks et al., 2018; Curran & Hussong, 2009). Related bridging approaches have modeled different measures of like constructs as distinct indicators of an underlying latent construct (Brown et al., 2018). These approaches are parallel to recent advances in psychometric score linking which demonstrate the comparability of a wide array of measures of similar constructs (Choi, Schalet, Cook, & Cella, 2014; Kaat et al., 2019). This enables harmonization but loses the precision of individual measurement and their standardized yield (e.g., t-scores). Of course, there is always the possibility that the use of the synthetic approach may fall short for mechanistic investigation because data originally collected for another purpose may not make nuanced significant distinctions underlying mental health risk (Bennett, Silverstein, & Niv, 2019). This will require empirical testing and may vary based on intent of the synthesized design. However, we underscore that the goal of risk calculation is an up or down decision about the probability of a young child developing impairing mental health risk *not* subtle mechanistic discovery.

3.2. Generalizability vs. population heterogeneity

Without question, the ability to harmonize and leverage multiple diverse cohorts with neuroimaging data enhances generalizability and enables risk enrichment. This is a particularly important future goal as neurodevelopmental imaging consortia are typically comprised of lower risk populations (Hanson et al., 2013; Howell et al., 2019). However, along with this greater representativeness comes the population heterogeneity introduced via combining samples with different inclusion/exclusion criteria, demographic, race/ethnicity, composition, sampling, regional and site characteristics (Curran & Hussong, 2009). One method to address this is to control for global cohort membership by including it in the models as a fixed effect covariate (Brincks et al., 2018; Curran & Hussong, 2009). Also of note is the increased prevalence of low base rate phenomena when samples are pooled (Curran & Hussong, 2009), which is of particularly high significance in the quest for more dimensionally defined psychopathologic spectra. It is also noteworthy that synthetic and integrative methods are inherently designed to examine the impact of sampling differences on risk prediction (Curran & Hussong, 2009). That is their purpose, not to establish that sampling differences are ignorable, but rather to adjust for this heterogeneity and have the opportunity to explicitly test their empirical salience (Curran & Hussong, 2009).

3.3. Expanded developmental scope vs. heterogeneity of design

Harmonizing cohorts unquestionably enables coverage of a far broader developmental span allowing for inference across the entire age period, overcoming the limitation of any individual study only following participants for a small fraction of that time (Curran & Hussong, 2009). However, inevitably, even studies broadly targeting the same developmental period will differ on follow-up range, and assessment time-points etc.

As noted above, rigorous validation checks are necessary to assess the influence of this on the validity of the synthetic cohort methods. To date, findings from CVD research utilizing these methods suggests that the loss of precision is minimal. However, this remains to be established in neurodevelopmental applications. As the field moves towards such big data, cumulative science approaches, it has also been suggested that increasing emphasis on at least a core of common measures in individual studies would substantially lessen obstacles to harmonization, facilitating synthetic endeavors (Blackwell, Wakschlag, Gershon, Cella, & Core, 2018; Brincks et al., 2018; Curran, 2009). This too has its disadvantages (in terms of loss of depth and novelty) and may be best as a first level, linkage, rather than an exclusive, measurement approach (Brincks et al., 2018; Curran & Hussong, 2009). We envision that these issues will continue to evolve and progress as computational methods gain traction. However, the imperative for the translational advances that may be brought about by the infant mental health risk calculator pushes us to work with the current state of the science so that the statistical perfect is not the enemy of the public health good.

4. Discussion

We propose that the use of a risk prediction framework that draws on multiple data sources and innovations in epidemiologic and computational data science is the next scientific frontier for generating infant mental health risk algorithms to predict which young children are at high likelihood of developing psychopathology by kindergarten age. This is necessary to understand the key precursors to guide prevention of mental disorders at the earliest possible developmental point. Such methods applied to early childhood neurodevelopmental data would greatly catalyze mental health prevention during early neuroplastic developmental periods. These methods which have only begun to be applied in adult psychiatric research and have not yet been employed in the first years of life, hold much promise for addressing the critical gap between neurodevelopmental discovery and clinical application. This translational research innovation may now be within reach based on advances in all necessary component scientific domains. Like any method, the synthetic cohort risk prediction computational approach brings challenges, but its potential benefits are also substantial (Bennett et al., 2019).

The challenges to clinical identification and prevention in early life are typically framed in terms of the risk of over- or under-identification (Luby, 2012). The potential for false positives and corollary stigmatization have impeded actualization of neurodevelopmental prevention. However, such early identification that has the potential for altering the lifespan trajectories of mental disorders (Wakschlag et al., 2019). Skepticism about the informativeness of developmentally-framed evaluations for clinical decision-making has also unnecessarily deterred clinical application. With increased recognition of pediatric mental disorders there have been escalation rates of psychopharmacologic intervention despite the absence of clear risk calculation parameters for the early childhood period (Pennap et al., 2018; Zito et al., 2007). Within this context, there is now consensus that the evidence base is “good enough” for clinical, real-world application of risk algorithms for early childhood identification of mental disorders (Luby, 2012; Shonkoff, 2016; Wakschlag et al., In Press).

Fundamentally, such application requires the generation of a developmentally-based risk calculator approach as employed in physical

disease (see Fig. 1). As described, this approach requires large pooled data sets for the necessary power and precision to account for developmental heterogeneity. The robust science base for CVD risk prediction relies on pooling dozens of cohorts and thousands of individuals (Allen et al., 2017). The novel synthetic cohort method employed more recently has distinct advantages for providing the level of developmental data key to differentiating extensive normative variation and individual differences in the first years of life from reliable clinical risk markers. We propose that generation of synthetic cohorts utilizing samples that contain behavioral measures and neuroimaging in early childhood is now a highly feasible and more cost-effective way to apply big data methods to the earliest detection of risk in mental health beginning from birth. We propose future studies that utilize samples that begin in infancy and ascertain neural as well as behavioral measures at multiple time points across early development that are enriched for clinical risk to best inform this area.

There has been increased interest and resource allocation to large multi-site neuroimaging consortia (Blackwell et al., 2018; Howell et al., 2019; Jernigan et al., 2016). These studies will provide an invaluable resource for mapping early brain and behavioral development. However, these designs are expensive and labor intensive and as currently conceived do not utilize clinical diagnostic measures, deep phenotyping or measures of impairment, and lack a roadmap to inform clinical care. The need for neuroimaging studies enriched for clinical risk that start in infancy and have comprehensive behavioral characterization and measures of environmental risk are necessary to bridge this gap. While screening measures may prove sufficient to detect global risk, more information is needed to more precisely inform early clinical risk. To begin to apply targeted prevention strategies it is necessary to determine which screening measures should be applied at what time-points during early development. To accomplish this, it is necessary to map early behaviors and neural markers frequently across time and investigate their association with clinical outcomes. By obtaining this kind of detailed fine-grained developmental data in early childhood samples, computational approaches can be applied to begin to derive algorithms to calculate risk for specific classes of psychopathology and/or disorders. Such an approach which is now highly feasible given advances in measurement of early childhood developmental psychopathology, making it now possible for computational science to revolutionize early intervention in mental disorders.

Currently, advances in the treatment of most mental disorders have stalled and established treatments for the vast majority of disorders are only moderately effective. Based on this, there is an urgent need for novel approaches to treatment. While there has been a new emphasis on treatments that are mechanistic, progress in this domain has been slow and incremental (Ferrante et al., 2018). There is increasing evidence that early interventions provide a time limited window of opportunity to achieve larger effects based on periods of high neuroplasticity and response to environmental experience. Early interventions in Autistic Spectrum Disorders and in Disruptive behavior provide cogent examples of the potential power of early intervention. The use of synthetic cohorts and computational approaches in samples that start in infancy and include deep phenotyping and neuroimaging provide an opportunity to advance this domain via the development of a pragmatic and generalizable risk calculator for preschool psychopathology. We propose that this is the critical next step to advancing neurodevelopmentally-oriented prevention and treatment of mental disorders with highest impact for reducing lifelong burden of mental illness by altering trajectories at their roots.

Funding

This work was supported via funding from the National Institute of Mental Health (NIMH), United States to Washington University School of Medicine from NIMH [R01MH113883, R01MH090786]; and funding to Northwestern University from NIMH [R01MH107652,

U01MH082830]; and the National Institute on Deafness and Other Communication Disorders, United States [NIDCD; R01DC016273].

References

- Allen, N. B., Ning, H., Jones, D. L., Zhao, L., & Siddique, J. (2017). Cardiovascular health across the lifespan: The development and validation of a synthetic cardiovascular cohort. *Circulation*, *135*(Suppl_1).
- Allen, N. B., Siddique, J., Wilkins, J. T., Shay, C., Lewis, C. E., Goff, D. C., et al. (2014). Blood pressure trajectories in early adulthood and subclinical atherosclerosis in middle age. *Jama*, *311*(5), 490–497.
- Barch, D. M., Belden, A. C., Tillman, R., Whalen, D., & Luby, J. L. (2018). Early childhood adverse experiences, inferior frontal gyrus connectivity, and the trajectory of externalizing psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(3), 183–190.
- Barch, D. M., Whalen, D., Gilbert, K., Kelly, D., Kappenman, E. S., Hajcak, G., et al. (2019). Neural indicators of anhedonia: Predictors and mechanisms of treatment change in a randomized clinical trial in early childhood depression. *Biological Psychiatry*, *85*(10), 863–871.
- Belden, A. C., Irvin, K., Hajcak, G., Kappenman, E. S., Kelly, D., Karlow, S., et al. (2016). Neural correlates of reward processing in depressed and healthy preschool-age children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(12), 1081–1089. <https://doi.org/10.1016/j.jaac.2016.09.503>.
- Bennett, D., Silverstein, S. M., & Niv, Y. (2019). The two cultures of computational psychiatry. *JAMA Psychiatry*, *76*(6), 563–564.
- Bernardini, F., Attademo, L., Cleary, S. D., Luther, C., Shim, R. S., Quartesan, R., et al. (2017). Risk prediction models in psychiatry: Toward a new frontier for the prevention of mental illnesses. *Journal of Clinical Psychiatry*, *78*(5), 572–583.
- Bick, J., & Nelson, C. A. (2016). Early adverse experiences and the developing brain. *Neuropsychopharmacology*, *41*(1), 177.
- Biedzio, D., & Wakschlag, L. S. (2018). Developmental emergence of disruptive behaviors beginning in infancy: Delineating normal: Abnormal boundaries to enhance early identification. *Handbook of infant mental health*, *4*.
- Bilgin, A., Baumann, N., Jaekel, J., Breeman, L. D., Bartmann, P., Bäuml, J. G., et al. (2018). *Early crying, sleeping, and feeding problems and trajectories of attention problems from childhood to adulthood*. Child development.
- Blackwell, C., Wakschlag, L. S., Gershon, R., Cella, D., & Core, E. P. (2018). Measurement framework for the environmental influences on children's health outcomes research program. *Current Opinion in Pediatrics*, *30*(2), 276–284.
- Bosl, W. J., Tager-Flusberg, H., & Nelson, C. A. (2018). EEG analytics for early detection of autism spectrum disorder: A data-driven approach. *Scientific Reports*, *8*(1), 6828.
- Briggs-Gowan, M. J., Carter, A. S., Bosson-Heenan, J., Guyer, A. E., & Horwitz, S. M. (2006). Are infant-toddler social-emotional and behavioral problems transient? *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(7), 849–858. <https://doi.org/10.1097/01.chi.0000220849.48650.59>.
- Brincks, A., Montag, S., Howe, G. W., Huang, S., Siddique, J., Ahn, S., et al. (2018). Addressing methodologic challenges and minimizing threats to validity in synthesizing findings from individual-level data across longitudinal randomized trials. *Prevention Science*, *19*(1), 60–73.
- Brown, C. H., Brincks, A., Huang, S., Perrino, T., Cruden, G., Pantin, H., et al. (2018). Two-year impact of prevention programs on adolescent depression: An integrative data analysis approach. *Prevention Science*, *19*(1), 74–94.
- Bufferd, S., Dyson, M., Hernandez, I., & Wakschlag, L. S. (2016). Explicating the “developmental” in preschool psychopathology. *Developmental Psychopathology, Maladaptation and Psychopathology*, *3*, 152.
- Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., et al. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478–1485. <https://doi.org/10.1126/science.1248429>.
- Casey, B., Oliveri, M. E., & Insel, T. (2014). A neurodevelopmental perspective on the research domain criteria (RDoC) framework. *Biological Psychiatry*, *76*(5), 350–353.
- Choi, S. W., Schalet, B., Cook, K. F., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychological Assessment*, *26*(2), 513.
- Chronis-Tusciano, A., Degnan, K. A., Pine, D. S., Perez-Edgar, K., Henderson, H. A., Diaz, Y., et al. (2009). Stable early maternal report of behavioral inhibition predicts lifetime social anxiety disorder in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(9), 928–935.
- Clauss, J. A., & Blackford, J. U. (2012). Behavioral inhibition and risk for developing social anxiety disorder: A meta-analytic study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(10), 1066–1075 e1061.
- Cole, P. M., Luby, J. L., & Sullivan, M. W. (2008). Emotions and the development of childhood depression: Bridging the gap. *Child Dev Perspect*, *2*(3), 141–148. <https://doi.org/10.1111/j.1750-8606.2008.00056.x>.
- Corbelli, J., Borrero, S., Bonnema, R., McNamara, M., Kraemer, K., Rubio, D., et al. (2014). Use of the Gail model and breast cancer preventive therapy among three primary care specialties. *Journal of Women's Health*, *23*(9), 746–752.
- Curran, P. J. (2009). The seemingly quixotic pursuit of a cumulative psychopathological science: Introduction to the special issue. *Psychological Methods*, *14*(2), 77.
- Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, *14*(2), 81.
- Dawson, G., Ashman, S. B., & Carver, L. J. (2000). The role of early experience in shaping behavioral and brain development and its implications for social policy. *Development and Psychopathology*, *12*(4), 695–712.
- Eggebrecht, A. T., Elison, J. T., Feczko, E., Todorov, A., Wolff, J. J., Kandala, S., et al. (2017). Joint attention and brain functional connectivity in infants and toddlers. *Cerebral Cortex*, *27*(3), 1709–1720.
- Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 313–337.
- Emerson, R. W., Adams, C., Nishino, T., Hazlett, H. C., Wolff, J. J., Zwaigenbaum, L., et al. (2017). Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. *Science Translational Medicine*, *9*(393) eaag2882.
- Ferrante, M., Redish, A. D., Oquendo, M. A., Averkbeck, B. B., Kinnane, M. E., & Gordon, J. A. (2018). *Computational psychiatry: A report from the 2017 NIMH workshop on opportunities and challenges* (In).
- Finlay-Jones, A., Varcin, K., Leonard, H., Bosco, A., Alvares, G., & Downs, J. (2019). *Very early identification and intervention for infants at risk of neurodevelopmental disorders: A transdiagnostic approach*. Child Development Perspectives.
- Gaffrey, M. S., & Luby, J. L. (2012). *Kiddie-schedule for affective disorders and schizophrenia - early childhood version, 2012 working draft (KSADS-EC)*. St. Louis, MO: Washington University School of Medicine.
- Glasgow, R. E. (2013). What does it mean to be pragmatic? Pragmatic methods, measures, and models to facilitate research translation. *Health Education & Behavior*, *40*(3), 257–265.
- Goldstein, B. L., Shankman, S. A., Kujawa, A., Torpey-Newman, D. C., Dyson, M. W., Olino, T. M., et al. (2018). Positive and negative emotionality at age 3 predicts change in frontal EEG asymmetry across early childhood. *Journal of Abnormal Child Psychology*, 1–11.
- Grabell, A. S., Li, Y., Barker, J. W., Wakschlag, L. S., Huppert, T. J., & Perlman, S. B. (2018). Evidence of non-linear associations between frustration-related prefrontal cortex activation and the normal: Abnormal spectrum of irritability in young children. *Journal of Abnormal Child Psychology*, *46*(1), 137–147.
- Graham, A. M., Buss, C., Rasmussen, J. M., Rudolph, M. D., Demeter, D. V., Gilmore, J. H., et al. (2016). Implications of newborn amygdala connectivity for fear and cognitive development at 6-months-of-age. *Dev Cogn Neurosci*, *18*, 12–25.
- Graham, A. M., Rasmussen, J. M., Entringer, S., Ward, E. B., Rudolph, M. D., Gilmore, J. H., et al. (2018). Maternal cortisol concentrations during pregnancy and sex-specific associations with neonatal amygdala connectivity and emerging internalizing behaviors. *Biological Psychiatry*, *2*(85), 172–181. <https://doi.org/10.1016/j.biopsych.2018.06.023>.
- Greenland, P., LaBree, L., Azen, S. P., Doherty, T. M., & Detrano, R. C. (2004). Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. *Jama*, *291*(2), 210–215.
- Halperin, J. M., & Marks, D. J. (2019). Practitioner Review: Assessment and treatment of preschool children with attention-deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, *60*(9), 930–943. <https://doi.org/10.1111/jcpp.13014>.
- Hanson, J. L., Hair, N., Shen, D. G., Shi, F., Gilmore, J. H., Wolfe, B. L., et al. (2013). Family poverty affects the rate of human infant brain growth. *PLoS One*, *8*(12), e80954. <https://doi.org/10.1371/journal.pone.0080954>.
- Hay, D. F., Waters, C. S., Perra, O., Swift, N., Kairis, V., Phillips, R., et al. (2014). Precursors to aggression are evident by 6 months of age. *Developmental Science*, *17*(3), 471–480.
- Howell, B. R., Styner, M. A., Gao, W., Yap, P.-T., Wang, L., Baluyot, K., et al. (2019). The UNC/UMN baby connectome project (BCP): An overview of the study design and protocol development. *NeuroImage*, *185*, 891–905. <https://doi.org/10.1016/j.neuroimage.2018.03.049>.
- Jernigan, T. L., Brown, T. T., Hagler, D. J., Jr., Akshoomoff, N., Bartsch, H., Newman, E., et al. (2016). The pediatric imaging, neurocognition, and genetics (PING) data repository. *NeuroImage*, *124*, 1149–1154.
- Kaat, A. J., Blackwell, C. K., Estabrook, R., Burns, J. L., Petittler, A., Briggs-Gowan, M. J., et al. (2019). Linking the child behavior checklist (CBCL) with the multidimensional assessment profile of disruptive behavior (MAP-DB): Advancing a dimensional spectrum approach to disruptive behavior. *Journal of Child and Family Studies*, *28*(2), 343–353.
- Kessel, E. M., Meyer, A., Hajcak, G., Dougherty, L. R., Torpey-Newman, D. C., Carlson, G. A., et al. (2016). Transdiagnostic factors and pathways to multifinality: The error-related negativity predicts whether preschool irritability is associated with internalizing versus externalizing symptoms at age 9. *Development and Psychopathology*, *28*(4pt1), 913–926.
- Kolen, M. J., & Brennan, R. L. (2014). Test equating, scaling, and linking: Methods and practices. *Springer Science & Business Media*.
- Lloyd-Jones, D. M. (2010). Cardiovascular risk prediction: Basic concepts, current status, and future directions. *Circulation*, *121*(15), 1768–1777.
- Lorber, M. F., Del Vecchio, T., & Slep, A. M. S. (2015). The emergence and evolution of infant externalizing behavior. *Development and Psychopathology*, *27*(3), 663–680.
- Luby, J. L. (2012). Dispelling the “they’ll grow out of it” myth: Implications for intervention. *Am J Psychiatric Assoc*, *169*, 1127–1129. <https://doi.org/10.1176/appi.ajp.2012.12081037>.
- Luby, J. L., Belden, A. C., Jackson, J. J., Lessov-Schlaggar, C. N., Harms, M. P., Tillman, R., et al. (2016). Early childhood depression and alterations in the trajectory of gray matter maturation in middle childhood and early adolescence. *JAMA Psychiatry*, *73*(1), 31–38. <https://doi.org/10.1001/jamapsychiatry.2015.2356>.
- McWeeny, S., Manning, B. L., Beach, S. D., Eddy, M. D., Gaab, N., Gabrieli, J. D. E., & Norton, E. S. (Manuscript in preparation). Reliability of the ERP mismatch negativity response in kindergartners.
- McKinnon, C. J., Eggebrecht, A. T., Todorov, A., Wolff, J. J., Elison, J. T., Adams, C. M., et al. (2019). Restricted and repetitive behavior and brain functional connectivity in infants at risk for developing autism spectrum disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(1), 50–61. <https://doi.org/10.1016/j.bpsc.2018.09.008>.

- Miller, M., Iosif, A.-M., Young, G. S., Hill, M. M., & Ozonoff, S. (2016). Early detection of ADHD: Insights from infant siblings of children with autism. *Journal of Clinical Child and Adolescent Psychology*, 1–8.
- Mittal, V. A., & Wakschlag, L. S. (2017). Research domain criteria (RDoC) grows up: Strengthening neurodevelopment investigation within the RDoC framework. *Journal of Affective Disorders*, 216, 30–35.
- Norton, E. S., Isaacs, S. H., Rogers, C., Smyer, C., Krogh-Jespersen, S., Briggs-Gowan, M., ... Wakschlag, L. S. (Manuscript in preparation). Reduced cortical thickness associated with irritability in preschool children: Replication across two diverse samples.
- Ozonoff, S. (2015). Early detection of mental health and neurodevelopmental disorders: The ethical challenges of a field in its infancy. *Journal of Child Psychology and Psychiatry*, 56(9), 933–935.
- Ozonoff, S., Young, G. S., Belding, A., Hill, M., Hill, A., Hutman, T., et al. (2014). The broader autism phenotype in infancy: When does it emerge? *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(4), 398–407 e392.
- Paulus, M. P., Huys, Q. J., & Maia, T. V. (2016). A roadmap for the development of applied computational psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 1(5), 386–392.
- Pencina, M. J., & D'Agostino Sr, R. B. (2012). Thoroughly modern risk prediction. *Science Translational Medicine*, 4(131) 131–110.
- Pennap, D., Zito, J. M., Santosh, P. J., Tom, S. E., Onukwugha, E., & Magder, L. S. (2018). Patterns of early mental health diagnosis and medication treatment in a medicaid-insured birth cohort. *JAMA pediatrics*, 172(6), 576–584.
- Pickett, K. E., Rathouz, P. J., Dukic, V., Kasza, K., Niessner, M., Wright, R. J., et al. (2009). The complex enterprise of modelling prenatal exposure to cigarettes: What is 'enough'? *Paediatric & Perinatal Epidemiology*, 23(2), 160–170.
- Pine, D. S. (2017). Clinical advances from a computational approach to anxiety. *Biological Psychiatry*, 82(6), 385–387.
- Pool, L. R., Ning, H., Wilkins, J., Lloyd-Jones, D. M., & Allen, N. B. (2018). Use of long-term cumulative blood pressure in cardiovascular risk prediction models. *JAMA cardiology*, 3(11), 1096–1100.
- Rogers, C. E., Sylvester, C. M., Mintz, C., Kenley, J. K., Shimony, J. S., Barch, D. M., et al. (2017). Neonatal amygdala functional connectivity at rest in healthy and preterm infants and early internalizing symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(2), 157–166. <https://doi.org/10.1016/j.jaac.2016.11.005>.
- Shaw, D. S. (2013). Future directions for research on the development and prevention of early conduct problems. *Journal of Clinical Child and Adolescent Psychology*, 42(3), 418–428.
- Shaw, D. S., & Gilliam, M. (2017). Early childhood predictors of low-income boys' pathways to antisocial behavior in childhood, adolescence, and early adulthood. *Infant Mental Health Journal*, 38(1), 68–82.
- Shonkoff, J. P. (2016). Capitalizing on advances in science to reduce the health consequences of early childhood adversity. *JAMA pediatrics*, 170(10), 1003–1007.
- Siddique, J., Reiter, J. P., Brincks, A., Gibbons, R. D., Crespi, C. M., & Brown, C. H. (2015). Multiple imputation for harmonizing longitudinal non-commensurate measures in individual participant data meta-analysis. *Statistics in Medicine*, 34(26), 3399–3414.
- Smith, J., Wakschlag, L. S., Walkup, J., Wilson, M., T, D., & Shaw, D. (in press). Dysregulated Irritability as a window on young children's psychiatric risk: Transdiagnostic impact via the Family Check-Up. *Development & Psychopathology*.
- Smyser, C. D., Dosenbach, N. U., Smyser, T. A., Snyder, A. Z., Rogers, C. E., Inder, T. E., et al. (2016). Prediction of brain maturity in infants using machine-learning algorithms. *NeuroImage*, 136, 1–9. <https://doi.org/10.1016/j.neuroimage.2016.05.029>.
- Sylvester, C. M., Smyser, C. D., Smyser, T., Kenley, J., Ackerman, J. J., Jr., Shimony, J. S., et al. (2017). Cortical functional connectivity evident after birth and behavioral inhibition at age 2. *American Journal of Psychiatry*, 175(2), 180–187.
- Wakschlag, L. S., Leventhal, B. L., Briggs-Gowan, M. J., Danis, B., Keenan, K., Hill, C., et al. (2005). Defining the "disruptive" in preschool behavior: What diagnostic observation can teach us. *Clinical Child and Family Psychology Review*, 8(3), 183–201. <https://doi.org/10.1007/s10567-005-6664-5>.
- Wakschlag, L. S., Perlman, S., Blair, R. J., Leibenluft, E., Briggs-Gowan, M., & Pine, D. (2018). The neurodevelopmental basis of early childhood disruptive behavior: Irritable and callous phenotypes as exemplars. *American Journal of Psychiatry*, 175(2), 114–130 10.1176/appi.ajp.2017.17010045.
- Wakschlag, L. S., Roberts, M., Flynn, R., Smith, J., Krogh-Jespersen, S., Kaat, A., et al. (2019). Future directions for early childhood prevention of mental disorders: A road map to mental health, earlier. *Journal of Clinical Child and Adolescent Psychology*, 1–16.
- Wakschlag, L. S., Tolan, P., & Leventhal, B. (2010). Research Review: 'Ain't misbehavin': Towards a developmentally specified nosology for preschool disruptive behavior. *Journal of Child Psychology and Psychiatry*, 51(1), 3–22.
- Zito, J. M., Safer, D. J., Valluri, S., Gardner, J. F., Korelitz, J. J., & Mattison, D. R. (2007). Psychotherapeutic medication prevalence in Medicaid-insured preschoolers. *Journal of Child and Adolescent Psychopharmacology*, 17(2), 195–203. <https://doi.org/10.1089/cap.2007.0006>.