SUPPLEMENTARY INFORMATION

Dense Sampling Approaches for Psychiatry Research: Combining Scanners and Smartphones

McGowan *et al.*

Supplemental Material

Systematic Review Methodology

Studies included in this review were selected to investigate dense sampling approaches that offer insight into brain-behavior relationships in real-world contexts. We performed our systematic review (Figure S1) in accordance with guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (1). In line with this special issue, our objective was to discuss the state-of-the-science of dense, longitudinal sampling of individual differences in dynamic brain-behavior relationships informing health and disease trajectories in real-world contexts. To do so, our search strategy involved identifying relevant search terms and databases, allocating a time frame (published within the last three years), and applying eligibility criteria (combining brain and ambulatory modalities in everyday life), and applying filters (English language articles). We then reviewed the literature using the inclusion and exclusion criteria, screened for duplicates, and conducted a critical analysis by extracting information from each article identified in our search. All papers were reviewed by the first author (ALM) and independent reviewers (ZMB, MJ, YK, FS, MS) split up the articles for review, collectively forming the second rater for all categories. Articles identified for review were analyzed using a table to sort and abstract selected information pertaining to our objective and readers of this special issue. We describe each step of this process in more detail below.

Search Method

An initial scoping literature review was conducted to identify records combining brain and ambulatory dense sampling approaches to examine brain-behavior relationships in real-world contexts. This initial review facilitated the further systematic review of records identified for abstracting. ALM performed a systematic search on 2/23/2022 in both Web of Science and PsychINFO using the following search terms:

(fMRI OR mri OR “functional magnetic resonance imag*” OR “magnetic resonance imag*” OR “brain imag*” OR “network neurosci*”) AND (smartphone OR “experience sampling” OR “experience-sampling” OR “daily diary” OR “ecological momentary assessment” OR “ambulatory assess*”)

In both databases, additional filters were selected during the search, including limiting records to English language, published during the dates of 01/01/2019 to 02/23/2022, and article document types.

Eligibility Criteria

Eligible records were required to have been published during the last three years to maintain relevance to the current special issue, which highlighted the advances in these methods in the past three years. Only records published in peer-reviewed journals were included, thus excluding conference abstracts and book chapters. Records had to be an empirical article, thus excluding protocol papers, literature reviews, and editorials, dissertations, commentaries, and opinion papers. Inclusion criteria for the review were English language, full text, peer-reviewed and scholarly empirical articles from the period 01/01/2019 to 02/23/2022.

Screening

See Figure S1 for a PRISMA flow diagram. All identified records from Web of Science (*n* = 113 records) and PsychINFO (*n* = 55 records) were exported to Excel files, including title, author, year of publication, and abstract information. The two databases were combined and initial screening was conducted by ALM. ALM screened the database for duplicates (*n* = 39 were removed). Next, the remaining records (*n* = 154) were screened based on title and abstract by ALM to exclude based on being on an unrelated topic, protocol paper, or review paper (*n* =46 records removed). ALM reviewed the reference lists of the excluded literature review records to identify any missing records to be screened included in screening (*n* = 25 records identified). A total of 108 full-text records were sought for retrieval and coded by 2 independent coders to extract the following information: sample size analyzed (*n*), mean age (years), population, duration of study and number of time points, neuropsychological/psychometric measures, ambulatory assessment modality, brain modality, rationale for combining the modalities, how the different modalities were combined, and main findings. At this step, additional records were excluded for not meeting eligibility criteria (*n* = 46). The Fleiss' kappa (121) was 0.76, indicating substantial agreement (88.2% agreement) between coders to exclude these records. We include the final *n* = 64 records in our literature review matrix (see Supplemental Table S1). In instances where there was disagreement between coders, the lead author (ALM) reviewed the information in the manuscript an additional time and a consensus was reached as to the correct coding with the other raters (ZMB, MJ, YK, FS, MS).

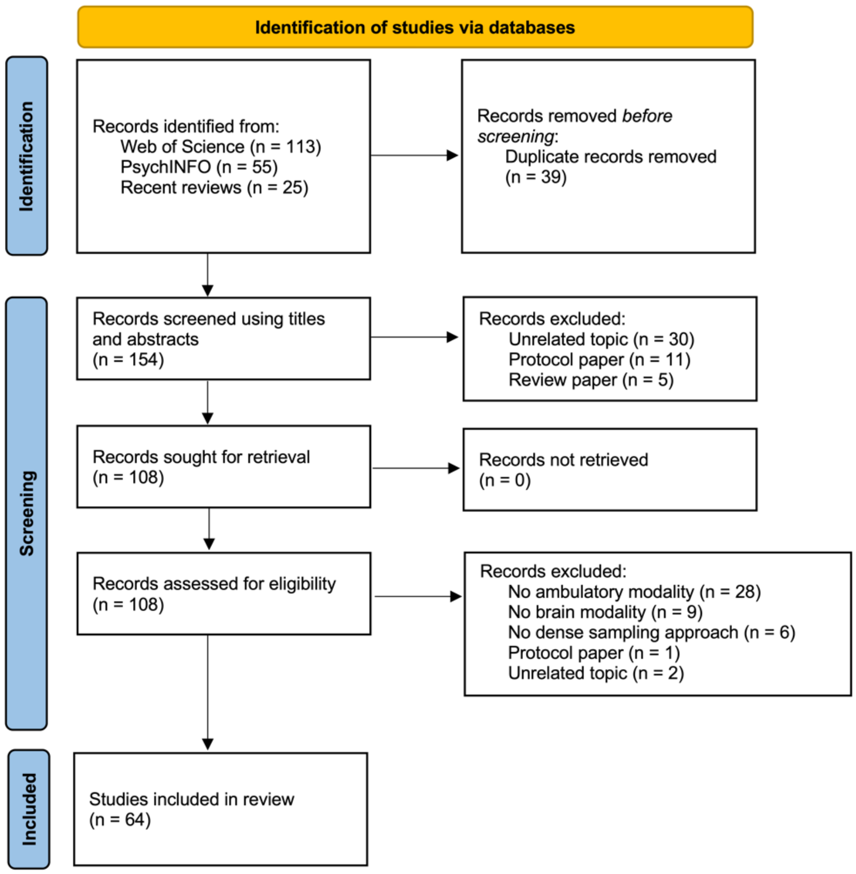
Review Summary

Here, we summarize the main findings from our systematic review (Table S1, Figure S2). The average study sample size was 85.4 (*SD* = 100.2) participants. The developmental period most commonly investigated was young adulthood (18-34 years old), with mean participant age of 26.7 (*SD* = 9.4) years. The majority of studies used a non-clinical community sample (*n* = 25), and two studies used electronic databases (*n* = 2; i.e., organ donors, Google User Experience portal) to recruit their community sample. Clinical samples spanned varied conditions, with depression (*n*=7) being the most represented. Studies employed a variety of ambulatory assessments, with the majority of studies (79%) using smartphone-based or email experience-sampling reports (e.g., text messaging, daily email survey links, or smartphone app collecting reports), 21% of studies using passive sensors (e.g., geolocation, mobility patterns, daily screen lock/unlock), 17% using other methods (e.g., smartphone stimulus presentation for cognitive tasks, other smartphone apps like guided meditation), and 9% using accelerometry (e.g., wrist-worn accelerometer). Study duration and time scale (i.e., seconds, hours, days, weeks) varied widely between studies as some studies incorporated interventions lasting one month to a year. The average number of time points captured for ambulatory assessment used in the study's primary analyses was 131.4 (*SD* = 552.74). The majority of studies (91%) used neuroimaging as the brain modality (i.e., fMRI, sMRI, rs-fMRI, qMRI, DWI, DTI), 9% used other brain modalities (i.e., icEEG, EEG, mobile EEG, eye gaze, PET-CT, neurofeedback), and two studies used multiple brain modalities (i.e., fMRI and neurofeedback; PET-CT and sMRI). Participants typically completed one MRI scanning session (83% of studies), yet one study reported collecting 35,714 hours of intracranial EEG (icEEG). Studies used various neuropsychological and psychometric measures, with 62% of studies using a psychiatric and/or clinical measure (e.g., Diagnostic and Statistical Manual of Mental Disorders interview, Center for Epidemiological Studies Depression Scale), 48% of studies measuring emotion, affect, and/or affect (e.g., Positive and Negative Affective Schedule), 47% of studies assessing a cognitive domain (e.g., working memory), 38% of studies measuring other domains (e.g., sleep, physical activity), 22% of studies measuring social functioning (e.g., social network survey, peer social interactions, parent-child interactions), and 12% of studies measuring contextual factors (e.g., exposure to green space, conversations about drinking in daily life). Studies differed in their rationales for combining scanners and smartphones, with studies endorsing multiple reasons. The majority of studies (70%) endorsed measuring real-life behaviors, 41% of studies sought to identify neural correlates or disease markers, 38% of studies desired to increase ecological validity of their findings, 17% cited overcoming measurement limitations of brain modalities and/or experience-sampling methods (i.e., retrospective report bias, highly controlled laboratory tasks), 17% sought to inform the design of therapeutics that could be personalized or incorporated into participants' daily lives, 15% described capturing the interaction between or dynamics of brain-behavior relationships, and 8% cited other reasons (e.g., capture individual variability in behavior). In most cases, data collection from scanners preceded data collection from smartphones (Figure S3).

Supplementary References

1. Moher D, Altman DG, Liberati A, Tetzlaff J (2011): PRISMA statement. *Epidemiology* 22: 128.

Figure S1. PRISMA flow diagram depicting systematic review.

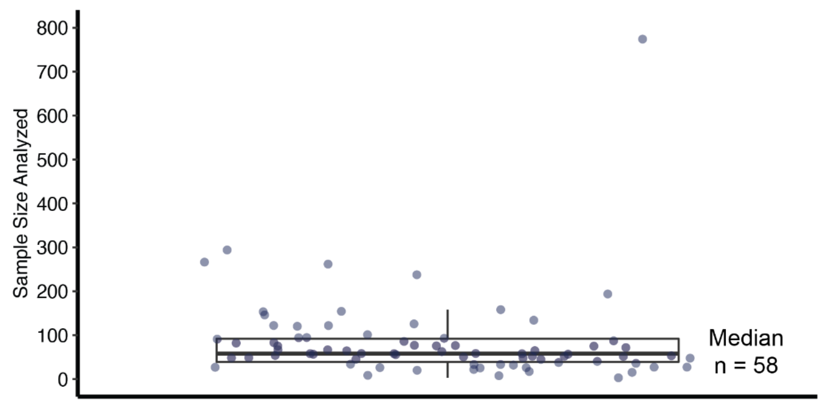


From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

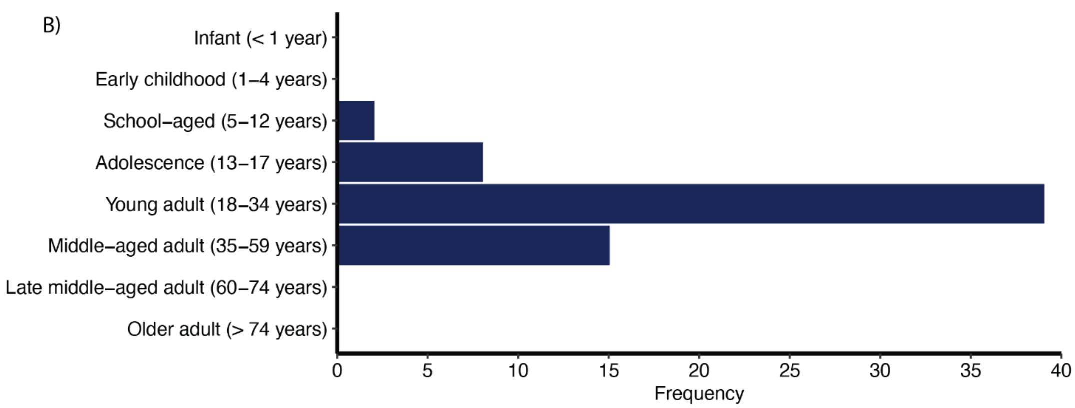
For more information, visit: http://www.prisma-statement.org/

Figure S2. Graphic illustration of (A) sample size analyzed. Each dot represents a study and the median sample analyzed (n= 58) represented as a black line. (B) Frequency of each age group included in studies. The majority of studies use young adults. (C) Population examined. Most studies use a community sample free from clinical diagnosis.

A)



B)



C)

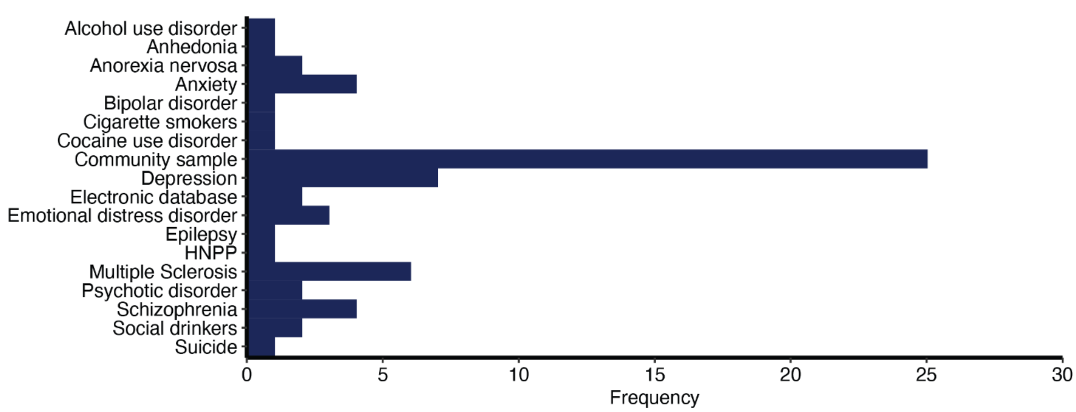


Figure S3. Temporal precedence of brain scan and smartphone sampling periods.