Archival Report

Biological Psychiatry

# Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies

Andre F. Marquand, Iead Rezek, Jan Buitelaar, and Christian F. Beckmann

**ABSTRACT**

**BACKGROUND:** Despite many successes, the case-control approach is problematic in biomedical science. It introduces an artificial symmetry whereby all clinical groups (e.g., patients and control subjects) are assumed to be well defined, when biologically they are often highly heterogeneous. By definition, it also precludes inference over the validity of the diagnostic labels. In response, the National Institute of Mental Health Research Domain Criteria proposes to map relationships between symptom dimensions and broad behavioral and biological domains, cutting across diagnostic categories. However, to date, Research Domain Criteria have prompted few methods to meaningfully stratify clinical cohorts.

**METHODS:** We introduce normative modeling for parsing heterogeneity in clinical cohorts, while allowing predictions at an individual subject level. This approach aims to map variation within the cohort and is distinct from, and complementary to, existing approaches that address heterogeneity by employing clustering techniques to fractionate cohorts. To demonstrate this approach, we mapped the relationship between trait impulsivity and reward-related brain activity in a large healthy cohort ($N = 491$).

**RESULTS:** We identify participants who are outliers within this distribution and show that the degree of deviation (outlier magnitude) relates to specific attention-deficit/hyperactivity disorder symptoms (hyperactivity, but not inattention) on the basis of individualized patterns of abnormality.

**CONCLUSIONS:** Normative modeling provides a natural framework to study disorders at the individual participant level without dichotomizing the cohort. Instead, disease can be considered as an extreme of the normal range or as —possibly idiosyncratic—deviation from normal functioning. It also enables inferences over the degree to which behavioral variables, including diagnostic labels, map onto biology.

*Keywords:* Gaussian process, Heterogeneity, Normative model, Outlier detection, Patient stratification, Research Domain Criteria

http://dx.doi.org/10.1016/j.biopsych.2015.12.023

The case-control approach to studying brain disorders has been successful for detecting group effects, for example, between patients and control subjects. However, it becomes problematic in domains such as psychiatry where disorders are diagnosed on the basis of symptoms that overlap between disorders, often yielding clinical groups that are heterogeneous and overlapping. This problem is particularly acute in psychiatry because biological tests to assist diagnosis or predict outcome have not been developed (1). Moreover, the case-control paradigm induces an artificial symmetry such that both cases and controls are assumed to be well-defined entities (Figure 1). This does not match the clinical view of disease, where disorders in individual patients manifest as deviations from a normal pattern of functioning.

In response to this problem, the National Institute of Mental Health launched the Research Domain Criteria (RDoC) initiative (2), which encourages researchers to link symptom dimensions with biological systems, cutting across diagnostic classifications. The ultimate aim of RDoC is to find "new ways of classifying psychiatric diseases based on multiple dimensions of biology and behavior" (http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml)—reducing heterogeneity in clinical cohorts; improving the neurobiological validity of disease classifications; and enabling more effective, personalized treatments. These objectives are also consistent with the European roadmap for mental health research (3). These objectives are difficult to achieve within the case-control paradigm, which, by definition, entails partitioning cohorts according to predefined labels, precluding later inferences about their validity.

The RDoC initiative has prompted considerable discussion (4,5) but to date has led to few methods to study heterogeneity within clinical cohorts. Of the reports published, nearly all have employed data-driven clustering methods aiming to fractionate clinical groups mostly on the basis of neuropsychological measures. For example, clustering methods have been applied




ISSN: 0006-3223

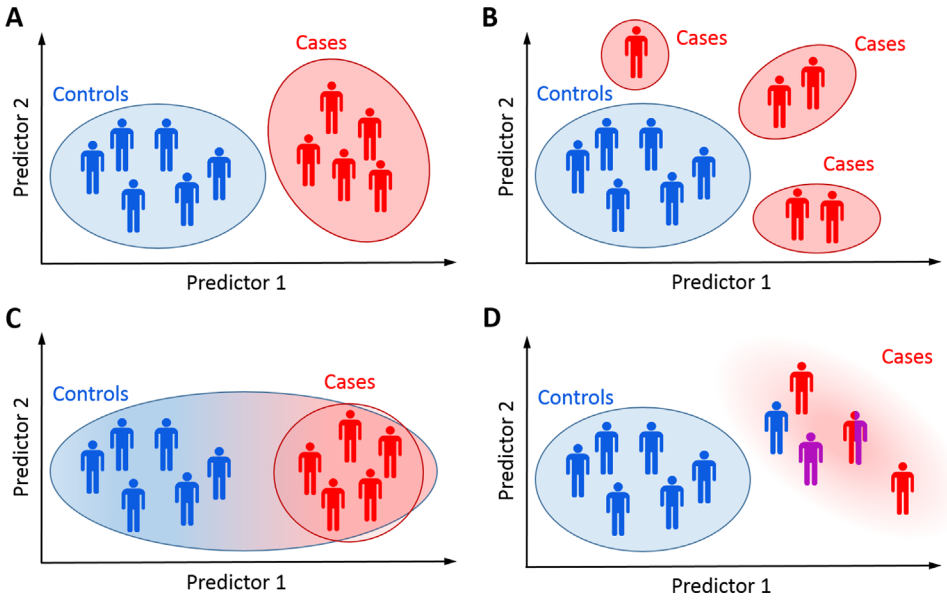


**Figure 1.** The classical case-control approach assumes that cases and controls each form a well-defined group **(A)**. This may often be a reasonable assumption, but in practice many other scenarios are possible. The clinical population may be composed of multiple groups, each having distinct pathology **(B)**; disease-related variation may be nested within healthy variation **(C)**; or the clinical group may be diffuse and heterogeneous as a result of misdiagnosis, comorbidities, or an aggregation of different pathologies **(D)**.

to subtype attention-deficit/hyperactivity disorder (ADHD) (6–9), mood disorders (10,11), and schizophrenia (12,13). Clustering is useful for identifying subgroups of participants at a particular time point but also has problems: 1) there are many different ways to partition clinical populations depending on the measures and clustering algorithm used; 2) some participants may not clearly belong to any class, or some classes may become unmanageably small (8); 3) patient subgroups may not be stable over time (14); 4) it may be difficult to choose a unique optimal number of clusters (e.g., different metrics may yield different optimal numbers of clusters or may not identify a unique maximum); 5) finally, it is unclear whether healthy participants should be clustered separately or in combination with patients. Some reports have suggested that disease variation may be nested within normal variation (7).

In this article, we propose an alternative conceptual advance for parsing heterogeneity in clinical and healthy cohorts. In contrast to clustering approaches, we propose a normative modeling approach that models biological variation across either 1) the entire study population (including all clinical groups) or 2) a large healthy sample. The intuition is that by mapping the full range of population variation, we can consider symptoms in individual patients as an extreme value within this distribution. This is analogous to the use of growth charts to map child development in terms of height and weight as a function of age, where deviations from a normal growth trajectory manifest as outliers within the normative range at each age. This approach is fundamentally different from, and complementary to, clustering (Figure 1). More concretely, we predict biological measures of brain function (e.g., neuroimaging) on the basis of clinically relevant covariates (e.g., trait measures). We build on preliminary work by ourselves and others (15–18) to introduce an analytical framework that allows us to 1) use data from large cohorts to learn a normative distribution that characterizes the study population; 2) make probabilistic statements about which participants deviate from the normative pattern; and 3) statistically map the brain regions underlying these deviations on a case-by-case basis, while permitting 4) diagnostic labels to be used as predictor variables, enabling inferences over the labels just as any other variable.

To illustrate, we map the relationship between trait impulsivity and reward-related brain activity in a large, healthy sample. This relationship is of high clinical relevance because impulsivity and impairments in reward processing are core features of many disorders, including ADHD (19,20) and addiction (21). We use delay discounting to quantify impulsivity, which measures the degree to which individuals devalue future rewards relative to immediate rewards (22) and is a stable measure of trait impulsivity (23). We then relate the model predictions to ADHD symptom dimensions to highlight specificity for particular symptom domains. Our approach is predicated on the assertions that 1) understanding healthy variation is a prerequisite to understanding disease variation and that this requires 2) the ability to determine where each subject lies within the population range because variation associated with most disorders overlaps with normal variation. We show that normative modeling provides a flexible and powerful means to operationalize these desiderata, to study variation in individual participants, and to highlight axes of variation relevant to clinical symptoms.

## METHODS AND MATERIALS

### Overview of Normative Modeling

Figure 2 shows an overview of the approach. First, we estimate a normative model that links clinical and biological variables. Specifically, we use Gaussian process regression (24) to predict a set of biological response variables (e.g., neuroimaging) from a set of clinically relevant covariates (e.g., trait scores), while estimating predictive confidence for every prediction. Measures of predictive confidence are important because they quantify the fit of each point to the normative