Propagating uncertainity in normative models

Richard Dinga and Marijn van Wingerden

Department of Cognitive Science and Artificial Intelligence,
Tilburg School for Humanities and Digital Sciences,
University of Tilburg, Tilburg, the Netherlands
Email: r.dinga@tilburguniversity.edu

Abstract

Normative modeling estimates population-level brain and psychometric phenotypes, adjusting for covariates like age and sex. However, these phenotypes often derive from noisy data, not direct measurements. Standard models applied to such data conflate true variation with measurement error, biasing estimates and distorting atypicality scores. Here we introduce a multilevel normative model that explicitly accounts for measurement uncertainty, enabling unbiased estimation of true phenotype distributions. Further, we propose new atypicality measures that incorporate uncertainity of the measurement thus reducing. Simulations confirm the method's accuracy in recovering true distributions. Applied to real data, it improves detection sensitivity for abnormal task-based brain activity and psychometric network deviations over conventional methods. By modeling measurement noise, this approach enhances the reliability of normative models, with implications for both research and clinical use, and potential applicability to other domains involving uncertain measurements.

1 Introduction

Normative models have been widely employed to estimate age- and sex-dependent normative ranges of brain-imaging phenotypes, with the goal of detecting brain atypicalities in individuals and linking these deviations to potential clinical characteristics (Dinga et al., 2021). These models aim to identify individuals whose brain phenotypes fall outside normative ranges, indicating atypical development or pathology.

Deviations from normative ranges of brain-imaging phenotypes have been linked to schizophrenia and bipolar disorder (Wolfers et al., 2018), ADHD (Wolfers et al., 2020), and autism (Bethlehem et al., 2020). This approach surpasses traditional case–control studies by not assuming uniform atypicalities across individuals with the same diagnosis (Marquand et al., 2016; Wolfers et al., 2020). Each individual can exhibit a unique form of brain atypicality detectable as a deviation from normative ranges.

A major challenge is that many brain variables are estimated with error rather than directly measured, biasing normative models and undermining atypicality detection. Task-related brain activity, for example, is represented by an estimated regression coefficient whose precision depends on design efficiency, BOLD time-series length, and residual variance. This measurement error varies across scanners, subjects, and sessions, distorting the true distribution and compromising atypicality estimates.

This study addresses these challenges using multilevel modeling to estimate the distribution of an unmeasured variable given the known uncertainty of its estimate. Building on established neuroimaging models (Friston et al., 2005; Mumford and Poldrack, 2007), we integrate them with normative modeling and propose new atypicality measures that account for measurement error.

We apply a mixed-effect location–scale meta-regression model (Viechtbauer and López-López, 2022) that treats subject-level measurements as random effects with known within-subject variance and unknown between-subject variance. The between-subject variance can depend on covariates, enabling heteroskedastic normative distributions. Simulations show unbiased parameter estimation even with subject-specific or covariate-correlated measurement errors. In real data, the proposed method markedly improves detection accuracy of atypical brain activity maps compared with traditional normative models.

2 Methods

The proposed normative model is based on a location–scale random-effects meta-regression model (Viechtbauer and López-López, 2022).

$$y_i = \mu + \eta_i + \epsilon_i, \qquad \epsilon_i \sim \mathcal{N}(0, \nu_i), \qquad \eta_i \sim \mathcal{N}(0, \tau^2).$$

This can be rewritten as a multilevel model:

$$\text{Level 1 (within-subjects):} \quad y_i = \mu_i + \epsilon_i, \quad (1)$$

$$\text{Level 2 (between-subjects):} \quad \mu_i = \mu + \eta_i. \quad (2)$$

Equivalently,

$$\mu_i \sim \mathcal{N}(\mu, \tau^2),$$

where μ is the mean and τ^2 is the variance of the estimated normative distribution. Estimating the normative model therefore involves estimating μ and τ for the unconditional case. In a full location–scale normative model, μ and τ can themselves be modeled as functions of covariates, enabling, for example, age-dependent centiles of brain activity.

The modified z-score quantifying the atypicality of a measurement relative to the norms is

$$z = \frac{y - \hat{y}}{\sqrt{\tau^2 + \epsilon^2}},$$

which scales the deviation by both the latent normative variance and the measurement uncertainty.

3 Results

In preliminary simulations, the proposed normative models were able to estimate the parameters of a normative distribution without bias, contrary to traditional methods, which overestimated the variance of the normative distribution due to unaccounted measurement error. The modified z-scores had higher discriminatory power in detecting samples sampled from the normative distributions and outliers. This finding was preliminary repeated in a real psychometric networks and functional MRI activation maps with simulated abnormalities.

4 Discussion

The proposed multi-level normative modeling framework builds on established statistical methodologies, including random-effects location–scale meta-regression (Viechtbauer and López-López, 2022) and mixed-effects neuroimaging modeling approaches (Friston et al., 2005). Framing normative modeling within a meta-analytic context generalizes existing methods to estimate population-level phenotype distributions while explicitly addressing measurement error in observed variables. This resolves a critical limitation in prior normative modeling work. Unlike neuroimaging multi-level models that focus on group-level estimates of brain activity or group differences (Friston et al., 2005; Mumford and Poldrack, 2007), our framework estimates covariate-adjusted normative distributions (e.g., accounting for age and sex) to enable individualized atypicality assessments.

The proposed normative model assumes that the modeled variables are conditionally normally distributed in the reference population. If the normal distribution does not sufficiently approximate the distribution of variables, transformation is required; otherwise, the resulting z-scores or other atypicality measures will be misleading. Consequently, the method is unsuitable for regularized psychometric or functional connectivity network models that are sparse—such as those employing LASSO—because the distribution of estimated variables will contain many zeros and violate normality.

The model's conditional normality assumption implies that variables must either conform to normality after suitable transformation or be modeled with alternative approaches; otherwise, the resulting z-scores and centiles remain unreliable. Hence, the framework is inappropriate for LASSO-regularized coefficients common in psychometric (Epskamp and Fried, 2018) and brain functional network estimation, where sparsity yields many zeros and precludes normality.

References

Bethlehem, R. A. I., Seidlitz, J., Romero-Garcia, R., Trakoshis, S., Dumas, G., and Lombardo, M. V. (2020). A normative modelling approach reveals age-atypical cortical thickness in a subgroup of males with autism spectrum disorder. *Communications Biology*, 3(1):1–10.

Dinga, R., Fraza, C. J., Bayer, J. M. M., Kia, S. M., Beckmann, C. F., and Marquand, A. F. (2021). Normative modeling of neuroimaging data using generalized additive models of location scale and shape. bioRxiv preprint. doi:10.1101/2021.06.14.448106.

Epskamp, S. and Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, 23(4):617–634.

Friston, K. J., Stephan, K. E., Lund, T. E., Morcom, A., and Kiebel, S. (2005). Mixed-effects and fmri studies. *NeuroImage*, 24(1):244–252.

Marquand, A. F., Rezek, I., Buitelaar, J., and Beckmann, C. F. (2016). Understanding heterogeneity in clinical cohorts using normative models: Beyond case-control studies. *Biological Psychiatry*, 80(7):552–561.

Mumford, J. A. and Poldrack, R. A. (2007). Modeling group fmri data. *Social Cognitive and Affective Neuroscience*, 2(3):251–257.

Viechtbauer, W. and López-López, J. A. (2022). Location-scale models for meta-analysis. *Research Synthesis Methods*, 13(6):697–715.

Wolfers, T., Beckmann, C. F., Hoogman, M., Buitelaar, J. K., Franke, B., and Marquand, A. F. (2020). Individual differences v. the average patient: Mapping the heterogeneity in adhd using normative models. *Psychological Medicine*, 50(2):314–323.

Wolfers, T., Doan, N. T., Kaufmann, T., Alnæs, D., Moberget, T., Agartz, I., Buitelaar, J. K., Ueland,

T., Melle, I., Franke, B., Andreassen, O. A., Beckmann, C. F., Westlye, L. T., and Marquand, A. F. (2018). Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry*, 75(11):1146–1155.