Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Transdiagnostic approaches to precision medicine: A Computational Psychiatry Primer

Benjamin Illingworth[1], Timothy Sandhu[1], Eleanor Smith[1], Claudia Lage[1], Brónagh McCoy[1], and Rebecca Lawson[1]

[1]Affiliation not available

October 26, 2022

Benjamin J. G. Illingworth MBBS BSc[a,b], Timothy R. Sandhu[b,c] BSc, Eleanor S. Smith PhD[b] , Claudia Lage MSc[b], Brónagh McCoy PhD[b], Rebecca P. Lawson PhD[b]

[a] Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom

[b] Department of Psychology, University of Cambridge, Cambridge, United Kingdom

[c] MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom

Word count for the body of the text: 1484

Declaration of Interest: None

Funding:

This work was supported by a 'Royal Society Wellcome Trust Henry Dale Fellowship' (RPL, grant number 10.35802/206691) and 'The Portuguese Foundation for Science and Technology (FCT)' (C.L., grant number SFRH/BD/144811/2019).

Ethical approval:

Ethical approval not required.

Correspondence to:

Dr Benjamin J. G. Illingworth

Department of Psychology, University of Cambridge

Downing Street, Cambridge, CB2 3EB, United Kingdom

bi232@cam.ac.uk

@DocIllingworth

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Abstract

Psychiatry's foremost challenge is that its symptom-based classification has insufficient predictive value for clinical outcomes. Precision psychiatry attempts to improve treatment response by searching for predictive biomarkers, which should be described in the context of the neurobiology of illness to help facilitate the development of new treatments.

This primer for clinicians aims to outline the problems that precision psychiatry faces, how computational methods can overcome them and how these methods can be used in clinical practice. Clinicians have an important role in translating this research and incorporating computationally-derived transdiagnostic biomarkers in clinical trials opening the possibility for personalised therapies.

Keywords:

Diagnosis; biomarkers; precision medicine;

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Why current diagnostic categorisation may impair the progress of precision psychiatry

Precision medicine aims to increase treatment success by tailoring therapies based on individual differences. As in other specialities, psychiatry has attempted to implement approaches from precision medicine by seeking biomarkers to inform treatment and therefore improve prognosis. There is a need for a precision approach in psychiatry as current treatments have been shown to have poor initial success rates for patients within well-defined diagnostic categories [1].

However, in contrast to other areas of medicine, psychiatric patients may not fit any diagnostic category clearly, or may have multiple diagnoses (co-morbidity) [1]. This implies that in psychiatry, diagnostic categories may not be sensitive enough to individual differences or representative of underlying pathologies. Using a symptom-based nosology may impair the search for biomarkers, as diagnostic groups are likely to be heterogeneous in terms of aetiology and pathogenesis. Here, multiple neurobiological mechanisms may lead to similar symptoms or a single mechanism may lead to different symptoms depending on their interaction with other biopsychosocial factors [2] (Fig 1A).

How this can be overcome by redefining groups around mechanisms of disease

Motivated by the National Institute of Mental Health Research Domain Criteria (RDoC) initiative, a recent approach to this problem considers mental health transdiagnostically. The aim is to better capture the complexities of individuals' mental health conditions and explain the prevalence of deficits across multiple disorders [3]. Computational methods have been utilised to better characterise variation in psychopathology and to define new sub-groups of mental illness [3]. These groups should be built around the presence of physiological or cognitive biomarkers and have roots in the neuro-biological mechanisms of mental illnesses (Fig 1A).

Factor analysis is one method of discovering new dimensions of psychopathology. To this end, large online population studies have illustrated novel and robust dimensions of psychiatric disease derived from a consistent set of self-report questionnaires which, independent of traditional scoring systems, are also associated with specific deficits on cognitive tasks (Fig. 1B) [3].

Theory-driven approaches: advancing understanding of disease processes

Cognitive tasks and neuroimaging are combined with theoretical approaches in computational psychiatry to learn more about differences in behaviour, which in extremis are associated with symptoms of mental illness. Computational models are created with the aim of explaining observed behaviour (e.g. responses on tasks) in a mathematical framework. These models are composed of free variables, i.e. parameters which vary on a per individual basis and represent unobservable neurobiological processes, such as learning or decision-making. When these models are applied to imaging and to cognitive tasks, the resulting parameters can illustrate individual differences in performance at a mechanistic level, capturing the heterogeneity that is a cardinal feature of complex conditions.

Reinforcement learning models represent a promising example of this approach. Under this framework, the relationship between two modes of learning about rewards can be assessed e.g., goal-directed vs. habitual. Fitting these models to each individual's data produces parameters, i.e., summaries of the learning process, which can be compared to reveal if there is a significant difference between two predefined groups or if novel groups exist within the population.

These models have a biological basis, and provide further mechanistic insight into the processes underlying aberrant behaviour. For example, mismatches between what is expected and what is actually observed by an individual generates prediction errors, which are encoded by dopaminergic neurons in subcortical regions [1]. Deficits in goal-directed learning, and therefore a tendency towards habits, have been found across disorders such as binge eating, addiction, social anxiety and obsessive-compulsive disorder (OCD), and have also been related to a specific transdiagnostic factor - compulsive behaviour and intrusive thought - which underlies all of these disorders [3].

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Data-driven approaches: Creating applications in clinical settings

Another approach in the search for predictive biomarkers, is the use of data-driven approaches such as machine learning. These methods have been used successfully on neuroimaging and clinical data to answer questions about prognosis and treatment outcomes. In one study, using existing clinical trial data, clinical variables were used to predict remission in depression following treatment with selective serotonin re-uptake inhibitors (SSRIs)[4]. Supervised dimensionality reduction enabled large numbers of clinical variables to be reduced to a select few predictive variables, which were then fed into a machine learning algorithm to predict remission in patients with depression. Applying this algorithm had a number needed to treat (NNT) of 14 to achieve remission in depression[1]. Ultimately, machine learning methods could provide the basis for the creation of tools to support clinical decision making.

Better together: A combined theory-driven and data-driven approach

To realise the full potential of parameters derived from computational models as biomarkers, theory-driven models will have to be combined with data-driven (machine learning) approaches. Generative embedding is a process which exemplifies this combined approach. Here, models of brain connectivity or cognitive tasks are used to fit parameters, representing physiological or cognitive processes, which can then be classified into subgroups by machine learning techniques. These groups are validated against clinical variables, and assessed for prognostic value. One example uses dynamic causal modelling, which infers synaptic coupling between large neuronal populations based on connection strengths between brain regions, to predict participant disease trajectory in depression [5].

There are several benefits to a combined theoretical and data-driven approach. Data-driven models are dependent on the quality of the input data, i.e. clean signals with little noise, and how sensitive they are to relevant biological, cognitive, and environmental processes that shape behaviour. Theory-driven models provide a specific mechanistic description for the input data, based on prior research and expertise. Parameters of a theory-driven model can therefore provide a framework that reduces the dimensionality, i.e. the total number of variables, of the input data, leading to more stable and accurate output from data-driven machine learning techniques.

Challenges facing this approach

Despite the aforementioned strengths, there are several challenges facing this approach. As time progresses it may be difficult to integrate knowledge across studies without a community-agreed standardisation. For example, methods such as factor analysis (used to derive transdiagnostic symptom dimensions), are dependent on the self-report questionnaires used [3]. Furthermore, computational methods also require significant expertise, and non-experts may find the tools difficult to navigate.

Using transdiagnostic measurements in clinical trials focusing on multiple related diagnoses is an approach that may prove the utility of computational biomarkers. As treatments may differ between diagnoses, transdiagnostic trials could help clinicians to identify therapies which work across related diagnoses, and identify comorbid patients who could benefit from these therapies (Fig 1B). However, despite studies demonstrating that behaviours, on which diagnosis rests, can be characterised by computational parameters with specific neurobiological correlates [6, 7], patients with similar computational parameters may not all respond to a particular pharmaceutical agent. This group of individuals may also be heterogeneous and large experimental medicine studies will be required to establish the mapping between neurobiology, pharmacology and computation in order for the computational approach to realise its potential to improve person-specific treatments.

Implementing transdiagnostic tools in clinical practice may also prove challenging. However, these tools have the potential to complement current nosological approaches. Although symptom-based diagnoses overlap and may only provide a rough approximation of underlying pathology, symptoms are still the primary concern of the patient [2]. The complexity of individual psychiatric conditions will mean that clinical assessment will remain vital to providing holistic treatment and will also help to determine which particular investigative

computational batteries e.g. questionnaires, tasks, imaging modality and specifications, would be most appropriate, in much the same way that a physician would run particular diagnostic biomarkers when there is suspicion of cancer. Because transdiagnostic factors are based on symptoms, they would augment diagnoses due to their demonstrable links to underlying psychopathology. As transdiagnostic factors and biomarkers develop, become standardised, and their predictive value is realised, greater weight can be placed on these investigative methods as diagnostic tools in psychiatry.

Finally, to realise the clinical potential of these tools, it will be imperative to make them both accessible and easy to operate in a low-resource environment. Factor analysis of self-report questionnaires may be one way to face this last challenge. Factor scores are robust and reliable if a consistent set of questions are used. If they are shown to be robustly related to specific computational descriptions of behaviour in cognitive tasks, the use of factors derived from self-report questionnaires could act as a proxy for computational biomarkers (Fig. 1B). This would preclude the need for time-consuming cognitive tasks and expensive imaging, making precision psychiatry more accessible. This is a particularly important consideration given the increased risk of mental illness in low socio-economic groups with limited access to well-funded mental health centres.

In summary, transdiagnostic computational phenotyping of mental illness may bridge the gap between biology and clinical practice, providing investigative and predictive tools which could precipitate a shift away from relatively crude group-average explanations that have had limited clinical utility in alleviating mental health symptomatology at the level of the individual.

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Figure 1:

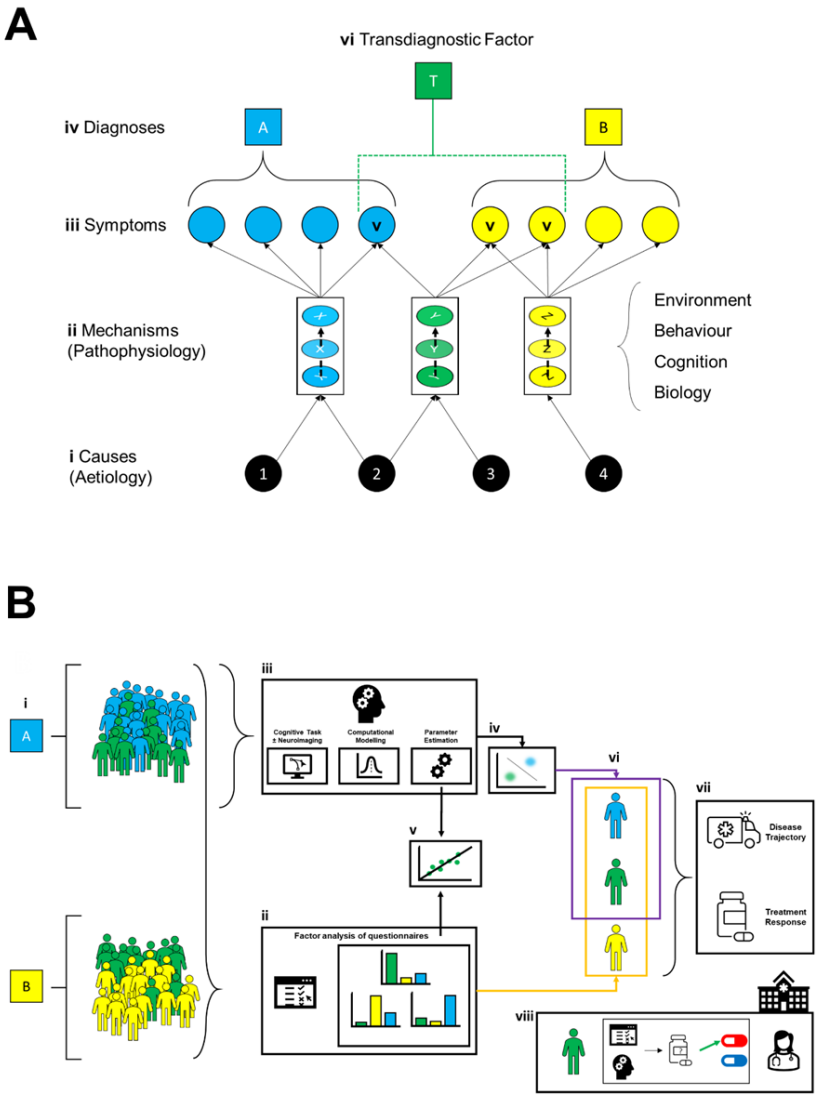

Figure 1.

(A) A simplified illustration of the different levels which may underlie psychiatric diagnoses. Traditional diagnoses are signified by blue and yellow branches, with a transdiagnostic approach denoted by the green branch. (i) Underlying causes (aetiology), e.g., genetic expression or predisposition, contribute to (ii) pathophysiological mechanisms, such as inflammation, aberrant synaptic gain, perceptual differences, or can result in psychopathology in response to the environment. For example, expression of a particular gene may lead to differences in synaptic gain, affecting output at the circuit level and ultimately influence behaviour through the interaction with other social and psychological factors (environment) to produce (iii) symptomatology. In this way, symptoms can be seen as the product of single or multiple maladaptive processes which can occur across multiple linked levels from biology through to the environment. Notably, many mechanisms can give rise to (v) the same symptom, as can the same mechanism give rise to many symptoms. Traditionally, symp-

toms are then categorized into a cluster known as a diagnosis (iv) , e.g., A or$B$. However, a transdiagnostic factor (vi) may recapitulate common symptoms or behaviours across disorders, and is related to a specific mechanism (*'Y'* denoted in green), which can be captured in cognitive experiments and/or neuroimaging (see Fig 1B).

(B) Computational psychiatry techniques implementing a precision approach. *(i)* Studies can include individuals with a particular diagnosis or across related diagnoses. (ii) Studies may involve factor analysis of questionnaires that aim to encompass a broad range of transdiagnostic behaviours and symptoms. (iii) The battery may involve cognitive tasks and neuroimaging from which model parameters can be estimated, thereby describing the behaviour/neurobiology of an individual. (iv) Model parameters can be used to find new subgroups within diagnoses, or alternatively (v) they can be correlated against factor scores to determine if the factors correspond to an underlying mechanism for disease across diagnoses. (vi) Using these data, subjects can be grouped either based on subgroups within existing disease categories (purple pathway) or across diagnoses (orange pathway). (vii) Longitudinal data from clinical trials can be used in conjunction with machine learning techniques to determine if the parameters produced from the models, or the factors associated with them, have any predictive value in terms of treatment response or prognosis. Looking for cognitive biomarkers within diagnoses is compatible with defining new dimensions of disease as cognitive biomarkers in existing disease subgroups may be common to multiple diagnoses associated with a single dimension. (viii) Data gained from these trials can be used in clinical practice in the form of a decision aid, where patients can undergo clinical investigation in the form of a questionnaire battery or undergo neuroimaging while performing a cognitive task to allow an informed decision about which treatment may be better. This may improve patient outcomes as it facilitates a quicker response to treatment and circumvents the trialling of multiple different therapies.

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Bibliography

1. Huys Q, Maia T, Frank M. Computational psychiatry as a bridge from neuroscience to clinical applications. Nature Neuroscience. 2016;19(3):404-413.

2. Stephan K, Binder E, Breakspear M, Dayan P, Johnstone E, Meyer-Lindenberg A et al. Charting the landscape of priority problems in psychiatry, part 2: pathogenesis and aetiology. The Lancet Psychiatry. 2016;3(1):84-90.

3. Gillan C, Seow T. Carving Out New Transdiagnostic Dimensions for Research in Mental Health. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging. 2020;5(10):932-934.

4. Chekroud A, Zotti R, Shehzad Z, Gueorguieva R, Johnson M, Trivedi M et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. The Lancet Psychiatry. 2016;3(3):243-250

5. Frässle S, Marquand A, Schmaal L, Dinga R, Veltman D, van der Wee N et al. Predicting individual clinical trajectories of depression with generative embedding. NeuroImage: Clinical. 2020;26:102213.[6]

6. Maia T, Conceição V. The Roles of Phasic and Tonic Dopamine in Tic Learning and Expression. Biological Psychiatry. 2017;82(6):401-412.

7. O'Callaghan C, Hezemans F, Ye R, Rua C, Jones P, Murley A et al. Locus coeruleus integrity and the effect of atomoxetine on response inhibition in Parkinson's disease. Brain. 2021;144(8):2513-2526.

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Supplementary Material

Table 1 Key of useful terms

Term	**Definition**
Precision Medicine	A form of practice using specific therapies which are selected ('personalised') for patients based on their individual characteristics or the characteristics of a group to which they belong [1]. The aim of personalised therapy is to maximise patient outcomes whilst reducing adverse effects. An example of precision psychiatry, could be taken as selecting a therapy for depression e.g. cognitive behavioural therapy (CBT) versus an SSRI based on the likelihood of success for that treatment given a patient's characteristics (clinical/biological).
Biomarkers	"A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes or responses to an exposure or intervention." [2] Biomarkers refer to substances which, found in the body, indicate information about a disease. These substances are usually component parts or bi-products of the disease process itself, and have traditionally been biological substrates such as proteins e.g. C-reactive protein as a biomarker for inflammation. As part of the disease process, biomarkers have value as they indicate the presence or prognosis of a disease. In psychiatric disease, due to a lack of traditional biomarkers for disease prognostication, increasing attention is being paid to computational parameters which can capture a behavioural process related to a particular disease.
Transdiagnostic Psychiatry	Transdiagnostic psychiatry aims to look across diagnoses to discover new dimensions of disease based on biological and behavioural mechanisms [3].
Nosology	Related to the classification of disease
Factor Analysis	Modelling observed variables as a weighted combination of a smaller number of latent variables (e.g. modelling scores from 9 questionnaires as 3 factors).
Reinforcement Learning	A framework for adaptive decision-making in the context of rewards and punishment.

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Computational Model	In neuroscientific terms, a computational model is a mathematical description that can be used to characterize complex cognitive processes, such as learning or decision-making. Parameters (see below) of the model can be estimated and quantify a specific part of the learning process, e.g. the weighting of new information compared to old. These parameters are estimated based on an individual's behavioural responses during a cognitive task.
Parameter	Models are composed of parameters which represent a specific part of the learning process. When models are fitted to data from a task, parameters can be used to describe an individual's performance mechanistically. For example, two people with major depressive disorder may show the same negative emotional bias on a cognitive task, but that behaviour may be caused by two different mechanisms - captured as differences in model parameters.
High dimensionality	Data can be described as highly dimensional when there are more measurable features or variables than there are independent samples. In these scenarios, machine learning algorithms perform poorly. Reducing the number of variables is important as it improves the performance of the algorithm. [4]
Machine Learning	Machine learning involves applying algorithms to data in order to make predictions or classifications based on input data, which either does (supervised) or does not (unsupervised) have known labels. A machine learning algorithm will produce an estimate about pattern or structure in the data.
Dynamic Causal Modelling (DCM)	A method commonly used for the quantification of effective connectivity (e.g. the influence that one neural population exerts over another), DCM allows comparison between models of interconnected networks of neuronal populations in order to explain data gained from dynamic imaging during cognitive tasks. [5]
Overfitting	Overfitting is a process in which models become extremely sensitive to noise when they are fitted to a training data set. The model inaccurately treats noise as signal of interest, so that it can better predict outcomes for the data that it is trained on. Highly dimensional data sets can lead to overfitting which in turn leads to poor predictions in new data (poor generalisability). [6]

Number needed to treat (NNT)	The number needed to treat describes the number of patients needing to receive a particular intervention so that one additional patient has a positive outcome. For example in a computational context, applying an algorithm which can predict remission in depression, the number needed to treat describes the number of patients the algorithm has to be applied to for the algorithm predict remission in an additional patient.

Supplementary Bibliography (Supporting References for the Glossary)

1. Manchia M, Pisanu C, Squassina A, Carpiniello B. Challenges and future prospects of precision medicine in psychiatry. Vol. 13, Pharmacogenomics and Personalized Medicine. Dove Medical Press Ltd; 2020. p. 127–40.
2. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet].
3. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: The seven pillars of RDoC. BMC Med. 2013 May 14;11(1).
4. Narisetty NN. Bayesian model selection for high-dimensional data. Handb Stat. 2019;
5. Brodersen KH, Deserno L, Schlagenhauf F, Lin Z, Penny WD, Buhmann JM, et al. Dissecting psychiatric spectrum disorders by generative embedding. NeuroImage Clin. 2014;4:98–111.
6. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. 2018; Available from: https://doi.org/10.1146/annurev-clinpsy-032816-
7. Paulus MP, Huys QJM, Maia T V. A Roadmap for the Development of Applied Computational Psychiatry. Vol. 1, Biological Psychiatry: Cognitive Neuroscience and Neuroimaging. Elsevier Inc; 2016. p. 386–92.
8. Wang XJ, Krystal JH. Computational psychiatry. Vol. 84, Neuron. Cell Press; 2014. p. 638–54.
9. Astle DE, Holmes J, Kievit R, Gathercole SE. Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. J Child Psychol Psychiatry. 2021 Jul 23;

Posted on 26 Oct 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166672691.16173006/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.