European Neuropsychopharmacology 88 (2024) 66–77

Contents lists available at ScienceDirect

European Neuropsychopharmacology

journal homepage: www.sciencedirect.com/journal/european-neuropsychopharmacology

The role of biomarkers in clinical development of drugs for neuropsychiatric disorders - A pragmatic guide

Daniel Umbricht [a,b,*], Martien J.H. Kas [c], Gerard R. Dawson [d]

[a] xperimed GmbH, Basel, Switzerland
[b] University of Zurich, Switzerland
[c] Groningen Institute for Evolutionary Life Sciences, University of Groningen, the Netherlands
[d] P1vital LTD, Oxfordshire, United Kingdom

A B S T R A C T

The failure rate of drugs being developed for neuropsychiatric indications remains high. Optimizing drug discovery and development requires not only a better neurobiological understanding of disease aetiology and development, but also the means by which we can measure relevant biological and clinical processes related to disease progression, drug target engagement, and sensitivity to treatment. Here we address the role and key considerations for the selection of biomarkers in clinical drug development for neuropsychiatric disorders. We do not provide an exhaustive list of biomarkers; rather we lay out a pragmatic, well-defined biomarker selection strategy that addresses the main goals for each of the phases in the drug development cycle. We discuss the key questions and issues that concern biomarker selection and implementation in each phase of development. For the better development of biomarkers, we emphasize the need to focus on discrete biological dysfunction and/or symptom domains rather than diagnoses. We also advocate the use of biomarker-based patient stratification in phase 2 and 3 to increase sensitivity and power and reduce costs. Our aim is to enhance precision and chances of success for these complex and heterogeneous brain disorders with a high unmet medical need.

1. Introduction

Drug development for neuropsychiatric indications remains a challenge. This is not for a lack of effort. Despite decades of investment into the translation of research findings to new treatments, the benefits for patients remain disappointingly low (Pangalos et al., 2007; Wong et al., 2019). This is not unique to neuropsychiatric drug development. Despite a prevailing view that drug development in neuropsychiatric disorders is particularly prone to failure, attrition rates in disease areas such as oncology are comparably low (Dowden and Munro, 2019; Pangalos et al., 2007; Wong et al., 2019). This attests to the fact that the pathobiology and pathophysiology of most human disorders are complex and our understanding of them is limited. However, the need for better and more effective treatments has never been greater. In recent years mental health illness has become the biggest cause of disability in the world (World Health Organization, 2017).

Specifically, the attrition in drug discovery and development of new treatments for neuropsychiatric disorders is due to many factors that include, but are not limited to: 1) the biological and physiological complexity of brain disorders; 2) the heterogeneity underlying clinical diagnostic entities; 3) the lack of biomarkers for diagnoses, patient stratification, response prediction and surrogate endpoints; 4) suboptimal endpoints that rely on subjective assessments by clinicians of the patients' symptoms and complaints; 5) a high placebo response in some indications; 6) a reluctance of the pharma industry to embrace truly innovative thinking and approaches in study design and conceptualizations of mental disorders. Adding to the challenge is the fact that the cost of neuropsychiatric clinical trials in comparison with other disease areas is relatively higher due to several factors including: (i) identifying the appropriate dose for phase 2 clinical trials; (ii) identifying the phenotype (stratification) for the inclusion of patients in these trials; (iii) reliable measures of drug efficacy and (iv) the high failure rate in phase 3 clinical trials (Gribkoff and Kaczmarek, 2017).

Clearly, biomarkers that help to improve the success rate from preclinical promise to efficacious treatments are urgently needed. However, the discussion on the development of biomarkers often remains focused on diagnostic criteria and existing paths to regulatory approval. This approach conjures up potentially unrealistic goals of precision psychiatry that are gleaned from oncology where biomarker-based diagnostic and therapeutic approaches have been very successful (Sarhadi and Armengol, 2022). However, the pursuit of diagnostic biomarkers for neuropsychiatric disorders, as currently defined by Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, (DSM-5), is in most cases a futile exercise as these disorders represent multifactorial,

* Corresponding author.
E-mail address: daniel.umbricht@xperimed.com (D. Umbricht).

https://doi.org/10.1016/j.euroneuro.2024.08.511
Received 24 April 2024; Received in revised form 12 August 2024; Accepted 14 August 2024
Available online 4 September 2024

multidimensional and heterogeneous disturbances of complex systems that any single unidimensional biomarker by definition cannot capture. Two recent studies provide strong support for this view. In the first (Winter et al., 2024) the authors applied machine learning to multivariate imaging data to develop a diagnostic biomarker for major depressive disorder (MDD). The diagnostic accuracy of the 4 million (!) models developed ranged from 48 to 61 %, essentially hovering around chance. The most plausible reason for this failure is the heterogeneity of the abnormalities in the biology, pharmacology and physiology underlying MDD rendering any attempt to find a unique imaging 'signature' futile. In contrast, Tozzi et al. (Tozzi et al., 2024) investigating the disease heterogeneity of patients with MDD and anxiety found six biotypes characterized by distinct profiles of connectivity both in resting state and task-activated functional magnetic resonance imaging (fMRI). These two studies strongly support the concept of heterogeneity regarding pathophysiology and plausibly pharmacology. We are not saying that biomarkers supporting a specific diagnosis, even if they are not pathognomonic or specific, cannot play a useful role in drug development, particularly in phases 2 and 3 (see below). However, for drug development we argue a key focus should be to develop and implement biomarkers that capture a specific pathophysiology and/or pharmacology that is critically operative in all or a subset of patients, that can be targeted pharmacologically and/or drives a distinct symptomatology or deficit.

Recent reviews of biomarkers in neuropsychiatric disorders (Abi-Dargham et al., 2023; Cortese et al., 2023; Garcia-Gutierrez et al., 2020) provide comprehensive overviews of biomarkers for neuropsychiatric indications; however, they do not focus specifically on their use and implementation in drug development. Therefore, in this article we lay out an overall pragmatic framework for the use of biomarkers in the development of new medical entities (NMEs) for neuropsychiatric indications. It is guided by a realistic appraisal of our current understanding of neuropsychiatric disorders, but also of the operational and methodological limits and challenges of implementing biomarkers in industry sponsored clinical trials. It is critical to note that the specific questions, and the needs that putative biomarkers have to address vary considerably by clinical phase and indication. Hence, the characteristics of biomarkers used in different clinical phases differ accordingly. Rather than attempting to identify specific biomarkers for symptoms or patient stratification and present an exhaustive list (see publications cited above), we instead consider which type of biomarker is likely to be the most appropriate for the specific needs and questions to be answered according to the development phase and indication and provide selected examples. This paper is focused more on the phase 1 and 2 stages of drug development as biomarkers play, and will continue to play, a bigger role in these phases than in phase 3, at least for the foreseeable future.

We assume that the increased availability of genetic and other molecular information of well characterized subjects, the developments in the ability to model the genetic architecture as well as the use of machine learning (ML) and artificial intelligence (AI) will be critical for the discovery, validation and deployment of useful biomarkers. While a comprehensive discussion of these rather technical aspects of biomarker development is beyond the scope of this paper, we will restrict our discussion of such outlooks to some critical remarks in the section 'Limitations, future directions and developments' and highlight selected biomarkers that exemplify approaches that we suggest are worth pursuing.

Although the term biomarker is often used in the context of neuropsychiatric drug development, most drug development programs conducted for indications such as MDD and schizophrenia have not used biomarkers beyond those measuring target engagement. We hope this review will help change this narrow approach, increase the attempts to develop and use biomarkers in drug development and subsequently increase the success rate of compounds for treatments aimed at neuropsychiatric disorders.

2. Definition and key questions

2.1. Definition of biomarker

Currently, the term biomarker is used widely, but with differing meanings. The need for a clear definition is important. This is particularly critical when seeking regulatory approval for the use of a biomarker for patient stratification or drug labelling. Thus a biomarker, as defined by the U.S. Food and Drug Administration (FDA)-National Institute of Health (NIH) Biomarker Working Group, is "a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or biological responses to an exposure or interventions" (FDA, 2016). Note that the FDA definition outlines seven biomarker categories including diagnostic, monitoring, predictive, prognostic, pharmacodynamic/response, safety, and susceptibility/risk; biomarkers discussed in this article fall into the categories: diagnostic, predictive and pharmacodynamic/response, but we do not specifically use these distinctions here.

However, many informative biomarkers being used, or recommended for use, in drug development may not need to meet this strict definition. The definition above suggests that a treatment induced change in biomarker reflects a change in the underlying biological system that results, or will result in, improve symptoms or outcomes for patients. From a pragmatic drug development,perspective the definition above may be too narrow to include all informative biomarkers. For example, citalopram reduces the accuracy with which healthy volunteers can identify negative emotional facial expressions reflecting a change in cognitive processing that, in patients, can result in an improvement of mood. Thus drug-induced changes in the biomarker 'accuracy for facial emotions' does not reflect a change in the symptom 'mood', but a change in cognitive processing that may result in a change in mood (Dawson et al., 2021). On the other hand, the definition above could also be considered as unspecific as it could be interpreted as including classical rating scales as they can indicate biological responses to an intervention.

Consequently, for this paper we would like to propose the pragmatic definition that a biomarker is any quantitative data indexing pharmacological, biological, physiological and behavioural processes/functions in the central nervous system (CNS) and their characteristics that cannot be clinically observed in signs and symptoms, examined directly or reported by clinicians or patients. For example, vital signs are not biomarkers whereas heart rate variability is a biomarker. Behavioural assays that capture reward functioning are biomarkers, whereas a rating on an anhedonia scale is not. Some assessments may fall into a grey zone of this definition. According to our definition classical pen and paper cognitive tests are not biomarkers for a state of the brain, while computerized assessments that capture reaction times are. In practice the exact boundary may be less relevant.

2.2. Key considerations and questions

Ideally, the outline of a biomarker plan is part of the initial drug target selection process. Even the most exciting target for which no existing or potential biomarkers can be identified should have lower priority as molecular target than those for which biomarkers exist or can be developed (note that this may apply more to large pharma where a number of molecular targets are usually considered, whereas a small biotech working on one or two molecules may not be in a position to choose). A clear biomarker strategy will inform and hopefully improve decision making as development progresses. The obvious benefit is the early detection of signals of efficacy, or the lack thereof, which will increase the speed of development or terminate ineffective treatments early in the process. As described above and elaborated on later, the questions which biomarkers to consider and implement (e.g. target engagement, efficacy, safety) vary according to the phase of development. Hence the biomarkers that can address the different questions

raised at each phase of drug development may differ or need to be modified to be informative. Thus, when it comes to the selection, development, and deployment of biomarkers it is critical to ask those questions that are appropriate for the development stage of the drug, and then design and deploy the biomarkers accordingly. Table 1 suggests which type of biomarker may be useful at each stage of the drug development process.

Consequently, the first step in the formulation of a biomarker strategy concerns the review and selection of biomarkers. Table 2 and Fig. 1 outline key aspects and critical questions that need to be evaluated when selecting biomarkers for different phases of development. A critical one for the scientific aspects of a biomarker is the evidence supporting its use. The weight each aspect is given varies by phase of drug development. Operational challenges of biomarker implementation, scalability and burden to patient are much more critical in phase 3 than in phase 1, whereas the translational aspect and/or pharmacological 'anchoring' of a biomarker are most critical in phase 1. While not exhaustive Table 2 also illustrates the potential for biomarkers to address specific aspects at different stages of the drug development process.

When considering the inclusion of a biomarker or biomarkers additional important related questions include:

(i) What question, or need, does the biomarker have to address?
(ii) What role does it play in the overall evaluation and decision-making process on whether to progress or terminate a molecule?
(iii) If a biomarker is used as pharmacodynamic readout what consequences do negative or positive biomarker results have?
(iv) Is the biomarker merely meant to support clinical findings in a small study where the clinical effects are small and the signal is weak?
(v) Is a Go/NoGo decision to the next phase dependent on the biomarker results?
(vi) If the results are non-consequential should the biomarker really be included in the study?

The clearer the answers to these questions, the better the characteristics and role of a biomarker can be defined. For instance, if a Go/NoGo decision depends on the biomarker then validity and reliability become tantamount, whereas when it is only meant to support clinical effects, one may be more liberal regarding these characteristics of the biomarker.

Table 1
Overview of which type of biomarker may be useful at each stage of the drug development process.

	Nature of Biomarker	Purpose of Biomarker
Phase 0	Translational biomarker	Characterize target engagement Demonstrate biologically relevant effects Demonstrate 'clinically' relevant effects
Phase 1a	Translational biomarker	Demonstrate target engagement Demonstrate biologically relevant effects
Phase 1b	Translational biomarker Efficacy biomarker	Demonstrate target engagement Demonstrate biologically relevant effects Demonstrate clinically relevant effects
Phase 2	Patient selection/ stratification biomarker Efficacy biomarker	Support diagnosis Identify patients with highest probability of response Support demonstration of clinical effects
Phase 3	Patient selection biomarker Efficacy biomarker Response predictor biomarker Relapse predictor/tracker biomarker	Support diagnosis Identify patients with highest probability of response Capture course/remission/relapse of disorder

Table 2
Key biomarker aspects and critical questions that need to be evaluated when selecting biomarkers for different phases of drug development.

	Biomarker Aspect	Relevant questions in the selection process of a specific biomarker
Translational	Translational	Is there actual data supporting the use of the biomarker in animals and humans? Are the key parameters equivalent in animals and humans? How good is the evidence?
Biology/ Physiology	Biological/ physiological anchoring	How well are biological/ physiological mechanisms underlying the biomarker understood? How good is the evidence?
Pharmacology	Pharmacological anchoring	How well is the pharmacological mechanism that modulates the biomarker understood? How good is the evidence?
	Pharmacological challenge	Are there pharmacological studies (challenge/treatment) supporting pharmacological anchoring?
Circuit engagement	Circuit engagement	Can the biomarker be linked to a specific circuit? Are findings in healthy controls and patients with a specific disorder equivalent?
Association with Phenotype	Association with mental/cognitive functions	Is the biomarker associated with and/or tracking specific functions and/or symptoms? How good is evidence?
	Association with specific disorders	Is the biomarker (its abnormalities) associated with specific diagnoses? How good is the evidence?
Validity/Test-Retest reliability	Metric qualities	How well has the biomarker been validated? What is its test-retest reliability?
Operations	Operational Ease	Operationally how easily can the biomarker be implemented? How difficult is the programming/ setup/administration of the paradigm required to support the implementation of the biomarker?
Paradigm/ Analytical Challenges	Paradigm challenges	How critical and/or challenging is the optimal paradigm set-up?
	Analytical Challenges	How challenging is the data management and analysis of the biomarker data?
Burden to Subjects	Burden to subjects	How burdensome to subjects is to perform the biomarker task? How large is the friction factor?

In neuropsychiatric drug development key biomarker modalities include neuroimaging, electroencephalogram (EEG) based assessments, quantitative behavioural assays (including some specific cognitive tasks) and passive physiological and behavioural monitoring. Table 3 provides a high-level qualification of different methods in these various modalities. A more granular description and characterization goes beyond the scope of this paper; rather Table 3 is meant to provide a high-level grid that highlights the aspects in which specific biomarker modalities and methods perform well and those in which they perform poorly. For instance, neuroimaging is a modality that cannot be easily and cost-effectively scaled up and implemented in large phase 2b/3 clinical trials (except for certain forms of imaging such as amyloid positron emission tomography (PET) which might be crucial for diagnosis and treatment selection in Alzheimer's Disease (AD) trials which are not considered here). Whereas neuroimaging can provide critical information on the functioning of brain circuits and their engagement in smaller proof of concept trials. On the other hand, resting state rsEEG can be easily scaled up to larger phase 2, and potentially phase 3 trials but provides little information on specific brain circuits and their engagement.

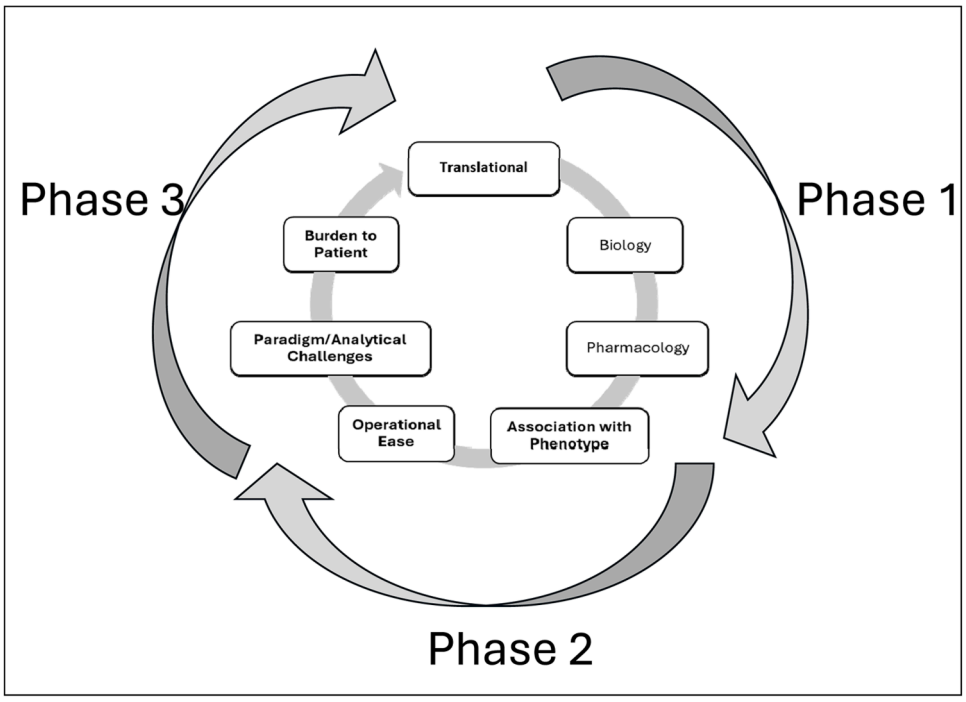

Fig. 1. Key Aspects of biomarkers and phases in which they are particularly important.

Table 3
A high-level qualification of different methods in different biomarker modalities.
Scoring: 0= Performs worst in this category; 1= performs moderately well in this category; 2= performs well in this category. Please note that for Paradigm/Analytical Challenges 0 is equal to high challenges while 2 indicates low challenges, for Burden to Subject and Operations 0 is equal to high subject burden and low ease/scalability, while 2 indicates low burden to subject and high scalability.

		Biomarker modality/methodology								
Biomarker Aspect		Imaging (rsMRI)	Imaging (task activated MRI)	Imaging (PET)	Sleep (PSG)	rsEEG	ERP (passive paradigms; N1; MMN)	ERP (active paradigms; P3)	Behavioural assays	Wearables/ Smart-phone based BM
Translational	**Translational**	**0**	**0**	**2**	**2**	**2**	**2**	**1**	**1**	**0**
Biology/ Physiology	**Biological/ Physiological anchoring**	**1**	**1**	**2**	**2**	**2**	**2**	**1**	**1**	**0**
Pharmacology	**Pharmacological anchoring**	**1**	**1**	**2**	**1**	**2**	**2**	**1**	**1**	**0**
	Pharmacological challenge	**1**	**1**	**2**	**1**	**2**	**2**	**1**	**1**	**0**
Circuit engagement		**2**	**2**	**0**	**1**	**0**	**0**	**1**	**1**	**0**
Association with Phenotype	**Association with mental/cognitive functions**	**2**	**2**	**0**	**1**	**1**	**1**	**1**	**2**	**2**
	Association with specific disorders	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**2**	**2**
Paradigm/ Analytical Challenges	**Paradigm challenges**	**2**	**0**	**2**	**2**	**2**	**1**	**1**	**2**	**2**
	Analytical Challenges	**1**	**0**	**2**	**2**	**2**	**1**	**1**	**2**	**2**
Burden to Subjects	**Burden to Subjects**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**2**	**2**
Operations	**Operational Ease (scalability)**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**2**	**2**
Reliability/ Validity		**0**	**1**	**2**	**2**	**2**	**2**	**1**	**2**	**2**

3. Biomarkers for each phase of development

3.1. Preclinical and phase 0

As mentioned above formulation of a biomarker plan, or at least a discussion of potential biomarkers should take place at the time of target selection. Ideally, a biomarker plan should be in place when a compound enters good laboratory practices (GLP) toxicology. Such a plan will inform the kind of preclinical studies should be conducted during preclinical stages and phase 0 (during GLP toxicology studies) to characterize compounds regarding translational biomarkers and address the question if clinical studies should be implemented to develop

biomarkers in humans (see below). Many animal behavioural assays may have construct validity; that is they measure a state that is relevant to the effect induced by the drug (good examples are models of anxiety that reliably predict outcomes in human trials when applied properly (Gurrell et al., 2023; McKernan et al., 2000). However, most are not translational in the sense that the same readout can be obtained in human subjects. For instance, a forced swim test may provide evidence of potential antidepressant effects of a drug, but there is no human equivalent of this test. Translational biomarkers, however, should be more or less homologous in key aspects both in animals and humans. They include but are not restricted to, rsEEG, some obligatory exogenous auditory event-related potentials such as P1, N1 and P2, auditory steady state response, reward learning behavioural assays, addiction assays, physiological responses and sleep parameters.

In instances of novel biomarkers, specific studies in healthy volunteers or patients may need to be conducted while the compound completes GLP toxicology assessments to characterize critical aspects of the biomarker such as test-retest reliability, validity etc. and address operational and implementation challenges. At the same time the modulation of the selected translational biomarker by the compound in development should be characterized in preclinical studies in rodents and, where appropriate, in non-human primates. Results of such studies then define a target biomarker profile that the compound should induce in humans as proof of functional target engagement and/or relevant biological effects (for instance on sleep biomarkers). This is particularly relevant for compounds for a new molecular target where knowledge of its physiological and behavioural effects may be scant or absent and for which no PET tracer exists. The simplest case is probably rsEEG which can be obtained easily in animals and humans. Indeed, for some targets EEG based biomarkers exist that are relatively specific, validated and supported by animal and/or human data. They can help evaluate novel compounds that modulate these targets (see Table 4).

Many translational biomarkers are not necessarily related to the target disorder but rather index basic physiological and/or behavioural functions that have been preserved through evolution. Research Domain Criteria (RDOC) (Insel et al., 2010) provide a framework for the use and development of translational biomarkers related to a target disorder. However, the utility of this approach is limited because many aspects of neuropsychiatric disorders cannot be modelled or observed in animals (e.g. delusions, suicidality) thus emphasizing the need for their validation in human trials.

In summary, the goals of a biomarker strategy in preclinical phases should be to implement the biomarker(s) in phase 1a, 1b PoM and PoC studies and a clear description of the expected effects of the compound on the biomarker parameters. This obviously requires a close collaboration between preclinical and clinical teams.

Table 4
Biomarkers that have been pharmacologically validated and likely to be informative in clinical studies for specific targets.

Target	Biomarker	Examples of Pharmacological MoA/Agent
NMDA receptor	• Mismatch Negativity • Gamma Frequency band in rsEEG • Auditory Steady State Response (40 Hz)	• GlyT1 inhibitors • DAAO inhibitors • NMDA PAMs/ NAMs • NMDA Antagonists (e.g. Ketamine)
GABA-A α1–5 subtype containing receptor Pharmacological approaches enhancing/ modulating PV interneuron functioning	• Beta Frequency Band in resting state EEG • All 40 Hz Paradigms (rsEEG gamma band, AASR, early auditory gamma response, early visual gamma response)	• GABAA α1–5 PAMs/NAMs • NMDA modulators • GABA-A & GABA-B modulators • KV3 modulators • Mitochondrial targets (oxidative chain) • Perineuronal net biology

3.2. Phase 1a

Target engagement and translational and biomarkers

The main goal of phase 1a trials is to characterize the safety and tolerability profile of a compound, determine the maximally tolerated dose, describe the pharmacokinetics of the compound and if possible, provide proof of target engagement. The need to demonstrate target engagement cannot be stressed enough as a lack of clear target engagement may account for up to 30 % of all drug failures (Cook et al., 2014).

The goals outlined above should guide the selection of biomarkers for phase 1 single and multiple ascending dose (SAD, MAD) studies. For example, if a PET ligand is available for the target in question, a dedicated PET study implemented after the completion of the SAD study is a must. However, for many targets a PET ligand is not available. In these cases, rsEEG may be a good option if animal studies have shown clear effects on EEG signals that can be used to determine functional target engagement in humans. In some cases, event-related potentials may also be useful. For instance, for compounds affecting N-methyl-D-aspartate (NMDA) receptor transmission specific event-related potentials (like mismatch negativity [MMN] and auditory steady state response (ASSR); see also Table 4) may be considered. These translational approaches underline the importance of characterizing EEG effects in animal models. Similarly, if a compound has shown demonstrable effects on sleep parameters preclinically a polysomnographic investigation may be appropriate (Note: novel methods and technologies are in development that promise the reliable acquisition of detailed sleep data at home). Other methods may include fMRI including blood oxygenation level dependent (BOLD) or arterial spin labelling (ASL) measures where the distribution of the target is well known and hence a plausible hypothesis of the likely drug effects can be formulated and tested. Biomarkers such as EEG may not only demonstrate target engagement, but also provide initial data on the drug exposure-effect relationship – critical information in the planning of phase 2 trials.

It is critical to know if a specific biomarker can provide useful information in healthy volunteers. If a biomarker measures the improvement of a function, it may not do so in healthy volunteers if the specific function is already operating at an optimal level. For instance, MMN is considered a good biomarker of NMDA receptor transmission where function is impaired. However, it would not be a good biomarker for a drug that improves NMDA receptor transmission in a study with healthy volunteers as it is unlikely to be enhanced in these subjects. The same applies to drugs that are targeting cognitive deficits. Most cognitive functions and hence biomarkers tracking them are not known to improve with pharmacological interventions in healthy volunteers. In contrast emotion processing and reward functions can be modulated in both directions in healthy volunteers (Cools et al., 2009; Harmer et al., 2003; Pizzagalli et al., 2020). On the other hand, if the biomarker is tracking an impairment, such as eye tracking as a measure of sedation, it works well in healthy volunteers. Consequently, if a biomarker is implemented that should track improvement of a function without clear evidence that it does so in healthy volunteers it should obviously be regarded a highly explorative and not have any bearing on decision making processes if no effects are observed in these subjects.

An option to be considered is to test the effects of a novel compound on pharmacologically or behaviourally (e.g. sleep deprivation, pharmacological challenge studies) induced cognitive and/or behavioural deficits/effects. However, the construct and predictive validity of such models regarding the pharmacological treatment of the targeted disorder is questionable as they have been developed to pinpoint potential pharmacological abnormalities, but not their correction. For instance, a negative result could be due to a mismatch between the disruptive effect of the challenge and the degree of corrective effects of the drug that

could not be differentiated from the lack of a therapeutic effect, that is the effect of the challenge could overwhelm any corrective effect.

Accurate and reliable pharmacodynamic (PD) readouts are often limited by the small sample sizes that are usually employed in phase 1a studies. To some extent modelling approaches that incorporate all cohorts can help mitigate this issue. Another possibility is to pivot a safety study to a study focused on PD readouts by expanding a sample in which an initial analysis has shown a drug effect on the biomarker of interest. This approach can increase the power of the statistical analyses. Nonetheless, the wish to implement PD biomarkers in SAD/MAD studies is often driven by the hope that they detect early signs of biological effects that support potential clinical efficacy, while the limitations of sample size and restrictions by the nature of the study (safety/tolerability) are neglected. The discussion on whether specific biomarkers should be implemented in the SAD/MAD studies should always include the option of conducting a dedicated phase 1b study that can be powered appropriately and for which the optimal participants for a given question can be included. While this may initially add time to the initial development program, the decision informed by the biomarker results will rest on much firmer grounds and reduce the overall development time. In addition, the inclusion and exclusion criteria and the study participants (healthy volunteers, patients) can be tailored specifically to the question addressed. In safety studies this cannot be achieved.

3.3. Phase 1b

Translational and efficacy biomarkers

In phase 1a determining target engagement is critical, whereas in phase 1b the goal is to acquire initial evidence of biologically and/or clinically relevant drug-induced effects. For most new compounds a proof of mechanism (PoM) study should normally be conducted in patients or in a relevant related phenotype that has the characteristics (e.g. mild symptoms) of the clinical population (see for instance (Perini et al., 2023). The selection of biomarkers should be guided by the pharmacological target of the compound as well as the symptomatology and/or deficit to be ameliorated. The conceptualization of the disorder/-symptomatology in question regarding circuitry, biology and neuropharmacology is thus highly critical. If the targeted symptomatology fits into a neurocircuitry based framework, then an imaging biomarker may yield the most useful information. Examples include striatal activation during a reward task or amygdala activation during an emotion processing task. Ideally, an imaging approach is combined with a behavioural assay that measures the same construct. In the case of reward functioning, effort based, probabilistic and reinforcement leaning tasks may be included. In case of compounds that affect the excitation/inhibition balance such as those targeting α-amino-3--hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), NDMA and Gamma-aminobutyric acid (GABA) receptor transmission specific EEG based biomarkers such as N1, MMN, ASSR may be useful (see Table 4).

It has to be emphasized that for phase 1b studies there is currently no biomarker available for any neuropsychiatric indication that provides a clear answer to the question '*Will the compound be effective in large phase 2 and 3 trials?*' nor to the question of how much a positive result will increase the probability of success in later trials and eventually lead to the compound becoming a licensed drug. Data to answer such questions are not and may never be available as it would require a populated confusion matrix (true positive, false positive, true negative and false negative cases). Rather, the desired effect should be considered as minimum proof that the compound exerts observable effects that have been associated with the symptom domain targeted. In other words, while a positive effect can be expected to reduce the risk in the next trial, the absence of such an effect makes it highly unlikely to see a clinically relevant effect in future studies. This is critical for Go/NoGo decisions. For instance, if a drug targeting anhedonia has no appreciable effect in imaging or behavioural reward function tasks, it can plausibly be concluded that the likelihood of an effect on negative symptoms in schizophrenia, or anhedonia in MDD patients is low. In other words, PoM studies are meant to identify drug failures early on ("kill early with confidence" or aptly termed "Fast fail" by NIMH (Krystal et al., 2018)), and reduce the risk for the next trial. Thus, even the most well validated biomarker will enhance the chance of, but never provide, proof of clinical efficacy. Such proof needs to be established in phase 2 and 3 trials.

For a biomarker to serve as basis for a Go/NoGo decision there must be good scientific evidence both from preclinical models and clinical studies that the biomarker captures a mechanism that is an essential or a prominent factor in driving the reductions or modulation of symptoms to be targeted. In most cases the requirement for preclinical evidence restricts this approach to those domains of symptoms and deficits that can be investigated in translational approaches in animals. The development and refinement of the RDOC framework has captured these domains at a basic level (Insel et al., 2010). However, it needs to be accepted that critical domains of psychopathology and functional deficits that are uniquely human cannot be reliably reproduced in animals and hence translational biomarkers do not exist.

Another seemingly trivial aspect is instead a critical one: Test-retest reliability and the reliable change index (Guhn et al., 2014), that is the change which exceeds the expected normal variance in repeated assessments. There may be very intriguing findings in patients and novel methods published by the leading experts in the field that could be used as biomarkers. However, in many cases the test-retest reliability and reproducibility are unknown. Understandably, the seemingly mundane, but important task of characterizing these metrics is not among the top goal of academia. If there is a lack of test-retest data then it is important to conduct such studies from the beginning or at the latest, during the conduct of phase 1a studies.

Recent examples of a PoM study as outlined above include a study of a phosphodiesterase 10 (PDE10) inhibitor that was being developed to treat negative symptoms in schizophrenia (Umbricht et al., 2021). It failed to demonstrate an effect on imaging and behavioural biomarkers of reward functioning in patients with schizophrenia and negative symptoms, consequently its development was abandoned. Another company tested a compound with the same mechanism of action as monotherapy in a larger phase 2 trial in patients with negative symptoms (Meyer-Lindenberg et al., 2022). The study was terminated after an interim analysis indicated a lack of an effect thus confirming the results of the smaller PoM study. Another example is a study with the kappa receptor antagonist, aticaprant, in development for anhedonia in depression. It had demonstrated positive effects in a PoM study using imaging and behavioural biomarkers of reward functioning that were very similar to those used in the PDE10 trial mentioned above (Krystal et al., 2020). A subsequent phase 2 trial also showed positive clinical effects. The compound is now in phase 3. There are also examples where the positive results of PoM study were not replicated. In a recent phase 2 trial an AMPA positive allosteric modulator (PAM) failed to improve cognitive deficits in schizophrenia (https://clinicaltrials.gov/study/NCT03745820?tab=results; https://www.biogentrialtransparency.com/content/dam/global-development/general/biogen-trial-link/educational/en-us/pdf/lls/english-master/NCT03745820-LLS.pdf). Previously it had reversed the effects of a ketamine challenge study in healthy volunteers (Ranganathan et al., 2017) and enhanced measures of cognition in a small study in patients with schizophrenia (Evans et al., 2016).

3.4. Phase 2

Biomarkers supporting patient selection, stratification and efficacy

The goal of a phase 2 trial is to demonstrate clinical efficacy in patients and determine the range of active doses. At this stage, biomarkers should ideally support the correct diagnosis of patients, the identification of patients with highest probability of response (addressing the known heterogeneity of patients), provide information for a hypothesis driven stratification and support demonstration of clinical effects.

Currently, patient selection in almost all phase 2 trials in the neuropsychiatric indications is based solely on diagnostic criteria and levels of the targeted symptomatology. We are not aware of any study that has used biomarkers that have been associated with the specific disorder and/or one of its dysfunctions or symptoms to support or corroborate the diagnosis. As mentioned above diagnostic biomarkers for currently used DSM-5 diagnoses are not and may never be available, but the use of biomarkers reliably identifying specific symptoms should be possible. Biomarkers exist that could be used to either support or render the diagnosis less likely. An example may be measures of anhedonia or the processing of emotion perceived in depression. They could be used to exclude patients who report symptoms but may not have the disorder. Although this approach may exclude patients who do have the disorder, it is balanced by excluding subjects who should not be enrolled thus reducing heterogeneity in the clinical trial population.

Even if such biomarkers are not used to corroborate a diagnosis, they could help to address the heterogeneity of the disorder by using a more biologically based characterization and stratification. Currently stratification factors used include demographics (age, sex) and/or illness characteristics (duration of illness, number of previous episodes etc.) but rarely biomarkers which might plausibly be associated with a different underlying biology and/or neurochemistry. To some extent this can be explained by the fact that, as yet, no definite biomarkers for drug-response have been identified for these disorders. This is likely the case because such universal biomarkers may not exist and may therefore never be discovered. It is more likely that for each mechanism of action (MoA) different biomarkers or a different biomarker profile may predict response. This of course results in a 'Catch 22′ situation for compounds with novel MoAs for which biomarkers of response are by definition not available. In such a case a careful characterization of patients with several biomarkers may be useful as they could reveal biomarker patterns associated with response that could be tested in later studies.

However, if such biomarkers are deemed essential then the only way to characterize them is to conduct relatively large studies in patients in which they are identified and validated before they are implemented in phase 2 trials. To our knowledge only one company (Altoneuroscience) is currently pursuing this approach. It conducted open label studies that identified and validated the biomarker profile associated with a favourable response to a compound labelled ALTO-300 (clinicaltrial.gov: NCT05118750, NCT05157945). The company has now started a phase 2 randomized controlled trial (RCT) with this compound using this biomarker profile for stratification with the hypothesis that biomarker positive patients will respond particularly strongly to the compound (clinicaltrial.gov: NCT05922878). This a nice example of a biomarker-based drug development program aimed at identifying patients who are responders to a particular drug thus increasing the sensitivity of the clinical trials and shortening the time to registration. However, this approach may not be considered a viable option for many companies given that it may add 1 to 2 years and $15 −20 m to phase 2 development costs.

An alternative approach could consist of collecting biomarker data in phase 2 trials with the hope of identifying predictive biomarkers before phase 3 trials are initiated. A related approach sitting between the two outlined above, is much less expensive than conducting large studies to develop a biomarker and more easily implemented. It consists of stratifying patients by specific biomarkers that can plausibly be associated with a biology and/or neuropharmacology potentially critical for a response to the compound in development. The underlying assumption is that patients who differ on such a biomarker also exhibit differences in critical pathophysiology and therefore will have a different response to the drug. As an example, if a compound targets the NMDA receptor such biomarkers may include gamma oscillations in rsEEG, mismatch negativity and auditory steady state responses as they have been linked to NMDA receptor mediated neurotransmission (Balla et al., 2020; Gil--da-Costa et al., 2013; Herzog et al., 2023; Javitt et al., 1996; Leishman et al., 2015; Rivolta et al., 2015; Sanacora et al., 2014; Schuelert et al., 2018; Sivarao et al., 2016; Sugiyama et al., 2021; Yan et al., 2023). It is conceivable that differences in these biomarkers may be associated with changes in response to compounds modulating NMDA transmission such as an antagonist to treat MDD or compounds enhancing transmission to treat schizophrenia.

In addition, there are a host of behavioural biomarkers measuring emotion perception and processing, reward functioning and intrinsic motivation, reaction time fluctuations and others that can be readily implemented in phase 1b or phase 2a trials as potential stratification factors. For some, aspects of the neurobiological and neuropharmacological underpinnings are known. It must be conceded that it is not known if differences in such biomarkers relate to biological and/or pharmacological differences that are relevant to the response to the drug investigated. Nonetheless, such a hypothesis is plausible and more importantly, such biomarkers should they be informative, can be easily implemented at a relatively low cost and readily be scaled to larger phase 2 trials.

For instance, the Bipolar-Schizophrenia Network for Intermediate Phenotypes Intermediate Phenotypes (B-SNIP) consortium studies (Tamminga et al., 2021) described and replicated the existence of three distinct biotypes among patients with schizophrenia, schizoaffective and bipolar disorder that cut across diagnoses. These biotypes differ regarding the profile of cognitive deficits and EEG markers. In particular, one biotype has a *hypoactive* EEG profile, while another a *hyperactive* EEG profile (Clementz et al., 2016). It is plausible that these differences manifest differences in underlying pathophysiology and pathopharmacology. It is reminiscent of the observation that both over and underactivation of the D_1 receptor prefrontally can lead to deficits in prefrontal functions (Durstewitz et al., 1999; Floresco, 2013; Vijayraghavan et al., 2007). Accordingly, compounds targeting cognitive deficits or other symptoms in these disorders may have very different or even opposite effects in subgroups defined by EEG hypo- or hyperactivation. Indeed, initial observations suggested that clozapine is only active in patients with EEG hypoactivation (personal communication, C. Tamminga). A similar finding, albeit in a much smaller sample, was reported by Webb et al. (2016). They found that patients with a diagnosis of MDD and similar levels of symptoms clustered into three distinct groups that were characterized by different degrees of neuroticism, attentional deficits and abnormal reward functioning related to distinct brain activation patterns. Again, it is a plausible hypothesis that such different profiles are associated with different constellations in underlying biology and pharmacology driving the symptoms and hence may confer a varying response to a specific compound.

There is clear evidence in support of this proposition. In challenge studies with the NMDA antagonist ketamine and the 5-HT2A agonist psilocyin in healthy volunteers both compounds induced similar degrees of altered states of consciousness; however, the magnitude of MMN at baseline was associated with the degree of acute behavioural effects only during the ketamine, but not the psilocybin challenge. Furthermore, while both compounds led to almost identical patterns in working memory deficits, their effects on MMN indexing more fundamental information processing differed significantly suggesting that abnormalities leading to higher level dysfunction can be very distinct. (Schmidt et al., 2012; Umbricht et al., 2002, 2000, 2003). More importantly, analyses of data obtained in the Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care (EMBARC) project have demonstrated an association of specific EEG signatures (Wu et al., 2020), deficits in reward functioning (Giles et al., 2023) and fMRI signatures in response to processing of emotionally conflicting information (Fonzo et al., 2019) with differential responses to selective serotonin reuptake inhibitors (SSRI) in patients with MDD indicating the existence of important differences in local and distributed network functioning that affect drug response. Importantly, in all these analytical models the severity of depression was not a relevant factor. Similarly, a recent study in patients with MDD from International Study to Predict Optimized Treatment in Depression (iSPOT-D) (Williams et al., 2011)

identified what the authors refer to as a 'cognitive biotype' characterized by impairments in executive function and response inhibition domains of cognitive control and associated with worse outcome to treatment with selective SSRIs and NSRIs (Hack et al., 2023).

Potentially, biomarkers could also be used to disambiguate studies in which the results are unclear due to an unexpectedly high placebo response. As an example, numerous studies have shown that antidepressants enhance the perception of positive emotions while decreasing the perception of negative emotions. Interestingly, there is a lack of a placebo response in this task (Huneke et al., 2017). Given that the use of traditional rating scales is associated with an ever-growing placebo response, a measure that can, in the presence of a high placebo response, disambiguate the question: *did the drug really work?,* would be very valuable (Dawson et al., 2021; Huneke et al., 2017).

Normative values often do not exist for these potential stratification biomarkers, the distribution in the population to be recruited and hence threshold values to assign patients to one or the other stratum are not known. In such cases a dynamic stratification can be used that uses a running median. It will stratify the sample into two groups that will be statistically significantly different regarding the biomarker in question and allow the testing of the hypothesis that these differences are associated with different responsivity to drugs. Although such an approach can be readily implemented, to our knowledge no phase 2 studies involving neuropsychiatric indications have done so.

While genetic or other molecular information has not been helpful or used in disentangling the heterogeneity of patients in neuropsychiatric drug development, it is worth highlighting recently described findings that can serve as examples of the kind of biomarker that may be useful for drug development. One concerns the discovery of two molecules extracted from peripherally collected exosomes that are derived from cortical parvalbumin-positive interneurons (PVIN) (Khadimallah et al., 2022). The concentration of these molecules provides information about existing mitochondrial dysfunction in PVINs. A dysfunction in PVIN has, for a long time, been implicated in the pathophysiology of schizophrenia (Curley and Lewis, 2012; Gonzalez-Burgos and Lewis, 2012; Lewis et al., 2012). However, as for many other dysfunctions, it may only be relevant, and of functional consequence, in a subset of patients. These two novel biomarkers may allow the identification of such patients. Indeed, it was shown that biomarker positive subjects in early stages of schizophrenia differed significantly on clinical measures and EEG-based biomarkers of PVIN functioning, thus identifying marked differences in clinically relevant underlying pathophysiology (Khadimallah et al., 2022). While the relevance for the response to a drug is not known, it is highly plausible that differences in PVIN physiology could be decisive. Hence, using these biomarkers to stratify patients for a clinical study in this way is intuitively plausible. Similarly, a recent finding of an association of peripherally measured levels of metalloproteinase-9- a molecule involved in regulating perineuronal nets which are critically involved in modulation of PVIN functioning (Bosiacki et al., 2019; Carceller et al., 2020) - with hippocampal volume and negative symptoms (Seitz-Holland et al., 2022) may provide another stratification biomarker. It could potentially help differentiate patients with and without a distinct pathophysiology involving PVINs in brain regions critically involved in schizophrenia. While reliable commercial methods to measure these biomarkers are, to our knowledge not commercially available yet, they are provided as examples of how a focused use of molecular biomarkers can be linked to potentially relevant differences in pathophysiology. These could be highly useful for drug development, whereas we consider a shotgun approach that collects genetic and other molecular information for 'potential use' as not the best use of time and resources. In our view blood-based biomarkers only make sense if they can be tied to specific, disease or symptom relevant dysfunctions and pathology.

An important factor influencing the acceptance and use of biomarkers relates to the measurement of functional effects. While the primary goal of phase 2 trial is to demonstrate clinically meaningful effects on key symptoms, evaluation of functional effects is important when preparing for phase 3 studies where demonstration of such effects is usually required by health authorities for marketing authorization. Digital biomarkers obtained mostly by passively collected data from smart phones that characterize functional aspects of the patients' behaviour should be considered (Jongs et al., 2020). In most cases these biomarkers are still considered exploratory as no clear set of such biomarkers has been defined that convincingly assesses the real-world functioning of patients. However, phase 2 data may be very helpful for the final validation of digital biomarkers.

3.5. Phase 3

The main goal of phase 3 clinical trials is the proof of clinical effectiveness and acquisition of such data to obtain a marketing authorization from regulators. Biomarkers could play a similar role as they do in phase 2 trials, namely their use for stratification or patient selection and adjudication of diagnosis. They probably cannot be used to support signals of clinical efficacy as regulators insist on a clear signal in standard rating scales and measures of functional outcomes. However, the biggest difference to phase 2 where companies are free to use biomarkers as they deem best. is that in phase 3 regulators would have to agree to using biomarkers for diagnostic, selection or stratification purposes. That means a strict validation process approved by regulators would need to be implemented prior to or concomitant with the phase 3 trial. We refer to the guidance issued by FDA and EMA (EMA, 2014; FDA, 2018). We are not aware of any successful phase 3 programs for neuropsychiatric indications in which biomarkers have been used for the purposes outlined above, but it is a natural progression if biomarkers gain traction in phase 2 studies.

Phase 3 studies are randomized controlled multicentre trials on large patient groups (300–3000 or more). Considering the size of these studies, biomarkers for phase 3 studies need to be easily implement from an operational perspective. Also, they should not depend on culturally defined or dependent behaviours as these trials usually span different continents and cultures. Similarly, biomarkers derived from language and speech analysis are likely not to be viable as they would require an immense validation effort across all languages. Presently, the use of digital biomarkers in phase 3 would be very useful as they could provide a scalable way forward for real-world longitudinal quantitative data collection, but clinical efficacy validation and regulatory approval is urgently needed to further speed up this process (Mantua et al., 2021).

There is one noticeable difference in potential use of biomarkers between phase 2 and 3 trials and this relates to maintenance or relapse prevention trials where the outcome measure is an observable worsening of the patient's condition. In contrast to biomarkers discussed previously such biomarkers can be quite unspecific and potentiality could just measure changes from a routine baseline behaviour. For instance, biomarkers assessing sleep efficiency may be of great use as sleep disturbances have been found to predict relapse both in schizophrenia (Gleeson et al., 2024) and MDD (Fang et al., 2019). Others may include biomarkers of social and behavioural activities.

4. Limitations, future directions and developments

Despite our optimism that biomarkers can play a useful role in the development of pharmaceutical therapeutics for neuropsychiatric indications, it must be conceded that the development and use of biomarkers for this purpose has been and remains challenging. These challenges are due to – amongst others - the polygenetic basis for all common neuropsychiatric disorders, the aetiological and pathophysiological heterogeneity of neuropsychiatric disorders, the complexity of the biological system that is the focus of the treatment (the brain), related to this our limited, simplistic and reductionistic conceptualization of normal brain functioning and disease states and finally a focus on diagnostic classification that is arguably not reflecting biological

categories, rather than biological functions and symptom dimensions. These aspects and those listed in Table 2 contribute in varying degrees to the potential limitations of specific biomarkers.

A detailed and exhaustive description of how biomarkers should be developed goes beyond the scope and goal of this paper and can be found elsewhere (Abi-Dargham et al., 2023). However, we would like to briefly address the role genetics, ML and AL might play in the development and analysis of biomarkers and the data they provide.

Although most neuropsychiatric disorders have a strong genetic basis, attempts to use genetic or other molecular information to derive informative biomarkers have not been successful - or have not been attempted. The reasons for this include the previously stated polygenic architecture of these disorders, the complex topographical and chronological expression patterns of risk genes, in some case small samples and the use of clinical diagnoses to identify the phenotypes. To develop biomarkers that will be useful for drug development genetic studies of psychiatric disease may need to abandon diagnoses as phenotype and use specific function/dysfunctions instead - aka deep phenotyping. To this end, either large population databases of well characterized subjects must be available similar to those provided by the UK Biobank or deep phenotyping data from disease-focused consortia such as the Consortium on the Genetics of Schizophrenia (COGS) (Swerdlow et al., 2015). Databases like the former allows the use of Mendelian randomization to unravel the genetic/molecular underpinnings of functions of interest, identify dysfunction-relevant genomic and proteomic profiles and/or clusters (see for instance (Bhattacharyya et al., 2024)). Similarly, COGS and similar initiatives in other disorders may identify genetic association with discrete functions or endophenotypes (Greenwood et al., 2019; Gur et al., 2007). They may not only help to identify new drug targets, but also biomarkers for patient stratification as specific genetic and/or molecular abnormalities may be seen at the population level, but may not be operative (that is involved in aetiology or symptom inducing) in each patient. In addition, future developments in modelling, ML and AI may help to develop models of genetic architecture that could identify relevant endophenotypes and associated biomarkers at a system level that captures enough of the variance of the targeted dysfunction. Moreover, ML and AI applied to objectively measurable data such as EEG, imaging and omics data are likely to be used more frequently to develop biomarkers and biomarker profiles for specific dysfunction and/or response to pharmacological interventions (see for example (Wu et al., 2020)).

To be truly successful all these approaches have to take the genetic and biological heterogeneity in the aetiology of targeted dysfunctions into account. In other words, a reductionistic approach that assumes a 1:1 relationship between phenotype and underlying genetic/biological abnormalities will likely fail. This may also be true regarding the neuropharmacology of dysfunctions – a heterogeneity must be assumed. That is disturbances of different neurotransmitter systems as well as opposite extreme states of neurotransmitter systems can lead to similar dysfunction in the circuitry of interest (see for instance (Durstewitz et al., 1999; Floresco, 2013; Rolls et al., 2008; Vijayraghavan et al., 2007). These issues present enormous challenges but without a conceptualization that takes the emergent and pleiotropic properties of complex biological systems into account ML and AI will most likely remain sophisticated data crunching approaches without much useful output.

It is worth highlighting a few points that are critical for future developments: a) for biomarkers to be useful for stratification, pharmacological subtyping or response prediction in developing new treatments for common neuropsychiatric disorders they have to capture a substantial amount of variance of the targeted dysfunction, otherwise they remain of academic interest; b) they can be measured and implemented operationally with relative ease; c) ideally, they should also be related to the neuropharmacology that is targeted and the biological functions they index should be well characterised. For these reasons both genetic and imaging findings have not yet generated many ‘actionable’ biomarkers. Interesting imaging, genetic and other data profiles associated with specific dysfunctions and/outcomes will have to be ‘packaged’ into actionable biomarkers that can be scaled for phase 2 and 3 trials. However, these likely developments are technical and methodological aspects that in themselves do not change the fundamental questions outlined in this paper that is - which biomarker should be used in different phases and for different purposes in drug development?

5. Discussion and concluding remarks

We have outlined the pragmatic use of biomarkers in clinical drug development with a focus on phase 1 and 2 trials. An approach (i.e. implementation of a phase 1b PoM study) that will help identify a drug likely to fail in phase 2 or 3 should be highly attractive to drug developers. However, the available data suggests otherwise. A high-level search of Trialtrove® for phase 1a, 1b and 2 trials conducted in schizophrenia and MDD between 2011 and 2022 yielded a surprisingly low number of phase 1b studies which we took as proxy for PoM studies. A total of 809 phase 1a trials were listed under the indication ‘schizophrenia’ and 757 under the indication ‘depression’. The corresponding numbers for phase 2 were 242 for the indication ‘schizophrenia’ and 303 for the indication ‘depression’, respectively. However, for both indications only 21 phase 1b trials each were identified. Even though this search is very high level, it nonetheless highlights the fact that in most cases phase 1b PoM studies are rarely conducted before initiating phase 2 clinical studies. Similarly, a more focused review of industry studies conducted in the last 10 years, categorized as phase 2 and listed in clinicaltrials.gov identified 101 studies in MDD and 71 in schizophrenia. Of these 24 and 28, respectively, provided study protocols, Among the studies in MDD, five sampled measures of the HPA axis in an exploratory fashion without a clearly stated hypothesis. One evaluated sleep as an exploratory outcome. None used biomarkers to stratify or support clinical readouts. A similar picture was observed amongst studies in schizophrenia: One used EEG readouts to support efficacy, one investigated the impact of allelic variation of the C4 complement gene that had been implicated in schizophrenia (Sekar et al., 2016) on outcome. Three more collected DNA and other molecular information without any specific hypothesis. Thus, the overwhelming majority of these studies in MDD and schizophrenia did not employ biomarkers in any of the ways outlined above.

We argue that if strategies were pursued to identify ineffective drugs early, then the failure rate in phase 3 should be much lower than the current 55 % for psychiatric indications which is among the highest compared to other disease areas (Wong et al., 2019). Recent estimates of the probability of launch for drugs that enter phase 1 range from 3 to 15 % (Dowden and Munro, 2019; Wong et al., 2019). Even though non-clinical factors are known to contribute to drug failures in phase 3, particularly in psychiatry (Agid et al., 2013; Umbricht et al., 2020; Undurraga and Baldessarini, 2012) a better early identification of drug failures should increase the success rate in phase 3.

Given these numbers it is worth considering what the barriers to early PoM studies are. The first is perhaps confidence: Rejecting a compound that might provide billions of dollars in sales based on a negative biomarker result could be very costly. Thus, stopping the development of a compound that has been sifted from a likely array of >10,000 potential candidates not long after it has entered development is likely to be met by some resistance. There may also be evidence suggesting that the compound is well tolerated and shows favourable PD characteristics which makes it a good candidate to ‘test the hypothesis in patients’. Unrealistic optimism - often the Achilles heel of drug development – may also be a contributing factor. Consequently, as outlined in detail it is paramount that the reliability, validity, biological anchoring of the biomarker and links between the biomarker and the clinical domain are clearly established before irreversible decisions regarding the development of compounds are made.

A second factor is the likely absence of biomarkers for a specific

indication and/or symptom domain and hence a PoM study cannot be used for a Go/NoGo decision. A third factor may be the argument that a PoM study only delays the proof of clinical efficacy in a phase 2 study in patients and hence going directly to phase 2 is a more straightforward approach. However, a negative PoM study in one indication may not necessarily be the death knell for a compound as it may be beneficial in other disorders not evaluated in a specific POM study. Even if a PoM study cannot provide the basis for a Go/NoGo decision, it still may be beneficial and worth considering as it can provide initial characterization of clinical effects, useful information on biomarkers that may be used in later studies and data on efficacious dose ranges, thus increasing the probability of success.

We acknowledge that critical questions can be raised against our proposed biomarker lead approaches. First, is there evidence that a biomarker-led drug development would shorten development time? We readily admit that this may not be the case. To the contrary, development time may be longer. However, we contend that the probability of success across the field would be increased by the possibility of identifying more failures early in development. Secondly, what are the examples that the proposed approach is more successful than the traditional one? We provide a few examples above supporting the view that biomarker focused PoM studies have provided both positive and negative signals that were later confirmed in phase 2 trials. We acknowledge that there is simply not enough data to reliably answer this question at the present time. However, given the long history of failures in neuropsychiatric drug development a continuation of the past approaches is difficult to defend.

In conclusion, while many conference talks are given and papers on the virtues of precision psychiatry and biomarker-based drug development are published, searches of databases indicate that in recent years most drug development programs have shied away from implementing a well thought through and comprehensive biomarker strategy. Given the high unmet need for developing new treatments for neuropsychiatric indications, we feel it is urgent to lay out a pragmatic path of how biomarkers may be usefully and efficiently implemented in drug development. The goal is to identify drugs with high efficacy and rejecting those with low or no efficacy as early as possible and to define subgroups of patients for whom a particular drug may be effective. Such an approach would enhance the probability of success of efficacious drugs meeting unmet needs by clearing the development pathway of resource sapping late-stage failures. In each phase of drug development different questions must be answered and biomarkers need to be chosen accordingly. While in phase 1 studies biomarkers that demonstrate proof of target engagement and initial evidence of clinical efficacy are critical, in phase 2 and 3 biomarkers should be implemented that support accurate diagnosis, provide tools for stratification and/or patient selection, support clinical efficacy and functional effects and detect early signs of relapse in maintenance trials. While the challenges in the development of NMEs for neuropsychiatric indications remain high, we are convinced that a well-defined biomarker strategy as outlined in this paper will help establish at least a ‘less imprecise psychiatry’ and enhance the probability of success to the benefit of underserved patient populations.

Conflicts of interests

DU is owner of xperimed LLC that provides consultation on all aspects of clinical drug development for neuropsychiatric indications, works on a contracting basis for Autifony Therapeutics and Gilgamesh pharmaceuticals and holds stocks of Roche, Novartis and Gilgamesh Pharmaceuticals. He has been consulting to Abbvie, Biogen, ERG, Forbion, Healthrhyhms, Kynexis, Psychogenic, Roche and Siesta.

MJHK declares no conflicts of interest.

GRD is owner of P1vital LTD that provides consultation on suitable biomarkers for inclusion in clinical clinicals and supplies biomarkers such as the Facial expression recognition task (FERT, described in the text to industry and academia.

Role of Funding

No funding

Acknowledgements

Key aspects of this paper were presented and discussed at the European College of Neuropsychopharmacology (ECNP) supported meeting of the ECNP Experimental Medicine Network “Validating new biomarkers for unmet needs in the management and treatment of depression”, held on 7–8 February 2023 in Frankfurt, Germany.

References

Abi-Dargham, A., Moeller, S.J., Ali, F., DeLorenzo, C., Domschke, K., Horga, G., Jutla, A., Kotov, R., Paulus, M.P., Rubio, J.M., Sanacora, G., Veenstra-VanderWeele, J., Krystal, J.H., 2023. Candidate biomarkers in psychiatric disorders: state of the field. World Psych 22, 236–262.

Anon. European Medicines Agency, 2014. Qualification of novel methodologies for medicine development.

Anon. U.S. Food and Drug Administration, 2016. FDA-NIH biomarker working group. best (biomarkers, endpoints, and other tools) resource.

Anon. U.S. Food and Drug Administration, 2018. Biomarker qualification: evidentiary framework; draft guidance for industry and FDA staff.

Agid, O., Siu, C.O., Potkin, S.G., Kapur, S., Watsky, E., Vanderburg, D., Zipursky, R.B., Remington, G., 2013. Meta-regression analysis of placebo response in antipsychotic trials, 1970-2010. Am. J. Psych 170, 1335–1344.

Balla, A., Ginsberg, S.D., Abbas, A.I., Sershen, H., Javitt, D.C., 2020. Translational neurophysiological biomarkers of N-methyl-D-aspartate receptor dysfunction in serine racemase knockout mice. Biomark. Neuropsych 2.

Bhattacharyya, U., John, J., Lam, M., Fisher, J., Sun, B., Baird, D., Chen, C.-Y., Lencz, T., 2024. Large-scale mendelian randomization study reveals circulating blood-based proteomic biomarkers for psychopathology and cognitive task performance. Cold Spring Harbor Lab.

Bosiacki, M., Gassowska-Dobrowolska, M., Kojder, K., Fabianska, M., Jezewski, D., Gutowska, I., Lubkowska, A., 2019. Perineuronal nets and their role in synaptic homeostasis. Int. J. Mol. Sci 20.

Carceller, H., Guirado, R., Ripolles-Campos, E., Teruel-Marti, V., Nacher, J., 2020. Perineuronal nets regulate the inhibitory perisomatic input onto parvalbumin interneurons and gamma activity in the prefrontal cortex. J. Neurosci 40, 5008–5018.

Clementz, B.A., Sweeney, J.A., Hamm, J.P., Ivleva, E.I., Ethridge, L.E., Pearlson, G.D., Keshavan, M.S., Tamminga, C.A., 2016. Identification of distinct psychosis biotypes using brain-based biomarkers. Am. J. Psych 173, 373–384.

Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., Pangalos, M.N., 2014. Lessons learned from the fate of AstraZeneca’s drug pipeline: a five-dimensional framework. Nat. Rev. Drug Discov 13, 419–431.

Cools, R., Frank, M.J., Gibbs, S.E., Miyakawa, A., Jagust, W., D’Esposito, M., 2009. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J. Neurosci 29, 1538–1543.

Cortese, S., Solmi, M., Michelini, G., Bellato, A., Blanner, C., Canozzi, A., Eudave, L., Farhat, L.C., Hojlund, M., Kohler-Forsberg, O., Leffa, D.T., Rohde, C., de Pablo, G.S., Vita, G., Wesselhoeft, R., Martin, J., Baumeister, S., Bozhilova, N.S., Carlisi, C.O., Leno, V.C., Floris, D.L., Holz, N.E., Kraaijenvanger, E.J., Sacu, S., Vainieri, I., Ostuzzi, G., Barbui, C., Correll, C.U., 2023. Candidate diagnostic biomarkers for neurodevelopmental disorders in children and adolescents: a systematic review. World Psych 22, 129–149.

Curley, A.A., Lewis, D.A., 2012. Cortical basket cell dysfunction in schizophrenia. J. Physiol 590, 715–724.

Dawson, G.R., Post, A., Smart, T.S., Browning, M., Harmer, C.J., 2021. Accuracy in recognising happy facial expressions is associated with antidepressant response to a NOP receptor antagonist but not placebo treatment. J. Psychopharmacol 35, 1473–1478.

Dowden, H., Munro, J., 2019. Trends in clinical success rates and therapeutic focus. Nat. Rev. Drug Discov 18, 495–496.

Durstewitz, D., Kelc, M., Güntürkün, O., 1999. A neurocomputational theory of the dopaminergic modulation of working memory functions. J. Neurosci 19, 2807–2822.

Evans, B., DeMartinis, N., Gaudreault, F., Mancuso, J., Zumpano, L., Walling, D., Kelley Erb, M., Bednar, M., Binneman, B., 2016. Safety, Pharmacokinetics (PK) and Clinical Effects of PF-04958242, an a-Amino-3-Hydroxy-5-Methyl-4-Isoxazolepropionic Acid (AMPA) Positive Allosteric Modulator (PAM) in Subjects With Stable Schizophrenia, ACNP 55th Annual Meeting. Neuropsychopharmacol 41 (Suppl 1), S415.

Fang, H., Tu, S., Sheng, J., Shao, A., 2019. Depression in sleep disturbance: a review on a bidirectional relationship, mechanisms and treatment. J. Cell. Mol. Med 23, 2324–2332.

Floresco, S.B., 2013. Prefrontal dopamine and behavioral flexibility: shifting from an "inverted-U" toward a family of functions. Front Neurosci 7, 62.

Fonzo, G.A., Etkin, A., Zhang, Y., Wu, W., Cooper, C., Chin-Fatt, C., Jha, M.K., Trombello, J., Deckersbach, T., Adams, P., McInnis, M., McGrath, P.J., Weissman, M.

M., Fava, M., Trivedi, M.H., 2019. Brain regulation of emotional conflict predicts antidepressant treatment response for depression. Nat. Hum. Behav 3, 1319–1331.
Garcia-Gutierrez, M.S., Navarrete, F., Sala, F., Gasparyan, A., Austrich-Olivares, A., Manzanares, J., 2020. Biomarkers in psychiatry: concept, definition, types and relevance to the clinical reality. Front Psych 11, 432.
Gil-da-Costa, R., Stoner, G.R., Fung, R., Albright, T.D., 2013. Nonhuman primate model of schizophrenia using a noninvasive EEG method. Proc. Natl. Acad. Sci. U. S. A 110, 15425–15430.
Giles, M.A., Cooper, C.M., Jha, M.K., Chin Fatt, C.R., Pizzagalli, D.A., Mayes, T.L., Webb, C.A., Greer, T.L., Etkin, A., Trombello, J.M., Chase, H.W., Phillips, M.L., McInnis, M.G., Carmody, T., Adams, P., Parsey, R.V., McGrath, P.J., Weissman, M., Kurian, B.T., Fava, M., Trivedi, M.H., 2023. Reward behavior disengagement, a neuroeconomic model-based objective measure of reward pathology in depression: findings from the EMBARC trial. Behav. Sci. (Basel) 13.
Gleeson, J.F., McGuckian, T.B., Fernandez, D.K., Fraser, M.I., Pepe, A., Taskis, R., Alvarez-Jimenez, M., Farhall, J.F., Gumley, A., 2024. Systematic review of early warning signs of relapse and behavioural antecedents of symptom worsening in people living with schizophrenia spectrum disorders. Clin. Psychol. Rev 107, 102357.
Gonzalez-Burgos, G., Lewis, D.A., 2012. NMDA receptor hypofunction, parvalbumin-positive neurons, and cortical gamma oscillations in schizophrenia. Schizophr. Bull 38, 950–957.
Greenwood, T.A., Lazzeroni, L.C., Maihofer, A.X., Swerdlow, N.R., Calkins, M.E., Freedman, R., Green, M.F., Light, G.A., Nievergelt, C.M., Nuechterlein, K.H., Radant, A.D., Siever, L.J., Silverman, J.M., Stone, W.S., Sugar, C.A., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Gur, R.C., Gur, R.E., Braff, D.L., 2019. Genome-wide association of endophenotypes for schizophrenia from the consortium on the genetics of schizophrenia (COGS) Study. JAMA. Psych 76, 1274–1284.
Gribkoff, V.K., Kaczmarek, L.K., 2017. The need for new approaches in CNS drug discovery: why drugs have failed, and what can be done to improve outcomes. Neuropharmacology 120, 11–19.
Guhn, M., Forer, B., Zumbo, B.D., 2014. Reliable change index. In: Michalos, A.C. (Ed.), Encyclopedia of Quality of Life and Well-Being Research. Springer Netherlands, Dordrecht, pp. 5459–5462.
Gur, R.E., Calkins, M.E., Gur, R.C., Horan, W.P., Nuechterlein, K.H., Seidman, L.J., Stone, W.S., 2007. The Consortium on the Genetics of Schizophrenia: neurocognitive Endophenotypes. Schizophr. Bull 33, 49–68.
Gurrell, R., Chang, I., Duvvuri, S., Giugliano, A., Pastino, G., Pham, T., Versavel, S., Jacobs, G., Pour, K.S., Zuiker, R., Sanchez, R., Renger, J., 2023. Darigabat reduces acute panic and fear symptoms induced by co2inhalation in healthy participants. Am. Acad. Neurol. Ann. Meet.
Hack, L.M., Tozzi, L., Zenteno, S., Olmsted, A.M., Hilton, R., Jubeir, J., Korgaonkar, M.S., Schatzberg, A.F., Yesavage, J.A., O'Hara, R., Williams, L.M., 2023. A cognitive biotype of depression and symptoms, behavior measures, neural circuits, and differential treatment outcomes: a prespecified secondary analysis of a randomized clinical trial. JAMA. Netw. Open 6, e2318411.
Harmer, C.J., Bhagwagar, Z., Perrett, D.I., Vollm, B.A., Cowen, P.J., Goodwin, G.M., 2003. Acute SSRI administration affects the processing of social cues in healthy volunteers. Neuropsychopharmacology 28, 148–152.
Herzog, L.E., Wang, L., Yu, E., Choi, S., Farsi, Z., Song, B.J., Pan, J.Q., Sheng, M., 2023. Mouse mutants in schizophrenia risk genes GRIN2A and AKAP11 show EEG abnormalities in common with schizophrenia patients. Transl. Psych 13, 92.
Huneke, N.T., Walsh, A.E., Brown, R., Browning, M., Harmer, C.J., 2017. No evidence for an acute placebo effect on emotional processing in healthy volunteers. J. Psychopharmacol 31, 1578–1587.
Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., Sanislow, C., Wang, P., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am. J. Psych 167, 748–751.
Javitt, D.C., Steinschneider, M., Schroeder, C.E., Arezzo, J.C., 1996. Role of cortical N-methyl-d-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. Proc. Nat. Acad. Sci 93, 11962–11967.
Jongs, N., Jagesar, R., van Haren, N.E.M., Penninx, B., Reus, L., Visser, P.J., van der Wee, N.J.A., Koning, I.M., Arango, C., Sommer, I.E.C., Eijkemans, M.J.C., Vorstman, J.A., Kas, M.J., 2020. A framework for assessing neuropsychiatric phenotypes by using smartphone-based location data. Transl. Psychiatry 10, 211.
Khadimallah, I., Jenni, R., Cabungcal, J.H., Cleusix, M., Fournier, M., Beard, E., Klauser, P., Knebel, J.F., Murray, M.M., Retsa, C., Siciliano, M., Spencer, K.M., Steullet, P., Cuenod, M., Conus, P., Do, K.Q., 2022. Mitochondrial, exosomal miR137-COX6A2 and gamma synchrony as biomarkers of parvalbumin interneurons, psychopathology, and neurocognition in schizophrenia. Mol. Psych 27, 1192–1204.
Krystal, A.D., Pizzagalli, D.A., Mathew, S.J., Sanacora, G., Keefe, R., Song, A., Calabrese, J., Goddard, A., Goodman, W., Lisanby, S.H., Smoski, M., Weiner, R., Iosifescu, D., Nurnberger Jr., J., Szabo, S., Murrough, J., Shekhar, A., Potter, W., 2018. The first implementation of the NIMH FAST-FAIL approach to psychiatric drug development. Nat. Rev. Drug Discov 18, 82–84.
Krystal, A.D., Pizzagalli, D.A., Smoski, M., Mathew, S.J., Nurnberger Jr., J., Lisanby, S.H., Iosifescu, D., Murrough, J.W., Yang, H., Weiner, R.D., Calabrese, J.R., Sanacora, G., Hermes, G., Keefe, R.S.E., Song, A., Goodman, W., Szabo, S.T., Whitton, A.E., Gao, K., Potter, W.Z., 2020. A randomized proof-of-mechanism trial applying the 'fast-fail' approach to evaluating kappa-opioid antagonism as a treatment for anhedonia. Nat. Med 26, 760–768.
Leishman, E., O'Donnell, B.F., Millward, J.B., Vohs, J.L., Rass, O., Krishnan, G.P., Bolbecker, A.R., Morzorati, S.L., 2015. Phencyclidine disrupts the auditory steady state response in rats. PLoS. ONE 10, e0134979.
Lewis, D.A., Curley, A.A., Glausier, J.R., Volk, D.W., 2012. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. Trends. Neurosci 35, 57–67.
Mantua, V., Arango, C., Balabanov, P., Butlen-Ducuing, F., 2021. Digital health technologies in clinical trials for central nervous system drugs: an EU regulatory perspective. Nature. Rev. Drug Discov 20, 83–84.
McKernan, R.M., Rosahl, T.W., Reynolds, D.S., Sur, C., Wafford, K.A., Atack, J.R., Farrar, S., Myers, J., Cook, G., Ferris, P., Garrett, L., Bristow, L., Marshall, G., Macaulay, A., Brown, N., Howell, O., Moore, K.W., Carling, R.W., Street, L.J., Castro, J.L., Ragan, C.I., Dawson, G.R., Whiting, P.J., 2000. Sedative but not anxiolytic properties of benzodiazepines are mediated by the GABA(A) receptor alpha1 subtype. Nat. Neurosci 3, 587–592.
Meyer-Lindenberg, A., Nielsen, J., Such, P., Lemming, O.M., Zambori, J., Buller, R., der Goltz, C.V., 2022. A double-blind, randomized, placebo-controlled proof of concept study of the efficacy and safety of Lu AF11167 for persistent negative symptoms in people with schizophrenia. Eur. Neuropsychopharmacol 61, 4–14.
Pangalos, M.N., Schechter, L.E., Hurko, O., 2007. Drug development for CNS disorders: strategies for balancing risk and reducing attrition. Nat. Rev. Drug Discov 6, 521–532.
Perini, F., Nazimek, J.M., McKie, S., Capitao, L.P., Scaife, J., Pal, D., Browning, M., Dawson, G.R., Nishikawa, H., Campbell, U., Hopkins, S.C., Loebel, A., Elliott, R., Harmer, C.J., Deakin, B., Koblan, K.S., 2023. Effects of ulotaront on brain circuits of reward, working memory, and emotion processing in healthy volunteers with high or low schizotypy. Schizophr. (Heidelb) 9, 49.
Pizzagalli, D.A., Smoski, M., Ang, Y.S., Whitton, A.E., Sanacora, G., Mathew, S.J., Nurnberger Jr., J., Lisanby, S.H., Iosifescu, D.V., Murrough, J.W., Yang, H., Weiner, R.D., Calabrese, J.R., Goodman, W., Potter, W.Z., Krystal, A.D., 2020. Selective kappa-opioid antagonism ameliorates anhedonic behavior: evidence from the fast-fail trial in mood and anxiety spectrum disorders (FAST-MAS). Neuropsychopharmacology 45, 1656–1663.
Ranganathan, M., DeMartinis, N., Huguenel, B., Gaudreault, F., Bednar, M.M., Shaffer, C. L., Gupta, S., Cahill, J., Sherif, M.A., Mancuso, J., Zumpano, L., D'Souza, D.C., 2017. Attenuation of ketamine-induced impairment in verbal learning and memory in healthy volunteers by the AMPA receptor potentiator PF-04958242. Mol. Psychiatry 22, 1633–1640.
Rivolta, D., Heidegger, T., Scheller, B., Sauer, A., Schaum, M., Birkner, K., Singer, W., Wibral, M., Uhlhaas, P.J., 2015. Ketamine Dysregulates the amplitude and connectivity of high-frequency oscillations in cortical-subcortical networks in humans: evidence from resting-state Magnetoencephalography-recordings. Schizophr. Bull 41, 1105–1114.
Rolls, E.T., Loh, M., Deco, G., Winterer, G., 2008. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. Nat. Rev. Neurosci 9, 696–709.
Sanacora, G., Smith, M.A., Pathak, S., Su, H.L., Boeijinga, P.H., McCarthy, D.J., Quirk, M. C., 2014. Lanicemine: a low-trapping NMDA channel blocker produces sustained antidepressant efficacy with minimal psychotomimetic adverse effects. Mol. Psych 19, 978–985.
Sarhadi, V.K., Armengol, G., 2022. Molecular Biomarkers in Cancer. Biomolecules 12, 1021.
Schmidt, A., Bachmann, R., Kometer, M., Csomor, P.A., Stephan, K.E., Seifritz, E., Vollenweider, F.X., 2012. Mismatch negativity encoding of prediction errors predicts S-ketamine-induced cognitive impairments. Neuropsychopharmacology 37, 865–875.
Schuelert, N., Dorner-Ciossek, C., Brendel, M., Rosenbrock, H., 2018. A comprehensive analysis of auditory event-related potentials and network oscillations in an NMDA receptor antagonist mouse model using a novel wireless recording technology. Physiol. Rep 6, e13782.
Seitz-Holland, J., Seethaler, M., Makris, N., Rushmore, J., Cho, K.K., Rizzoni, E., Vangel, M., Sahin, O.S., Heller, C., Pasternak, O., Szczepankiewicz, F., Westin, C.F., Losak, J., Ustohal, L., Tomandl, J., Vojtisek, L., Kudlicka, P., Jani, M., Woo, T.W., Kasparek, T., Kikinis, Z., Kubicki, M., 2022. The association of matrix metalloproteinase 9 (MMP9) with hippocampal volume in schizophrenia: a preliminary MRI study. Neuropsychopharmacology 47, 524–530.
Sekar, A., Bialas, A.R., de Rivera, H., Davis, A., Hammond, T.R., Kamitaki, N., Tooley, K., Presumey, J., Baum, M., Van Doren, V., Genovese, G., Rose, S.A., Handsaker, R.E., Daly, M.J., Carroll, M.C., Stevens, B., McCarroll, S.A., Schizophrenia Working Group of the Psychiatric Genomics, 2016. Schizophrenia risk from complex variation of complement component 4. Nature 530, 177–183.
Sivarao, D.V., Chen, P., Senapati, A., Yang, Y., Fernandes, A., Benitex, Y., Whiterock, V., Li, Y.W., Ahlijanian, M.K., 2016. 40 Hz auditory steady-state response is a pharmacodynamic biomarker for cortical NMDA receptors. Neuropsychopharmacology 41, 2232–2240.
Sugiyama, S., Ohi, K., Kuramitsu, A., Takai, K., Muto, Y., Taniguchi, T., Kinukawa, T., Takeuchi, N., Motomura, E., Nishihara, M., Shioiri, T., Inui, K., 2021. The auditory steady-state response: electrophysiological index for sensory processing dysfunction in psychiatric disorders. Front Psych 12, 644541.
Swerdlow, N.R., Gur, R.E., Braff, D.L., 2015. Consortium on the Genetics of Schizophrenia (COGS) assessment of endophenotypes for schizophrenia: an introduction to this Special Issue of Schizophrenia Research. Schizophr. Res 163, 9–16.
Tamminga, C.A., Clementz, B.A., Pearlson, G., Keshavan, M., Gershon, E.S., Ivleva, E.I., McDowell, J., Meda, S.A., Keedy, S., Calhoun, V.D., Lizano, P., Bishop, J.R., Hudgens-Haney, M., Alliey-Rodriguez, N., Asif, H., Gibbons, R., 2021. Biotyping in psychosis: using multiple computational approaches with one data set. Neuropsychopharmacology 46, 143–155.
Tozzi, L., Zhang, X., Pines, A., Olmsted, A.M., Zhai, E.S., Anene, E.T., Chesnut, M., Holt-Gosselin, B., Chang, S., Stetz, P.C., Ramirez, C.A., Hack, L.M., Korgaonkar, M.S.,

Wintermark, M., Gotlib, I.H., Ma, J., Williams, L.M., 2024. Personalized brain circuit scores identify clinically distinct biotypes in depression and anxiety. Nat. Med.
Umbricht, D., Abt, M., Tamburri, P., Chatham, C., Holiga, S., Frank, M.J., Collins, A.G.E., Walling, D.P., Mofsen, R., Gruener, D., Gertsik, L., Sevigny, J., Keswani, S., Dukart, J., 2021. Proof-of-mechanism study of the phosphodiesterase 10 inhibitor RG7203 in patients with schizophrenia and negative symptoms. Biol. Psych. Glob. Open. Sci 1, 70–77.
Umbricht, D., Koller, R., Vollenweider, F.X., Schmid, L., 2002. Mismatch negativity predicts psychotic experiences induced by NMDA receptor antagonist in healthy volunteers. Biol. Psych 51, 400–406.
Umbricht, D., Kott, A., Daniel, D.G., 2020. The effects of erratic ratings on placebo response and signal detection in the roche bitopertin phase 3 negative symptom studies—a post hoc analysis. Schizophr. Bullet. Open 1.
Umbricht, D., Schmid, L., Koller, R., Vollenweider, F.X., Hell, D., Javitt, D.C., 2000. Ketamine-induced deficits in auditory and visual context-dependent processing in healthy volunteers: implications for models of cognitive deficits in schizophrenia. Arch. Gen. Psych 57, 1139–1147.
Umbricht, D., Vollenweider, F.X., Schmid, L., Grubel, C., Skrabo, A., Huber, T., Koller, R., 2003. Effects of the 5-HT2A agonist psilocybin on mismatch negativity generation and AX-continuous performance task: implications for the neuropharmacology of cognitive deficits in schizophrenia. Neuropsychopharmacology 28, 170–181.
Undurraga, J., Baldessarini, R.J., 2012. Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. Neuropsychopharmacology 37, 851–864.
Vijayraghavan, S., Wang, M., Birnbaum, S.G., Williams, G.V., Arnsten, A.F., 2007. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. Nat. Neurosci 10, 376–384.
Webb, C.A., Dillon, D.G., Pechtel, P., Goer, F.K., Murray, L., Huys, Q.J., Fava, M., McGrath, P.J., Weissman, M., Parsey, R., Kurian, B.T., Adams, P., Weyandt, S., Trombello, J.M., Grannemann, B., Cooper, C.M., Deldin, P., Tenke, C., Trivedi, M., Bruder, G., Pizzagalli, D.A., 2016. Neural correlates of three promising endophenotypes of depression: evidence from the EMBARC study. Neuropsychopharmacology 41, 454–463.
Williams, L.M., Rush, A.J., Koslow, S.H., Wisniewski, S.R., Cooper, N.J., Nemeroff, C.B., Schatzberg, A.F., Gordon, E., 2011. International study to predict optimized treatment for depression (iSPOT-D), a randomized clinical trial: rationale and protocol. Trials 12, 4.
Winter, N.R., Blanke, J., Leenings, R., Ernsting, J., Fisch, L., Sarink, K., Barkhau, C., Emden, D., Thiel, K., Flinkenflugel, K., Winter, A., Goltermann, J., Meinert, S., Dohm, K., Repple, J., Gruber, M., Leehr, E.J., Opel, N., Grotegerd, D., Redlich, R., Nitsch, R., Bauer, J., Heindel, W., Gross, J., Risse, B., Andlauer, T.F.M., Forstner, A. J., Nothen, M.M., Rietschel, M., Hofmann, S.G., Pfarr, J.K., Teutenberg, L., Usemann, P., Thomas-Odenthal, F., Wroblewski, A., Brosch, K., Stein, F., Jansen, A., Jamalabadi, H., Alexander, N., Straube, B., Nenadic, I., Kircher, T., Dannlowski, U., Hahn, T., 2024. A systematic evaluation of machine learning-based biomarkers for major depressive Disorder. JAMA. Psychiatry.
Wong, C.H., Siah, K.W., Lo, A.W., 2019. Estimation of clinical trial success rates and related parameters. Biostatistics 20, 273–286.
World Health Organization, 2017. World health statistics 2017: monitoring health for the SDGs, Sustainable Development Goals.
Wu, W., Zhang, Y., Jiang, J., Lucas, M.V., Fonzo, G.A., Rolle, C.E., Cooper, C., Chin-Fatt, C., Krepel, N., Cornelssen, C.A., Wright, R., Toll, R.T., Trivedi, H.M., Monuszko, K., Caudle, T.L., Sarhadi, K., Jha, M.K., Trombello, J.M., Deckersbach, T., Adams, P., McGrath, P.J., Weissman, M.M., Fava, M., Pizzagalli, D.A., Arns, M., Trivedi, M.H., Etkin, A., 2020. An electroencephalographic signature predicts antidepressant response in major depression. Nat. Biotechnol. 38, 439–447.
Yan, T., Suzuki, K., Kameda, S., Kuratomi, T., Mihara, M., Maeda, M., Hirata, M., 2023. Intracranial EEG recordings of high-frequency activity from a wireless implantable bmi device in awake nonhuman primates. IEEE. Trans. Biomed. Eng. 70, 1107–1113.