Date of Revision: 4 September 2023 Word count: 3403

25 Years of fMRI Cue-Reactivity—Overview of Parameter Space and Potential for Biomarker Development

Addiction Cue-Reactivity Initiative (ACRI) Group

#Corresponding Author:

Hamed Ekhtiari, MD, PhD
Laureate Institute for Brain Research
6655 S Yale Ave, Tulsa, OK 74136-3326
Phone: +1 918 502 5120; email: hekhtiari@laureateinstitute.org

Key Points

Importance: We assessed potentials for fMRI drug-cue-reactivity (FDCR) derived biomarkers to improve intervention development and clinical care for substance use disorders (SUDs) and identified key challenges.

Findings: 415 FDCR studies are assessed with a systematic review. Results from 357 studies could potentially help develop diagnostic, prognostic, susceptibility, severity, monitoring, predictive or response biomarkers. We also identify substantial heterogeneity in task and study design that can hinder biomarker development.

Meaning: A sizable literature supports the development of FDCR-derived biomarkers, but moving forward requires large-scale collaboration, methodological harmonization and optimization, and clinical and analytical validation.

Abstract

Importance: In the last 25 years, fMRI drug cue-reactivity (FDCR) studies have characterized the neurobiology of drug cue-reactivity. However, no FDCR-derived biomarkers have been approved yet for treatment development or clinical adoption. Traversing this translational gap requires a systematic assessment of the FDCR literature evidence and its heterogeneity and an evaluation of possible clinical uses of FDCR-derived biomarkers.

Objective: We use a systematic review of FDCR studies to summarize the state of the field, assess their potential for biomarker development, and outline a clear process for biomarker qualification to guide future research and validation efforts.

Evidence Review: We reviewed every original FDCR investigation published until the end of 2022. Collected data cover study design, participant characteristics, FDCR task design, and whether each study provided evidence that might potentially help develop susceptibility, diagnostic, response, prognostic, predictive, or severity biomarkers for one or more addictive disorders.

Findings: There were 415 FDCR studies published between 1998-2022. Most focus on nicotine (29.6%), alcohol (29.2%), or cocaine (11.1%), and most utilize visual cues (85.3%). Together, these studies recruited 19,311 participants, including 13,812 individuals with past or current SUDs. Most studies could potentially support biomarker development, including diagnostic (32.7%), treatment response (32.3%), severity (19.2%), prognostic (6.9%), predictive (5.7%), monitoring (2.7%), and susceptibility (0.5%) biomarkers. One hundred and fifty-five interventional studies used FDCR, mostly to investigate pharmacological (43.2%) or cognitive/behavioral (33.5%) interventions. 141 studies used FDCR as a response measure and 134 (88.7%) reported significant interventional FDCR alterations. Twenty-five studies used FDCR as an intervention outcome predictor, with 96% of these studies finding significant associations between FDCR markers and treatment outcomes.

Conclusions and Relevance: Based on this systematic review and the proposed biomarker development framework, we outline a pathway for the development and regulatory qualification of FDCR-based biomarkers of addiction and recovery. Further validation could support the use of FDCR-derived measures, potentially accelerating treatment development and improving diagnostic, prognostic, and predictive clinical judgments.

Introduction

The evaluation of substance use disorders (SUDs) is currently reliant on interviews, self-reported measures, and biological assays of drug metabolites which mostly reflect substance use and confound the distinction between markers of substance use and the complex pathophysiology underlying SUDs[1]. Growing recognition of this issue has led to recent interest in identifying the neurobiological underpinnings of SUDs[2] and translating this knowledge to facilitate the development of novel treatment targets and interventions and theoretically grounded, empirically sound, and clinically relevant "biomarkers" for patient-tailored care[3]. A particularly impactful paradigm in addiction medicine has been fMRI drug cue reactivity (FDCR), where brain activation patterns during an individual's exposure to addiction-related sensory stimuli are measured as a potential marker of underlying neuropathology[4]. FDCR has consistently shown that SUDs are associated with remarkable aberrations in the neural circuitry underpinning incentive salience, reward evaluation, interoception, memory, habit formation, and executive control[5,6]. See box 1 for a general overview of biomarkers in psychiatry and addiction medicine, for an introduction to FDCR along with eFigure1.

Box 1. Biomarkers in psychiatry and addiction medicine

(*Refer to the bibliography in online-only materials for items cited in boxes)

The FDA-National Institutes of Health (NIH) Biomarker Working Group defines a "biomarker" as "a defined characteristic measured as an indicator of normal or pathogenic biological processes, or biological responses to an exposure or intervention, including therapeutic interventions" [BEST (Biomarkers, EndpointS, and other Tools) Resource[1*]]. The development of clinically relevant biomarkers is a major goal of addiction neuroscience and translational psychiatry. Regulating agencies have shown increasing interest in validated biomarkers, with the FDA's biomarker qualification program, among others, working to provide formal endorsement of biomarkers to facilitate their use in drug development and regulatory decisions[2]. Recent reviews and opinions have outlined the potential for an expanding group of central and peripheral biomarkers of major psychiatric conditions, including genomic, epigenetic, and transcriptomic biomarkers[3], proteomic biomarkers[4], inflammatory markers[5], non-inflammatory chemokines[6], cardiovascular biomarkers[7], hormonal and neurotransmitter profiles[8], cognitive and behavioral markers[9], biomarkers derived from neuroimaging paradigms[10,11], and multi-modal biomarkers[12]. Several neuroimaging biomarkers are also at varying stages of validation by the FDA for neurological or psychiatric disorders. These include baseline hippocampal volume assessed by Magnetic Resonance Imaging (MRI) in Alzheimer's disease and Glx (Glutamine+ Glutamate) measured in the brain by Magnetic Resonance Spectroscopy (MRS) in depression. Notably, the NIMH "Fast-Fail" trial initiative supports the use of functional MRI (fMRI) in early-phase drug development to lower the risk of failure in large clinical trials: in the first implementation of the approach, task-related fMRI revealed that kappa opioid receptor antagonism can enhance reward-related ventral striatal activation, supporting larger trials for cross-diagnostic treatment of anhedonia[13,14].

Commensurate with broader progress in biomarker development across various psychiatric disorders, different types of brain-based markers with potential for clinical translation have been proposed for addictive disorders, but their clinical and analytical validation remains limited[15]. Objective biological metrics of SUDs are currently limited to measures of substance use - mainly testing for psychoactive substances or their metabolites in biological samples[16,17] - or measures that reflect the toxic effects of use[18]. Notably, these biomarkers reflect endpoints of substance use and toxicity and are not informed by the dynamic processes that underlie how drug use behaviors relate to addiction. This limitation hampers the clinical use of intermediate phenotypes and the development of biomarkers to identify at-risk

individuals and to mechanistically inform, predict, and monitor interventions[19]. Relatedly, although the DSM-5 proposed diagnostic criteria for behavioral addictions (BAs), including gambling disorder and internet gaming disorder, no biomarkers are included for BAs[20]. According to the FDA website (visited December 15th, 2020), there are no qualified biomarkers or ongoing qualification processes for biomarkers in addiction medicine/psychiatry [21]. The only submitted biomarker qualification proposal - covering a single nucleotide polymorphism in the delta opioid receptor 1 gene - appears to have been rejected at an early phase[21].

Theoretical Background on fMRI Drug Cue-Reactivity (FDCR)

A popular paradigm to assess brain function in individuals with SUDs is data acquisition with fMRI during the administration of a drug cue-reactivity task[22]. Similar paradigms have been developed to investigate reactivity to addiction-relevant cues in BAs[23–25]. These paradigms (referred to collectively as "cue-reactivity paradigms") involve the presentation of a variety of conditioned cues, associated with the availability or use of substances or other similarly desirable experiences to participants who have had prior experiences with them. The cue-reactivity paradigm rests on the understanding that addictive disorders involve sensitization to addiction-relevant cues[26], which can trigger behavioral and physiological responses associated with craving and anticipation[27]. Cue-reactivity tasks had been developed and validated extensively before the advent of fMRI and have been readily modified and adopted in fMRI research[28,29]. Engagement with addiction-relevant cues under fMRI scanning enables the exploration of the neural mechanisms that are associated with the response to addiction-relevant cues[30], and fMRI drug cue-reactivity (FDCR) has demonstrated that SUDs are associated with aberrations in the neural circuitry underpinning incentive salience, reward evaluation, interoception, memory, habit formation, and executive control[31,32].

If variations in FDCR signal are associated with the existence and severity of addiction-related processes, the development of FDCR-derived biomarkers could aid in diagnostic classification and sub-grouping, assessing disease severity, identifying at-risk individuals, understanding the neural mechanisms involved in effective interventions, targeting disrupted neural function with novel interventions, early evaluation of new interventions based on surrogate endpoints such as target-engagement, and monitoring treatment effectiveness[13,32–34]. More recent avenues of research have combined cue-reactivity with other paradigms during fMRI acquisition[35–37] and investigated the interaction of FDCR and genomic[38], epigenetic[39], metabolic[40], physiological[41], developmental[42], behavioral[43], cognitive[44–47], personality[48] and psychiatric[49,50] correlates of SUDs. Considering the multi-faceted and multi-causal nature of these disorders and their frequent co-occurrence with other mental and physical health conditions, such studies establish the etiological importance of FDCR in SUDs and lead to better characterizations of addictive processes, ultimately enabling the development of multi-domain biomarkers[51,52] For example, neuro-genetics studies have shown that the A118G single nucleotide polymorphism of the mu opioid receptor (OPRM1) gene may result in higher levels of FDCR[53] (Ray et al., 2014) and also impact the clinical response to naltrexone (a μ-opioid antagonist medication)[54].

The road to FDCR-derived biomarkers

In the third decade of FDCR research, with consistently observed correlations between FDCR and important clinical outcomes[7,8], biomarkers derived from FDCR paradigms could inform intervention development or clinical care of people with SUDs. Given the expense and technical difficulty of qualifying biomarkers for use in regulatory decision-making, for example to support the approval of specific interventions, frameworks have been developed to facilitate the validation of biomarkers. According to the biomarker validation frameworks developed by organizations such as the European Medicines Agency (EMA)[9] and the FDA[10], an initial step in developing FDCR-derived biomarkers with regulatory approval

would be the specification of precise "contexts of use" (COU). Different methods and standards of validation might be required, for example, for an FDCR-derived biomarker developed to classify individuals with SUDs into different subtypes compared to one used to predict individual responses to a specific intervention. Just as crucially, the methodological details of any FDCR-derived biomarker would need to be carefully considered and clearly specified since they may influence the FDCR signal and the interpretation of the biomarker[11,12].

In the next stage, the defined biomarker will need to be characterized and validated within the COU. A principal step is "analytical" validation, establishing appropriate accuracy, repeatability, and reproducibility of the biomarker within the proposed COU[13]. Demonstrating "clinical" validity requires elucidating the etiological link of an FDCR biomarker to SUD symptoms and establishing that the biomarker appropriately measures a clinical feature of a disease, disease outcome, or treatment outcome[14]. Finally, the practical use of FDCR-derived biomarkers in clinical or drug development contexts requires demonstration of cost-effectiveness. These validation steps require a combination of systematic reviews and meta- and mega-analyses, expert consensus, and new studies to address potential evidentiary gaps. An overview of the overall FDCR biomarker development framework is provided in Figure 1.

Moving towards the development of clinically relevant FDCR-derived biomarkers necessitates taking stock of the current state and evolution of FDCR as a research field. While many useful systematic reviews and meta-analyses of cue-reactivity fMRI studies are available[7,15–18], these efforts have largely focused on estimating neuroimaging effect sizes rather than systematically investigating the methodological characteristics of FDCR studies and the potential of FDCR for biomarker development. We present a systematic review and synthesis of the FDCR literature, covering basic study design features, studied substances and behaviors, and methodological parameters, to outline the degree of methodological heterogeneity and to identify outstanding gaps in the evidence. We then provide a systematic assessment of the potential of FDCR studies for biomarker development under the NIH framework in translational addiction science and discuss exemplar FDCR indices. We finally highlight a set of concrete actions and future directions in the translation of FDCR-derived biomarkers from the bench to the bedside, based on the outlined biomarker development framework and the systematic review.

Methods and Results

Detailed methods and results of the systematic review sections are presented in the online-only materials (eMethods and eResults sections) and the search terms and syntax can be found in eTables 1 and 2. The methods and results are organized according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist and the protocol for this systematic review was pre-registered[19]. While we refer to fMRI "drug" cue-reactivity (including alcohol) throughout the manuscript, BA studies focusing on gambling and gaming were not excluded as they constitute a small portion of the cue-reactivity literature and involve cue-reactivity paradigms similar to drug cue-reactivity studies. Separate analyses of substance and behavioral addictions can be found in eFigure 10.

The final database includes 415 studies, from 19 countries (eFigures 2 and 3) and will be continually updated, according to a registered protocol, to provide an up-to-date repository of FDCR studies and facilitate future investigations. Our results indicate a growing interest in the FDCR paradigm, with 307 of

the 415 FDCR studies in our database published in the last ten years of the systematic review (eFigure 4). We will first consider the methodological aspects of reviewed studies.

Methodological heterogeneity and biomarker specification

A central element of an FDCR experiment is the selection of cues used to elicit neural reactivity, with a wide array of options available: while 85.3% of reviewed studies used visual cues, others used a variety of auditory, semantic, gustatory, olfactory or tactile reminders of drugs or drug use, alone or in various combinations (Figure 2 and eFigure 5). The impact of cue sensory modality in FDCR remains under-explored, but cues in different sensory modalities likely induce markedly different neural activations[20] and multi-sensory cues or delivering drug cues together with other rewarding stimuli may improve ecological validity and FDCR signal [21,22].

Basic task-design elements also vary considerably between studies (Figure 2). Sixty-two percent of studies used blocked designs, which are popular since repeated presentations of drug-relevant stimuli may constitute more robust exposure and subsequent activation. However, event-related designs may be better able to optimally characterize the shape of the BOLD response to drug cues[11], and more sophisticated mixed designs could model interactions between cue exposure and context. Furthermore, FDCR has been combined with other task modalities to probe the interaction of cue exposure and different cognitive processes (52 studies). Such combined paradigms are attempted to increase ecological validity since drug cue-reactivity engages with multiple neurocognitive processes. For example, FDCR during response inhibition was able to predict tobacco abstinence[23].

Methodological parameters should ideally be chosen based on evidence from meta- and mega-analyses or at least empirical results, with alternative sources such as structured expert opinion used to address knowledge gaps[11]. Such choices also involve trade-offs: for example, simple visual FDCR paradigms may be selected since they are relatively inexpensive and already widely used[24], while complex interactional designs and multisensory stimuli with greater ecological validity may be technically challenging and more difficult to standardize between studies[25]. On the other hand, multisensory stimuli may improve signal-to-noise ratio to increase reliability at the same scanning duration[26]. Overall, since methodological heterogeneity between studies can hamper the comparison of findings[27] and complicate meta-analyses for biomarker development[28], it is important to promote standardized best practices and methodological harmonization to the extent that is practical. Appropriate reporting and explanation of key methodological elements and harmonized reporting standards is essential regardless of what choices are made, for example using the COBIDAS guideline[29] and the recently developed ENIGMA-ACRI reporting checklist[11].

Participant characteristics

There is evidence that participant characteristics substantially impact the FDCR signal, highlighting the importance of specifying target populations for FDCR biomarkers and ensuring the diversity of populations used to develop such biomarkers. Overall, 19,311 individuals participated in FDCR studies from 1998-2022, including 12,950 (67.1%) men and 5,130 (26.5%) women, with the sex of 1231 participants (6.4%) not explicitly specified (eFigure 6 The fact that only 26.5% of participants in FDCR studies have been women raises questions about the generalizability of findings and potential biomarkers informed by this literature, since men and women may have markedly distinct neural activation patterns during drug cue exposure[30,31]

While outside the scope of the present review, other demographic factors such as age, socio-economic status and social determinants of health, medical and psychiatric comorbidities, and cultural background likely impact the FDCR signal as well[11].

Future studies would benefit from complex multivariate modelling techniques which can disambiguate the influence of various participant characteristics and other methodological choices and investigate complex FDCR patterns. Further, the median sample size of FDCR studies in our database is only 37, which may be too small to discover replicable FDCR markers[32]. Larger samples as well as meta- and mega-analyses are important for developing valid and generalizable biomarkers. This systematic review aims to provide a comprehensive overview of the entire FDCR field, and the broad inclusion criteria for study participants included studies of individuals who met SUD diagnostic criteria and those who used substances without meeting such criteria, and did not exclude studies of participants with various comorbidities. These and the methodological heterogeneities reported in this systematic review prevent us from performing a meta-analysis across studies, but future meta- and mega-analyses of clusters of studies in the database are possible and facilitated by our ongoing effort to catalog and share FDCR studies[19].

Contexts of use of FDCR biomarkers

Another principal consideration when developing an FDCR biomarker is its context of use. First, it should be clear for what SUD(s) the biomarker is developed. This choice hinges on considering both the burden of a disorder and the extent of the FDCR literature on that disorder. To provide two promising examples, nicotine and alcohol use disorders are both major contributors to morbidity and mortality worldwide[5,33] and have been extensively investigated with FDCR paradigms, comprising 29.6% and 29.2% of our database, respectively (eFigure 4). Then, the COU specification should clarify whether the FDCR-derived biomarker is to be used for diagnostic or prognostic purposes, to select or assess interventions, or as an intervention target (see Table 1 and eFigure 7 for the biomarker taxonomy and examples). This choice should guide the design and interpretation of the biomarker and ultimately its validation.

Studies with relevant evidence for developing “diagnostic” biomarkers constitute the largest category in our review with 143 examples, of which 93.7% have reported significant findings (Figure 3). These studies have mostly investigated differences in FDCR between individuals with SUDs and healthy controls, though some have assessed differences between clinically-relevant SUD subtypes. The diagnostic studies in our database have all essentially conducted statistical comparisons of the FDCR signal between participant groups defined a priori, though in principle, researchers could start from the other end, i.e., with data-driven identification of “neurotypes” using the fMRI data. While these provide insights into the neural correlates of SUDs, the diagnosis of SUDs currently relies on relatively inexpensive clinical interviews and drug tests and it is unlikely that FDCR-derived biomarkers would find clinical use in identifying SUDs. Another non-interventional context of use is susceptibility assessment, where there have been promising results for example in assessing adolescent susceptibility to SUDs based on FDCR in reward-related regions[6,34]. The other two, and likely most promising non-interventional COUs for FDCR biomarkers, constitute prognostic evaluation and monitoring of individuals diagnosed with SUDs: there is evidence that baseline nucleus accumbens drug cue-reactivity, for example, can statistically predict relapse better than conventional clinical measures[35]. These latter classes of FDCR biomarkers could add to the limited repertoire of tools available to meaningfully predict the course of SUDs and monitor their progression, but

their development requires expensive longitudinal studies. Only 21.2% of studies in our database include more than one timepoint (Figure 2).

Using FDCR biomarkers to develop, select, implement, or monitor the impact of interventions may be more cost-effective. There are 155 interventional studies in our database, most using FDCR in the context of pharmacological (43.2%; most commonly naltrexone in 10 studies) or cognitive/behavioral (33.5%) interventions. These studies form a sizable evidence-base to support the development of multiple types of interventional biomarkers for some SUDs, particularly alcohol and nicotine use disorders which constitute 34.4% and 32.4% of the 155 interventional studies in our database, respectively (Figure 4). Individuals with SUDs are highly heterogeneous in their responses to different treatments[36], Partly since different interventions target distinct mechanisms of disease which vary between individuals. "Predictive" FDCR biomarkers could reflect underlying neural pathology and may predict treatment response could guide treatment planning and reduce poor outcomes: For example, higher ventral striatal FDCR may predict greater efficacy of naltrexone than acamprosate for alcohol use disorder, possibly since ventral striatal FDCR may reflect reward-related craving and naltrexone has craving-suppressing effects[37]. Our review indicates that the "predictive" biomarker category is under-investigated, however, with only 25 relevant studies. Much more common are "response" biomarker studies, where post-intervention FDCR or intervention-induced changes in FDCR are thought to reflect an intervention's neurophysiological effect. There are 141 supporting pieces of evidence for response biomarker development across the 155 interventional studies in our review and growing evidence demonstrates the sensitivity of FDCR signals to detect intervention effects in the striatum[38,39], amygdala[40,41] , prefrontal cortical regions[42,43], insula[43] and cingulate cortices[44,45] - all regions widely implicated in SUDs. Given the importance of interventional FDCR studies, a more detailed breakdown of intervention types is presented in eFigure 8.

Finally, an FDCR biomarker could be validated as a "surrogate endpoint" if it can be shown that FDCR causally mediates the therapeutic impact of an intervention on clinical outcomes[10]. Particularly salient examples from drug development are the use of blood pressure reduction to assess the effectiveness of anti-hypertensive medication, or the reduction of hemoglobin A1C as a surrogate marker for the effectiveness of diabetes treatments[10]. Surrogate FDCR endpoints would accelerate drug development as a candidate therapeutic could be approved based on its immediate impact on the FDCR signal without the need to measure clinical outcomes over much longer time spans. Such FDCR markers may at least serve in the rapid screening of candidate therapeutics, for example in the context of aforementioned "FAST-FAIL" trials. Relatedly, FDCR markers that are linked to clinically relevant outcomes such as craving may provide direct and personalized targets for direct intervention. Ten studies in our database used neurofeedback where participants learned to directly reduce their cue-reactivity in regions where they showed high FDCR, such as the striatum[46] or highly reactive cortical areas[47]. Our review includes only twelve neuromodulation studies that used FDCR. However, none used FDCR for target selection directly, which is possible in principle since the modulation of FDCR signal by brain stimulation has been shown to predict craving reduction after stimulation[48]. Indeed, one retrospective analysis (published shortly after the period of coverage of this systematic review) suggests that TMS might be more clinically effective in treating alcohol use disorder if the TMS-induced electric field overlaps with an individual's endogenous alcohol cue-reactivity map[49].

Validation of FDCR biomarkers

Specified FDCR biomarkers need validation for regulatory approval[9,14]. "Clinical validation" requires demonstrating etiological links between the FDCR signal and an SUD. Our reviewed studies have investigated relationships between cue-exposure-associated neural activation patterns and other facets of SUDs, and this converging evidence helps buttress the clinical validity of FDCR by showing that it is linked to self-reported measures of craving (128 studies, see eFigure 5) and behaviors such as attentional bias and reward responsiveness[50,51], physiological responses such as increased skin conductance during drug cue exposure[52], and polymorphisms in genes related to glutamate, opioid, and dopamine signaling[17,53] thought to be involved in addiction. For example, neuro-genetic studies suggest that the A118G single nucleotide polymorphism of the mu opioid receptor (OPRM1) gene and the 9R allele of the dopamine transporter gene (DAT1) may result in higher levels of FDCR[54,55], and a large clinical experiment showed subsequently that both alleles interact to influence both FDCR and its reduction following naltrexone administration in alcohol-dependent individuals[56]. This body of literature can be leveraged, together with future FDCR investigations using robust longitudinal designs and extensive phenotypic and clinical profiling to establish the clinical validity of an FDCR biomarker.

Next, "analytical validation" requires establishing that an FDCR biomarker has appropriate accuracy and reliability within the proposed context of use[13]. While some recent evidence supports the reproducibility[57] and predictive accuracy[58] of certain FDCR patterns, many fMRI tasks suffer from low test-retest reliability[59,60] and recent findings point to a similar challenge for FDCR[61]. This highlights the need to systematically improve FDCR measurement and identify signal patterns optimal for biomarker development. Further, moving from group-level effects to biomarkers for individual-level decision-making requires the definition of normative signal ranges across contexts and groups: for example, some FDCR studies define "high FDCR" individuals as those whose FDCR value is greater than the median of study participants[62]. Such studies support further investigation to systematically establish a normative range to determine which individuals have abnormally high or low regional FDCR.

One way to establish normative FDCR bounds and design FDCR biomarkers with optimal analytic properties would be meta- and mega-analysis across previous studies, exemplified by a meta-analysis which demonstrated that short-duration cues in event-related designs may induce more reliable FDCR than longer cue presentations in blocked designs[63]. However, meta-analyses of previous studies should account for publication bias, flexible reporting and interpretation of results, and the fact that published findings may be the result of post-hoc, exploratory investigation. The very low rate of non-significant results in our database (Figure 3 and eFigure 9) is likely in part driven by these factors, which affect neuroscience research more broadly[64]. More insight into the analytic properties of various FDCR-derived measures would also enable appropriate task design: for example, without estimates of effect size and power analysis it's unclear whether the median FDCR task duration of 720 seconds in our database is sufficient given usual repetition times.

Finally, practical use of FDCR-derived biomarkers in clinical or drug development contexts requires that their cost-effectiveness be demonstrated. Given the costs of fMRI and potential harms of false negative or positive results, FDCR-derived biomarkers should be capable of feasibly and meaningfully complementing indicators that are often less expensive to measure, such as self-reported addiction severity or behavioral

phenotypes. This requires explicit cost-benefit modeling in future FDCR biomarker development studies and attempts to make FDCR more cost-effective by optimizing study designs for sample sizes, scanning procedures, and scan durations. It is also important to select biomarker types likely to offer the greatest utility. For example, diagnostic biomarker development may be foundational but unlikely to offer clinical utility outweighing the costs, and the gold standard of diagnosis will likely remain clinical interviewing. FDCR biomarkers may be much more cost-effective for prognosis, treatment selection, and intervention development, for which alternative markers are less available.

We discuss two particularly promising FDCR markers in Box 2, one reflecting global cue-related brain activity and the other local activation. Both examples demonstrate how validating evidence can converge across contexts of use.

Box 2. Local and global FDCR: Two exemplar cases

(*Refer to the bibliography in online-only materials for items cited in boxes)

We highlight two examples of promising FDCR signals across contexts of use. A robust FDCR biomarker would likely be useful across multiple contexts of use and would also be supported by converging avenues of validating evidence. A promising regional marker is striatal FDCR, which meets several important characteristics of a putative neural biomarker in alcohol use disorder (AUD). In a diagnostic context, several studies have reported significant differences in striatal FDCR between individuals with and without AUD[55,56*] and a ventral to dorsal striatum FDCR shift with more compulsive alcohol use[57]. There is support for the prognostic potential of striatal FDCR, with several studies demonstrating significant associations with subsequent alcohol use and relapse in AUD[58–61] and increases in relapse prediction accuracy of machine-learning models, over and above clinical variables[62]. In addition, converging evidence indicates that striatal FDCR is sensitive to behavioral AUD treatments such as cue-exposure therapy or drugs such as naltrexone [59,63]or nalmefene[64], illustrating that longitudinal assessment of striatal FDCR can monitor treatment effects. Further, acquiring striatal FDCR before treatment predicts naltrexone treatment response, such that individuals with high striatal FDCR benefited more from naltrexone[65], supporting the predictive potential of striatal FDCR. This finding was replicated in an independent sample[66] and could be expanded to positive (i.e., higher response to alcohol cues) versus negative (i.e., higher response to neutral cues) FDCR in striatal regions[59], indicating that absolute levels of striatal FDCR can be used to predict treatment efficacy across datasets.

With the advent of machine learning techniques capable of discovering robust patterns of activity distributed across the brain, it is possible to develop FDCR biomarkers that reflect neural processes involved in FDCR beyond a single region. This would be in line with the growing understanding that neural processes are often undergirded by distributed brain networks [67], and that multivariate brainwide association studies may require smaller samples to discover brain-behavior relationships[68]. There have been a few attempts to date to use FDCR to create and validate a whole-brain-based biomarker in SUDs[69]. In a recent example, machine learning on FDCR data from individuals with alcohol, cocaine, and tobacco use disorders identified a multivariate whole-brain marker that reliably associated drug craving, accurately classified individuals with SUDs from healthy controls, detected responses to interventions, and mediated the effects of intrinsic visual craving features on craving ratings [70]. While additional validation is required and ongoing as the authors note, current evidence supports the clinical and analytical validity of this multivariate marker as a diagnostic and response biomarker.

Conclusion

A growing number of biomarkers are widely used in biomedical research and clinical practice, but their role remains mostly limited in addiction medicine and psychiatry more broadly[65]. This paper provides an overview of fMRI drug cue-reactivity (FDCR) research, a promising paradigm for biomarker development for addictive disorders. FDCR biomarkers could classify patients, have prognostic value, improve treatment selection, and facilitate intervention development and personalized care. While the field faces numerous challenges — from methodological heterogeneity and small sample sizes to a lack of systematic biomarker development and validation efforts — under-utilized resources to overcome them exist. Ultimately, however, biomarker specification and validation efforts will likely require moving beyond traditional single-site studies and may involve mega-analyses using infrastructure developed by initiatives such as the Enhanced NeuroImaging Genetics through Meta-Analyses (ENIGMA) International Consortium[66] or multi-site collaborations and harmonized, longitudinal assessment following examples such as the Human Connectome Project and the Adolescent Brain Cognitive Development (ABCD) project[67,68], with expert consensus to address remaining gaps (see eFigure 11 for a summary of systematic review results and these future directions). Towards this aim, several authors of the present manuscript have formed the steering committee of the ENIGMA Addiction Cue-Reactivity Initiative (ACRI) within the ENIGMA Addiction working group to facilitate consensus development, methodological harmonization, and data sharing for mega-analyses[69]. Large-scale biomarker definition and validation studies would require substantial funding and resources often difficult to secure or justify for a single research institution or pharmaceutical company. This endeavor necessitates formation of diverse consortia to pool resources and guide validation efforts, develop best practices in study design and reporting, and engage in ongoing dialogue with commercial and public health stakeholders. Ultimately, there will be a need to form public/private partnerships that inform future biomarker development studies and systematically approach the arduous task of translating FDCR-derived biomarkers to clinical use.

Authors Information

Department of Psychiatry & Behavioral Sciences, University of Minnesota, Minneapolis, MN, USA (Zare-Bidoky, Ekhtiari, Zilverstand, Soleimani); Laureate Institute for Brain Research, Tulsa, OK, USA (Ekhtiari, Kuplicki, Paulus); Iranian National Center for Addiction Studies (INCAS), Tehran University of Medical Sciences, Iran (Zare-Bidoky, Fathi Jouzdani, Khojasteh Zonooni, Ebrahimi, Rafei); University of Melbourne, Australia (Sangchooli); Department of Psychiatry, University of Colorado School of Medicine, Aurora, CO, 80045, USA (Schacht); Department of Psychiatry, Institute for Drug and Alcohol Studies, Virginia Commonwealth University, Richmond, VA, USA (Bjork); Department of Biobehavioral Health, The Pennsylvania State University, University Park, PA, USA (Claus); McLean Hospital, Harvard Medical School, Belmont, MA, USA (Kaufman), Department of Psychiatry and Behavioral Sciences, Addiction Sciences Division, Medical University of South Carolina, Charleston, SC, USA (Prisciandaro, Brady, Mellick); Department of Psychology, The Pennsylvania State University, State College, Pennsylvania, USA (Wilson); Heidelberg University, Field of Focus IV, Core Facility for Neuroscience of Self-Regulation (CNSR), Heidelberg, Germany (Wüstenberg); Department of Psychiatry and Addiction, Université de Montréal, Montréal, Canada (Potvin); Department of Addictive Behaviour and Addiction Medicine, Central Institute of Mental Health (CIMH), Medical Faculty Mannheim, Heidelberg University, Germany (Bach, Kiefer, Vollstädt-Klein); Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada (Ahmadi); Faculty of Health, Health and Medical University, Potsdam, Germany (Beck); School of Medicine, University of St Andrews, St Andrews, Scotland, UK (Baldacchino); Department of Psychiatry and Neurosciences, Charité Campus Mitte, Charité-Universitätsmedizin Berlin, Berlin,

Germany (Beck, Heinz); Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, USA (Brewer); Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA (Childress, Wetherill); Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA (Courtney, Tapert); Center for BrainHealth, School of Behavioral and Brain Sciences, University of Texas at Dallas, Dallas, TX, USA (Filbey); Departments of Psychiatry, University of Vermont, Burlington, VT, USA (Garavan, Juliano, Mackey, Potenza); Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, USA (Ghahremani, Grodin, London, Ray); Departments of Psychiatry & Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, USA (Goldstein); Department of Psychiatry, Amsterdam University Medical Center, University of Amsterdam, the Netherlands and Amsterdam Neuroscience, Amsterdam, The Netherlands (Goudriaan); Department of Cancer Biology, Wake Forest School of Medicine, Winston-Salem, NC, USA (Hanlon); Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric University Hospital Zurich, University of Zurich, Zürich, Switzerland (Haugg); Center for Social and Affective Neuroscience, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden (Heilig, Perini); Department of Neurology, Albert Szent-Györgyi Health Centre, University of Szeged, Hungary (Holczer), Amsterdam Institute for Addiction Research, Amsterdam UMC, Department of Psychiatry, University of Amsterdam, The Netherlands (Holst); Department of Neuroscience, Medical University of South Carolina, Charleston, SC, USA (Joseph); Cognitive and Pharmacological Neuroimaging Unit, National Institute on Drug Abuse, Baltimore, MD 21224, United States (Janes); Department of Psychiatry, McGill University, Montreal, QC, Canada (Leyton); Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA (McClernon); Specialty of Addiction Medicine, Sydney Medical School, Faculty of Medicine and Health, University of Sydney, NSW, Australia (Morley); McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA (Noori); Neuroimaging and Analysis Group, Research Center for Molecular and Cellular Imaging, Tehran University of Medical Sciences, Tehran, Iran (Oghabian); TSET Health Promotion Research Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA (Oliver); Department of Psychiatry, University of Vermont, Burlington, VT, USA (Owens); Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA (Sinha, Steele, Kober); Department of Psychiatry, Technische Universität Dresden, Dresden, Germany (Smolka); Institute of Psychopharmacology, Central Institute of Mental Health, Mannheim, Germany (Spanagel); Department of Psychology, University of New Mexico, Albuquerque, NM, USA (Witkiewitz); School of Life Science and Technology, Xidian University, Xi'an, China (Yuan); Department of Psychology, School of Humanities and Social Science, University of Science and Technology of China, Anhui, China (Zhang); Turner Institute for Brain and Mental Health, Monash University, Clayton, Australia (Verdejo-Garcia);

Competing Interests

Dr Potenza has consulted for and advised Opiant Pharmaceuticals, Idorsia Pharmaceuticals, Baria-Tek, AXA, Game Day Data and the Addiction Policy Forum; has been involved in a patent application with Yale University and Novartis; has received research support from the Mohegan Sun Casino, Children and Screens, and Connecticut Council on Problem Gambling; has been a Board member for multiple organizations including the International Society of Addiction Medicine, Addiction Policy Forum, and National Council on Problem Gambling; has participated in surveys, mailings or telephone consultations related to drug addiction, impulse control disorders or other health topics; and has consulted for law offices and gambling entities on issues related to impulse control, internet use or addictive disorders. Dr Oliver is co-developer on a US patent for a device designed to predict behavioral risks from everyday environmental images. Dr Verdejo-Garcia has received funding from Elsevier for editorial work and from Servier for consultancy work. He is also on the Scientific Advisory Board of Brainwell although he does not receive honorarium. Other authors report no conflicts of interest.

Funding/Support

Dr Bach, Dr Beck, Dr Spanagel and Dr Heinz are supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant Project ID 402170461-TRR 265. Dr Bjork is supported by national institute of health (NIH) UG1DA050207. Dr Brewer is supported by National Center For Complementary

and Alternative Medicine (NCCIH) grants R01 AT007922, R61AT009337, UH2 AT008145, R34 AT008948, National Institute of Mental Health (NIMH) grant R41MH118130, national cancer institute (NCI) grant R21CA184254, national institute on drug abuse (NIDA) grants R34 DA037886, R03 DA029163, K12 DA00167, P50 DA09241, national institute on aging (NIA) grant (R21AG062004). Dr Claus is supported by NIH grant R01AA023665. Dr Courtney is supported by Tobacco Research Disease Related Program (TRDRP) grant T30IP0962 and NIH grant R21 AA030284. Dr. Filbey is supported by NIDA grant 1R01DA042490; Dr Goldstein is supported by NCCIH grant R01AT010627 and NIDA grants 1R01DA041528; 1R01DA047851; R01DA048301; 1R01DA049547; 1R21DA054281; and subcontract 271201800035C. Dr Grodin is supported by National Institute on Alcohol Abuse and Alcoholism (NIAAA) grant 1K01AA029712. Dr Haugg is supported by Forschungskredit of the University of Zurich (FK-18-030). Dr. Kaufman is supported by NIH grant R01 DA041866. Dr London is supported by The Thomas P and Katherine K Pike Chair in Addiction Studies and grants R37 DA044467, R01 DA045162, UG3 DA048388, and R03 DA052719. Dr Mackey is supported by NIH grant R01DA047119. Dr. Mellick is supported by NIAAA grant K23AA028535. Dr Morley is supported by National Health and Medical Research Council, Medical Research Future Fund. Dr Oliver is supported by NIH grants K23DA042898 and R01DA053342. Dr. Potenza is supported by NIH grants including R01 DK121551, R01AT01050, R01 AA029611 and RF1 MH128614. Dr. Prisciandaro is supported by NIH grants R01AA025365 and R01DA054275. Dr Schacht is supported by NIH grants R01 AA027765 and R01 AA026859. Dr Verdejo-Garcia is supported by National Health and Medical Research Council (NHMRC) Investigator Grant 2009464. Dr Vollstädt-Klein is supported DFG, German Research Foundation grants TRR 265 Project ID-402170461, Project ID-421888313, Project-ID 437718741 and GRK2350 Project-ID 324164820. Dr Wilson is supported by NIH grants R01DA055774, R21DA045853. Dr Leyton is supported by Canadian Institutes of Health Research grants 133537 and 152910 and Fonds de Recherche Québec-Santé grant 30614.Dr Potvin is supported by Canadian Institutes of Health Research (148561) and Fonds de Recherche du Québec-Santé grants (327571). Dr. Janes is supported by the National Institute on Drug Abuse -Intramural Research Program. Dr Zhang is supported by The Chinese National Programs for Brain Science and Brain-like Intelligence Technology (2021ZD0202101). Dr. Zilverstand is supported by the NIH grant R01 AA029406. Dr. Ekhtiari is supported by funds from Laureate Institute for Brain Research (LIBR), Tulsa, OK, and Medical Discovery Team on Addiction (MDTA), University of Minnesota, Minneapolis, MN and Brain and Behavior Foundation (NARSAD Young Investigator Award #27305).

Role of Funder/Sponsor Statement

Funders and Sponsors did not have any roles in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Access to Data and Data Analysis

Dr. Ekhtiari had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Data Sharing Statement

All the data used in this systematic review are publicly available on open science frame work (OSF) webpage of this project. (https://osf.io/eb972/)

416

417

418

Bibliography

1. Volkow ND, Koob G, Baler R. Biomarkers in substance use disorders. *ACS Chem Neurosci*. 2015;6(4):522-525. doi:10.1021/acschemneuro.5b00067

2. Koob GF, Volkow ND. Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry*. 2016;3(8):760-773. doi:10.1016/S2215-0366(16)00104-8

3. Carmichael O, Schwarz AJ, Chatham CH, et al. The role of fMRI in drug development. *Drug Discov Today*. 2018;23(2):333-348. doi:10.1016/j.drudis.2017.11.012

4. Ekhtiari H, Nasseri P, Yavari F, Mokri A, Monterosso J. Neuroscience of drug craving for addiction medicine: From circuits to therapies. *Prog Brain Res*. 2016;223:115-141. doi:10.1016/bs.pbr.2015.10.002

5. He H, Pan Z, Wu J, Hu C, Bai L, Lyu J. Health Effects of Tobacco at the Global, Regional, and National Levels: Results From the 2019 Global Burden of Disease Study. *Nicotine Tob Res*. 2022;24(6):864-870. doi:10.1093/ntr/ntab265

6. Heitzeg MM, Cope LM, Martz ME, Hardee JE. Neuroimaging Risk Markers for Substance Abuse: Recent Findings on Inhibitory Control and Reward System Functioning. *Curr Addict Rep*. 2015;2(2):91-103. doi:10.1007/s40429-015-0048-9

7. Zilverstand A, Huang AS, Alia-Klein N, Goldstein RZ. Neuroimaging Impaired Response Inhibition and Salience Attribution in Human Drug Addiction: A Systematic Review. *Neuron*. 2018;98(5):886-903. doi:10.1016/j.neuron.2018.03.048

8. Beck A, Wüstenberg T, Genauck A, et al. Effect of brain structure, brain function, and brain connectivity on relapse in alcohol-dependent patients. *Arch Gen Psychiatry*. 2012;69(8):842-852. doi:10.1001/archgenpsychiatry.2011.2026

9. Bakker E, Hendrikse NM, Ehmann F, et al. Biomarker Qualification at the European Medicines Agency: A Review of Biomarker Qualification Procedures From 2008 to 2020. *Clinical Pharmacology & Therapeutics*. 2022;112(1):69-80. doi:10.1002/cpt.2554

10. Gromova M, Vaggelas A, Dallmann G, Seimetz D. Biomarkers: Opportunities and Challenges for Drug Development in the Current Regulatory Landscape. *Biomarker Insights*. 2020;15. doi:10.1177/1177271920974652

11. Ekhtiari H, Zare-Bidoky M, Sangchooli A, et al. A methodological checklist for fMRI drug cue reactivity studies: development and expert consensus. *Nat Protoc*. Published online February 4, 2022:1-31. doi:10.1038/s41596-021-00649-4

12. Wilson SJ, Sayette MA. Neuroimaging craving: urge intensity matters. *Addiction*. 2015;110(2):195-203. doi:10.1111/add.12676

13. Abramson RG, Burton KR, Yu JPJ, et al. Methods and Challenges in Quantitative Imaging Biomarker Development. *Academic Radiology*. 2015;22(1):25-32. doi:10.1016/j.acra.2014.09.001

14. Kraus VB. Biomarkers as drug development tools: discovery, validation, qualification and use. *Nature Reviews Rheumatology*. 2018;14(6):354-362.

15. Hill-Bowen LD, Riedel MC, Poudel R, et al. *The Cue-Reactivity Paradigm: An Ensemble of Networks Driving Attention and Cognition When Viewing Drug-Related and Natural-Reward Stimuli*.; 2020:2020.02.26.966549. Accessed August 13, 2021. https://www.biorxiv.org/content/10.1101/2020.02.26.966549v1

16. Noori HR, Cosa Linan A, Spanagel R. Largely overlapping neuronal substrates of reactivity to drug, gambling, food and sexual cues: A comprehensive meta-analysis. *Eur Neuropsychopharmacol*. 2016;26(9):1419-1430. doi:10.1016/j.euroneuro.2016.06.013

17. Schacht JP, Anton RF, Myrick H. Functional neuroimaging studies of alcohol cue reactivity: A quantitative meta-analysis and systematic review. *Addict Biol*. 2013;18(1):121-133. doi:10.1111/j.1369-1600.2012.00464.x

18. Tang DW, Fellows LK, Small DM, Dagher A. Food and drug cues activate similar brain regions: a meta-analysis of functional MRI studies. *Physiol Behav*. 2012;106(3):317-324. doi:10.1016/j.physbeh.2012.03.009

19. Ekhtiari H, ACRI Secretariat. A Systematic Review on fMRI Drug Cue Reactivity Studies. Published online May 18, 2020. Accessed November 29, 2020. https://osf.io/eb972/

20. Yalachkov Y, Kaiser J, Görres A, Seehaus A, Naumer MJ. Sensory modality of smoking cues modulates neural cue reactivity. *Psychopharmacology (Berl)*. 2013;225(2):461-471. doi:10.1007/s00213-012-2830-x

21. Cortese BM, Uhde TW, Brady KT, et al. The fMRI BOLD response to unisensory and multisensory smoking cues in nicotine-dependent adults. *Psychiatry Res*. 2015;234(3):321-327. doi:10.1016/j.pscychresns.2015.10.008

22. Kroemer NB, Veldhuizen MG, Delvy R, Patel BP, O'Malley SS, Small DM. Sweet taste potentiates the reinforcing effects of e-cigarettes. *Eur Neuropsychopharmacol*. 2018;28(10):1089-1102. doi:10.1016/j.euroneuro.2018.07.102

23. Gilman JM, Radoman M, Schuster RM, et al. Anterior insula activation during inhibition to smoking cues is associated with ability to maintain tobacco abstinence. *Addictive Behaviors Reports*. 2018;7:40-46. doi:10.1016/j.abrep.2018.01.002

24. Jasinska AJ, Stein EA, Kaiser J, Naumer MJ, Yalachkov Y. Factors modulating neural reactivity to drug cues in addiction: a survey of human neuroimaging studies. *Neurosci Biobehav Rev*. 2014;38:1-16. doi:10.1016/j.neubiorev.2013.10.013

25. Goodyear K. Multisensory Environments to Measure Craving During Functional Magnetic Resonance Imaging. *Alcohol Alcohol*. 2019;54(3):193-195. doi:10.1093/alcalc/agz021

26. Murphy K, Bodurka J, Bandettini PA. How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage*. 2007;34(2):565-574. doi:10.1016/j.neuroimage.2006.09.032

27. Comes AL, Papiol S, Mueller T, Geyer PE, Mann M, Schulze TG. Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. *Transl Psychiatry*. 2018;8(1):1-15. doi:10.1038/s41398-018-0219-2

28. Carvalho AF, Solmi M, Sanches M, et al. Evidence-based umbrella review of 162 peripheral biomarkers for major mental disorders. *Transl Psychiatry*. 2020;10(1):1-13. doi:10.1038/s41398-020-0835-5

29. Nichols TE, Das S, Eickhoff SB, et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nature Neuroscience*. 2017;20(3):299-303. doi:10.1038/nn.4500

30. Wetherill RR, Young KA, Jagannathan K, et al. The impact of sex on brain responses to smoking cues: a perfusion fMRI study. *Biology of Sex Differences*. 2013;4(1):9. doi:10.1186/2042-6410-4-9

31. Kaag AM, Wiers RW, de Vries TJ, Pattij T, Goudriaan AE. Striatal alcohol cue-reactivity is stronger in male than female problem drinkers. *European Journal of Neuroscience*. 2019;50(3):2264-2273. doi:10.1111/ejn.13991

32. Turner BO, Paul EJ, Miller MB, Barbey AK. Small sample sizes reduce the replicability of task-based fMRI studies. *Communications Biology*. 2018;1(1):62. doi:10.1038/s42003-018-0073-z

33. Degenhardt L, Charlson F, Ferrari A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry*. 2018;5(12):987-1012. doi:10.1016/S2215-0366(18)30337-7

34. Tervo-Clemmens B, Quach A, Calabro FJ, Foran W, Luna B. Meta-analysis and review of functional neuroimaging differences underlying adolescent vulnerability to substance use. *NeuroImage*. 2020;209:116476. doi:10.1016/j.neuroimage.2019.116476

35. MacNiven KH, Jensen EL, Borg N, Padula CB, Humphreys K, Knutson B. Association of neural responses to drug cues with subsequent relapse to stimulant use. *JAMA network open*. 2018;1(8):e186466-e186466.

36. Carroll KM. The profound heterogeneity of substance use disorders: Implications for treatment development. *Curr Dir Psychol Sci*. 2021;30(4):358-364. doi:10.1177/09637214211026984

37. Mann K, Vollstädt-Klein S, Reinhard I, et al. Predicting naltrexone response in alcohol-dependent patients: the contribution of functional magnetic resonance imaging. *Alcohol Clin Exp Res*. 2014;38(11):2754-2762. doi:10.1111/acer.12546

38. Schacht JP, Randall PK, Latham PK, et al. Predictors of naltrexone response in a randomized trial: reward-related brain activation, OPRM1 genotype, and smoking status. *Neuropsychopharmacology*. 2017;42(13):2640-2653.

39. Bach P, Weil G, Pompili E, et al. Incubation of neural alcohol cue reactivity after withdrawal and its blockade by naltrexone. *Addiction Biology*. 2020;25(1):e12717.

40. Machielsen MW, Veltman DJ, van den Brink W, de Haan L. Comparing the effect of clozapine and risperidone on cue reactivity in male patients with schizophrenia and a cannabis use disorder: A randomized fMRI study. *Schizophrenia Research*. 2018;194:32-38.

41. Young KA, Franklin TR, Roberts DCS, et al. Nipping cue reactivity in the bud: baclofen prevents limbic activation elicited by subliminal drug cues. *J Neurosci*. 2014;34(14):5038-5043. doi:10.1523/JNEUROSCI.4977-13.2014

42. Franklin T, Wang Z, Suh JJ, et al. Effects of varenicline on smoking cue–triggered neural and craving responses. *Arch Gen Psychiatry*. 2011;68(5):516-526. doi:10.1001/archgenpsychiatry.2010.190

43. Li X, Chen L, Ma R, et al. The top-down regulation from the prefrontal cortex to insula via hypnotic aversion suggestions reduces smoking craving. *Hum Brain Mapp*. 2019;40(6):1718-1728. doi:10.1002/hbm.24483

44. Herremans SC, De Raedt R, Van Schuerbeek P, et al. Accelerated HF-rTMS protocol has a rate-dependent effect on dacc activation in alcohol-dependent patients: an open-label feasibility study. *Alcoholism: Clinical and Experimental Research*. 2016;40(1):196-205.

45. Mondino M, Luck D, Grot S, et al. Effects of repeated transcranial direct current stimulation on smoking, craving and brain reactivity to smoking cues. *Scientific reports*. 2018;8(1):1-11.

46. Kirsch M, Gruber I, Ruf M, Kiefer F, Kirsch P. Real-time functional magnetic resonance imaging neurofeedback can reduce striatal cue-reactivity to alcohol stimuli. *Addict Biol*. 2016;21(4):982-992. doi:10.1111/adb.12278

47. Hartwell KJ, Hanlon CA, Li X, et al. Individualized real-time fMRI neurofeedback to attenuate craving in nicotine-dependent smokers. *J Psychiatry Neurosci*. 2016;41(1):48-55. doi:10.1503/jpn.140200

48. Yang LZ, Shi B, Li H, et al. Electrical stimulation reduces smokers' craving by modulating the coupling between dorsal lateral prefrontal cortex and parahippocampal gyrus. *Social Cognitive and Affective Neuroscience*. 2017;12(8):1296-1302.

49. McCalley DM, Hanlon CA. The importance of overlap: A retrospective analysis of electrical field maps, alcohol cue-reactivity patterns, and treatment outcomes for alcohol use disorder. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*. 2023;16(3):724-726. doi:10.1016/j.brs.2023.04.015

50. Groefsema MM, Mies GW, Cousijn J, Engels RCME, Sescousse G, Luijten M. Brain responses and approach bias to social alcohol cues and their association with drinking in a social setting in young adult males. *Eur J Neurosci*. 2020;51(6):1491-1503. doi:10.1111/ejn.14574

51. Molokotos E, Peechatka AL, Wang KS, Pizzagalli DA, Janes AC. Caudate reactivity to smoking cues is associated with increased responding to monetary reward in nicotine-dependent individuals. *Drug Alcohol Depend*. 2020;209:107951. doi:10.1016/j.drugalcdep.2020.107951

52. Wang W, Zhornitsky S, Le TM, et al. Cue-elicited craving, thalamic activity, and physiological arousal in adult non-dependent drinkers. *Journal of Psychiatric Research*. 2019;116:74-82. doi:10.1016/j.jpsychires.2019.06.005

53. Bach P, Kirsch M, Hoffmann S, et al. The effects of single nucleotide polymorphisms in glutamatergic neurotransmission genes on neural response to alcohol cues and craving. *Addict Biol*. 2015;20(6):1022-1032. doi:10.1111/adb.12291

565 54. Ray LA, Courtney KE, Hutchison KE, Mackillop J, Galvan A, Ghahremani DG. Initial evidence that
566 OPRM1 genotype moderates ventral and dorsal striatum functional connectivity during alcohol cues.
567 *Alcohol Clin Exp Res*. 2014;38(1):78-89. doi:10.1111/acer.12136

568 55. Moeller SJ, Parvaz MA, Shumay E, et al. Gene x Abstinence Effects on Drug Cue Reactivity in
569 Addiction: Multimodal Evidence. *Journal of Neuroscience*. 2013;33(24):10027-10036.
570 doi:10.1523/JNEUROSCI.0695-13.2013

571 56. Schacht JP, Anton RF, Voronin KE, et al. Interacting effects of naltrexone and OPRM1 and DAT1
572 variation on the neural response to alcohol cues. *Neuropsychopharmacology*. 2013;38(3):414-422.
573 doi:10.1038/npp.2012.195

574 57. Ekhtiari H, Kuplicki R, Aupperle RL, Paulus MP. It is Never as Good the Second Time Around: Brain
575 Areas Involved in Salience Processing Habituate During Repeated Drug Cue Exposure in Treatment
576 Engaged Abstinent Methamphetamine and Opioid Users. *NeuroImage*. Published online May 19,
577 2021:118180. doi:10.1016/j.neuroimage.2021.118180

578 58. Seo S, Mohr J, Beck A, Wüstenberg T, Heinz A, Obermayer K. Predicting the future relapse of
579 alcohol-dependent patients from structural and functional brain images. *Addiction Biology*.
580 2015;20(6):1042-1055. doi:10.1111/adb.12302

581 59. Elliott ML, Knodt AR, Ireland D, et al. What Is the Test-Retest Reliability of Common Task-Functional
582 MRI Measures? New Empirical Evidence and a Meta-Analysis. *Psychol Sci*. 2020;31(7):792-806.
583 doi:10.1177/0956797620916786

584 60. Kennedy JT, Harms MP, Korucuoglu O, et al. Reliability and Stability Challenges in ABCD Task fMRI
585 Data. *Neuroimage*. 2022;252:119046. doi:10.1016/j.neuroimage.2022.119046

586 61. Bach P, Reinhard I, Koopmann A, et al. Test–retest reliability of neural alcohol cue-reactivity: Is
587 there light at the end of the magnetic resonance imaging tube? *Addiction Biology*. 2022;27(1):e13069.
588 doi:10.1111/adb.13069

589 62. Bach P, Weil G, Pompili E, et al. FMRI-based prediction of naltrexone response in alcohol use
590 disorder: a replication study. *Eur Arch Psychiatry Clin Neurosci*. 2021;271(5):915-927.
591 doi:10.1007/s00406-021-01259-7

592 63. Engelmann JM, Versace F, Robinson JD, et al. Neural substrates of smoking cue reactivity: A meta-
593 analysis of fMRI studies. *NeuroImage*. 2012;60(1):252-262. doi:10.1016/j.neuroimage.2011.12.024

594 64. Hong YW, Yoo Y, Han J, Wager TD, Woo CW. False-positive neuroimaging: Undisclosed flexibility in
595 testing spatial hypotheses allows presenting anything as a replicated finding. *NeuroImage*.
596 2019;195:384-395. doi:10.1016/j.neuroimage.2019.03.070

597 65. Lozupone M, La Montagna M, D'Urso F, et al. The role of biomarkers in psychiatry. *Reviews on*
598 *biomarker studies in psychiatric and neurodegenerative disorders*. Published online 2019:135-162.

599 66. Mackey S, Kan KJ, Chaarani B, et al. Chapter 10 - Genetic imaging consortium for addiction
600 medicine: From neuroimaging to genes. In: Ekhtiari H, Paulus MP, Ekhtiari H, Paulus MP, eds. *Progress in*
601 *Brain Research*. Vol 224. Neuroscience for Addiction Medicine: From Prevention to Rehabilitation -

Methods and Interventions. ; 2016:203-223. Accessed June 28, 2021. https://www.sciencedirect.com/science/article/pii/S0079612315001326

67. Garavan H, Bartsch H, Conway K, et al. Recruiting the ABCD sample: Design considerations and procedures. *Developmental cognitive neuroscience*. 2018;32:16-22.

68. Van Essen DC, Ugurbil K, Auerbach E, et al. The Human Connectome Project: A data acquisition perspective. *NeuroImage*. 2012;62(4):2222-2231. doi:10.1016/j.neuroimage.2012.02.018

69. ENIGMA Addiction | University of Vermont. Accessed September 5, 2023. https://www.enigmaaddictionconsortium.com/

70. Dager AD, Anderson BM, Rosen R, et al. Functional magnetic resonance imaging (fMRI) response to alcohol pictures predicts subsequent transition to heavy drinking in college students. *Addiction*. 2014;109(4):585-595.

71. Costumero V, Rosell-Negre P, Bustamante JC, et al. Left frontoparietal network activity is modulated by drug stimuli in cocaine addiction. *Brain Imaging Behav*. 2018;12(5):1259-1270. doi:10.1007/s11682-017-9799-3

72. Smith DG, Simon Jones P, Bullmore ET, Robbins TW, Ersche KD. Enhanced orbitofrontal cortex function and lack of attentional bias to cocaine cues in recreational stimulant users. *Biol Psychiatry*. 2014;75(2):124-131. doi:10.1016/j.biopsych.2013.05.019

73. Burnette EM, Grodin EN, Schacht JP, Ray LA. Clinical and Neural Correlates of Reward and Relief Drinking. *Alcohol Clin Exp Res*. 2021;45(1):194-203. doi:10.1111/acer.14495

74. Wetherill RR, Hager N, Jagannathan K, et al. Early Versus Late Onset of Cannabis Use: Differences in Striatal Response to Cannabis Cues. *Cannabis and Cannabinoid Research*. 2016;1(1):229-233. doi:10.1089/can.2016.0026

75. Shi Z, Jagannathan K, Padley JH, et al. The role of withdrawal in mesocorticolimbic drug cue reactivity in opioid use disorder. *Addict Biol*. 2021;26(4):e12977. doi:10.1111/adb.12977

76. Mendrek A, Dinh-Williams L, Bourque J, Potvin S. Sex differences and menstrual cycle phase-dependent modulation of craving for cigarette: an FMRI pilot study. *Psychiatry J*. 2014;2014:723632. doi:10.1155/2014/723632

77. Dong GH, Wang M, Zhang J, Du X, Potenza MN. Functional neural changes and altered cortical-subcortical connectivity associated with recovery from Internet gaming disorder. *J Behav Addict*. 2019;8(4):692-702. doi:10.1556/2006.8.2019.75

78. Goudriaan AE, Veltman DJ, van den Brink W, Dom G, Schmaal L. Neurophysiological effects of modafinil on cue-exposure in cocaine dependence: a randomized placebo-controlled cross-over study using pharmacological fMRI. *Addict Behav*. 2013;38(2):1509-1517. doi:10.1016/j.addbeh.2012.04.006

79. Nichols TT, Foulds J, Yingst JM, et al. Cue-reactivity in experienced electronic cigarette users: Novel stimulus videos and a pilot fMRI study. *Brain Res Bull*. 2016;123:23-32. doi:10.1016/j.brainresbull.2015.10.003

Tables and Figures

Table 1: Potential fMRI drug cue-reactivity (FDRC)-based biomarker domains, their definitions, and sample studies that provide supporting evidence for biomarker development. *Note that potential FDCR-derived safety biomarkers were very rare in the database, and thus have not been included as a separate category in other tables and figures. All the definitions for biomarkers have been directly adapted from the BEST Glossary, except for "Severity" biomarkers (defined based on previous biomarker literature as discussed).*

Biomarker Type	***Description***	***Examples of studies that can provide supporting evidence for biomarker development***
Susceptibility	Indicates the potential for developing a disease or medical condition in an individual who does not currently have the clinically apparent disease or the medical condition	Baseline cue-reactivity in the ventromedial prefrontal cortex, orbitofrontal cortex, anterior cingulate cortex, striatum, and insula was greater in individuals who subsequently transitioned from moderate to heavy drinking compared to people who did not transition[70]
Diagnostic	Detects or confirms the presence of a disease or condition of interest, or identifies individuals with a subtype of the disease	**SUDs vs. Healthy:** *Individuals with cocaine use disorder showed higher FDCR compared to controls in a frontoparietal network[71] *Individuals with cocaine use disorder compared to people with recreational stimulant use showed greater orbitofrontal and anterior cingulate FDCR during a cocaine-cue Stroop task[72] **SUD Subtyping:** *Among people with heavy alcohol consumption, "relief" drinking (due to negative reinforcement or habit) compared to reward drinking (due to positive reinforcement) was associated with greater dorsal striatal FDCR[73] *Individuals with cannabis use disorder and early-onset cannabis use showed FDCR in the dorsal striatum, while those with late-onset use showed FDCR in the ventral striatum[74]
Severity	Is correlated with greater intensity of the disease	In individuals with opioid use disorder, baseline FDCR in the nucleus accumbens, orbitofrontal cortex, and amygdala was associated with drug use severity (Addiction Severity Index Drug Composite Score), and withdrawal symptoms mediated the relationship between nucleus accumbens FDCR and drug use severity[75]
Prognostic	Identifies the likelihood of a clinical event, disease recurrence, or progression in patients who have the disease or medical condition of	Among individuals with stimulant use disorder, baseline FDCR in the nucleus accumbens was prospectively associated with time to relapse and could classify individuals into those who would relapse and those would not at 3 months after the scan, with an accuracy outperforming predictions using self-reported and clinical measures[35]

	interest	
Monitoring	Is measured repeatedly for assessing the status of a disease or medical condition or for evidence of exposure to (or effect of) a medical product or environmental agent	**Naturalistic Monitoring:** *Among women with tobacco use disorder, frontal, temporal, and parietal regions showed FDCR during the follicular phase of the menstrual cycle but not the luteal phase[76] *Among participants with internet gaming disorder (IGD) followed over one year, natural recovery from IGD was associated with decreased anterior cingulate and lentiform FDCR and an increase in cue-related effective anterior cingulate cortex-lentiform connectivity[77] **Treatment Monitoring:** In a randomized placebo-controlled trial of individuals with alcohol use disorder, naltrexone lowered ventral striatal FDCR from baseline, and more FDCR reduction was associated with greater clinical response[38]
Predictive	The existence or intensity of the biomarker reflects the propensity of individuals to experience favorable or unfavorable effects from exposure to a medical product or environmental agent	*In individuals with alcohol use disorder, the existence of left putamen FDCR at baseline and the reduction of left putamen FDCR early during treatment predicted the effectiveness of naltrexone[39] *In individuals with alcohol use disorder, high baseline FDCR in the ventral striatum statistically predicted response to naltrexone[37]. Notably, this finding has been directly replicated[62]
Response	Shows that a biological effect has occurred in an individual exposed to a medical product or environmental agent	**Biological Response:** In a randomized placebo-controlled trial of individuals with cocaine use disorder, modafinil acutely reduced FDCR in the ventral tegmental area and increased FDCR in the anterior cingulate and putamen, eliminating differences between participants with cocaine use disorder and healthy control participants[78] **Potential Surrogate Endpoint:** In a randomized sham-controlled trial involving people with tobacco use disorder, active versus sham tDCS over the dorsolateral prefrontal cortex (DLPFC) increased cue-related functional connectivity between the DLPFC and the parahippocampus, and this increase was correlated with decreased cigarette craving[48]
Safety	Is measured before or after an exposure to a medical product or an environmental agent to indicate the likelihood, presence, or extent of toxicity as an adverse effect	Contributing to discussions on the safety of electronic cigarettes, FDCR showed that e-cigarette smoking may immediately increase FDCR[79] Furthermore, sweet taste and nicotine content may synergistically influence the nucleus accumbens FDCR to the sight and smell of e-cigarettes[22]. Safety FDCR biomarkers may overlap with prognostic or response biomarkers in the context of SUDs since SUDs involve the use of substances whose safety may be assessed using FDCR

645

646 **Figure 1. Four major steps in the validation of potential fMRI drug cue-reactivity-derived biomarkers.**
647 Initially, a context of use for an FDCR-derived biomarker is specified and the potential biomarker is
648 precisely defined. Following analytical and clinical validation and cost-benefit analysis, the compiled
649 evidence is presented for regulatory approval. The FDA evaluates the use of biomarkers for drug
650 development through a biomarker qualification process involving submission of a Letter of Intent, a
651 Qualification Plan, and a Full Qualification Package, though a Letter of Support may be issued by the FDA to
652 indicate its support for a biomarker before formal qualification. The use of FDCR-derived biomarkers in
653 clinical contexts requires the endorsement of a constellation of other institutions. Surr. Endpoint:
654 Surrogate Endpoint.
655

Figure 2. Task and study design features of fMRI drug cue-reactivity studies. a. Number of time points in FDCR studies. Eighty-one studies scanned participants at two time points, six at three time points, and one with four time-points. b. Boxplot representing the distribution of median inter-scan intervals (in days) for FDCR studies with more than one scanning session. Ten studies scanned individuals more than once within the same day (interval = 0 days). c. Main FDCR task design type. d. Boxplot of the distribution of FDCR task durations. e. Paradigms combined with FDCR tasks in 52 studies in the database. f. FDCR studies, broken down by stimulus and substance/behavior type. “Multiple” stands for those studies including more than one type of addictive substance/behavior The "other" category includes inhalants and betel-quid.

Figure 3. Seven fMRI drug cue-reactivity study types. a. FDCR studies which, by virtue of their study design, could theoretically support the development of each biomarker type, broken down by substance or behavior of interest. Note that all cells do not sum to 415 since some studies do not fit the biomarker framework and some studies fit multiple biomarker types. b. The number of significant and non-significant supporting biomarker-related findings. The "other" category includes inhalants and betel-quid.

673

674 **Figure 4. fMRI drug cue-reactivity studies with an intervention or manipulation**. a. Types of interventional
675 FDCR studies each year, including randomized controlled trials (RCTs), controlled trials without
676 randomization, single-arm trials, and retrospective studies. b. Types of interventions in interventional FDCR
677 studies. c. Role of FDCR in interventional studies: FDCR can be measured before an intervention to predict
678 intervention results or measured after an intervention to assess impact with or without a comparison to
679 baseline FDCR.

680

681

682

Sources of Evidence

1 - Systematic Reviews
2 - Expert Elicitation/Consensus
3 - Meta/Mega-Analyses
4 - New/Prospective Research

Discovery & Definition

Biomarker Specification[1,2]

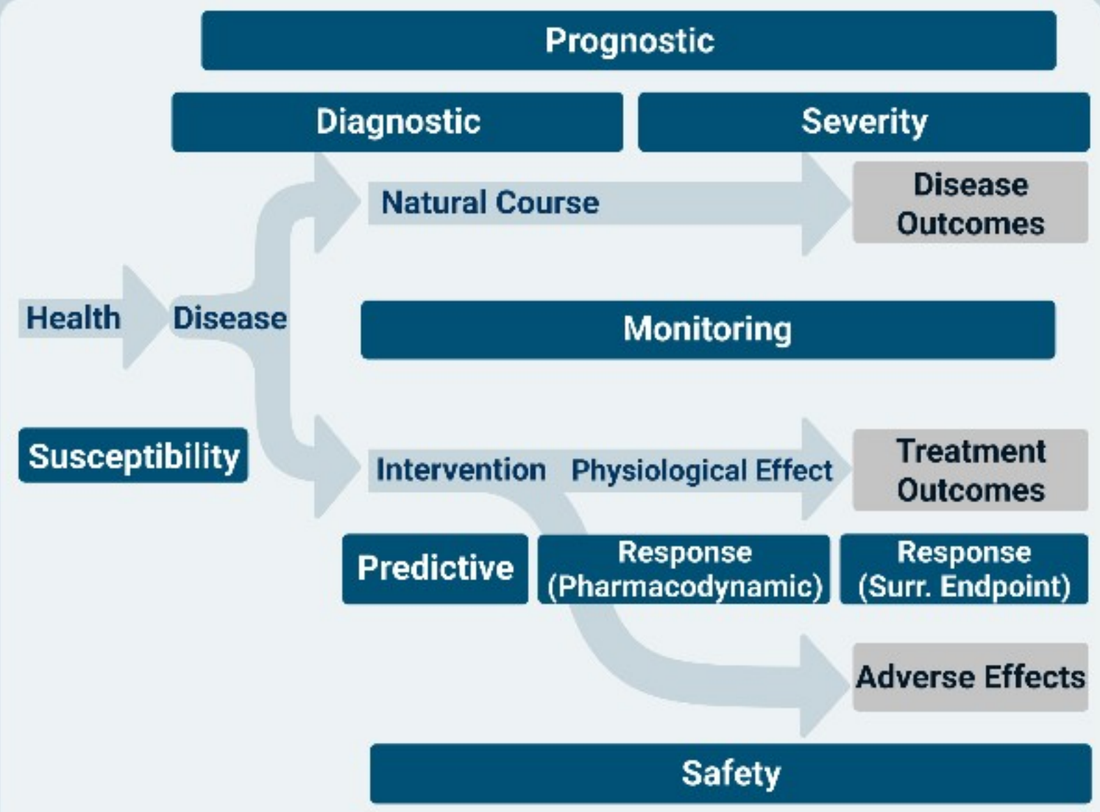
Target Population
- Demographic Criteria
- Drug Use Profile
- Clinical Characteristics
- MRI and Task-related Criteria

Standard Operating Procedure
- Abstinence status
- Fasting
- Familiarization procedures
- Time of Day

Image Acquisition
- MRI Device Properties
- Imaging Protocol
- Image Quality Control Criteria
- Task Parameters

Analysis & Utilization
- FCR Measure Specification
- Image Preprocessing
- fMRI Analysis Framework
- Clinical Interpretation

Context of Use Specification[1,2]



Validation

Analytical Validation[1,3,4]
- Accuracy/Precision
- Subject, Operator, and Scanner Reliability
- Natural FDCR Variation
- Instrument/Random Error Modelling

Clinical Validation
- Etiological Grounding of FDCR in Addiction[1]
- Retrospective/Prospective Study Design[3,4]
- Statistical Analysis Plan[2]

Cost-Benefit Analysis[1-3]
- Drug Development Acceleration
- Preclinical/Clinical Benefit
- False Positive/Negative Cost
- Scan Cost

Regulatory Qualification

Letter of Support

Drug Development Pathway

FDA Qualification Framework
- Letter of Intent (LOI)
- Qualification Plan (QP)
- Full Qualification Package (FQP)

Qualification Announcement

Clinical Use Pathway

NIA/NIMH/NIAAA
SAMHSA
APA/ASAM
- Institutional Review
- Endorsement, Recommendation and Inclusion in Guidelines

Practice Guidelines

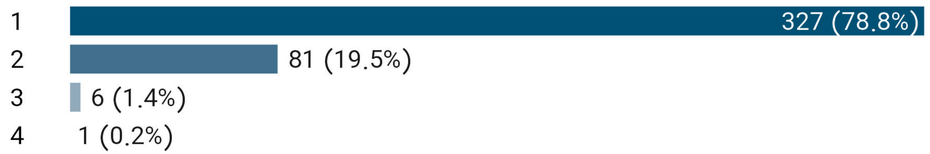
Application

FDCR as Drug Development Tool

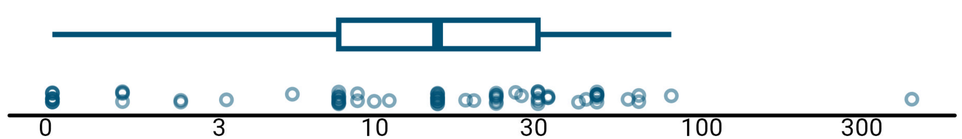
FDCR in Direct Clinical Use

Use and Ongoing Evaluation/Optimization

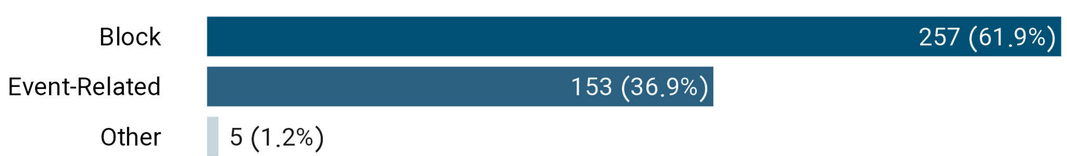
a. fMRI Time Points



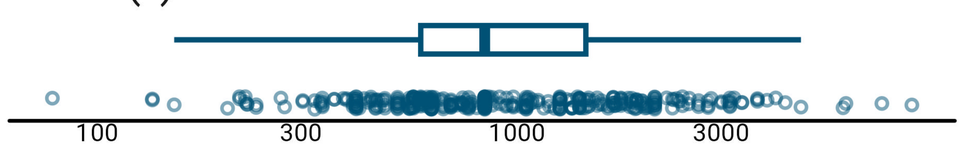
b. Interval (days) (n = 88)



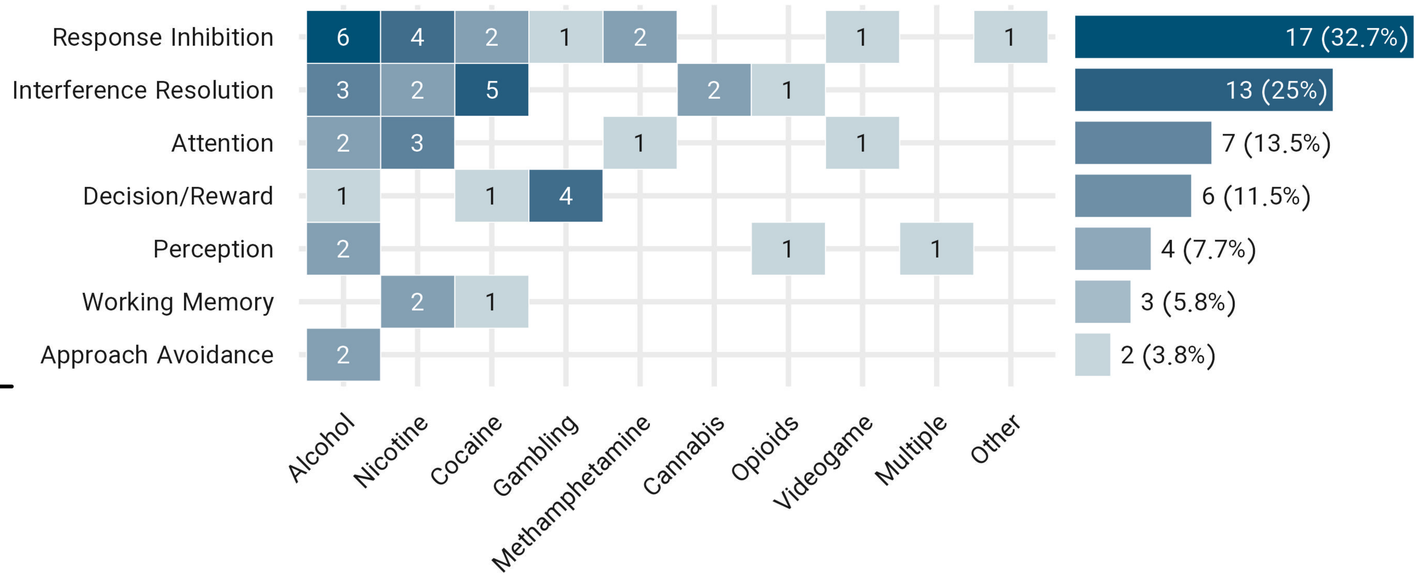
c. Design



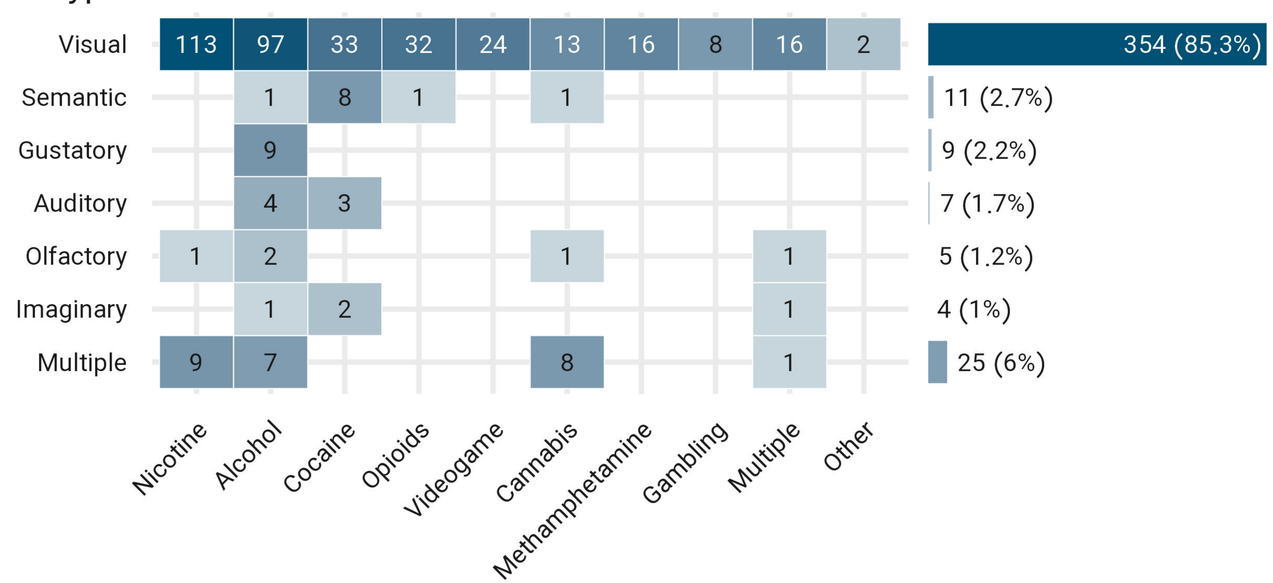
d. Task Duration (s)

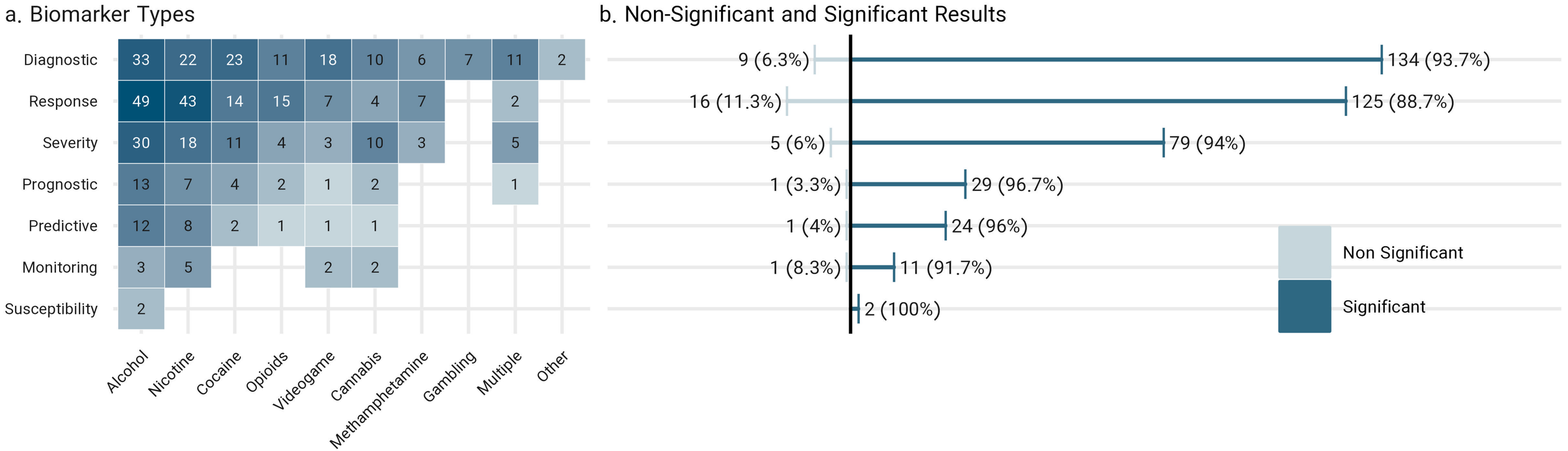


e. Combined Task (n = 52)

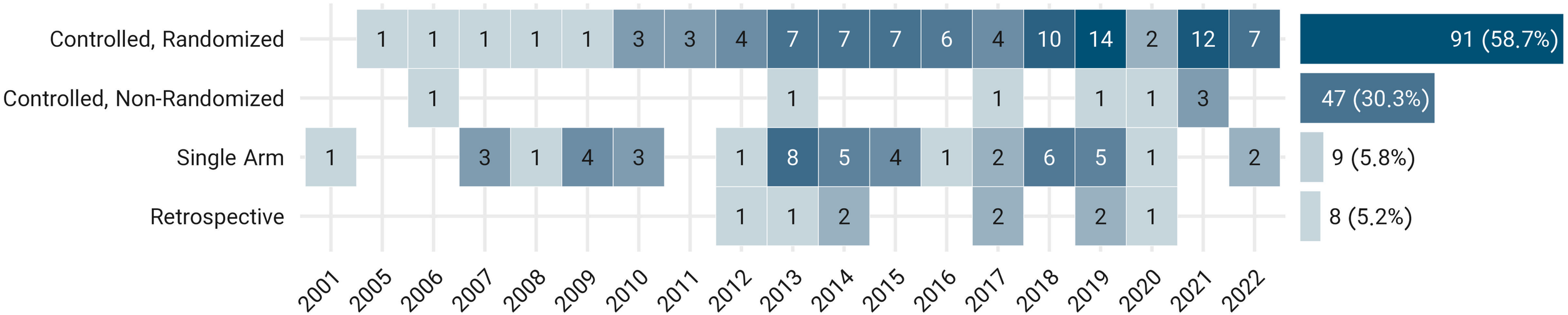


f. Stimulus Type

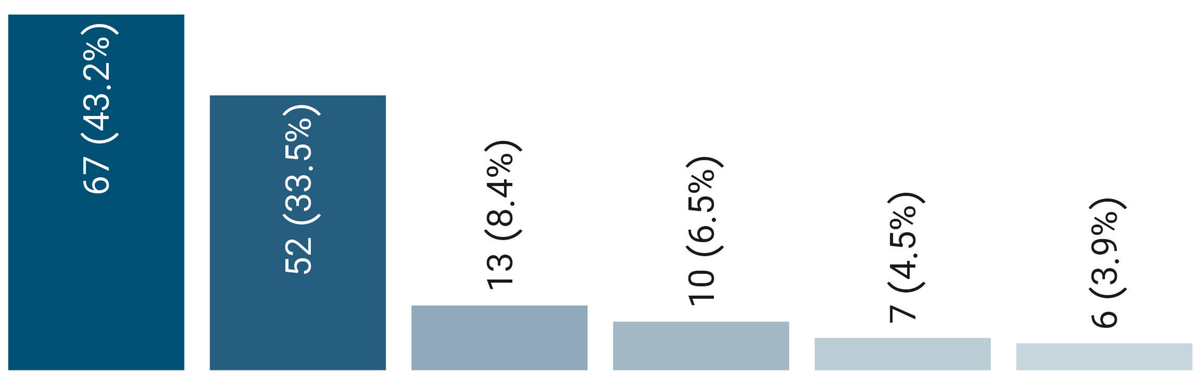



a. Biomarker Types
Diagnostic
Response
Severity
Prognostic
Predictive
Monitoring
Susceptibility
33 22 23 11 18 10 6 7 11 2
49 43 14 15 7 4 7 2
30 18 11 4 3 10 3 5
13 7 4 2 1 2 1
12 8 2 1 1 1
3 5 2 2
2
Alcohol
Nicotine
Cocaine
Opioids
Videogame
Cannabis
Methamphetamine
Gambling
Multiple
Other
b. Non-Significant and Significant Results
9 (6.3%)
134 (93.7%)
16 (11.3%)
125 (88.7%)
5 (6%)
79 (94%)
1 (3.3%)
29 (96.7%)
1 (4%)
24 (96%)
1 (8.3%)
11 (91.7%)
2 (100%)
Non Significant
Significant

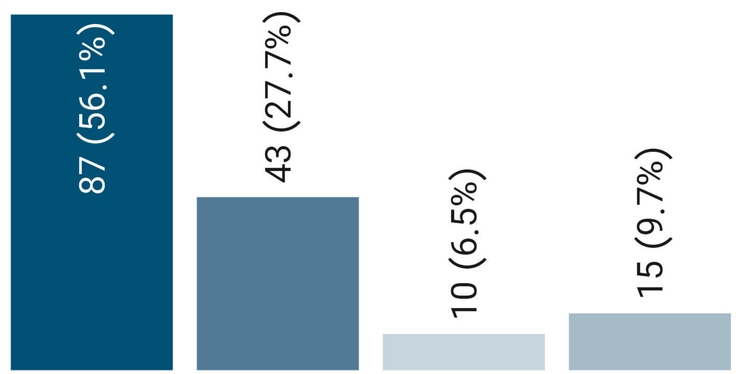
a. Interventional Study Types

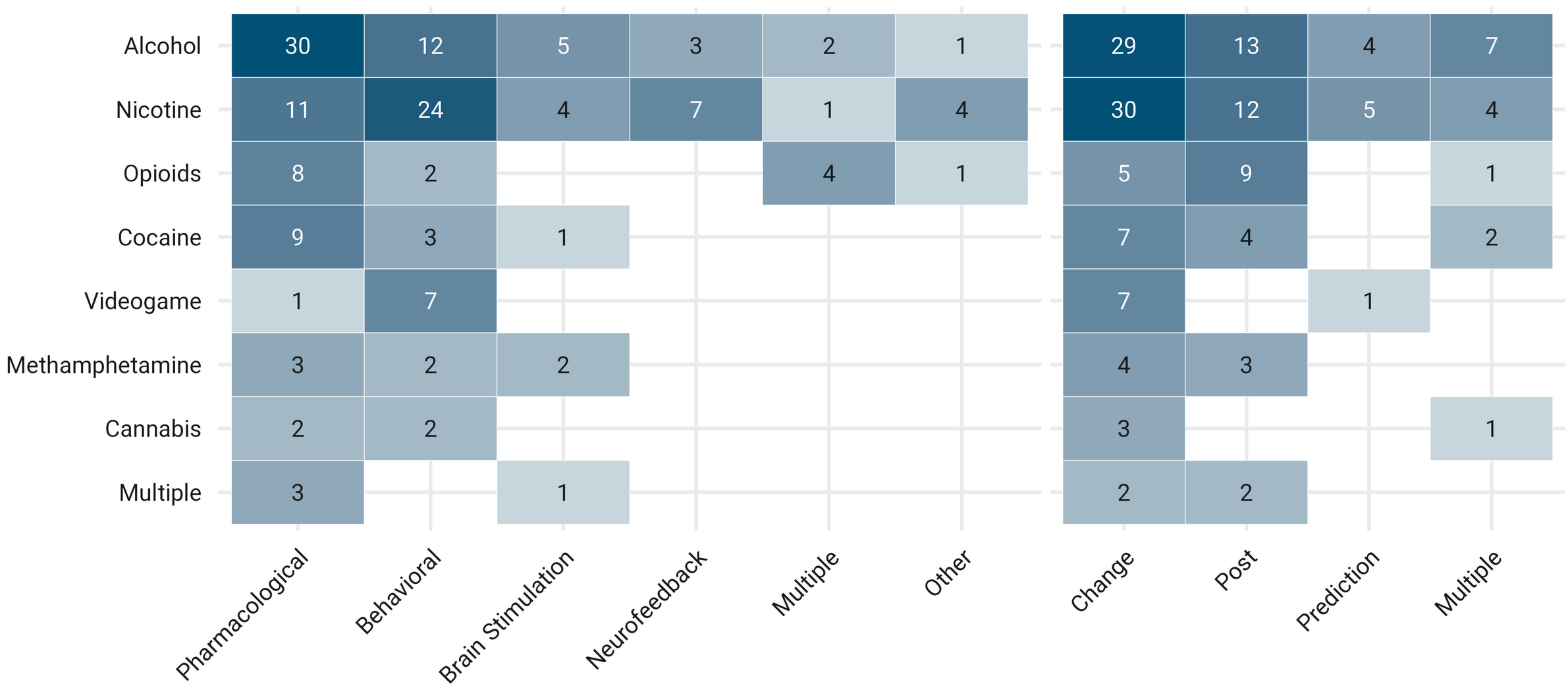


b. Intervention Type (n = 155)



c. fMRI Role (n = 155)

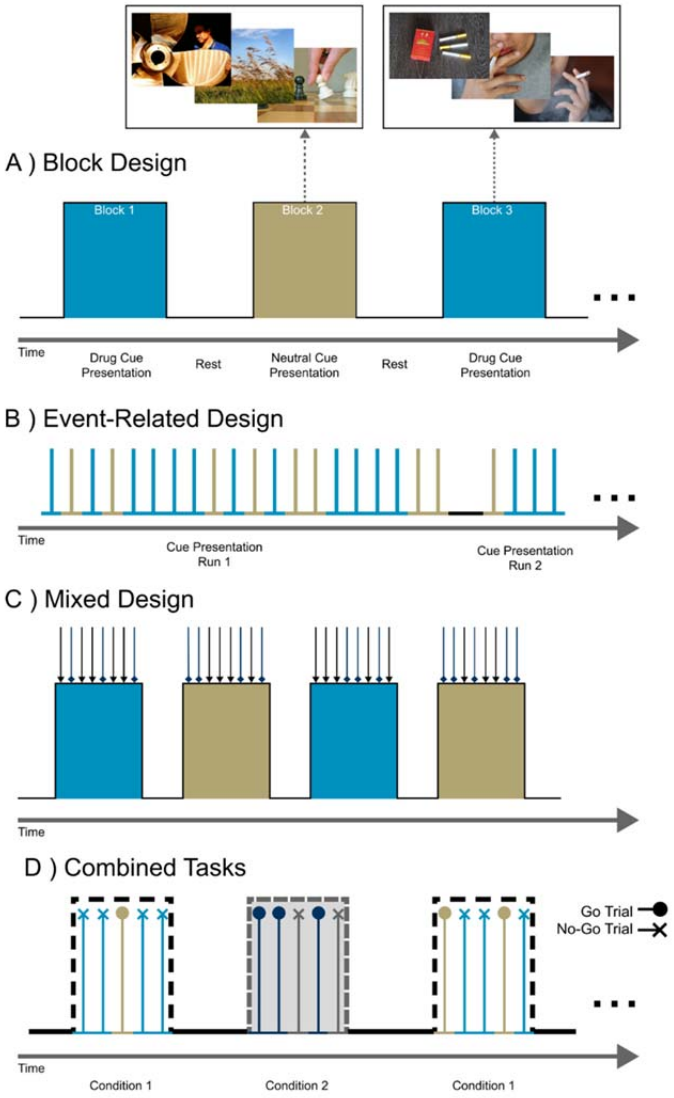




Online-Only Materials:

2 Tables, 11 Figures

Contents



eFigure 1. Overview of fMRI Drug Cue-Reactivity Task Designs. Overview of fMRI Drug Cue Reactivity Task Designs A) Cues/Stimuli are presented in groups or "blocks" containing a series of similarly conditioned cues which are then separated by a delay from the next block. B) Stimuli are presented in succession with or without a delay, without being arranged by their type. Tasks may have a few sections or "runs" where a delay separates each run from the next without the participant exiting the scanner. C) Mixed design tasks may borrow elements (like grouping or sequence randomization) from either a block design or an event-related design with the addition of another set of changing conditions or events that occur concurrently with the task D) Combined tasks use cue-reactivity concurrently with another cognitive task (e.g.,

Go/No-Go task). Designs can incorporate stimuli presented in various modalities, including visual (static or dynamic), auditory, olfactory, or tactile.

Methods

The methods section is organized based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. The protocol for this systematic review was pre-registered[1]. While we refer to fMRI "drug" cue-reactivity (including alcohol) throughout the manuscript, behavioral addiction studies focusing on problematic videogame playing or gambling were not excluded as they constitute a small portion of the cue-reactivity literature and involve cue-reactivity paradigms similar to drug cue-reactivity studies. Since behavioral addictions (BAs) have recently been added to the widely used nomenclature system, and the pathophysiology may not be completely the same as SUDs, one should be cautious in analyzing these data together. Therefore, information on BAs can be seen separately in our database for future use.

Eligibility criteria: Original studies were selected according to the criteria outlined below.

Study design and methodology: We employed a broad perspective in the inclusion of studies, including all types of original research (e.g., basic research, observational studies, and interventional studies). Only peer-reviewed studies were included.

Of interest were original studies that included one or more fMRI-based investigations as a major part of their methodology, whether as an observational or as a treatment tool (e.g. in fMRI neurofeedback). For at least some of the study population, the fMRI investigation had to include a cue-reactivity task, including the presentation of substance- or problematic behavior-related cues and at least one other class of cues (i.e., neutral or non-substance-related) for comparison. Acute challenge studies involving direct administration of substances of use were not included unless cues associated with the substance/object of use were explicitly presented as well. Cues also had to be ecologically valid; i.e., they needed to be associated with routine drug-taking behaviors and not be novel conditioned cues associated with the substance/object of use for the first time during the experiment. We excluded studies that did not provide details about the fMRI protocol, setting and tasks, outcomes of interest used in the analysis, and basic fMRI measures. There were no further exclusions, and both whole-brain and region of interest (ROI)-based fMRI studies were included.

Participants: Every study required at least one human population or sub-population with more than one member, for which at least one of the following needed to be true:

At least one circumscribed group of participants had a diagnosis of at least one SUD or BA, either manifest as active use or in remission; with the diagnosis made either before the study, as part of the study protocol during the investigation, or by the end of the study (i.e., with the diagnosis serving as an outcome measure).

At least one group of participants was included explicitly because they regularly consumed a potential object of addiction (substance or behavior) and/or had a risky pattern of consumption that might lead to addiction, and the study focused on their reactivity to cues of that substance or behavior.

At least one group of participants had been assigned a score for an addiction-related phenomenon (such as addiction or drug-use severity) with or without an explicit diagnosis of an SUD or BA, and the relationship of this score to important outcomes in the study had been investigated.

No restriction was placed on study participants based on demographic, ethnic, biological, or clinical factors (such as any co-occurring disorders).

Language: Only publications with their full text in English were included.

Information source: Existing research was identified and retrieved using PubMed. Relevant articles were identified using a comprehensive search strategy for all terms related to addiction, fMRI, and cue-reactivity, as detailed below.

Search strategy: Considering the subject of the review, a list of three sets of keywords was compiled (eTable 1). These terms were adapted for use in PubMed (exact search syntax and search results are outlined in eTable 2). The first set included synonyms of “functional magnetic resonance imaging”, the second included terms related to cue-reactivity, and the third included synonyms of “addiction” and various terms related to SUDs and BAs and addiction medicine. To help widen the search, no filters were used. The exclusion of systematic reviews and other non-original research and the application of other inclusion/exclusion criteria were handled manually. Given the large volume of relevant literature on PubMed, other search engines or grey literature were not used.

eTable 1. Search terms used for this systematic review

fMRI	"functional MRI"
"functional magnetic resonance"	"cue-reactivity"
"cue exposure"	"craving"
"cue induced"	"drug cue"
"drug cues"	
addict*	dependence
"substance use"	"substance abuse"
"drug abuse"	"drug use"
nicotine	smoker
tobacco	opioid
opiate	heroin
marijuana	cannabis
"thc"	alcohol*
cocaine	amphetamine
methamphetamine	"behavioral addiction"
"internet addiction"	"problematic gaming"
"gaming disorder"	"gambling disorder
"problem gambling"	
fMRI search terms	1 OR 2 OR 3
Cue-reactivity search terms	4 OR 5 OR 6 OR 7 OR 8 OR 9
Addiction search terms	10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR

	31 OR 32 OR 33 OR 34
Final search	35 AND 36 AND 37

114

eTable 2. Final syntax of PubMed search, and number of raw search results

Term Group	Search	Number of results on 5 Jan 2023
fMRI search	fMRI OR "functional MRI" OR "functional magnetic resonance"	573243
Cue-reactivity search	"cue reactivity" OR "cue exposure" OR craving OR "cue induced" OR "drug cue"	9841
Addiction search	addict* OR dependence OR "substance use" OR "substance abuse" OR "drug abuse" OR "drug use" OR nicotine OR smoker OR tobacco OR opioid OR opiate OR heroin OR marijuana OR cannabis OR "THC" OR alcohol* OR cocaine OR amphetamine OR methamphetamine OR "behavioral addiction" OR "behavioral addiction" OR "internet addiction" OR "problematic gaming" OR "gaming disorder" OR "gambling disorder" OR "problem gambling"	1424082
Final search	(fMRI OR "functional MRI" OR "functional magnetic resonance") AND ("cue reactivity" OR "cue exposure" OR craving OR "cue induced" OR "drug cue" OR "drug cues") AND (addict* OR dependence OR "substance use" OR "substance abuse" OR "drug abuse" OR "drug use" OR nicotine OR smoker OR tobacco OR opioid OR opiate OR heroin OR marijuana OR cannabis OR "THC" OR alcohol* OR cocaine OR amphetamine OR methamphetamine OR "behavioral addiction" OR "behavioral addiction" OR "internet addiction" OR "problematic gaming" OR "gaming disorder" OR "gambling disorder" OR "problem gambling")	952

Study records

Data management: Literature search results were imported to Excel. Screening of articles for relevance was performed by reviewing the title and abstract sections of candidate texts, and full texts were obtained for studies that passed preliminary screening.

Study selection: Screening forms were developed for title/abstract and full-text assessment and studies were checked by two authors (MZB and AS). The authors initially checked the eligibility of fifty randomly chosen studies under the supervision of the corresponding author (HE) as a calibration exercise to ensure eligibility criteria were applied consistently [2]. After title and abstract screening, the two authors screened the full texts of papers that either met the eligibility criteria or had an uncertain status. Any papers with an uncertain eligibility status after full-text screening were then discussed with HE until a consensus on their inclusion was reached. Reasons for the exclusion of articles at the title and abstract or full-text screening stages were recorded, according to the PRISMA framework[3]. Neither of the review authors was blind to the journal titles, study authors, or institutions.

Data collection: Data were filled into a spreadsheet by PA, AFJ, AH, and AKZ. Consistency between the authors was honed through a calibration exercise in which all authors evaluated and discussed their ratings for 20 randomly chosen studies [2]. AS, MZB and HE further refined the data extraction form to reduce inconsistency and ambiguity after the exercise. Data on study design features and basic methodological parameters were extracted first, and each article was reviewed independently by two authors in two separate spreadsheets, with inconsistencies resolved in discussions with MZB and AS with HE's supervision. To check whether any study samples overlapped with other studies (e.g. in the case of re-analysis studies), a single rater (AFJ) screened the methods sections of all studies.

Data items: We extracted publication details, publication country (where the first affiliation of the first study author is located or the affiliation of the majority of the authors in case country was not clear), publication year (based on PubMed's indexing), the substance or behavior (main substance(s) and/or behavior(s) of interest in the study), main experimental task design type (whether cues were presented in blocked, event-related, or mixed forms), stimulus type (sensory modality of cues), combined tasks (whether cue-reactivity was paired with other tasks; and what tasks were used), task duration (seconds, excluding other paradigms that may have been implemented in the scanner), study sample characteristics (number of participants of each sex; number of participants with untreated or treated addictive disorders, drug-using individuals who did not meet SUD criteria, individuals in long-term abstinence, and healthy non-using participants), intervention (if included, type of intervention), association with a future event (a non-fMRI variable measured at a later point in time based on fMRI results), number of fMRI sessions (times each participant was scanned), and interval between fMRI sessions (if participants were scanned more than once for a study, the average time interval between the scans). Yes/No ratings were used to classify whether the design of each FDCR study allowed for it to be potentially used to develop susceptibility, diagnostic, response, prognostic, predictive, or severity biomarkers for one or more SUDs/BAs. Yes/No ratings were also used to specify whether a study investigated relationships between FDCR-derived parameters and subjective

159 craving, demographic variables, behavioral measures, biochemical assays, participant genetics,
160 non-FDCR structural or functional neural markers, physiological parameters, or psychiatric
161 assessments. For each study investigating use of FDCR as a biomarker type or assessing FDCR
162 correlates, it was also rated as to whether significant test results were observed. However, we
163 elected to use the relatively simple metric of "significance" given the extreme heterogeneity of
164 analyses and reported statistics in the field, which would complicate further quantitative
165 synthesis. The scope of this work is to provide an overview of the status of the field and address
166 the current heterogeneities to provide a roadmap to support the development of evidence that
167 can be used in higher quality quantitative metrics in the future.

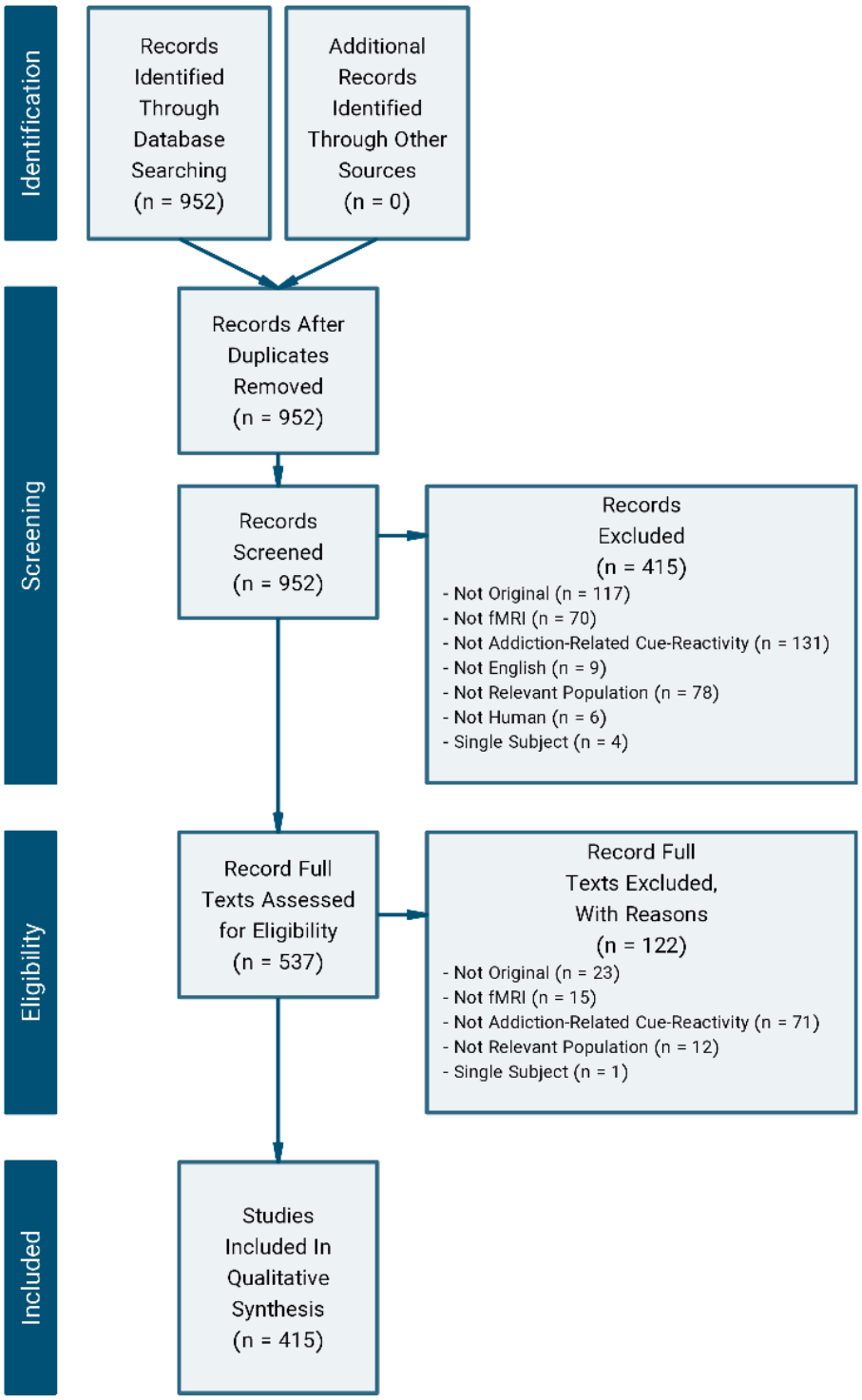
168 **Software**

169 The PubMed search engine from the National Library of Medicine's online portal
170 (www.ncbi.nlm.nih.gov/pubmed/) was used to conduct the search. Endnote X9[4] was used for
171 reference management. Google Sheets from Google's Google Docs Editors suite was used to
172 design tables for data extraction and sharing among authors. Data analyses and illustrations
173 were conducted using R version 4.0.5[5].

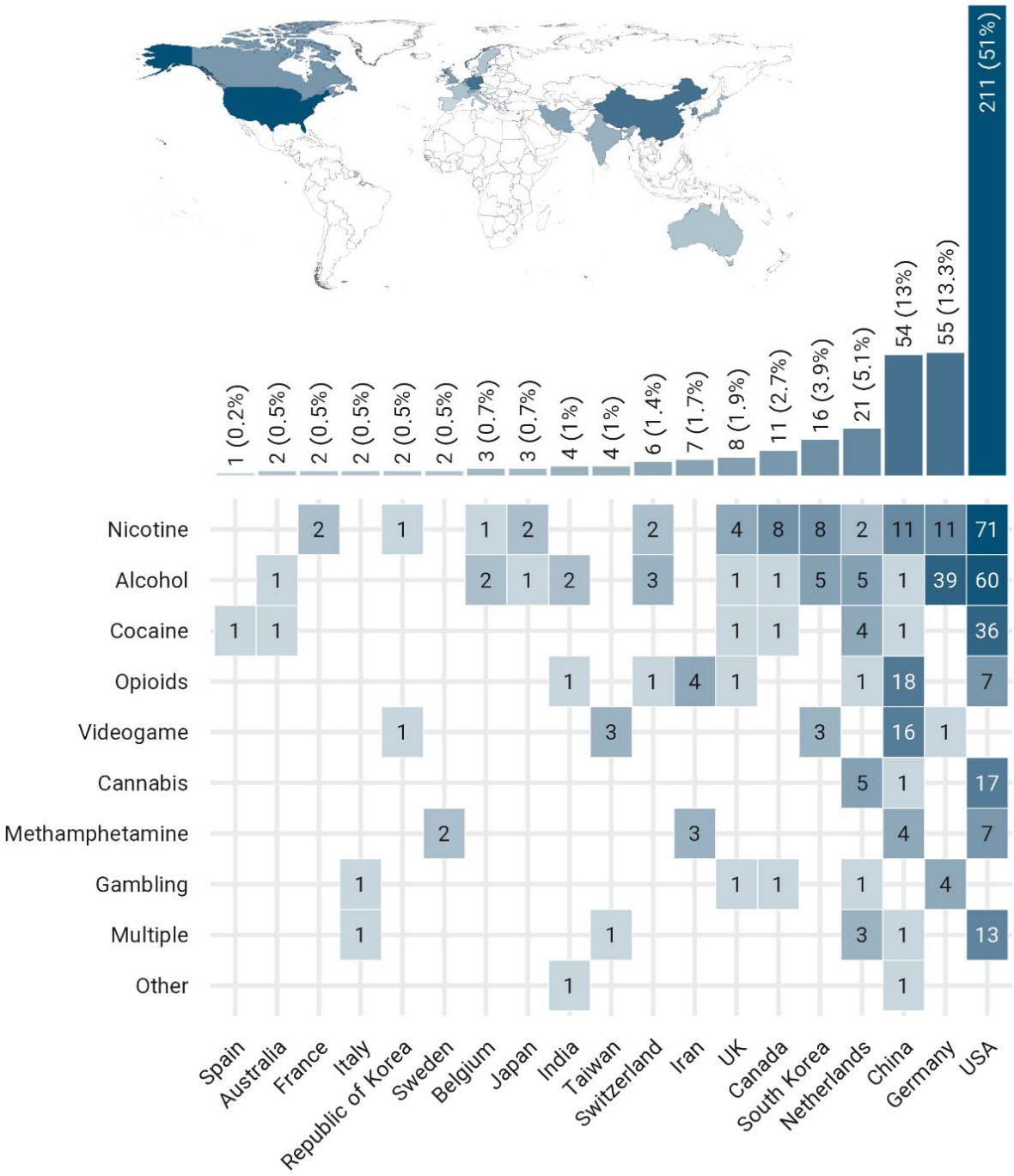
174 The protocol for this systematic review was developed throughout 2019 and was first
175 registered on the Open Science Framework (OSF) website on May 18th, 2020. The current
176 extracted database is available publicly in the OSF page (https://osf.io/eb972/). As this is an
177 ongoing systematic review, we recommend viewing the OSF page of this project for the latest
178 developments and updates[1].

Results

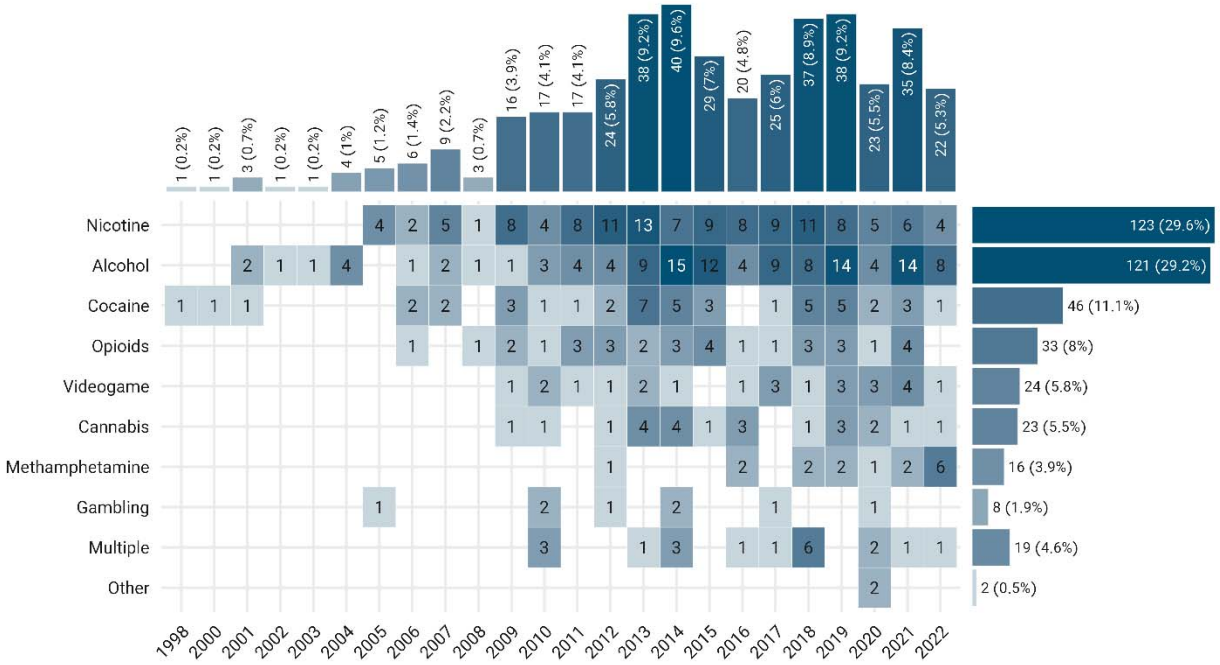
The search was performed on January 5, 2023, yielding 952 results. Of these, 415 were excluded at the title-and-abstract screening stage, and 122 were excluded after full-text screening, yielding a total of 415 FDCR publications that were included in the data extraction phase of the systematic review. The PRISMA flowchart is presented in eFigure 2. Most studies are from the US (51.0%) followed by Germany (13.3%) and China (13.0%) (eFigure 3). A breakdown of papers by the substance or behavior of interest shows that most studies have been conducted on various forms of either nicotine (29.6%), alcohol (29.2%), or cocaine (11.1%) use/use disorders, overall accounting for 69.9% of the papers in the database. Earlier studies in the database were all focused on cocaine and alcohol, with the first studies on cannabis and video games published in 2009 and the first on methamphetamine published in 2012 (eFigure 4). There is an overall yearly increase in the number of FDCR studies, with the vast majority of studies (303, 74.0 %) published in the last 10 years.



eFigure 2. PRISMA Flowchart. The titles and abstracts of 952 records from the start of 1998 until the end of 2022 were screened, and 415 were excluded during preliminary screening. The full texts of 537 records were extracted and assessed for eligibility. Ultimately, 415 records were included in the systematic review.



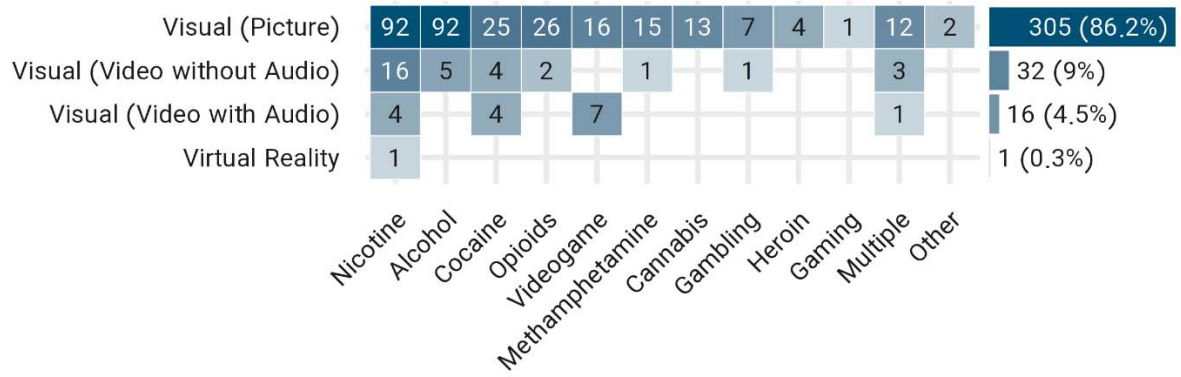
eFigure 3. Global contribution to fMRI drug cue-reactivity (FDCR) studies. Number of FDCR studies in each country, broken down by the type of addictive substance/behavior. “Multiple” stands for studies including more than one type of addictive substance/behavior. The "other" category includes inhalants and betel-quid. Note that only papers whose full-text was in English were included, potentially leading to a relative over-representation of majority English-speaking countries.



eFigure 4. fMRI drug cue-reactivity studies (1998-2022). Number of FDCR studies each year from 1998 till the end of 2022, broken down by the type of addictive substance/behavior. “Multiple” stands for those studies including more than one type of addictive substance/behavior. The "other" category includes inhalants and betel-quid.

Study and task design

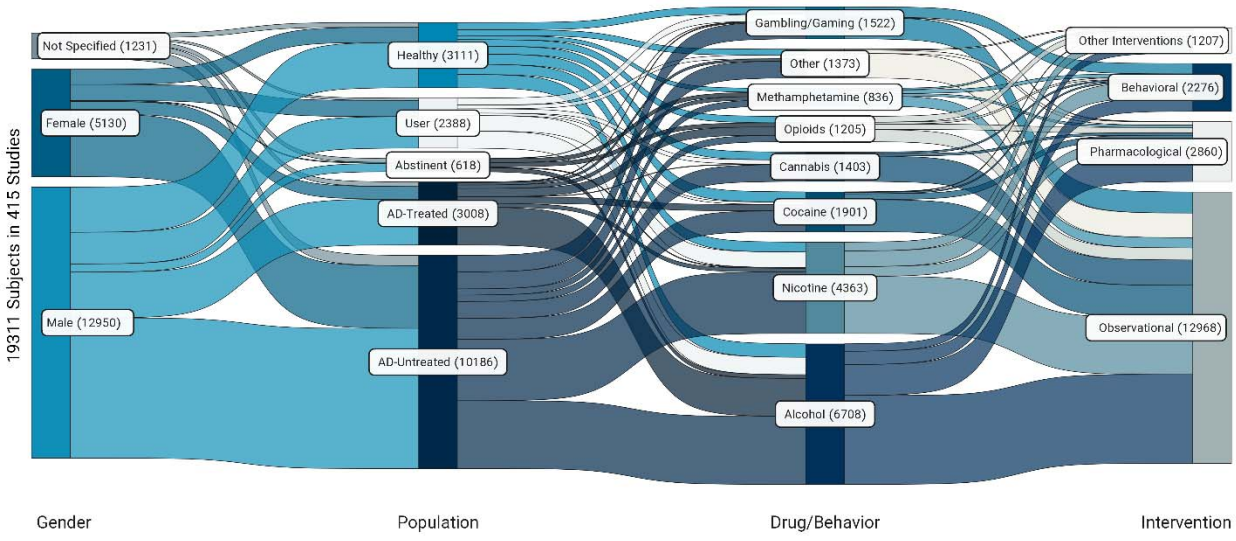
Most FDCR studies scanned participants at a single time point (78.8%). For the 88 studies with more than one scanning time point, the median inter-scan interval was 14 days, though a relatively wide distribution was observed (IQR = 21) (Figure 2b). The vast majority of studies (85.3%) used visual stimuli (for a detailed breakdown, see eFigure 5), with a minority using other stimulus types such as semantic (2.7%), gustatory (2.2%), auditory (1.7%), olfactory (1.2%), and imaginary (1%) stimuli. Another 25 multi-sensory studies (6%) used various combinations of stimuli (Figure 2f). Cues have been commonly presented in a block (61.9%) or event-related (36.9%) design, with only 1.2% of studies using other designs or both event-related and blocked-design FDCR tasks within a single study (Figure 2c). The median FDCR task duration was 720 seconds (IQR = 800) (Figure 2d), and 52 FDCR studies used combined FDCR tasks: these are tasks in which the presentation of addiction-relevant cues is paired with another concurrent task component to probe cognitive functions such as response inhibition (32.7% of the 52 studies), interference resolution (25.0%), attention (13.5%), decision-making and reward processing (11.5%), perception (7.7%), working memory (5.8%), or approach/avoidance (3.8%) (Figure 2f).



eFigure 5. Breakdown of visual cues. Among 354 sets of visual cues used in FDCR studies, they are broken down into pictures and videos with audio and without audio.

Participants in FDCR studies

Overall, 19,311 individuals participated in FDCR studies from 1998-2022. Of these, 12,950 were male (66.1%) and 5,130 were female (26.5%), with the sex of 1231 participants (6.4%) not explicitly specified. The median sample size of FDCR studies was 37. The 19,311 participants can be divided into 10,186 individuals with untreated addictive disorders (52.7%), 3,008 individuals with addictive disorders undergoing treatment at recruitment (15.6%), 2,388 individuals who used potentially addictive substances or engaged in potentially addictive behaviors without necessarily meeting addictive disorder criteria (12.4%), 618 individuals in long-term abstinence (3.2%) and 3,111 participants (16.1%) who were not using substances (i.e., “healthy controls”). A plurality of the participants (6708, 34.7%) were recruited to investigate alcohol use/use disorders with the following statistics for other use/use disorders: nicotine (4363, 22.6%), cocaine (1901, 10.0%), cannabis (1403, 7.2%), opioid (1205, 6.2%), and methamphetamine (836, 4.3%). Of the remaining participants, 1373 (7.1%) used betel-quid, inhalants, or multiple substances, and 1522 (7.9%) were recruited in studies focusing on gambling or video game playing. While most participants (13037, 67.5%) were recruited in observational studies, a substantial portion participated in trials or experimental studies involving pharmacological (2897, 15.0%), behavioral (2257, 11.7%), or other interventions (1120, 5.8%), such as neurofeedback or non-invasive brain stimulation (eFigure 6). No duplicated samples across studies were discovered in the database based on a screening.



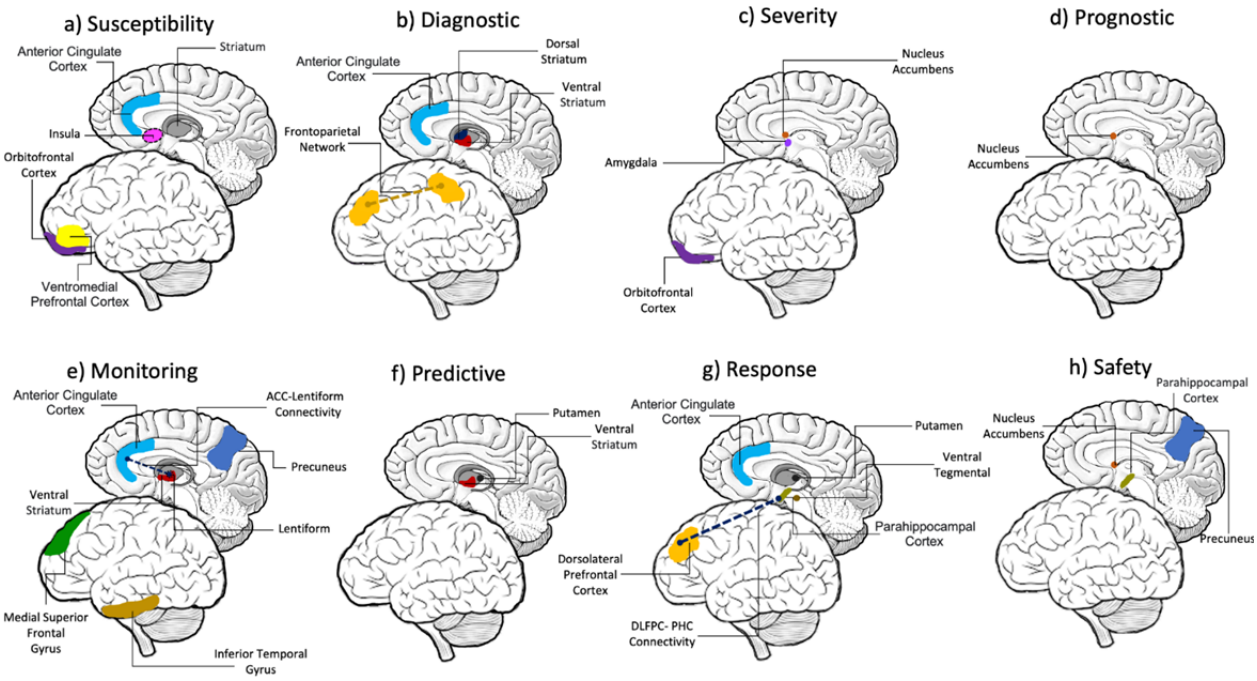
eFigure 6. Participants in fMRI drug cue-reactivity studies (N = 19,311). The Sankey diagram represents the number of participants in FDCR studies divided by sex, population type, potentially addictive drugs and behaviors, and interventions. The width of the boxes in each column represents the relative prevalence of each category in the column, while the width of the ribbons connecting the categories across columns represents the proportion of participants shared between each of the two categories. AD: Addictive Disorder (including both SUDs and BAs, diagnosed formally based on widely used criteria such as the Diagnostic and Statistical Manual (DSM) or International Classification of Diseases (ICD)). Participants who used substances without necessarily meeting diagnostic criteria are termed "User".

Study design types and relevance for potential biomarker development

It is important to note that none of the FDCR indices used by studies in the systematic review constitute fully validated biomarkers at this time. As detailed previously, any biological signal needs to undergo an extensive validation process to qualify as an actual biomarker of disease or recovery, which is not the case for any of the FDCR-derived measures in our included studies. However, the evidence presented in 335 of the studies in our database (75.9%) could potentially support the development of at least one future FDCR biomarker, by virtue of their study designs. We defined seven types of biomarkers based on their context of use. These biomarker types have all been directly adapted from the BEST Glossary[6], with the exception of "severity" biomarkers which are indices that reflect latent disease severity and were defined based on previous biomarker literature[7–9]. None of the studies in our database explicitly used FDCR as an index of "safety" and thus we removed the BEST *safety* biomarkers category.

Nevertheless, we provide two examples of studies that we think point to contexts in which FDCR-derived safety indices might prove useful.

These studies tested a total of 437 relationships (404 significant and 33 non-significant), across contexts of use, between FDCR-derived and clinical measures in 7 different biomarker categories: (1) In *diagnostic* studies, the FDCR signal reflects differences between populations (143 (32.7%) of the included studies, 134 studies reporting a significant association of FDCR and a grouping variable and nine reporting a non-significant association). (2) A *response* index might reflect the neural impact of an intervention (141 (32.3%) studies, 125 reporting significant and 16 non-significant results). (3) In a *severity* context, it would be tested whether an FDCR signal co-varies with addiction severity indices (such covariations were reported in 84 (19.2%) of the studies, 79 significant and five non-significant). (4) A *prognostic* measure should link to future disease course (30 (6.9%) studies, 29 significant and one non-significant). (5) A *predictive* index should explain a significant portion of variance in intervention outcomes (investigated in 25 (5.7%) studies, 24 significant and one non-significant). (6) A *monitoring* index should explain a significant portion of the variance of changes in clinically-relevant variables over time (reported in 12 (2.7%) studies, 11 significant, and one non-significant). Note that “monitoring” measures are only distinguished from “response” markers (in interventional contexts) and “severity” markers (in observational contexts) in that they can be measured repeatedly over time, and their variation over time within one individual is clinically meaningful. (7) A *susceptibility* index would assess the link between FDCR and the progression of non-addictive to addictive use (such links were reported in only 2 (0.5%) studies, both significant) (Figure 3). These biomarkers are defined in Table 1, and related example findings for each are presented in Table 1 and eFigure 7.



eFigure 7. Examples of brain regions in fMRI drug cue-reactivity (FDCR) studies with supporting evidence for potential biomarker development. Visual representation of regions with significant FDCR signal in example studies presented in Table 1. Each panel presents significant findings from studies whose results could support the development of one of the biomarker types in the modified BEST biomarker taxonomy, with each region presented with a unique color across panels. Note that these are example findings, and do not necessarily generalize beyond the context of the studies referenced in Table 1.

Interventional FDCR studies

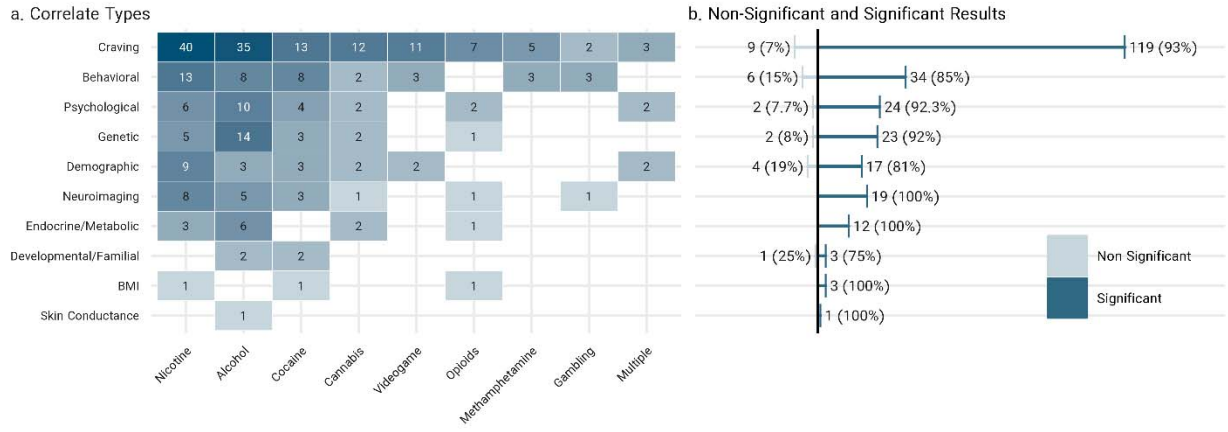
Given the importance of interventional studies and the potential of FDCR to develop response or predictive biomarkers, we present a separate summary of interventional studies in the database. Overall, 155 studies (37.3%) used FDCR in the context of a therapeutic intervention or experimental manipulation. Most commonly, interventional studies used target and control interventions with random assignment (91 studies, 58.7% of interventional studies). Eight studies (5.2%) included a control group without random assignment, 47 (30.3%) included only a single intervention arm without a control condition, and 9 (5.8%) investigated individuals who had been treated retrospectively, for example by comparing them to individuals with untreated SUDs or by comparing individuals who had undergone treatment for different lengths of time (Figure 4a). Most interventional FDCR studies investigated pharmacological agents (67 studies, 43.2% of the 155 interventional FDCR studies) and cognitive or behavioral interventions (52

studies, 33.5%) (Figure 4b). The most commonly investigated pharmacological agents were naltrexone (10 studies), varenicline (4 studies), baclofen (4 studies), oxytocin (3 studies), and methadone (3 studies). Four studies investigated the impacts of administering a potentially addictive substance, rather than a therapeutic one. Among cognitive and behavioral interventions, the most common were simple abstinence (10 studies) and instructed craving regulation (4 studies). Seven studies used mixes of interventions in different modalities (eFigure 8). Besides pharmacological and behavioral interventions, 12 studies (7.7% of interventional studies) used brain stimulation technologies (7 TMS, 4 tDCS, and 1 DBS), and 10 (6.5%) employed neurofeedback (Figure 4b) (For a detailed breakdown of interventional FDCR studies, see eFigure 6). A majority of the interventional studies (141 out of 155, 91%) used FDCR as a response biomarker, and 125 reported significant FDCR alterations as a result of treatment. Twenty-five studies (16.1%) used FDCR as a predictive biomarker, with 24 observing significant correlations between baseline FDCR and treatment outcomes. Among the 130 studies using FDCR as an outcome measure, 87 measured pre- to post-intervention changes in FDCR as an index of intervention effect (66.9%), and 43 (33.1%) measured only post-intervention cue-reactivity (Figure 4c).

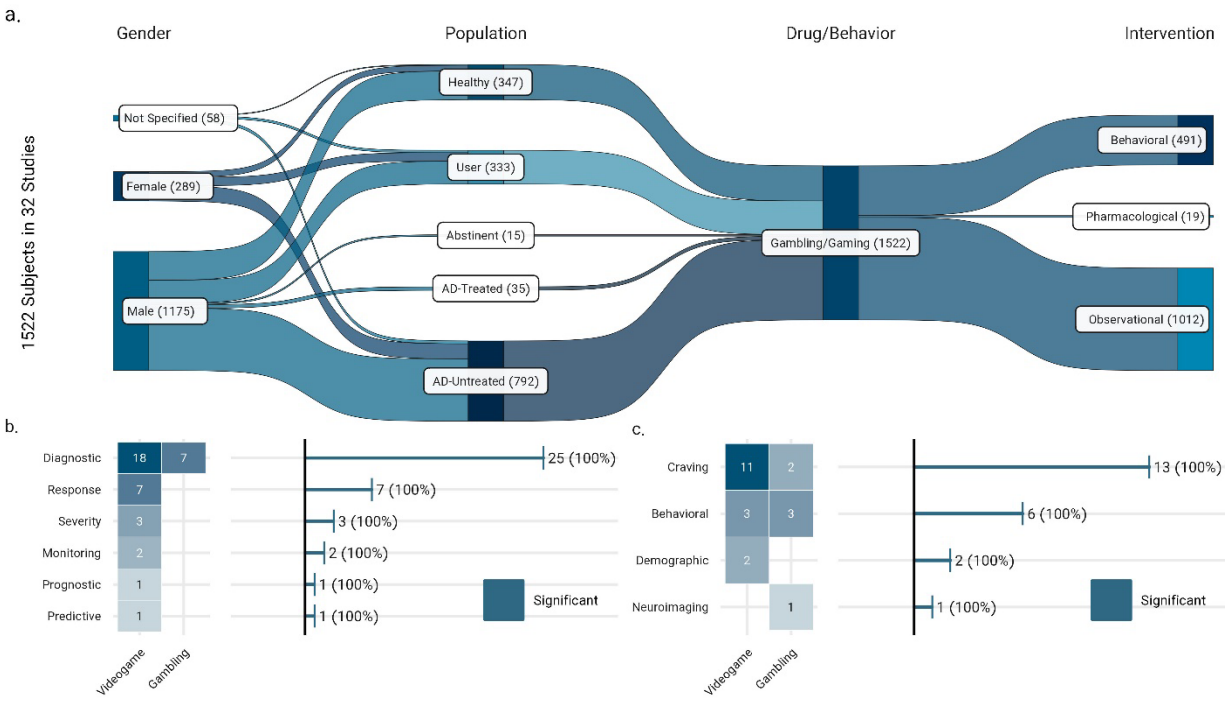
eFigure 8. Detailed breakdown of interventional FDCR studies with pharmacological (N = 67) or behavioral (N = 52) interventions. The "Multiple" column stands for those studies that included more than one type of addictive substance/behavior, while the "Multiple" rows stand for those FDCR studies which used multiple pharmacological interventions or multiple behavioral interventions.

Cross-modal Correlations

Further, 278 studies in the database also tested the relationship between one or more FDCR-derived parameters and non-FDCR variables (other than direct measures of disease severity) such as craving, impulsivity, physiological markers of cue-reactivity, hormonal profiles, and gene variants, with 255 significant and 23 non-significant test results (eFigure 5). Such investigations could be helpful to demonstrate links between FDCR and different aspects of SUDs and to clinically validate FDCR markers by supporting their etiological relevance in SUDs.



eFigure 9. Multi-modal correlations in FDCR studies. a. Studies which investigated correlations between FDCR results and other types of measures, broken down by substance or behavior of interest in each study. “Multiple” stands for those studies that included more than one type of addictive substance/ behavior. The "other" category includes inhalants and betel-quid. Note that numbers do not sum to 415 since some studies investigated no multi-modal correlations, while some fit multiple categories. b. Dumbbell plot showing the number of significant and non-significant tests of multi-modal correlation.

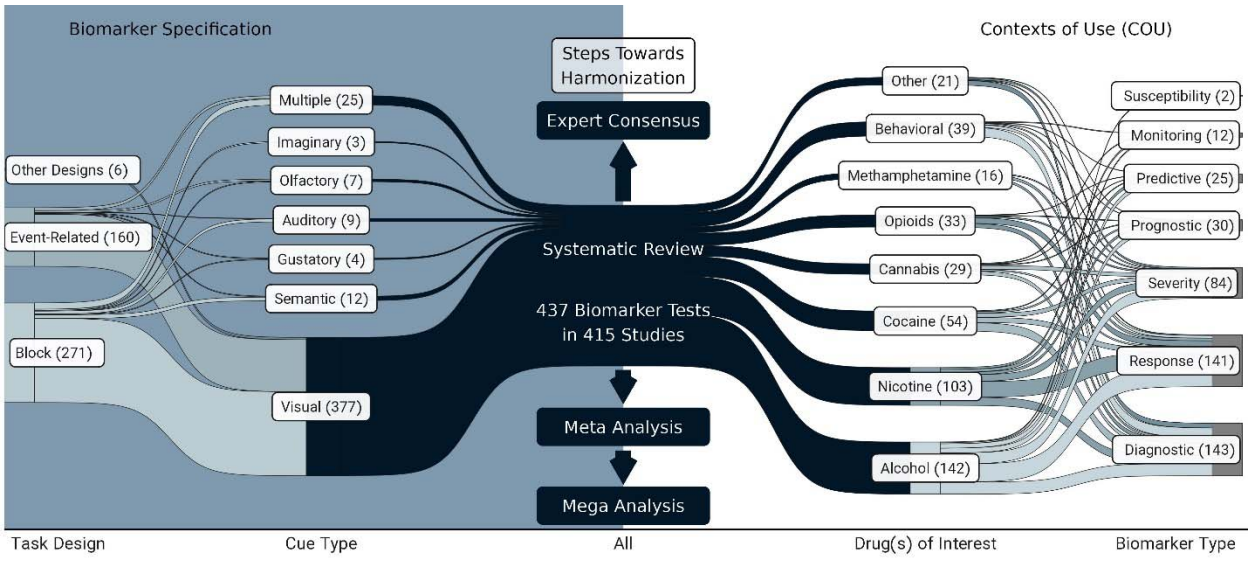


358

359

360 **eFigure 10. Separate analyses for behavioral addictions.** a. Participants in behavioral addiction
361 studies. The Sankey diagram represents the number of participants in fMRI cue-reactivity
362 studies divided by sex, population type, potentially addictive drugs and behaviors, and
363 interventions. b. Seven fMRI cue-reactivity study types for behavioral addictions. The dumbbell
364 plot shows 100% significant supporting biomarker-related findings for each biomarker
365 categories. c. Multi-modal correlations in fMRI cue-reactivity studies in behavioral addictions.
366 The dumbbell plot shows 100% significant test of multi-modal correlations.

367



eFigure 11. Preliminary map of the evidence and future directions in biomarker development. The Sankey diagram presents a summary of the methodological parameters and contexts of use (COUs) across the 437 potential biomarkers in the systematic review. Moving forward, expert consensus and meta- and mega-analyses may be used to facilitate harmonization and the development of optimal FDCR biomarkers which would undergo analytical and clinical validation and cost-benefit analysis before regulatory qualification for drug development or clinical use.

Bibliography for Boxes in the Manuscript:

BEST (Biomarkers E and other Tools) Resource. BEST (Biomarkers, EndpointS, and other Tools) resource. Published online 2016.

2. Kraus VB. Biomarkers as drug development tools: discovery, validation, qualification and use. *Nature Reviews Rheumatology*. 2018;14(6):354-362.

3. García-Gutiérrez MS, Navarrete F, Sala F, Gasparyan A, Austrich-Olivares A, Manzanares J. Biomarkers in Psychiatry: Concept, Definition, Types and Relevance to the Clinical Reality. *Front Psychiatry*. 2020;11. doi:10.3389/fpsyt.2020.00432

4. Comes AL, Papiol S, Mueller T, Geyer PE, Mann M, Schulze TG. Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. *Transl Psychiatry*. 2018;8(1):1-15. doi:10.1038/s41398-018-0219-2

5. Yuan N, Chen Y, Xia Y, Dai J, Liu C. Inflammation-related biomarkers in major psychiatric disorders: a cross-disorder assessment of reproducibility and specificity in 43 meta-analyses. *Transl Psychiatry*. 2019;9(1):1-13. doi:10.1038/s41398-019-0570-y

6. Stuart MJ, Baune BT. Chemokines and chemokine receptors in mood disorders, schizophrenia, and cognitive impairment: A systematic review of biomarker studies. *Neuroscience & Biobehavioral Reviews*. 2014;42:93-115. doi:10.1016/j.neubiorev.2014.02.001

7. Bough KJ, Amur S, Lao G, et al. Biomarkers for the Development of New Medications for Cocaine Dependence. *Neuropsychopharmacology*. 2014;39(1):202-219. doi:10.1038/npp.2013.210

8. Carvalho AF, Solmi M, Sanches M, et al. Evidence-based umbrella review of 162 peripheral biomarkers for major mental disorders. *Transl Psychiatry*. 2020;10(1):1-13. doi:10.1038/s41398-020-0835-5

9. McGorry P, Keshavan M, Goldstone S, et al. Biomarkers and clinical staging in psychiatry. *World Psychiatry*. 2014;13(3):211-223. doi:10.1002/wps.20144

10. Fu CHY, Costafreda SG. Neuroimaging-based biomarkers in psychiatry: clinical opportunities of a paradigm shift. *Can J Psychiatry*. 2013;58(9):499-508. doi:10.1177/070674371305800904

11. Goldstein RZ. Neuropsychoimaging Measures as Alternatives to Drug Use Outcomes in Clinical Trials for Addiction. *JAMA Psychiatry*. 2022;79(9):843-844. doi:10.1001/jamapsychiatry.2022.1970

12. Fernandes BS, Karmakar C, Tamouza R, et al. Precision psychiatry with immunological and cognitive biomarkers: a multi-domain prediction for the diagnosis of bipolar disorder or schizophrenia using machine learning. *Transl Psychiatry*. 2020;10(1):1-13. doi:10.1038/s41398-020-0836-4

13. Grabb MC, Hillefors M, Potter WZ. The NIMH "Fast-Fail Trials" (FAST) Initiative: Rationale, Promise, and Progress. *Pharmaceut Med*. 2020;34(4):233-245. doi:10.1007/s40290-020-00343-y

14. Krystal AD, Pizzagalli DA, Mathew SJ, et al. The first implementation of the NIMH FAST-FAIL approach to psychiatric drug development. *Nature Reviews Drug Discovery*. 2019;18(1):82.

15. Bahji A, Brietzke E, Soares C, Stuart H. Recent Advances in Biomarkers of Addiction: A Narrative Review. *Canadian Journal of Addiction*. 2021;12(1):6-12. doi:10.1097/CXA.0000000000000107

16. National Institutes of Health. National Institute on Drug Abuse. Resource Guide: Screening for Drug Use in General Medical Settings. Published online 2020. Accessed May 20, 2021. https://archives.drugabuse.gov/publications/resource-guide-screening-drug-use-in-general-medical-settings/biological-specimen-testing

17. Neumann T, Spies C. Use of biomarkers for alcohol use disorders in clinical practice. *Addiction*. 2003;98 Suppl 2:81-91. doi:10.1046/j.1359-6357.2003.00587.x

18. Zakhari S, Li TK. Determinants of alcohol use and abuse: Impact of quantity and frequency patterns on liver disease. *Hepatology*. 2007;46(6):2032-2039. doi:10.1002/hep.22010

19. Kwako LE, Bickel WK, Goldman D. Addiction Biomarkers: Dimensional Approaches to Understanding Addiction. *Trends Mol Med*. 2018;24(2):121-128. doi:10.1016/j.molmed.2017.12.007

20. Petry N. *Behavioral Addictions: DSM-5® and Beyond*.; 2015.

21. Center for Drug Evaluation and Research. Biomarker Qualification Submissions. *FDA*. Published online October 21, 2020. Accessed December 16, 2020. https://www.fda.gov/drugs/biomarker-qualification-program/biomarker-qualification-submissions

22. Ekhtiari H, Nasseri P, Yavari F, Mokri A, Monterosso J. Neuroscience of drug craving for addiction medicine: From circuits to therapies. *Prog Brain Res*. 2016;223:115-141. doi:10.1016/bs.pbr.2015.10.002

23. Starcke K, Antons S, Trotzke P, Brand M. Cue-reactivity in behavioral addictions: A meta-analysis and methodological considerations. *J Behav Addict*. 2018;7(2):227-238. doi:10.1556/2006.7.2018.39

24. Boswell RG, Kober H. Food cue reactivity and craving predict eating and weight gain: a meta-analytic review. *Obes Rev*. 2016;17(2):159-177. doi:10.1111/obr.12354

25. Lorenz RC, Krüger JK, Neumann B, et al. Cue reactivity and its inhibition in pathological computer game players. *Addict Biol*. 2013;18(1):134-146. doi:10.1111/j.1369-1600.2012.00491.x

26. Berridge KC, Robinson TE. Liking, Wanting and the Incentive-Sensitization Theory of Addiction. *Am Psychol*. 2016;71(8):670-679. doi:10.1037/amp0000059

27. Drummond DC. What does cue-reactivity have to offer clinical research? *Addiction*. 2000;95(8s2):129-144. doi:10.1046/j.1360-0443.95.8s2.2.x

28. Carter BL, Tiffany ST. Meta-analysis of cue-reactivity in addiction research. *Addiction*. 1999;94(3):327-340. doi:10.1046/j.1360-0443.1999.9433273.x

29. Rohsenow DJ, Childress AR, Monti PM, Niaura RS, Abrams DB. Cue reactivity in addictive behaviors: theoretical and treatment implications. *International Journal of the Addictions*. 1991;25(sup7):957-993.

30. Ekhtiari H, Faghiri A, Oghabian MA, Paulus MP. Chapter 7 - Functional neuroimaging for addiction medicine: From mechanisms to practical considerations. In: Ekhtiari H, Paulus MP, Ekhtiari H, Paulus MP, eds. *Progress in Brain Research*. Vol 224. Neuroscience for Addiction Medicine: From Prevention to Rehabilitation - Methods and Interventions. ; 2016:129-153. Accessed February 22, 2019. http://www.sciencedirect.com/science/article/pii/S0079612315001508

31. Goldstein RZ, Volkow ND. Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. *American Journal of Psychiatry*. 2002;159(10):1642-1652.

32. Zilverstand A, Huang AS, Alia-Klein N, Goldstein RZ. Neuroimaging Impaired Response Inhibition and Salience Attribution in Human Drug Addiction: A Systematic Review. *Neuron*. 2018;98(5):886-903. doi:10.1016/j.neuron.2018.03.048

33. Paulus MP, Stewart JL. Neurobiology, Clinical Presentation, and Treatment of Methamphetamine Use Disorder: A Review. *JAMA Psychiatry*. 2020;77(9):959-966. doi:10.1001/jamapsychiatry.2020.0246

34. Garrison KA, Potenza MN. Neuroimaging and Biomarkers in Addiction Treatment. *Curr Psychiatry Rep*. 2014;16(12):513. doi:10.1007/s11920-014-0513-5

35. Genauck A, Matthis C, Andrejevic M, et al. Neural correlates of cue-induced changes in decision-making distinguish subjects with gambling disorder from healthy controls. *Addiction Biology*. 2021;26(3):e12951. doi:10.1111/adb.12951

36. Liu GC, Yen JY, Chen CY, et al. Brain activation for response inhibition under gaming cue distraction in internet gaming disorder. *Kaohsiung J Med Sci*. 2014;30(1):43-51. doi:10.1016/j.kjms.2013.08.005

37. Miedl SF, Büchel C, Peters J. Cue-induced craving increases impulsivity via changes in striatal value signals in problem gamblers. *J Neurosci*. 2014;34(13):4750-4755. doi:10.1523/JNEUROSCI.5020-13.2014

38. Blaine S, Claus E, Harlaar N, Hutchison K. TACR1 genotypes predict fMRI response to alcohol cues and level of alcohol dependence. *Alcohol Clin Exp Res*. 2013;37 Suppl 1:E125-130. doi:10.1111/j.1530-0277.2012.01923.x

39. Wiers CE, Shumay E, Volkow ND, et al. Effects of depressive symptoms and peripheral DAT methylation on neural reactivity to alcohol cues in alcoholism. *Transl Psychiatry*. 2015;5:e648. doi:10.1038/tp.2015.141

40. Kroemer NB, Wuttig F, Bidlingmaier M, Zimmermann US, Smolka MN. Nicotine enhances modulation of food-cue reactivity by leptin and ghrelin in the ventromedial prefrontal cortex. *Addict Biol*. 2015;20(4):832-844. doi:10.1111/adb.12167

41. Wang W, Zhornitsky S, Le TM, et al. Cue-elicited craving, thalamic activity, and physiological arousal in adult non-dependent drinkers. *Journal of Psychiatric Research*. 2019;116:74-82. doi:10.1016/j.jpsychires.2019.06.005

42. Elton A, Smitherman S, Young J, Kilts CD. Effects of childhood maltreatment on the neural correlates of stress- and drug cue-induced cocaine craving. *Addict Biol*. 2015;20(4):820-831. doi:10.1111/adb.12162

43. Groefsema MM, Mies GW, Cousijn J, Engels RCME, Sescousse G, Luijten M. Brain responses and approach bias to social alcohol cues and their association with drinking in a social setting in young adult males. *Eur J Neurosci*. 2020;51(6):1491-1503. doi:10.1111/ejn.14574

44. Burnette EM, Grodin EN, Lim AC, MacKillop J, Karno MP, Ray LA. Association between impulsivity and neural activation to alcohol cues in heavy drinkers. *Psychiatry Res Neuroimaging*. 2019;293:110986. doi:10.1016/j.pscychresns.2019.110986

45. Ceceli AO, King SG, McClain N, Alia-Klein N, Goldstein RZ. The Neural Signature of Impaired Inhibitory Control in Individuals with Heroin Use Disorder. *J Neurosci*. 2023;43(1):173-182. doi:10.1523/JNEUROSCI.1237-22.2022

46. Goldstein RZ, Tomasi D, Alia-Klein N, et al. Dopaminergic Response to Drug Words in Cocaine Addiction. *J Neurosci*. 2009;29(18):6001-6006. doi:10.1523/JNEUROSCI.4247-08.2009

47. Huang Y, Ceceli A, Kronberg G, et al. Cortico-striatal engagement during cue-reactivity, reappraisal, and savoring of drug and non-drug stimuli predicts craving in heroin addictio. Published online 2022. doi:10.1101/2022.05.27.22275628

48. Vincent GM, Cope LM, King J, Nyalakanti P, Kiehl KA. Callous-Unemotional Traits Modulate Brain Drug Craving Response in High-Risk Young Offenders. *J Abnorm Child Psychol*. 2018;46(5):993-1009. doi:10.1007/s10802-017-0364-8

49. Feldstein Ewing SW, Filbey FM, Chandler LD, Hutchison KE. Exploring the Relationship Between Depressive and Anxiety Symptoms and Neuronal Response to Alcohol Cues. *Alcohol Clin Exp Res*. 2010;34(3):396-403. doi:10.1111/j.1530-0277.2009.01104.x

50. Moran LV, Betts JM, Ongur D, Janes AC. Neural Responses to Smoking Cues in Schizophrenia. *Schizophr Bull*. 2018;44(3):525-534. doi:10.1093/schbul/sbx085

51. Baurley JW, McMahan CS, Ervin CM, Pardamean B, Bergen AW. Biosignature Discovery for Substance Use Disorders using Statistical Learning. *Trends Mol Med*. 2018;24(2):221-235. doi:10.1016/j.molmed.2017.12.008

52. Rawls E, Kummerfeld E, Zilverstand A. An integrated multimodal model of alcohol use disorder generated by data-driven causal discovery analysis. *Communications Biology*. 2021;4(1):1-12. doi:10.1038/s42003-021-01955-z

53. Ray LA, Courtney KE, Hutchison KE, Mackillop J, Galvan A, Ghahremani DG. Initial evidence that OPRM1 genotype moderates ventral and dorsal striatum functional connectivity during alcohol cues. *Alcohol Clin Exp Res*. 2014;38(1):78-89. doi:10.1111/acer.12136

54. Schacht JP, Anton RF, Voronin KE, et al. Interacting effects of naltrexone and OPRM1 and DAT1 variation on the neural response to alcohol cues. *Neuropsychopharmacology*. 2013;38(3):414-422. doi:10.1038/npp.2012.195

55. Braus DF, Wrase J, Grüsser S, et al. Alcohol-associated stimuli activate the ventral striatum in abstinent alcoholics. *J Neural Transm*. 2001;108(7):887-894. doi:10.1007/s007020170038

56. Sjoerds Z, Brink W van den, Beekman ATF, Penninx BWJH, Veltman DJ. Cue Reactivity Is Associated with Duration and Severity of Alcohol Dependence: An fMRI Study. *PLOS ONE*. 2014;9(1):e84560. doi:10.1371/journal.pone.0084560

57. Vollstädt-Klein S, Wichert S, Rabinstein J, et al. Initial, habitual and compulsive alcohol use is characterized by a shift of cue processing from ventral to dorsal striatum. *Addiction*. 2010;105(10):1741-1749. doi:10.1111/j.1360-0443.2010.03022.x

58. Beck A, Wüstenberg T, Genauck A, et al. Effect of brain structure, brain function, and brain connectivity on relapse in alcohol-dependent patients. *Arch Gen Psychiatry*. 2012;69(8):842-852. doi:10.1001/archgenpsychiatry.2011.2026

59. Bach P, Weil G, Pompili E, et al. Incubation of neural alcohol cue reactivity after withdrawal and its blockade by naltrexone. *Addiction Biology*. 2020;25(1):e12717.

60. Bach P, Kirsch M, Hoffmann S, et al. The effects of single nucleotide polymorphisms in glutamatergic neurotransmission genes on neural response to alcohol cues and craving. *Addict Biol*. 2015;20(6):1022-1032. doi:10.1111/adb.12291

61. Reinhard I, Leménager T, Fauth-Bühler M, et al. A comparison of region-of-interest measures for extracting whole brain data using survival analysis in alcoholism as an example. *Journal of Neuroscience Methods*. 2015;242:58-64. doi:10.1016/j.jneumeth.2015.01.001

62. Seo S, Mohr J, Beck A, Wüstenberg T, Heinz A, Obermayer K. Predicting the future relapse of alcohol-dependent patients from structural and functional brain images. *Addiction Biology*. 2015;20(6):1042-1055. doi:10.1111/adb.12302

63. Schacht JP, Randall PK, Latham PK, et al. Predictors of naltrexone response in a randomized trial: reward-related brain activation, OPRM1 genotype, and smoking status. *Neuropsychopharmacology*. 2017;42(13):2640-2653.

64. Karl D, Bumb JM, Bach P, et al. Nalmefene attenuates neural alcohol cue-reactivity in the ventral striatum and subjective alcohol craving in patients with alcohol use disorder. *Psychopharmacology*. 2021;238(8):2179-2189. doi:10.1007/s00213-021-05842-7

65. Mann K, Vollstädt-Klein S, Reinhard I, et al. Predicting naltrexone response in alcohol-dependent patients: the contribution of functional magnetic resonance imaging. *Alcohol Clin Exp Res*. 2014;38(11):2754-2762. doi:10.1111/acer.12546

66. Bach P, Weil G, Pompili E, et al. FMRI-based prediction of naltrexone response in alcohol use disorder: a replication study. *Eur Arch Psychiatry Clin Neurosci*. 2021;271(5):915-927. doi:10.1007/s00406-021-01259-7

67. Krishnan A, Woo CW, Chang LJ, et al. Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *eLife*. 5:e15166. doi:10.7554/eLife.15166

68. Spisak T, Bingel U, Wager TD. Multivariate BWAS can be replicable with moderate sample sizes. *Nature*. 2023;615(7951):E4-E7. doi:10.1038/s41586-023-05745-x

69. Konova AB, Zilverstand A. Deriving Generalizable and Interpretable Brain-Behavior Phenotypes of Cannabis Use. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2023;8(3):238-240. doi:10.1016/j.bpsc.2023.01.003

70. Koban L, Wager TD, Kober H. A neuromarker for drug and food craving distinguishes drug users from non-users. *Nat Neurosci*. 2023;26(2):316-325. doi:10.1038/s41593-022-01228-w

Bibliography for Online-Only Materials:

1. Ekhtiari H, ACRI Secretariat. A Systematic Review on fMRI Drug Cue Reactivity Studies. Published online May 18, 2020. Accessed November 29, 2020. https://osf.io/eb972/

2. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/M18-0850

3. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*. 2015;4:1. doi:10.1186/2046-4053-4-1

4. The EndNote Team. EndNote. Published online 2013.

5. R Core Team. R: A language and environment for statistical computing. Published online 2013.

6. BEST (Biomarkers E and other Tools) Resource. BEST (Biomarkers, EndpointS, and other Tools) resource. Published online 2016.

7. Bivona G, Agnello L, Ciaccio M. Biomarkers for Prognosis and Treatment Response in COVID-19 Patients. *Ann Lab Med*. 2021;41(6):540-548. doi:10.3343/alm.2021.41.6.540

8. Ligezka AN, Radenkovic S, Saraswat M, et al. Sorbitol Is a Severity Biomarker for PMM2-CDG with Therapeutic Implications. *Ann Neurol*. 2021;90(6):887-900. doi:10.1002/ana.26245

9. de Mendonça EB, Schmaltz CA, Sant'Anna FM, et al. Anemia in tuberculosis cases: A biomarker of severity? *PLoS One*. 2021;16(2):e0245458. doi:10.1371/journal.pone.0245458

Review of PSY23-0829 (Ekhtiari et al.)

Editorial comments

Thank you for the careful and substantive review of the manuscript, and for considering it for publication in JAMA Psychiatry. We have thoroughly revised the manuscript in response to both the editorial requirements and the reviewer's comments. In summary, we have:

1) substantially reduced the word count and references in line with the journal's requirements. Some of the material has been removed or summarized, and most has been restructured and moved to online-only materials,
2) moved several figures and tables to online-only materials,
3) added a "key points" section,
4) moved all authors to a group byline, "Addiction Cue Reactivity Initiative (ACRI) group",
5) added the required additional sections (Role of Funder/Sponsor Statement, Access to Data and Data Analysis, Data Sharing Statement)
6) changed the styling format in line with the AMA and JAMA requirements,
7) revised the manuscript substantially in response to reviewer comments, with responses to each comment provided below.

Reviewer #1

1. This is a very interesting systematic review of the potential role of fMRI-Derived Cue Reactivity (FDCR) in serving as a biomarker for addictions. The authors include many of the leaders in this field, and the review uses excellent methodology and analyses associated with systematic reviews. They find over 400 such relevant research articles that have documented studies of FDCR in association with prediction of substance use, SUD risk and treatment response.
   We appreciate the reviewer's careful review of the manuscript and their recognition of the impact and high quality of the manuscript.

2. Notably about 60% of these FDCR studies are in nicotine (tobacco) and alcohol populations, where we have FDA-approved pharmacotherapes for potential biomarker validation. An important issue that is raised in interpretation of this large number of studies is the heterogeneity of the methods used (e.g. scanning times, single versus multiple FDCR sessions, differences in substances studied and whether they meet criteria for an SUD or not, psychiatric and medical co-morbidities) and the relatively small sample size per study (~35) which make viable conclusions from a systematic review of this topic very difficult.
   As the reviewer correctly notes, there are substantial methodological and conceptual heterogeneities between the reviewed studies and many FDCR studies have small and unbalanced sample sizes. This was indeed one aim of the systematic review: to systematically evaluate the current state of the FDCR literature and identify challenges and potentials for biomarker development. We further emphasize these objectives, methodological heterogeneity and sample size limitations of previous research in the

"Participant Characteristics" section, page 8, lines 153-166, and online-only materials, in the "eMethods" section, page 10, lines 162-167.

3. In addition, there is no attempt to quantify outcomes of FDCR with specific substances or paradigms, and we are told only what percentage of studied are "significant".
   We considered the use of formal meta-analytic techniques, but ultimately elected to use the relatively simple metric of "significance" given the extreme heterogeneity in the field. The scope of this work is to survey the status of the field broadly and provide a roadmap to support FDCR biomarker development, and also facilitate future meta-analyses of homogenous subsets of studies in the database. We have added a discussion of this point in the "Participant Characteristics" section, page 8, lines 163-166.

4. By the end of the manuscript, one is left with the impression that while much work has been done in the field, actionable items for biomarker development are not well-established. It is evident that in order to develop viable and validated treatment biomarkers, larger studies that incorporate standardized FDCR procedures that are replicable and cost-effective are clearly needed. Thus, the value of the present systematic review for the field of addictions biomarker development, beyond what has been recently published (e.g. Ekhtiari et al., 2022. Nature Protocols; Goldstein, 2022. JAMA Psychiatry), is not entirely clear.
   The mentioned papers focus specifically on: (A) The development of a tool/checklist to improve reporting quality standards; and B) The potential utility of FDCR as a clinical outcome measure (as only one of many types of biomarkers). Here, we attempt to provide an overview of the FDCR literature, assess methodological heterogeneity, and develop a comprehensive biomarker development framework to identify key challenges. We agree with the reviewer that actionable items for biomarker development have not been well-established in the field, and indeed one aim of the present manuscript is to highlight and specify actions such as larger collaborations and funding, methodological harmonization, biomarker specification and clinical/analytical validation (see biomarker specification and validation sections, as well as the "Methodological heterogeneity and biomarker specification" section, page 7, lines 139-144, the "Participant Characteristics" section, page 8, lines 159-160 and 165-166 , the "Validation of FDCR biomarkers" section, page 10, lines 253-257, and the "Conclusion" section, page 12, lines 301-310) . Thus, we view this paper as complementary – not redundant with – the important works mentioned. Some of these points have been made more explicitly in the revised manuscript.

Reviewer #2

5. This is a tour-de-force effort by a large group of alcohol and substance abuse neuroimaging experts to summarize, quantify, and prescribe the future use of functional

neuroimaging (Fmri) research over the last 20 or so years. The paper describes the review and data abstraction method utilized in this effort, which seems appropriate. The primary authors should be congratulated on the scope of this effort, and the detail in which they proceeded in collapsing the diverse data across multiple substances of abuse and various domains. Beyond the extraordinary catalogue of the relevant published material, they venture beyond to begin the important process of characterizing domains in which the totality of this work might benefit future therapeutic drug discovery and clinical care. In particular, they, and their expert co-authors, should be commended for their thoughtful analytic and prescriptive approach for the potential of functional neuroimaging of alcohol and substance abuse disorders (AUD and SUD) to be qualified by the FDA as a biomarker(s) for drug development. This is an extremely important and relevant undertaking, which parenthetically is not trivial in scope, technology, and cost.
We appreciate the reviewer's positive comment, and indeed hope that this and other efforts can help facilitate the development of clinically relevant biomarkers.

6. Inclusion criteria for study participants was rather broad from those who met SUD/AUD diagnostic criteria, to those who used substances excessively or had some rating-scale Indication of addictive severity. Those with co-occurring disorders were also included. This can either be seen as a plus or a minus. Since the effort in this paper was to "catalogue and order" germane studies in this area, that is a plus, however, further work needs to clarify the overlap or distinction of these co-occurring disorders from the major findings from primary independent AUD and SUD where the biology might be a bit less complex – an important distinction to make.
   We agree with the reviewer's comment about the tension between comprehensiveness and homogeneity. Given that we focused on reporting the status of the FDCR field so far, we elected for broadness as the reviewer has noticed. We have further clarified the aim and limitations of the study and touch on future directions as suggested in the "Participant Characteristics" section on page 7-8.

7. Importantly, the data base forming the substrate of this paper was made publicly available. This group and others can utilize the work presented in this paper (and more) to address many more questions such as: how specific cues (within and between substances), prestation times of cues, various brain regional reactivities, sample size effects, and many others might impact various choices for standardization and eventually drug development. A comment might be made about this in the conclusion or elsewhere.
   We are glad that the reviewer has noticed the importance of a central repository of FDCR studies. Sentences are now added to the "Participant Characteristics" section, page 8, lines 159-166 to highlight this further as well.

8. It is good to see that some reference is made to the role of genetics and epigenetics in influencing brain induced cue reactivity. In this reviewer's opinion, the marriage of functional molecular genetics and brain neuroimaging is one of the few investigative pathways to untangle the biological and treatment aspects of AUD and SUD. Perhaps a

bit more emphasis should be placed in this area.
We definitely agree with this important comment and believe that multimodal integration is essential as the field moves towards developing clinically relevant and biologically interpretable biomarkers. This is discussed both in the revised "Validation of FDCR Biomarkers" section on page 10, lines 241-245 and in online-only materials. Several authors of the present manuscript have formed the steering committee of the ENIGMA Addiction Cue-Reactivity Initiative (ACRI) within the ENIGMA Addiction working group to facilitate consensus development, methodological harmonization, and data sharing for mega-analyses (now mentioned in the "Conclusion" section, page 12, lines 301-310)

9. The authors might consider commenting on what looks like "a peak or flattening" of the number of published studies over time. Irrespective of COVID-19 impact, is there a reason for that? Cost, divergence of results, need for new paradigms? It might speak to the need for consilience on data and methods. In fact, more emphasis should be placed on the need for likely commercial, or governmental, (or both) development of a tool box or standards for investigators to use across studies. Perhaps NIAAA or NIDA could take the lead in this standardization.
We agree with the reviewer that this trend is intriguing, though we do not believe our data can clarify the causes. As suggested, we further emphasized the need for accessible resources to enable further harmonization and synthesis in the field in the "Conclusion" section, on page 12, lines 301-305. We have actually taken steps towards this goal with a recent "design and reporting standards" checklist (https://www.nature.com/articles/s41596-021-00649-4) and are undertaking an effort to develop a toolbox to help implement and modify an FDCR task (https://osf.io/fbeu8/).

10. Please explain the difference between a block vs. event related imaging design. While this might be obvious to experienced neuroimagers it is likely not obvious to the more naïve reader. Perhaps a bit more clarity is needed about the main "dependent variable" for those studies where only a direct substance cue without additional/combined tasks were done. The assumption is that the "cue reactivity" had some increased (or perhaps decreased) salience, reward potential, etc. A bit of time is spent describing the combined tasks without first defining the essence/meaning of the simple cue reactivity task.
We appreciate the reviewer bringing these issues to our attention, especially since we aim for the manuscript to be accessible to a wide audience of scientists and practitioners. A few sentences are now added in the "Introduction" section, on page 4, lines 57-63, along with a new figure (eFigure 1), to familiarize readers with the basic design and varieties of an FDCR task.

11. While figure 4 is visually appealing, it is hard to follow and takes too long to decipher. It would be best to present the data in tabular form which can be done with more clarity.
We see the reviewer's point about the figure's decipherability. We have moved the figure

to online-only materials (eFigure 6) and the information presented in this figure is further discussed in the “Participants in FDCR studies” section (online-only materials pages 15-16).

12. Regarding the “predictive index” which can also be considered a “therapeutic response index/marker/indicator”, it would be wise to mention that differential drug mechanisms might act differentially on established brain reactivity and potentially brain regional reactivity. For instance, opiate antagonists might block reward-mediated pathways but other drugs like acamprosate or gabapentin might not do so. Others might work through cognitive enhancing or impulse control cortical pathways etc. This is alluded to later but should be emphasized a bit more in this section.
    This is an important point. This is now further discussed in the discussion section in “Contexts of use of FDCR biomarkers” section, page 9, lines 201-206, to briefly illustrate and discuss differential intervention mechanisms.

13. Good job on Figure 5, very creative use of graphics (heat map) and adapted odds ratio type graphics. However, it would be helpful to add (%) after the numbers in graph 5b to better control for overall number of studies. One could alternatively only provide % and not N. This is also true off Supplementary Figure 4b. One other important point needs to be made about this section (perhaps more important for supplementary figure 4b) is publication bias and post-hoc exploratory analyses. It seems a bit unusual that across the multitude of domains explored in that figure there were “very few” non-significant findings. This is likely accounted for by publication bias and post-hoc exploratory observations. This should be more made explicit.
    We appreciate the reviewer’s kind note, and percentages are now added to Figure 3 and eFigure 9. The issue of potential publication bias and post-hoc exploration is also discussed in greater detail in the “Validation of FDCR biomarkers” section of the revised manuscript on page 10, lines 261-265.

14. Figure 6 probably should be removed since it is showing only the results from the studies referenced in Figure 5. In some ways it is a graphic rehash of what is written in figure 5, but more importantly it represents the results of only one or two studies, while perhaps well done and representative, are trivial in the context of what this article is trying to do i.e., aggregate data across a number of studies. It would have been better to have utilized the relevant studies detailed under the substance categories in Figure 5. And in tabular (or perhaps brain graphic) form show what regions were most reported most frequently affected in each biomarker-type by substance-category. Perhaps this can be limited to the top two-three substance categories. These can be included as either primary or supplemental tables. If the authors choose not to do that, they might comment that this needs to be done in a future manuscript but still removing figure 6 which seems premature.

This was included to provide a visual companion to the table of exemplar studies, but we agree with this comment and the figure is now moved to the online-only materials (eFigure 7).

15. While recognizing that this paper contains considerable material and contains significant depth of content, it would be useful to provide a bit more information on the interventional studies in regards to actual predictors (brain region by drug). To interest pharma companies, and to "jump start" more commercial exploration and interest, more detail in this regard would be useful. Perhaps adding a supplementary tables or at least indicate which drugs showed the most promise for a treatment response being predicted by a FDCR biomarker. This is done in a summary way in the discussion but more depth in the results appears warranted.
    Thanks for this critical suggestion by the reviewer. Besides further highlighting the potential uses of FDCR biomarkers in drug development (for example with Fast-Fail trials), we have provided some statistics about the frequency of various intervention types and interventions in the discussion of interventional biomarkers in "Contexts of use of FDCR biomarkers" section, on page 9, lines 196-200 and lines 219-223. Further detail is provided in the online-only materials of the revised manuscript on pages 18-19.

Discussion:

16. While a discussion of "combined paradigms" seems logical in a compendium of this work, if word space is needed, this can be shortened to two sentences with the highlight on the last sentence (lines 564-565) being the most important.
    The change is now made as suggested in the revised manuscript.

17. The section on methodological design is very important and mostly accurate. However, if I were a pharma company going to invest in this area, I would be "scared to death" about the highlighted/perceived complexity of the issue. Considerable "parametric work" might be needed prior to validation in a qualification plan for FDA approval. Unfortunately, within typical R01 review and funding, parametric studies are not highly valued. Some statement regarding the need for appropriate parametric studies (as indicated in this section) and funding sources (like perhaps NIH (NIAAA, NIDA) contracts, RFA's etc.) need to be highlighted. The conclusion alludes to some of this, but I think it would be worth being more explicit with a larger "call to arms" for NIH the FDA and pharma companies to cooperate in this effort.
    This issue had been previously raised among co-authors as well, and we agree with the reviewer that a more forceful "call to arms" is warranted. Funding issues are now discussed more explicitly in the revised manuscript as suggested in the "Conclusion" section, page 12, lines 301-310, and we hope that our outlining of concrete and incremental steps towards tackling the most substantial challenges will help allay fears of insurmountability.

18. In the section under population differences, I think it might be useful to reduce the discussion to acknowledging that various demographic and disease specific differences (sex, race, severity of SUD, recency of use, co-morbid conditions etc.) should be further explored using the data derived from this review or others. One wonders whether machine learning, AI, or large multivariate models might be of particular use here and encouraged.
    We had previously briefly touched on the potential of multivariate and deep learning models. We agree with the reviewer's note and the discussion is now slightly extended (within word limit bounds) to highlight that such models hold immense potential to aid the development of useful neuromarkers and disambiguate the impact of demographic factors ("Participant characteristics" section, page 8 lines 153-158 and Box 2).

19. The section on FDCR intervention biomarkers is particularly important and some relevant literature is added. A further, more refined, analysis of this area from the data available from the studies identified in this review is in order and should be noted as suggested above.
    We agree with the importance of interventional FDCR biomarkers, particularly biomarkers of treatment response. We have highlighted the need for further quantitative syntheses of studies in the database for biomarker development in the "Participant characteristics" section, on page 8, lines 159-166.

20. The idea of the FDCR biomarker as a surrogate endpoint is intriguing. This idea could be enhanced by pointing to several other diseases such as brain amyloid scanning for new Alzheimer drugs, and cholesterol monitoring for heart disease. These analogies could draw attention from pharma companies to encourage translational thinking. This a laudable goal but likely a way off in FDA regulatory thinking.
    We agree with the reviewer's observation, and a sentence is now added mentioning the useful analogies that the reviewer recommended in the "Contexts of use of FDCR biomarkers" section, on page 9, lines 216-219.

21. There should be a small section detailing "future impediments" to developing and validating a FDCR qualification plan that should include: sufficient expertise within the FDA internal Qualification Section, and the significant costs of doing qualification studies and who will bear those costs. It is likely that the perceived low return-on-investment for Pharma companies to invest in a costly FDCR qualification plan will be a large impediment. It would seem, that the most likely mechanism for such an event would be through an NIH funded mechanism – perhaps in a public/private partnership. The article should address this crucial point that goes a bit beyond what was already mentioned.
    We completely agree with the reviewer. We discussed this important point further in the "Conclusion" section, on page 12, lines 304-310, highlighting the importance of a suitable funding mechanism in the revised manuscript.

Tables/figures:

22. Figure 4 is hard to understand despite having visual appeal. It should be turned into a table.
    We have moved the figure to online-only materials (eFigure 6) and the information of this figure is further discussed in the “Participants in FDCR studies” section in online-only materials, page 15, lines 230-348.

Reviewer #3

23. This manuscript describes the potential of using fMRI drug cue reactivity (FDCR) as a biomarker for addictions research. The authors detail FDA criteria for evaluating biomarkers and conducted a systematic review of prior FDCR work. As they note, the latter part of this is not very novel, as there are multiple recent reviews of the FDCR literature. Presumably there are also reviews detailing the FDA criteria for biomarker development, however combining the two things here and proposing future directions to move the field forward is a potentially important contribution to the literature. However, I do note several limitations.
    We thank the reviewer for their positive comment. We have highlighted the fact that, unlike recent reviews and systematic reviews of the FDCR literature, we focus on the methodological and study design characteristics of FDCR studies in the “Introduction” section, page 6, lines 91-97.

24. A primary limitation of this review is the lack of detail and consideration of data quality. For example, the first study in Table 1 (Dager et al., 2016) had a small sample size (N<20 per group) which is a weakness. The fact that small sample sizes used to be standard in fMRI does not mean that findings from such studies should not be viewed within the context of this limitation. Similarly, it is highly likely that many of the studies included used approaches for multiple comparison correction that are unacceptable by today’s standards. I also did not see any references to excess motion or whether the quality of preprocessing steps was evaluated. Some discussion of these critical methodological issues is warranted.
    We agree with the reviewer that there are substantial methodological heterogeneities between the reviewed studies, and many FDCR studies suffer from limitations such as small sample sizes and improper preprocessing or analysis. We have now more explicitly highlighted several methodological and analytical shortcomings in the “Participant characteristics” section non page 8, lines 158-160 and in “Validation of FDCR biomarkers” section, page 10 lines 253-261, and 265-268, though we did not collect data on multiple comparisons correction and head motion correction methods since the scope of the present manuscript is already very broad.

25. Based on the definition of ‘predictive’ biomarkers in Table 1 (“existence or intensity of the biomarker reflects the propensity of individuals to experience favorable or unfavorable

effects"), it is not clear that the example studies necessarily meet this definition. For example, the Bach 2021 paper describes a median split of data that is then used to predict relapse. Again, some discussion of the methodological limitations of prior work seems missing here. This is not to criticize the existing studies but to provide appropriate context.

We appreciate the reviewer's highlighting of this issue. A principal aim of the present manuscript is the explication of challenges in developing FDCR-derived biomarkers. We have now added a discussion of this limitation in the context of "Validation of FDCR biomarkers" section on page 10, lines 253-257, and online-only materials eFigure 7, page 18, lines 303-304.

26. Including subtyping studies under 'diagnostic' seems overly simplistic (especially as the subtypes described, e.g., 'relief drinking' are not themselves diagnostic categories). Were these studies actual data-driven subtyping studies or just comparisons of individuals with or without a given characteristic?

    We use the term "diagnostic" for this class of biomarkers in line with the cited FDA-NIH handbook from which we have adapted this category, and have now endeavored to further clarify its scope in the revised manuscript in the "Contexts of use of FDCR biomarkers" section on page 8, lines 178-183. The diagnostic studies in our database have all essentially conducted statistical comparisons of the FDCR signal between participant groups defined a priori, though in principle, researchers could start from the other end, i.e., with data-driven identification of "neurotypes" using the fMRI data.

27. Figure 6 only describes findings from studies highlighted in Table 1, which is not ultimately very helpful in the context of the much larger number of studies included in the review. It is also potentially misleading, as it seems to imply to that, for example, the amygdala is important for severity but not for prognosis. This is particularly problematic if some of the studies used to create this figure did not use whole-brain approaches (e.g., may not have tested the amygdala at all).

    We agree with the reviewer's comment and can see how this figure might be viewed as a summation of "what is known" rather than a visual illustration of exemplar findings. The figure is now moved to online-only materials (eFigure 7) with further explanation in its caption to avoid any misreading.

28. The authors note that the studies reviewed 'may have included partially or fully overlapping samples'. This is a pretty big limitation. It is best practice for authors to report when they are publishing a re-analysis of a prior sample, so presumably it would have been possible to remove such studies or to at least quantify how many studies have overlapping samples?

    We agree that this is an important limitation when estimating the number of participants across FDCR studies, and that our provided numbers are effectively an "upper bound". We did seek to address this issue by having one rater review the methods sections of all included studies and no sample duplications were discovered in the database based on

what study authors have explicitly reported in papers. This is now noted in the online-only “eResults” section, page 15, lines 247-248.

29. Figure 4 is very hard to follow. Suggest creating a table or different figure other than a Sankey diagram.
    We agree with this concern (also raised by another reviewer). The figure may be confusing since it’s an attempt at summarizing much of the data in a single visualization. We have moved the figure to online-only materials (eFigure 6) and the information of this figure is further discussed in the “Participants in FDCR studies” section in online-only materials.

30. Inclusion of individuals with behavioral addictions needs better justification and it would also be helpful to see these studies analyzed separately.
    Thanks to the reviewer for raising this important issue. To make our systematic review as broad as possible, we decided to include all the addiction types that are defined in the DSM 5, including gambling disorder and internet gaming disorder. We do acknowledge its importance and have included a discussion in the “Methods and Results” section, page 6, lines 106-110, and online-only eMethods page, lines 61-65. Furthermore, we have added eFigure 10 presenting our central analyses for behavioral addiction studies only.

31. In Fig. 4, are all “AD” individuals those with a confirmed DSM diagnosis or are some of these individuals who were included on the basis of a rating scale (as indicated in the methods)?
    These are all individuals diagnosed with an addictive disorder based on widely accepted diagnostic criteria, such as those outlined in the DSM or ICD manuals. Individuals who use substances but are not formally diagnosed with an addictive disorder are termed “Users”. Further details are added to the legend of eFigure 6, page 16, lines 256-259.

BYLINE

Addiction Cue Reactivity Initiative (ACRI) Group

LIST OF AUTHORS

Arshiya Sangchooli*, MD, Mehran Zare-Bidoky*, MD, Ali Fathi Jouzdani, MD, Joseph Schacht, PhD, James M. Bjork, PhD, Eric Claus, PhD, James J Prisciandaro, PhD, Stephen J. Wilson, PhD, Torsten Wüstenberg, PhD, Stéephane Potvin, PhD, Pooria Ahmadi, MD, Patrick Bach, PhD, Alex Baldacchino, MD.PhD, Anne Beck, PhD, Kathleen Brady, MD.PhD, Judson Brewer MD.PhD, Anna Rose Childress, PhD, Kelly E Courtney, PhD, Mohsen Ebrahimi, MSc, Francesca M Filbey, PhD, Hugh Garavan, PhD, Dara G Ghahremani, PhD, Rita Z Goldstein, PhD, Annaeke E Goudriaan, PhD, Erica N Grodin, PhD, Colleen A. Hanlon, PhD, Amelie Haugg, PhD, Markus Heilig, MD.PhD, Andreas Heinz, MD.PhD, Adrienn Holczer, PhD, Ruth J Van Holst, PhD, Jane E. Joseph, PhD, Anthony Juliano, PhD, Marc J. Kaufman, PhD, Falk Kiefer, PhD, Arash Khojasteh Zonooni, MD, Rayus Kuplicki, PhD, Marco Leyton, PhD, Edythe D. London, PhD, Scott Mackey, PhD , F. Joseph McClernon, PhD, William Mellick, PhD, Kirsten Morley, PhD, Hamid R Noori, PhD, Mohammad Ali Oghabian, PhD, Jason A. Oliver, PhD, Max Owens, PhD, Martin Paulus, MD.PhD, Irene Perini, PhD, Parnian Rafei, MSc, Lara Ray, PhD, Rajita Sinha, PhD, Michael N. Smolka, MD.PhD, Ghazaleh Soleimani, PhD, Rainer Spanagel, PhD, Vaughn Steele, PhD, Susan F. Tapert, PhD, Sabine Vollstäedt-Klein, PhD, Reagan R Wetherill, PhD, Katie Witkiewitz, PhD, Kai Yuan,PhD, Xiaochu Zhang, PhD, Antonio Verdejo-Garcia, PhD, Marc N. Potenza, MD.PhD, Amy C. Janes, PhD, Hedy Kober, PhD, Anna Zilverstand, PhD, Hamed Ekhtiari, MD.PhD

AFFILIATIONS

Authors Information

Department of Psychiatry & Behavioral Sciences, University of Minnesota, Minneapolis, MN, USA (Zare-Bidoky,Zilverstand, Soleimani, Ekhtiari); Laureate Institute for Brain Research, Tulsa, OK, USA (Kuplicki, Paulus, Ekhtiari); University of Melbourne, Australia (Sangchooli); Department of Psychiatry, University of Colorado School of Medicine, Aurora, CO, 80045, USA (Schacht); Department of Psychiatry, Institute for Drug and Alcohol Studies, Virginia Commonwealth University, Richmond, VA, USA (Bjork); Department of Biobehavioral Health, The Pennsylvania State University, University Park, PA, USA (Claus); Department of Psychiatry and Behavioral Sciences, Addiction Sciences Division, Medical University of South Carolina, Charleston, SC, USA (Prisciandaro, Brady, Mellick); Department of Psychology, The Pennsylvania State University, State College, Pennsylvania, USA (Wilson); Heidelberg University, Field of Focus IV, Core Facility for Neuroscience of Self-Regulation (CNSR), Heidelberg, Germany (Wüstenberg); Department of Psychiatry and Addiction, Université de Montréal, Montréal, Canada (Potvin); Department of Addictive Behaviour and Addiction Medicine, Central Institute of Mental Health (CIMH), Medical Faculty Mannheim, Heidelberg University, Germany (Bach, Kiefer, Vollstädt-Klein); Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada (Ahmadi); Faculty of Health, Health and Medical University, Potsdam, Germany (Beck); School of Medicine, University of St Andrews, St Andrews, Scotland, UK (Baldacchino); Department of Psychiatry and Neurosciences, Charité Campus Mitte, Charité-Universitätsmedizin Berlin, Berlin, Germany (Beck, Heinz); Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, USA (Brewer); Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA (Childress, Wetherill); Department

of Psychiatry, University of California, San Diego, La Jolla, CA, USA (Courtney, Tapert); Center for BrainHealth, School of Behavioral and Brain Sciences, University of Texas at Dallas, Dallas, TX, USA (Filbey); Departments of Psychiatry, University of Vermont, Burlington, VT, USA (Garavan, Juliano, Mackey, Potenza); Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, USA (Ghahremani, Grodin, London, Ray); Departments of Psychiatry & Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, USA (Goldstein); Department of Psychiatry, Amsterdam University Medical Center, University of Amsterdam, the Netherlands and Amsterdam Neuroscience, Amsterdam, The Netherlands (Goudriaan); Department of Cancer Biology, Wake Forest School of Medicine, Winston-Salem, NC, USA (Hanlon); Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric University Hospital Zurich, University of Zurich, Zürich, Switzerland (Haugg); Center for Social and Affective Neuroscience, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden (Heilig, Perini); Department of Neurology, Albert Szent-Györgyi Health Centre, University of Szeged, Hungary (Holczer), Amsterdam Institute for Addiction Research, Amsterdam UMC, Department of Psychiatry, University of Amsterdam, The Netherlands (Holst); McLean Hospital, Harvard Medical School, Belmont, MA, USA (Kaufman), Iranian National Center for Addiction Studies (INCAS), Tehran University of Medical Sciences, Iran (Fathi Jouzdani, Khojasteh Zonooni, Ebrahimi, Rafei); Department of Neuroscience, Medical University of South Carolina, Charleston, SC, USA (Joseph); Cognitive and Pharmacological Neuroimaging Unit, National Institute on Drug Abuse, Baltimore, MD 21224, United States (Janes); Department of Psychiatry, McGill University, Montreal, QC, Canada (Leyton); Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA (McClernon); Specialty of Addiction Medicine, Sydney Medical School, Faculty of Medicine and Health, University of Sydney, NSW, Australia (Morley); McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA (Noori); Neuroimaging and Analysis Group, Research Center for Molecular and Cellular Imaging, Tehran University of Medical Sciences, Tehran, Iran (Oghabian); TSET Health Promotion Research Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA (Oliver); Department of Psychiatry, University of Vermont, Burlington, VT, USA (Owens); Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA (Sinha, Steele, Kober); Department of Psychiatry, Technische Universität Dresden, Dresden, Germany (Smolka); Institute of Psychopharmacology, Central Institute of Mental Health, Mannheim, Germany (Spanagel); Department of Psychology, University of New Mexico, Albuquerque, NM, USA (Witkiewitz); School of Life Science and Technology, Xidian University, Xi'an, China (Yuan); Department of Psychology, School of Humanities and Social Science, University of Science and Technology of China, Anhui, China (Zhang); Turner Institute for Brain and Mental Health, Monash University, Clayton, Australia (Verdejo-Garcia);

* Arshiya Sangchooli and Mehran Zare-Bidoky contributed equally to this work.

Supplemental Online Content

Addiction Cue-Reactivity Initiative (ACRI) Network. Parameter space and potential for biomarker development in 25 years of fMRI drug cue reactivity: a systematic review. *JAMA Psychiatry*. Published online February 7, 2024. doi:10.1001/jamapsychiatry.2023.5483

This supplementary material has been provided by the authors to give readers additional information about their work.

eBox. Biomarkers in Psychiatry and Addiction Medicine

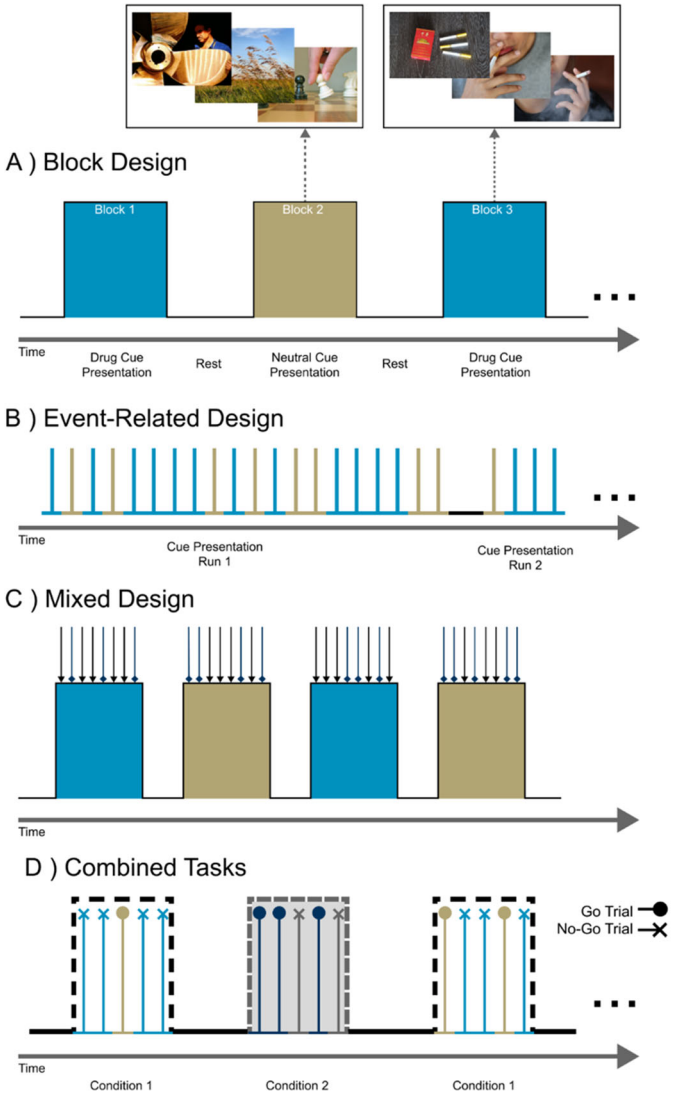
The FDA-National Institutes of Health (NIH) Biomarker Working Group defines a "biomarker" as "a defined characteristic measured as an indicator of normal or pathogenic biological processes, or biological responses to an exposure or intervention, including therapeutic interventions" [BEST (Biomarkers, EndpointS, and other Tools) Resource[1]]*. The development of clinically relevant biomarkers is a major goal of addiction neuroscience and translational psychiatry. Regulating agencies have shown increasing interest in validated biomarkers, with the FDA's biomarker qualification program, among others, working to provide formal endorsement of biomarkers to facilitate their use in drug development and regulatory decisions[2]. Recent reviews and opinions have outlined the potential for an expanding group of central and peripheral biomarkers of major psychiatric conditions, including genomic, epigenetic, and transcriptomic biomarkers[3], proteomic biomarkers[4], inflammatory markers[5], non-inflammatory chemokines[6], cardiovascular biomarkers[7], hormonal and neurotransmitter profiles[8], cognitive and behavioral markers[9], biomarkers derived from neuroimaging paradigms[10,11], and multi-modal biomarkers[12]. Several neuroimaging biomarkers are also at varying stages of validation by the FDA for neurological or psychiatric disorders. These include baseline hippocampal volume assessed by Magnetic Resonance Imaging (MRI) in Alzheimer's disease and Glx (Glutamine+ Glutamate) measured in the brain by Magnetic Resonance Spectroscopy (MRS) in depression. Notably, the NIMH "Fast-Fail" trial initiative supports the use of functional MRI (fMRI) in early-phase drug development to lower the risk of failure in large clinical trials: in the first implementation of the approach, task-related fMRI revealed that kappa opioid receptor antagonism can enhance reward-related ventral striatal activation, supporting larger trials for cross-diagnostic treatment of anhedonia[13,14].

Commensurate with broader progress in biomarker development across various psychiatric disorders, different types of brain-based markers with potential for clinical translation have been proposed for addictive disorders, but their clinical and analytical validation remains limited[15]. Objective biological metrics of SUDs are currently limited to measures of substance use - mainly testing for psychoactive substances or their metabolites in biological samples[16,17] - or measures that reflect the toxic effects of use[18]. Notably, these biomarkers reflect endpoints of substance use and toxicity and are not informed by the dynamic processes that underlie how drug use behaviors relate to addiction. This limitation hampers the clinical use of intermediate phenotypes and the development of biomarkers to identify at-risk individuals and to mechanistically inform, predict, and monitor interventions[19]. Relatedly, although the DSM-5 proposed diagnostic criteria for behavioral addictions (BAs), including gambling disorder and internet gaming disorder, no biomarkers are included for BAs[20]. According to the FDA website (visited December 15th, 2020), there are no qualified biomarkers or ongoing qualification processes for biomarkers in addiction medicine/psychiatry[21]. The only submitted

biomarker qualification proposal - covering a single nucleotide polymorphism in the delta opioid receptor 1 gene - appears to have been rejected at an early phase.

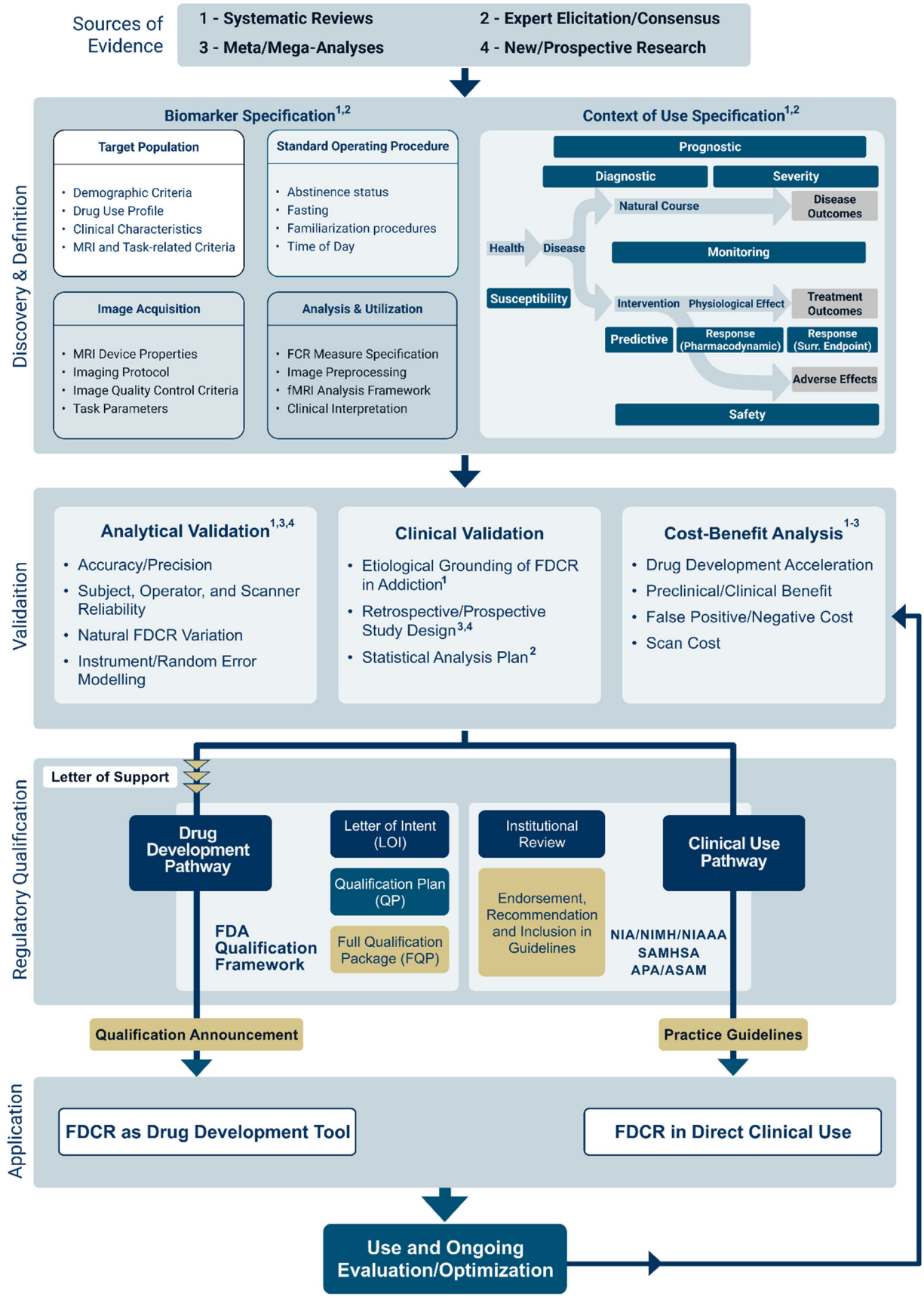
Reactivity to addiction-relevant cues

A popular paradigm to assess brain function in individuals with SUDs is data acquisition with fMRI during the administration of a drug cue-reactivity task[22]. Similar paradigms have been developed to investigate reactivity to addiction-relevant cues in BAs[23–25]. These paradigms (referred to collectively as "cue-reactivity paradigms") involve the presentation of a variety of conditioned cues, associated with the availability or use of substances or other similarly desirable experiences to participants who have had prior experiences with them. The cue-reactivity paradigm rests on the understanding that addictive disorders involve sensitization to addiction-relevant cues[26], which can trigger behavioral and physiological responses associated with craving and anticipation[27]. Cue-reactivity tasks had been developed and validated extensively before the advent of fMRI and have been readily modified and adopted in fMRI research[28,29]. Engagement with addiction-relevant cues under fMRI scanning enables the exploration of the neural mechanisms that are associated with the response to addiction-relevant cues[30], and fMRI drug cue-reactivity (FDCR) has demonstrated that SUDs are associated with aberrations in the neural circuitry underpinning incentive salience, reward evaluation, interoception, memory, habit formation, and executive control[31,32]. If variations in FDCR signal are associated with the existence and severity of addiction-related processes, the development of FDCR-derived biomarkers could aid in diagnostic classification and sub-grouping, assessing disease severity, identifying at-risk individuals, understanding the neural mechanisms involved in effective interventions, targeting disrupted neural function with novel interventions, early evaluation of new interventions based on surrogate endpoints such as target-engagement, and monitoring treatment effectiveness[13,32–34]. More recent avenues of research have combined cue-reactivity with other paradigms during fMRI acquisition[35–37] and investigated the interaction of FDCR and genomic[38], epigenetic[39], metabolic[40], physiological[41], developmental[42], behavioral[43], cognitive[44–47], personality[48] and psychiatric[49,50] correlates of addiction. Considering the multi-faceted and multi-causal nature of these disorders and their frequent co-occurrence with other mental and physical health conditions, such studies establish the etiological importance of FDCR in SUDs and lead to better characterizations of addictive processes, ultimately enabling the development of multi-domain biomarkers[51,52] For example, neuro-genetics studies have shown that the A118G single nucleotide polymorphism of the mu opioid receptor (OPRM1) gene may result in higher levels of FDCR[53] (Ray et al., 2014) and also impact the clinical response to naltrexone (a μ-opioid antagonist medication)[54].



eFigure 1. Overview of fMRI Drug Cue Reactivity Task Designs

Overview of fMRI Drug Cue Reactivity Task Designs A) Cues/Stimuli are presented in groups or "blocks" containing a series of similarly conditioned cues which are then separated by a delay from the next block. B) Stimuli are presented in succession with or without a delay, without being arranged by their type. Tasks may have a few sections or "runs" where a delay separates each run from the next without the participant exiting the scanner. C) Mixed design tasks may borrow elements (like grouping or sequence randomization) from either a block design or an event-related design with the addition of another set of changing conditions or events that occur concurrently with the task D) Combined tasks use cue-reactivity concurrently with another cognitive task (e.g., Go/No-Go task). Designs can incorporate stimuli presented in various modalities, including visual (static or dynamic), auditory, olfactory, or tactile.



eFigure 2. Four Major Steps in the Validation of Potential fMRI Drug Cue Reactivity–Derived Biomarkers

Initially, a context of use for an FDCR-derived biomarker is specified and the potential biomarker is precisely defined. Following analytical and clinical validation and cost-benefit analysis, the compiled evidence is presented for regulatory approval. The FDA evaluates the use of biomarkers for drug development through a biomarker qualification process involving submission of a Letter of Intent, a Qualification Plan, and a Full Qualification Package, though a Letter of Support may be issued by the FDA to indicate its support for a biomarker before formal qualification. The use of FDCR-derived biomarkers in clinical contexts requires the endorsement of a constellation of other institutions. Surr. Endpoint: Surrogate Endpoint.

eMethods.

The methods section is organized based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. The protocol for this systematic review was pre-registered[55]. While we refer to fMRI "drug" cue-reactivity (including alcohol) throughout the manuscript, behavioral addiction studies focusing on problematic videogame playing or gambling were not excluded as they constitute a small portion of the cue-reactivity literature and involve cue-reactivity paradigms similar to drug cue-reactivity studies. Since behavioral addictions (BAs) have recently been added to the widely used nomenclature system, and the pathophysiology may not be completely the same as SUDs, one should be cautious in analyzing these data together. Therefore, information on BAs can be seen separately in our database for future use.

Eligibility criteria: Original studies were selected according to the criteria outlined below.

Study design and methodology: We employed a broad perspective in the inclusion of studies, including all types of original research (e.g., basic research, observational studies, and interventional studies). Only peer-reviewed studies were included.

Of interest were original studies that included one or more fMRI-based investigations as a major part of their methodology, whether as an observational or as a treatment tool (e.g. in fMRI neurofeedback). For at least some of the study population, the fMRI investigation had to include a cue-reactivity task, including the presentation of substance- or problematic behavior-related cues and at least one other class of cues (i.e., neutral or non-substance-related) for comparison. Acute challenge studies involving direct administration of substances of use were not included unless cues associated with the substance/object of use were explicitly presented as well. Cues also had to be ecologically valid; i.e., they needed to be associated with routine drug-taking behaviors and not be novel conditioned cues associated with the substance/object of use for the first time during the experiment. We excluded studies that did not provide details about the fMRI protocol, setting and tasks, outcomes of interest used in the analysis, and basic fMRI measures. There were no further exclusions, and both whole-brain and region of interest (ROI)-based fMRI studies were included.

Participants: Every study required at least one human population or sub-population with more than one member, for which at least one of the following needed to be true:

At least one circumscribed group of participants had a diagnosis of at least one SUD or BA, either manifest as active use or in remission; with the diagnosis made either before the study, as part of the study protocol during the investigation, or by the end of the study (i.e., with the diagnosis serving as an outcome measure).

At least one group of participants was included explicitly because they regularly consumed a potential object of addiction (substance or behavior) and/or had a risky pattern of consumption that might lead to addiction, and the study focused on their reactivity to cues of that substance or behavior.

At least one group of participants had been assigned a score for an addiction-related phenomenon (such as addiction or drug-use severity) with or without an explicit diagnosis of an SUD or BA, and the relationship of this score to important outcomes in the study had been investigated.

No restriction was placed on study participants based on demographic, ethnic, biological, or clinical factors (such as any co-occurring disorders).

Language: Only publications with their full text in English were included.

Information source: Existing research was identified and retrieved using PubMed. Relevant articles were identified using a comprehensive search strategy for all terms related to addiction, fMRI, and cue-reactivity, as detailed below.

Search strategy: Considering the subject of the review, a list of three sets of keywords was compiled (eTable 1). These terms were adapted for use in PubMed (exact search syntax and search results are outlined in eTable 2). The first set included synonyms of “functional magnetic resonance imaging”, the second included terms related to cue-reactivity, and the third included synonyms of “addiction” and various terms related to SUDs and BAs and addiction medicine. To help widen the search, no filters were used. The exclusion of systematic reviews and other non-original research and the application of other inclusion/exclusion criteria were handled manually. Given the large volume of relevant literature on PubMed, other search engines or grey literature were not used.

eTable 1. Search Terms Used for this Systematic Review

fMRI	"functional MRI"
"functional magnetic resonance"	"cue-reactivity"
"cue exposure"	"craving"
"cue induced"	"drug cue"
"drug cues"	
addict*	dependence
"substance use"	"substance abuse"
"drug abuse"	"drug use"
nicotine	smoker
tobacco	opioid
opiate	heroin
marijuana	cannabis
"thc"	alcohol*
cocaine	amphetamine
methamphetamine	"behavioral addiction"
"internet addiction"	"problematic gaming"
"gaming disorder"	"gambling disorder
"problem gambling"	
fMRI search terms	1 OR 2 OR 3
Cue-reactivity search terms	4 OR 5 OR 6 OR 7 OR 8 OR 9

Addiction search terms	10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34
Final search	35 AND 36 AND 37

eTable 2. Final Syntax of PubMed Search, and Number of Raw Search Results

Term Group	Search	Number of results on 5 Jan 2023
fMRI search	fMRI OR "functional MRI" OR "functional magnetic resonance"	573243
Cue-reactivity search	"cue reactivity" OR "cue exposure" OR craving OR "cue induced" OR "drug cue"	9841
Addiction search	addict* OR dependence OR "substance use" OR "substance abuse" OR "drug abuse" OR "drug use" OR nicotine OR smoker OR tobacco OR opioid OR opiate OR heroin OR marijuana OR cannabis OR "THC" OR alcohol* OR cocaine OR amphetamine OR methamphetamine OR "behavioral addiction" OR "behavioral addiction" OR "internet addiction" OR "problematic gaming" OR "gaming disorder" OR "gambling disorder" OR "problem gambling"	1424082
Final search	(fMRI OR "functional MRI" OR "functional magnetic resonance") AND ("cue reactivity" OR "cue exposure" OR craving OR "cue induced" OR "drug cue" OR "drug cues") AND (addict* OR dependence OR "substance use" OR "substance abuse" OR "drug abuse" OR "drug use" OR nicotine OR smoker OR tobacco OR opioid OR opiate OR heroin OR marijuana OR cannabis OR "THC" OR alcohol* OR cocaine OR amphetamine OR methamphetamine OR "behavioral addiction" OR "behavioral addiction" OR "internet addiction" OR "problematic gaming" OR "gaming disorder" OR "gambling disorder" OR "problem gambling")	952

Study records

Data management: Literature search results were imported to Excel. Screening of articles for relevance was performed by reviewing the title and abstract sections of candidate texts, and full texts were obtained for studies that passed preliminary screening.

Study selection: Screening forms were developed for title/abstract and full-text assessment and studies were checked by two authors (MZB and AS). The authors initially checked the eligibility of fifty randomly chosen studies under the supervision of the corresponding author (HE) as a calibration exercise to ensure eligibility criteria were applied consistently [56]. After title and abstract screening, the two authors screened the full texts of papers that either met the eligibility criteria or had an uncertain status. Any papers with an uncertain eligibility status after full-text screening were then discussed with HE until a consensus on their inclusion was reached. Reasons for the exclusion of articles at the title and abstract or full-text screening stages were recorded, according to the PRISMA framework[57]. Neither of the review authors was blind to the journal titles, study authors, or institutions.

Data collection: Data were filled into a spreadsheet by PA, AFJ, AH, and AKZ. Consistency between the authors was honed through a calibration exercise in which all authors evaluated and discussed their ratings for 20 randomly chosen studies [56]. AS, MZB and HE further refined the data extraction form to reduce inconsistency and ambiguity after the exercise. Data on study design features and basic methodological parameters were extracted first, and each article was reviewed independently by two authors in two separate spreadsheets, with inconsistencies resolved in discussions with MZB and AS with HE's supervision. To check whether any study samples overlapped with other studies (e.g. in the case of re-analysis studies), a single rater (AFJ) screened the methods sections of all studies.

Data items: We extracted publication details, publication country (where the first affiliation of the first study author is located or the affiliation of the majority of the authors in case country was not clear), publication year (based on PubMed's indexing), the substance or behavior (main substance(s) and/or behavior(s) of interest in the study), main experimental task design type (whether cues were presented in blocked, event-related, or mixed forms), stimulus type (sensory modality of cues), combined tasks (whether cue-reactivity was paired with other tasks; and what tasks were used), task duration (seconds, excluding other paradigms that may have been implemented in the scanner), study sample characteristics (number of participants of each sex; number of participants with untreated or treated addictive disorders, drug-using individuals who did not meet SUD criteria, individuals in long-term abstinence, and healthy non-using participants), intervention (if included, type of intervention), association with a future event (a non-fMRI variable measured at a later point in time based on fMRI results), number of fMRI sessions (times each participant was scanned), and interval between fMRI sessions (if participants were scanned more than once for a study, the average time interval between the scans). Yes/No ratings were used to classify whether the design of each FDCR study allowed for it to be potentially used to develop susceptibility, diagnostic, response, prognostic, predictive, or severity biomarkers for one or more SUDs/BAs. Yes/No ratings were also used to specify whether a study investigated relationships between FDCR-derived parameters and subjective craving,

demographic variables, behavioral measures, biochemical assays, participant genetics, non-FDCR structural or functional neural markers, physiological parameters, or psychiatric assessments. For each study investigating use of FDCR as a biomarker type or assessing FDCR correlates, it was also rated as to whether significant test results were observed. However, we elected to use the relatively simple metric of "significance" given the extreme heterogeneity of analyses and reported statistics in the field, which would complicate further quantitative synthesis. The scope of this work is to provide an overview of the status of the field and address the current heterogeneities to provide a roadmap to support the development of evidence that can be used in higher quality quantitative metrics in the future.

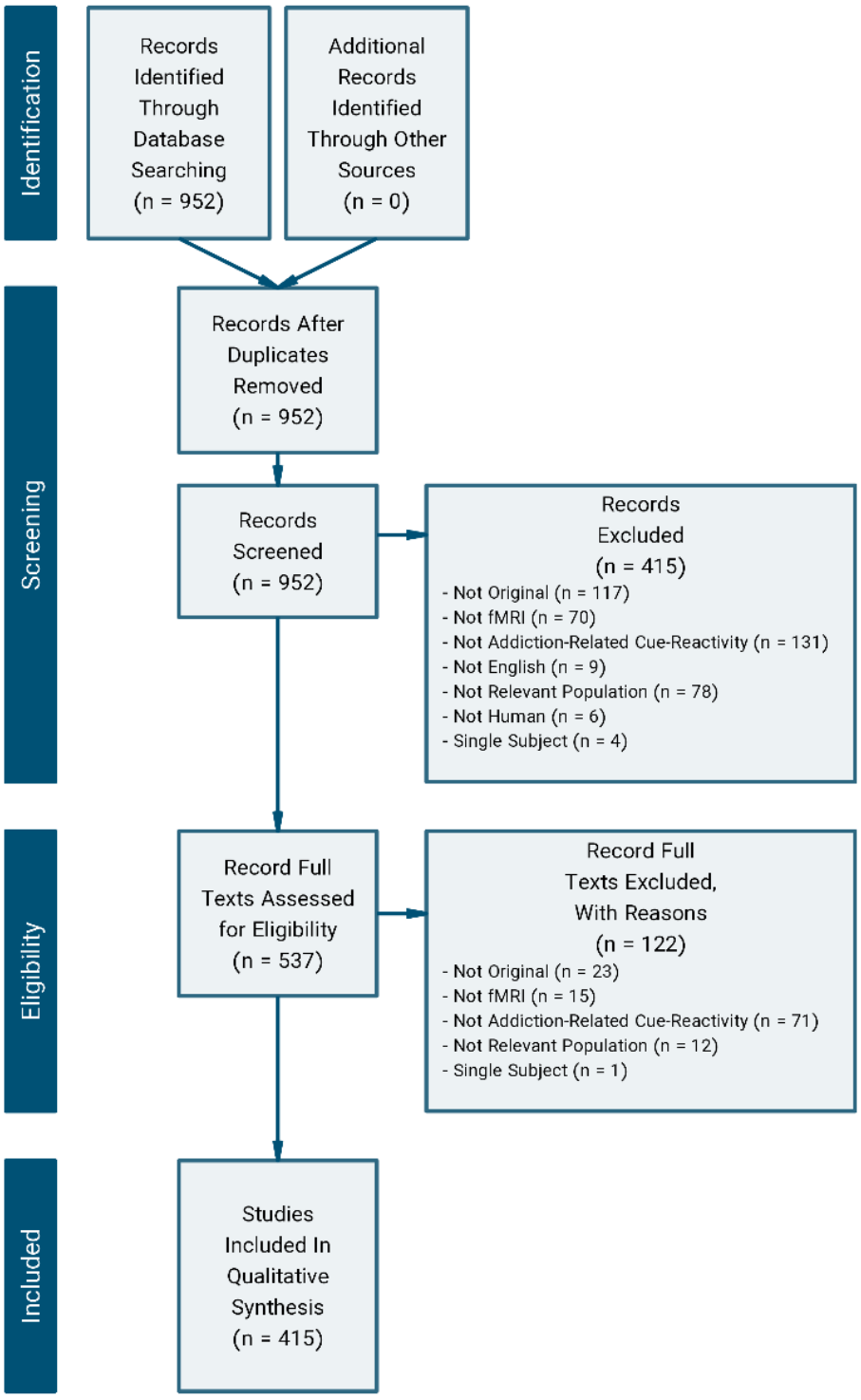
Software

The PubMed search engine from the National Library of Medicine's online portal (www.ncbi.nlm.nih.gov/pubmed/) was used to conduct the search. Zotero[58] was used for reference management. Google Sheets from Google's Google Docs Editors suite was used to design tables for data extraction and sharing among authors. Data analyses and illustrations were conducted using R version 4.0.5[59].

The protocol for this systematic review was developed throughout 2019 and was first registered on the Open Science Framework (OSF) website on May 18th, 2020. The current extracted database is available publicly in the OSF page (https://osf.io/eb972/). As this is an ongoing systematic review, we recommend viewing the OSF page of this project for the latest developments and updates[55].

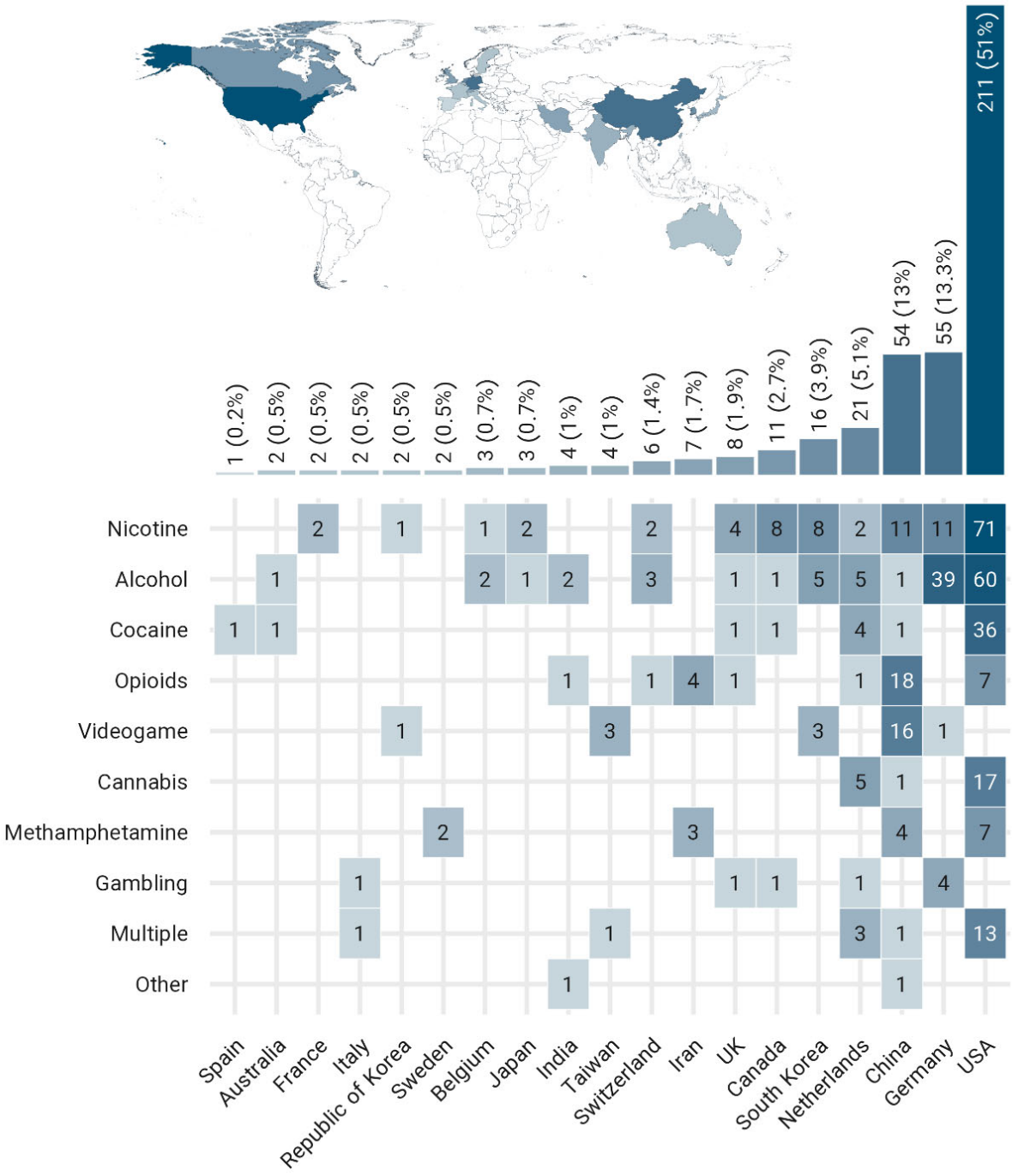
eResults.

The search was performed on January 5, 2023, yielding 952 results. Of these, 415 were excluded at the title-and-abstract screening stage, and 122 were excluded after full-text screening, yielding a total of 415 FDCR publications that were included in the data extraction phase of the systematic review. The PRISMA flowchart is presented in eFigure 3. Most studies are from the US (51.0%) followed by Germany (13.3%) and China (13.0%) (eFigure 4). A breakdown of papers by the substance or behavior of interest shows that most studies have been conducted on various forms of either nicotine (29.6%), alcohol (29.2%), or cocaine (11.1%) use/use disorders, overall accounting for 69.9% of the papers in the database. Earlier studies in the database were all focused on cocaine and alcohol, with the first studies on cannabis and video games published in 2009 and the first on methamphetamine published in 2012 (eFigure 5). There is an overall yearly increase in the number of FDCR studies, with the vast majority of studies (303, 74.0 %) published in the last 10 years.



eFigure 3. PRISMA Flowchart

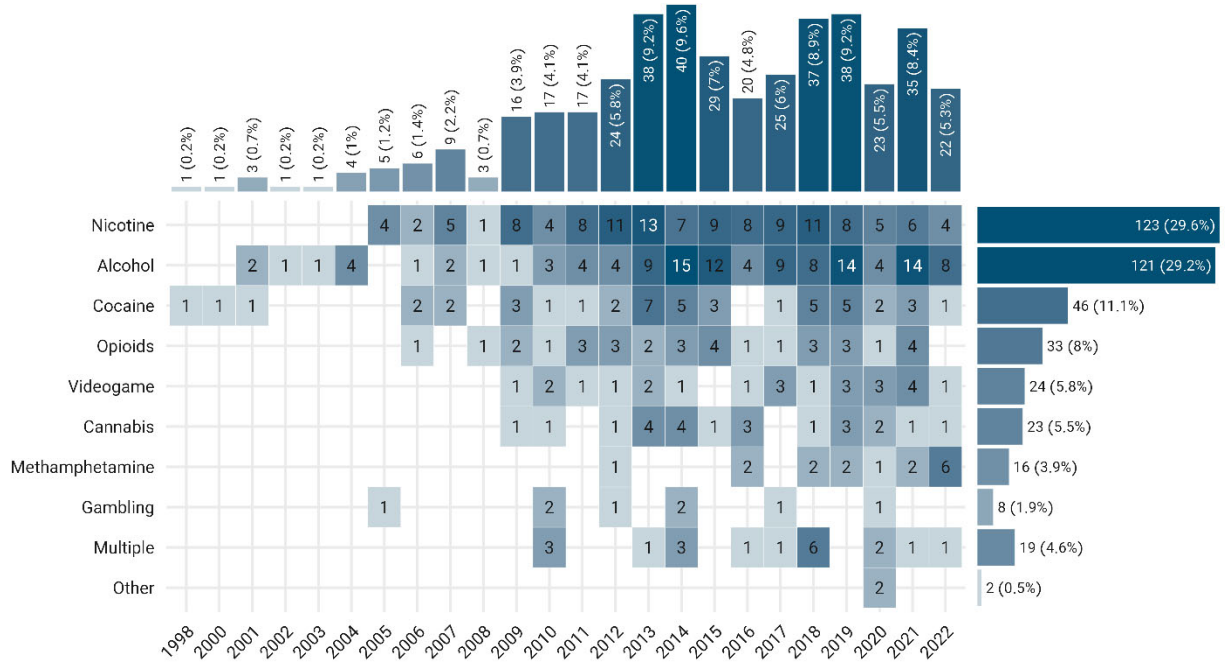
The titles and abstracts of 952 records from the start of 1998 until the end of 2022 were screened, and 415 were excluded during preliminary screening. The full texts of 537 records were extracted and assessed for eligibility. Ultimately, 415 records were included in the systematic review.



eFigure 4. Global Contribution to fMRI Drug Cue Reactivity (FDCR) Studies

Number of FDCR studies in each country, broken down by the type of addictive substance/behavior. “Multiple” stands for studies including more than one type of addictive substance/behavior. The "other" category includes inhalants and betel-quid. Note that only

papers whose full-text was in English were included, potentially leading to a relative over-representation of majority English-speaking countries.

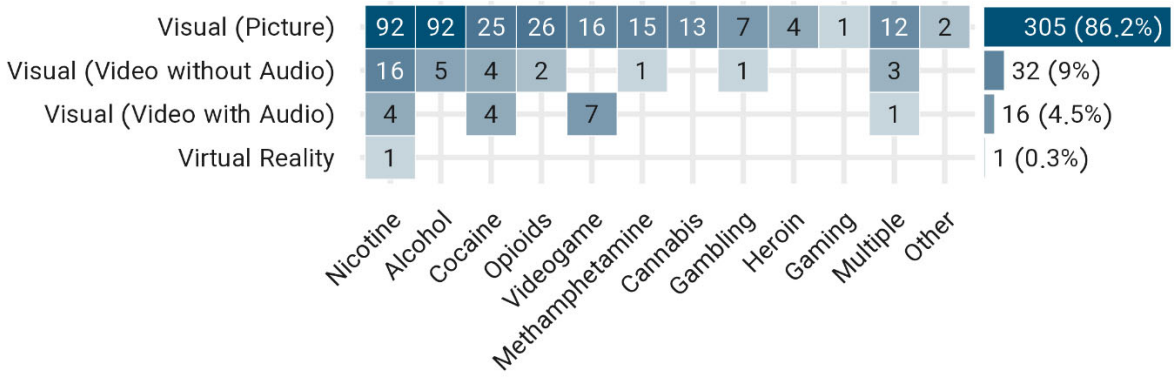


eFigure 5. fMRI Drug Cue Reactivity Studies (1998-2022)

Number of FDCR studies each year from 1998 till the end of 2022, broken down by the type of addictive substance/behavior. “Multiple” stands for those studies including more than one type of addictive substance/behavior. The "other" category includes inhalants and betel-quid.

Study and task design

Most FDCR studies scanned participants at a single time point (78.8%). For the 88 studies with more than one scanning time point, the median inter-scan interval was 14 days, though a relatively wide distribution was observed (IQR = 21) (Figure 1b). The vast majority of studies (85.3%) used visual stimuli (for a detailed breakdown, see eFigure 6), with a minority using other stimulus types such as semantic (2.7%), gustatory (2.2%), auditory (1.7%), olfactory (1.2%), and imaginary (1%) stimuli. Another 25 multi-sensory studies (6%) used various combinations of stimuli (Figure 1f). Cues have been commonly presented in a block (61.9%) or event-related (36.9%) design, with only 1.2% of studies using other designs or both event-related and blocked-design FDCR tasks within a single study (Figure 1c). The median FDCR task duration was 720 seconds (IQR = 800) (Figure 1d), and 52 FDCR studies used combined FDCR tasks: these are tasks in which the presentation of addiction-relevant cues is paired with another concurrent task component to probe cognitive functions such as response inhibition (32.7% of the 52 studies), interference resolution (25.0%), attention (13.5%), decision-making and reward processing (11.5%), perception (7.7%), working memory (5.8%), or approach/avoidance (3.8%) (Figure 1f).

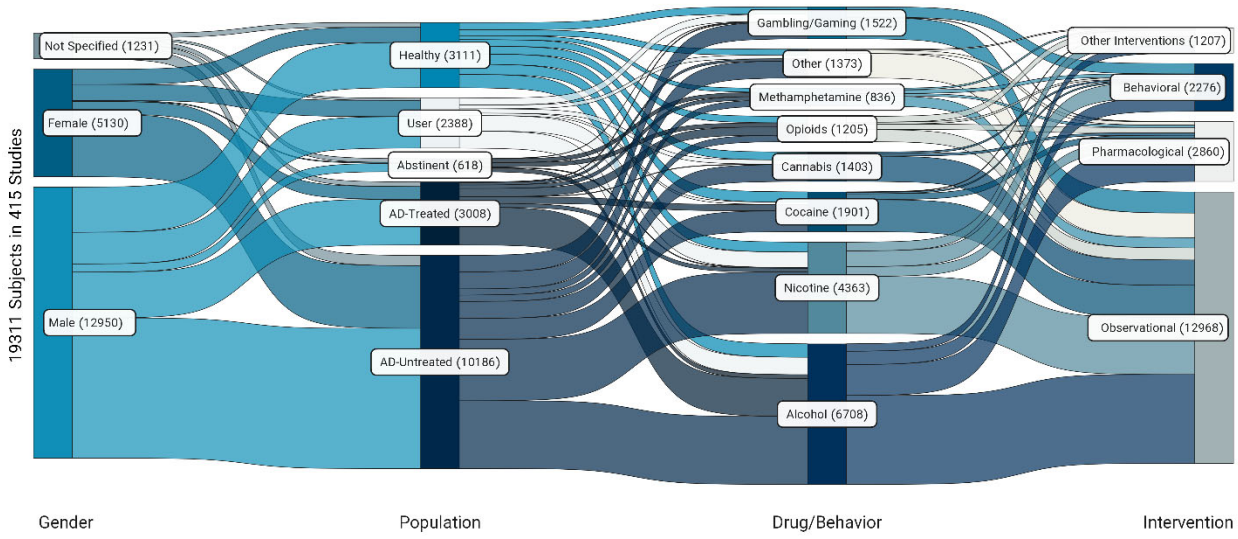


eFigure 6. Breakdown of Visual Cues

Among 354 sets of visual cues used in FDCR studies, they are broken down into pictures and videos with audio and without audio.

Participants in FDCR studies

Overall, 19,311 individuals participated in FDCR studies from 1998-2022. Of these, 12,950 were male (66.1%) and 5,130 were female (26.5%), with the sex of 1231 participants (6.4%) not explicitly specified. The median sample size of FDCR studies was 37. The 19,311 participants can be divided into 10,186 individuals with untreated addictive disorders (52.7%), 3,008 individuals with addictive disorders undergoing treatment at recruitment (15.6%), 2,388 individuals who used potentially addictive substances or engaged in potentially addictive behaviors without necessarily meeting addictive disorder criteria (12.4%), 618 individuals in long-term abstinence (3.2%) and 3,111 participants (16.1%) who were not using substances (i.e., “healthy controls”). A plurality of the participants (6708, 34.7%) were recruited to investigate alcohol use/use disorders with the following statistics for other use/use disorders: nicotine (4363, 22.6%), cocaine (1901, 10.0%), cannabis (1403, 7.2%), opioid (1205, 6.2%), and methamphetamine (836, 4.3%). Of the remaining participants, 1373 (7.1%) used betel-quid, inhalants, or multiple substances, and 1522 (7.9%) were recruited in studies focusing on gambling or video game playing. While most participants (13037, 67.5%) were recruited in observational studies, a substantial portion participated in trials or experimental studies involving pharmacological (2897, 15.0%), behavioral (2257, 11.7%), or other interventions (1120, 5.8%), such as neurofeedback or non-invasive brain stimulation (eFigure 7). No duplicated samples across studies were discovered in the database based on a screening.



eFigure 7. Participants in fMRI Drug Cue-Reactivity Studies (N = 19 311)

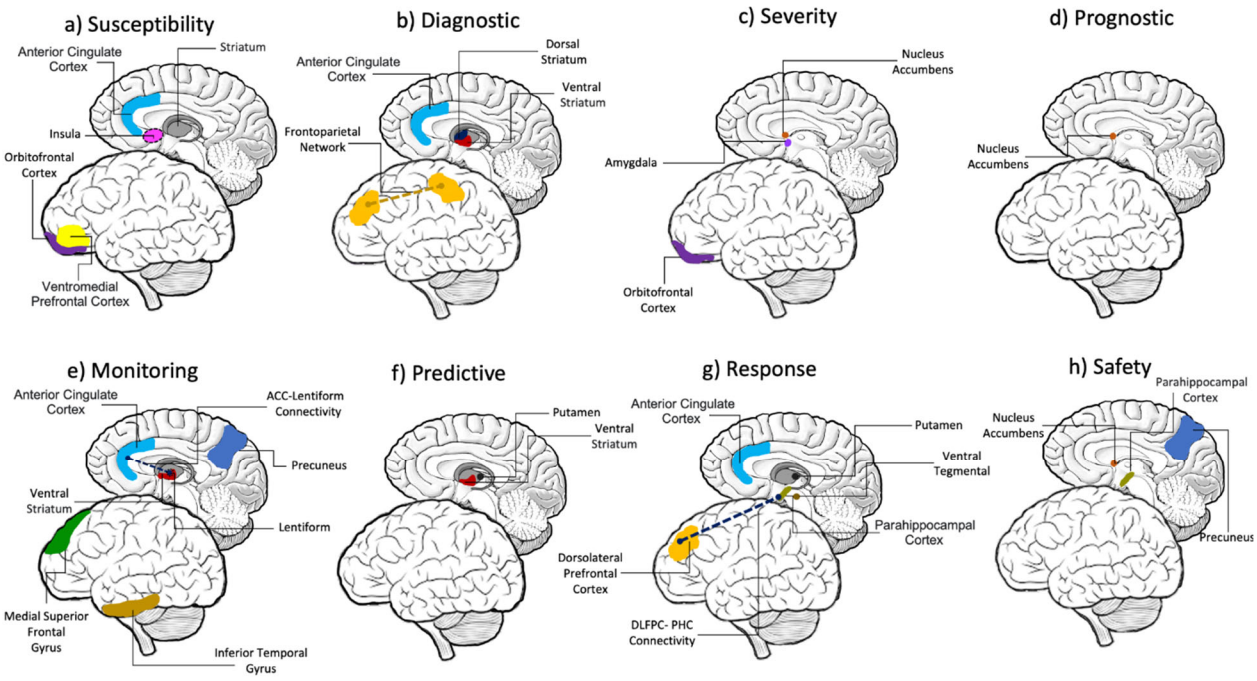
The Sankey diagram represents the number of participants in FDCR studies divided by sex, population type, potentially addictive drugs and behaviors, and interventions. The width of the boxes in each column represents the relative prevalence of each category in the column, while the width of the ribbons connecting the categories across columns represents the proportion of participants shared between each of the two categories. AD: Addictive Disorder (including both SUDs and BAs, diagnosed formally based on widely used criteria such as the Diagnostic and Statistical Manual (DSM) or International Classification of Diseases (ICD)). Participants who used substances without necessarily meeting diagnostic criteria are termed “User”.

Study design types and relevance for potential biomarker development

It is important to note that none of the FDCR indices used by studies in the systematic review constitute fully validated biomarkers at this time. As detailed previously, any biological signal needs to undergo an extensive validation process to qualify as an actual biomarker of disease or recovery, which is not the case for any of the FDCR-derived measures in our included studies. However, the evidence presented in 335 of the studies in our database (75.9%) could potentially support the development of at least one future FDCR biomarker, by virtue of their study designs. We defined seven types of biomarkers based on their context of use. These biomarker types have all been directly adapted from the BEST Glossary[1], with the exception of “severity” biomarkers which are indices that reflect latent disease severity and were defined based on previous biomarker literature[60–62]. None of the studies in our database explicitly used FDCR as an index of

"safety" and thus we removed the BEST *safety* biomarkers category. Nevertheless, we provide two examples of studies that we think point to contexts in which FDCR-derived safety indices might prove useful.

These studies tested a total of 437 relationships (404 significant and 33 non-significant), across contexts of use, between FDCR-derived and clinical measures in 7 different biomarker categories: (1) In *diagnostic* studies, the FDCR signal reflects differences between populations (143 (32.7%) of the included studies, 134 studies reporting a significant association of FDCR and a grouping variable and nine reporting a non-significant association). (2) A *response* index might reflect the neural impact of an intervention (141 (32.3%) studies, 125 reporting significant and 16 non-significant results). (3) In a *severity* context, it would be tested whether an FDCR signal co-varies with addiction severity indices (such covariations were reported in 84 (19.2%) of the studies, 79 significant and five non-significant). (4) A *prognostic* measure should link to future disease course (30 (6.9%) studies, 29 significant and one non-significant). (5) A *predictive* index should explain a significant portion of variance in intervention outcomes (investigated in 25 (5.7%) studies, 24 significant and one non-significant). (6) A *monitoring* index should explain a significant portion of the variance of changes in clinically-relevant variables over time (reported in 12 (2.7%) studies, 11 significant, and one non-significant). Note that "monitoring" measures are only distinguished from "response" markers (in interventional contexts) and "severity" markers (in observational contexts) in that they can be measured repeatedly over time, and their variation over time within one individual is clinically meaningful. (7) A *susceptibility* index would assess the link between FDCR and the progression of non-addictive to addictive use (such links were reported in only 2 (0.5%) studies, both significant) (Figure 2). These biomarkers are defined in Table 1, and related example findings for each are presented in Table 1 and eFigure 8.



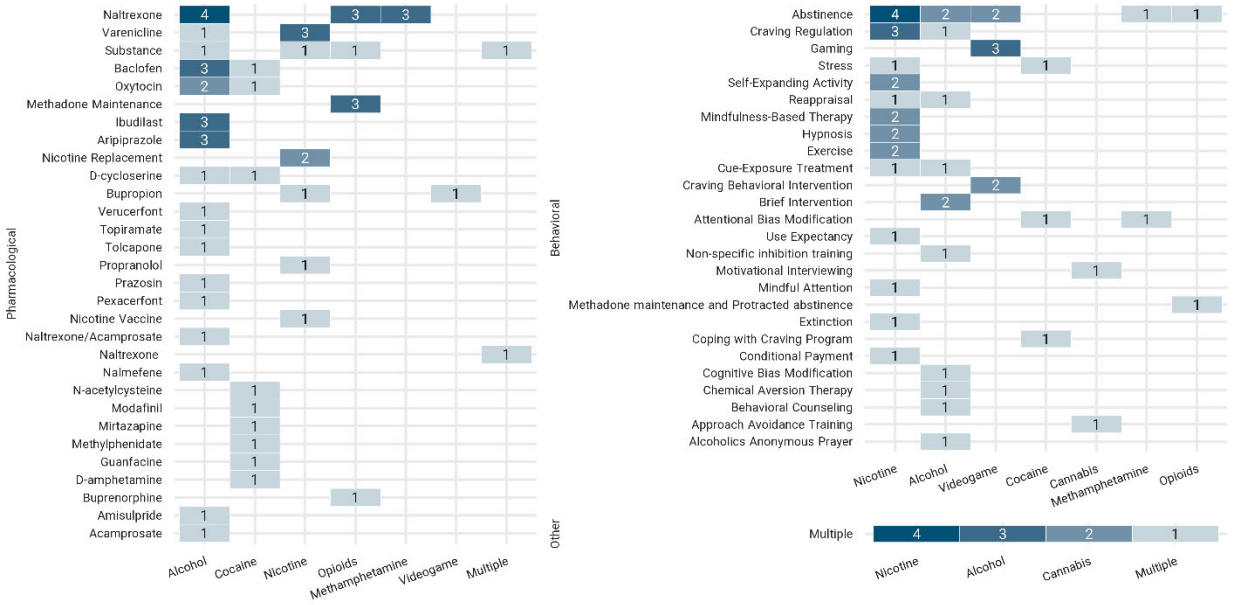
eFigure 8. Examples of Brain Regions in fMRI Drug Cue Reactivity (FDCR) Studies With Supporting Evidence for Potential Biomarker Development

Visual representation of regions with significant FDCR signal in example studies presented in Table 1. Each panel presents significant findings from studies whose results could support the development of one of the biomarker types in the modified BEST biomarker taxonomy, with each region presented with a unique color across panels. Note that these are example findings, and do not necessarily generalize beyond the context of the studies referenced in Table 1.

Interventional FDCR studies

Given the importance of interventional studies and the potential of FDCR to develop response or predictive biomarkers, we present a separate summary of interventional studies in the database. Overall, 155 studies (37.3%) used FDCR in the context of a therapeutic intervention or experimental manipulation. Most commonly, interventional studies used target and control interventions with random assignment (91 studies, 58.7% of interventional studies). Eight studies (5.2%) included a control group without random assignment, 47 (30.3%) included only a single intervention arm without a control condition, and 9 (5.8%) investigated individuals who had been treated retrospectively, for example by comparing them to individuals with untreated SUDs or by comparing individuals who had undergone treatment for different lengths of time (Figure 3a).

Most interventional FDCR studies investigated pharmacological agents (67 studies, 43.2% of the 155 interventional FDCR studies) and cognitive or behavioral interventions (52 studies, 33.5%) (Figure 3b). The most commonly investigated pharmacological agents were naltrexone (10 studies), varenicline (4 studies), baclofen (4 studies), oxytocin (3 studies), and methadone (3 studies). Four studies investigated the impacts of administering a potentially addictive substance, rather than a therapeutic one. Among cognitive and behavioral interventions, the most common were simple abstinence (10 studies) and instructed craving regulation (4 studies). Seven studies used mixes of interventions in different modalities (eFigure 9). Besides pharmacological and behavioral interventions, 12 studies (7.7% of interventional studies) used brain stimulation technologies (7 TMS, 4 tDCS, and 1 DBS), and 10 (6.5%) employed neurofeedback (Figure 3b) (For a detailed breakdown of interventional FDCR studies, see eFigure 9). A majority of the interventional studies (141 out of 155, 91%) used FDCR as a response biomarker, and 125 reported significant FDCR alterations as a result of treatment. Twenty-five studies (16.1%) used FDCR as a predictive biomarker, with 24 observing significant correlations between baseline FDCR and treatment outcomes. Among the 130 studies using FDCR as an outcome measure, 87 measured pre- to post-intervention changes in FDCR as an index of intervention effect (66.9%), and 43 (33.1%) measured only post-intervention cue-reactivity (Figure 3c).

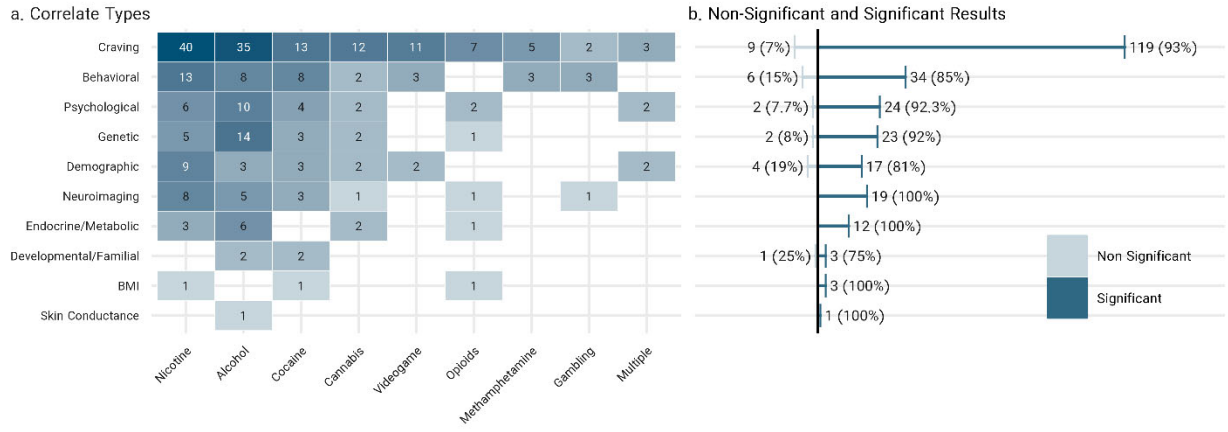


eFigure 9. Detailed Breakdown of Interventional FDCR Studies With Pharmacological (n = 67) or Behavioral (n = 51) Interventions

The “Multiple” column stands for those studies that included more than one type of addictive substance/behavior, while the “Multiple” rows stand for those FDCR studies which used multiple pharmacological interventions or multiple behavioral interventions.

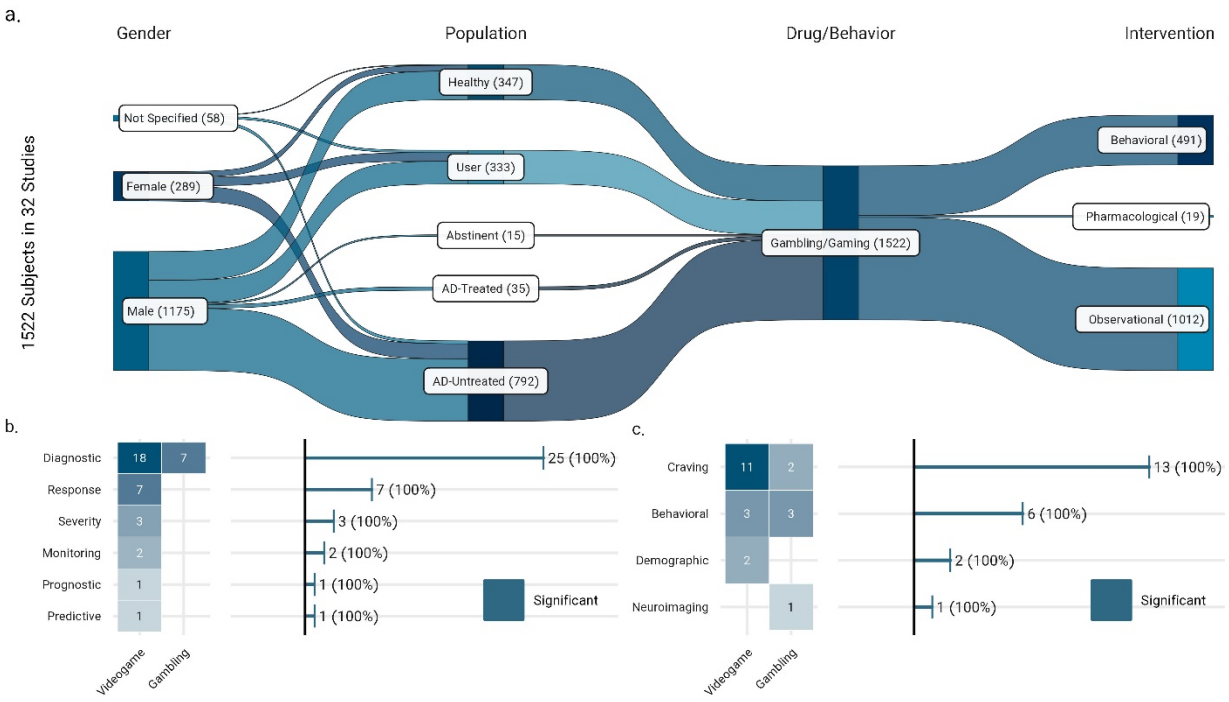
Cross-modal Correlations

Further, 278 studies in the database also tested the relationship between one or more FDCR-derived parameters and non-FDCR variables (other than direct measures of disease severity) such as craving, impulsivity, physiological markers of cue-reactivity, hormonal profiles, and gene variants, with 255 significant and 23 non-significant test results (eFigure 10). Such investigations could be helpful to demonstrate links between FDCR and different aspects of SUDs and to clinically validate FDCR markers by supporting their etiological relevance in SUDs.



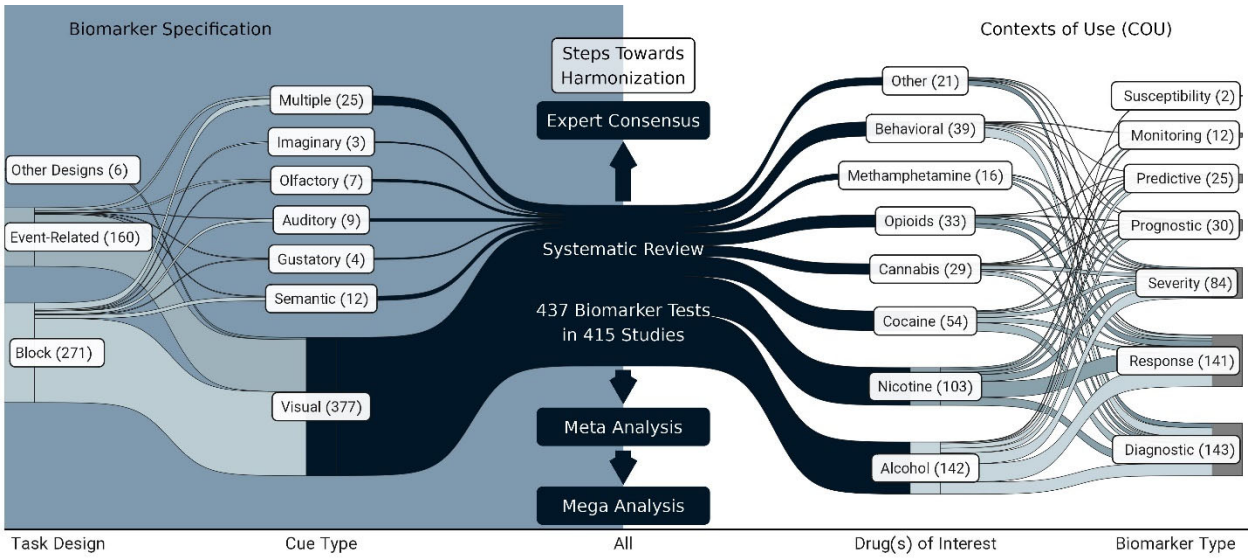
eFigure 10. Multimodal Correlations in FDCR Studies

a. Studies which investigated correlations between FDCR results and other types of measures, broken down by substance or behavior of interest in each study. “Multiple” stands for those studies that included more than one type of addictive substance/ behavior. The "other" category includes inhalants and betel-quid. Note that numbers do not sum to 415 since some studies investigated no multi-modal correlations, while some fit multiple categories. b. Dumbbell plot showing the number of significant and non-significant tests of multi-modal correlation.



eFigure 11. Separate Analyses for Behavioral Addictions

a. Participants in behavioral addiction studies. The Sankey diagram represents the number of participants in fMRI cue-reactivity studies divided by sex, population type, potentially addictive drugs and behaviors, and interventions. b. Seven fMRI cue-reactivity study types for behavioral addictions. The dumbbell plot shows 100% significant supporting biomarker-related findings for each biomarker categories. c. Multi-modal correlations in fMRI cue-reactivity studies in behavioral addictions. The dumbbell plot shows 100% significant test of multi-modal correlations.



eFigure 12. Preliminary Map of the Evidence and Future Directions in Biomarker Development

The Sankey diagram presents a summary of the methodological parameters and contexts of use (COUs) across the 437 potential biomarkers in the systematic review. Moving forward, expert consensus and meta- and mega-analyses may be used to facilitate harmonization and the development of optimal FDCR biomarkers which would undergo analytical and clinical validation and cost-benefit analysis before regulatory qualification for drug development or clinical use.

eReferences.

1. BEST (Biomarkers E and other Tools) Resource. BEST (Biomarkers, EndpointS, and other Tools) resource. Published online 2016.

2. Kraus VB. Biomarkers as drug development tools: discovery, validation, qualification and use. *Nature Reviews Rheumatology*. 2018;14(6):354-362.

3. García-Gutiérrez MS, Navarrete F, Sala F, Gasparyan A, Austrich-Olivares A, Manzanares J. Biomarkers in Psychiatry: Concept, Definition, Types and Relevance to the Clinical Reality. *Front Psychiatry*. 2020;11. doi:10.3389/fpsyt.2020.00432

4. Comes AL, Papiol S, Mueller T, Geyer PE, Mann M, Schulze TG. Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. *Transl Psychiatry*. 2018;8(1):1-15. doi:10.1038/s41398-018-0219-2

5. Yuan N, Chen Y, Xia Y, Dai J, Liu C. Inflammation-related biomarkers in major psychiatric disorders: a cross-disorder assessment of reproducibility and specificity in 43 meta-analyses. *Transl Psychiatry*. 2019;9(1):1-13. doi:10.1038/s41398-019-0570-y

6. Stuart MJ, Baune BT. Chemokines and chemokine receptors in mood disorders, schizophrenia, and cognitive impairment: A systematic review of biomarker studies. *Neuroscience & Biobehavioral Reviews*. 2014;42:93-115. doi:10.1016/j.neubiorev.2014.02.001

7. Bough KJ, Amur S, Lao G, et al. Biomarkers for the Development of New Medications for Cocaine Dependence. *Neuropsychopharmacology*. 2014;39(1):202-219. doi:10.1038/npp.2013.210

8. Carvalho AF, Solmi M, Sanches M, et al. Evidence-based umbrella review of 162 peripheral biomarkers for major mental disorders. *Transl Psychiatry*. 2020;10(1):1-13. doi:10.1038/s41398-020-0835-5

9. McGorry P, Keshavan M, Goldstone S, et al. Biomarkers and clinical staging in psychiatry. *World Psychiatry*. 2014;13(3):211-223. doi:10.1002/wps.20144

10. Fu CHY, Costafreda SG. Neuroimaging-based biomarkers in psychiatry: clinical opportunities of a paradigm shift. *Can J Psychiatry*. 2013;58(9):499-508. doi:10.1177/070674371305800904

11. Goldstein RZ. Neuropsychoimaging Measures as Alternatives to Drug Use Outcomes in Clinical Trials for Addiction. *JAMA Psychiatry*. 2022;79(9):843-844. doi:10.1001/jamapsychiatry.2022.1970

12. Fernandes BS, Karmakar C, Tamouza R, et al. Precision psychiatry with immunological and cognitive biomarkers: a multi-domain prediction for the diagnosis of bipolar disorder or

schizophrenia using machine learning. *Transl Psychiatry*. 2020;10(1):1-13. doi:10.1038/s41398-020-0836-4

13. Grabb MC, Hillefors M, Potter WZ. The NIMH "Fast-Fail Trials" (FAST) Initiative: Rationale, Promise, and Progress. *Pharmaceut Med*. 2020;34(4):233-245. doi:10.1007/s40290-020-00343-y

14. Krystal AD, Pizzagalli DA, Mathew SJ, et al. The first implementation of the NIMH FAST-FAIL approach to psychiatric drug development. *Nature Reviews Drug Discovery*. 2019;18(1):82.

15. Bahji A, Brietzke E, Soares C, Stuart H. Recent Advances in Biomarkers of Addiction: A Narrative Review. *Canadian Journal of Addiction*. 2021;12(1):6-12. doi:10.1097/CXA.0000000000000107

16. National Institutes of Health. National Institute on Drug Abuse. Resource Guide: Screening for Drug Use in General Medical Settings. Published online 2020. Accessed May 20, 2021. https://archives.drugabuse.gov/publications/resource-guide-screening-drug-use-in-general-medical-settings/biological-specimen-testing

17. Neumann T, Spies C. Use of biomarkers for alcohol use disorders in clinical practice. *Addiction*. 2003;98 Suppl 2:81-91. doi:10.1046/j.1359-6357.2003.00587.x

18. Zakhari S, Li TK. Determinants of alcohol use and abuse: Impact of quantity and frequency patterns on liver disease. *Hepatology*. 2007;46(6):2032-2039. doi:10.1002/hep.22010

19. Kwako LE, Bickel WK, Goldman D. Addiction Biomarkers: Dimensional Approaches to Understanding Addiction. *Trends Mol Med*. 2018;24(2):121-128. doi:10.1016/j.molmed.2017.12.007

20. Petry N. *Behavioral Addictions: DSM-5® and Beyond*.; 2015.

21. Center for Drug Evaluation and Research. Biomarker Qualification Submissions. *FDA*. Published online October 21, 2020. Accessed December 16, 2020. https://www.fda.gov/drugs/biomarker-qualification-program/biomarker-qualification-submissions

22. Ekhtiari H, Nasseri P, Yavari F, Mokri A, Monterosso J. Neuroscience of drug craving for addiction medicine: From circuits to therapies. *Prog Brain Res*. 2016;223:115-141. doi:10.1016/bs.pbr.2015.10.002

23. Starcke K, Antons S, Trotzke P, Brand M. Cue-reactivity in behavioral addictions: A meta-analysis and methodological considerations. *J Behav Addict*. 2018;7(2):227-238. doi:10.1556/2006.7.2018.39

24. Boswell RG, Kober H. Food cue reactivity and craving predict eating and weight gain: a meta-analytic review. *Obes Rev*. 2016;17(2):159-177. doi:10.1111/obr.12354

25. Lorenz RC, Krüger JK, Neumann B, et al. Cue reactivity and its inhibition in pathological computer game players. *Addict Biol*. 2013;18(1):134-146. doi:10.1111/j.1369-1600.2012.00491.x

26. Berridge KC, Robinson TE. Liking, Wanting and the Incentive-Sensitization Theory of Addiction. *Am Psychol*. 2016;71(8):670-679. doi:10.1037/amp0000059

27. Drummond DC. What does cue-reactivity have to offer clinical research? *Addiction*. 2000;95(8s2):129-144. doi:10.1046/j.1360-0443.95.8s2.2.x

28. Carter BL, Tiffany ST. Meta-analysis of cue-reactivity in addiction research. *Addiction*. 1999;94(3):327-340. doi:10.1046/j.1360-0443.1999.9433273.x

29. Rohsenow DJ, Childress AR, Monti PM, Niaura RS, Abrams DB. Cue reactivity in addictive behaviors: theoretical and treatment implications. *International Journal of the Addictions*. 1991;25(sup7):957-993.

30. Ekhtiari H, Faghiri A, Oghabian MA, Paulus MP. Chapter 7 - Functional neuroimaging for addiction medicine: From mechanisms to practical considerations. In: Ekhtiari H, Paulus MP, Ekhtiari H, Paulus MP, eds. *Progress in Brain Research*. Vol 224. Neuroscience for Addiction Medicine: From Prevention to Rehabilitation - Methods and Interventions. ; 2016:129-153. Accessed February 22, 2019. http://www.sciencedirect.com/science/article/pii/S0079612315001508

31. Goldstein RZ, Volkow ND. Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. *American Journal of Psychiatry*. 2002;159(10):1642-1652.

32. Zilverstand A, Huang AS, Alia-Klein N, Goldstein RZ. Neuroimaging Impaired Response Inhibition and Salience Attribution in Human Drug Addiction: A Systematic Review. *Neuron*. 2018;98(5):886-903. doi:10.1016/j.neuron.2018.03.048

33. Paulus MP, Stewart JL. Neurobiology, Clinical Presentation, and Treatment of Methamphetamine Use Disorder: A Review. *JAMA Psychiatry*. 2020;77(9):959-966. doi:10.1001/jamapsychiatry.2020.0246

34. Garrison KA, Potenza MN. Neuroimaging and Biomarkers in Addiction Treatment. *Curr Psychiatry Rep*. 2014;16(12):513. doi:10.1007/s11920-014-0513-5

35. Genauck A, Matthis C, Andrejevic M, et al. Neural correlates of cue-induced changes in decision-making distinguish subjects with gambling disorder from healthy controls. *Addiction Biology*. 2021;26(3):e12951. doi:10.1111/adb.12951

36. Liu GC, Yen JY, Chen CY, et al. Brain activation for response inhibition under gaming cue distraction in internet gaming disorder. *Kaohsiung J Med Sci*. 2014;30(1):43-51. doi:10.1016/j.kjms.2013.08.005

37. Miedl SF, Büchel C, Peters J. Cue-induced craving increases impulsivity via changes in striatal value signals in problem gamblers. *J Neurosci*. 2014;34(13):4750-4755. doi:10.1523/JNEUROSCI.5020-13.2014

38. Blaine S, Claus E, Harlaar N, Hutchison K. TACR1 genotypes predict fMRI response to alcohol cues and level of alcohol dependence. *Alcohol Clin Exp Res*. 2013;37 Suppl 1:E125-130. doi:10.1111/j.1530-0277.2012.01923.x

39. Wiers CE, Shumay E, Volkow ND, et al. Effects of depressive symptoms and peripheral DAT methylation on neural reactivity to alcohol cues in alcoholism. *Transl Psychiatry*. 2015;5:e648. doi:10.1038/tp.2015.141

40. Kroemer NB, Wuttig F, Bidlingmaier M, Zimmermann US, Smolka MN. Nicotine enhances modulation of food-cue reactivity by leptin and ghrelin in the ventromedial prefrontal cortex. *Addict Biol*. 2015;20(4):832-844. doi:10.1111/adb.12167

41. Wang W, Zhornitsky S, Le TM, et al. Cue-elicited craving, thalamic activity, and physiological arousal in adult non-dependent drinkers. *Journal of Psychiatric Research*. 2019;116:74-82. doi:10.1016/j.jpsychires.2019.06.005

42. Elton A, Smitherman S, Young J, Kilts CD. Effects of childhood maltreatment on the neural correlates of stress- and drug cue-induced cocaine craving. *Addict Biol*. 2015;20(4):820-831. doi:10.1111/adb.12162

43. Groefsema MM, Mies GW, Cousijn J, Engels RCME, Sescousse G, Luijten M. Brain responses and approach bias to social alcohol cues and their association with drinking in a social setting in young adult males. *Eur J Neurosci*. 2020;51(6):1491-1503. doi:10.1111/ejn.14574

44. Burnette EM, Grodin EN, Lim AC, MacKillop J, Karno MP, Ray LA. Association between impulsivity and neural activation to alcohol cues in heavy drinkers. *Psychiatry Res Neuroimaging*. 2019;293:110986. doi:10.1016/j.pscychresns.2019.110986

45. Ceceli AO, King SG, McClain N, Alia-Klein N, Goldstein RZ. The Neural Signature of Impaired Inhibitory Control in Individuals with Heroin Use Disorder. *J Neurosci*. 2023;43(1):173-182. doi:10.1523/JNEUROSCI.1237-22.2022

46. Goldstein RZ, Tomasi D, Alia-Klein N, et al. Dopaminergic Response to Drug Words in Cocaine Addiction. *J Neurosci*. 2009;29(18):6001-6006. doi:10.1523/JNEUROSCI.4247-08.2009

47. Huang Y, Ceceli A, Kronberg G, et al. Cortico-striatal engagement during cue-reactivity, reappraisal, and savoring of drug and non-drug stimuli predicts craving in heroin addictio. Published online 2022. doi:10.1101/2022.05.27.22275628

48. Vincent GM, Cope LM, King J, Nyalakanti P, Kiehl KA. Callous-Unemotional Traits Modulate Brain Drug Craving Response in High-Risk Young Offenders. *J Abnorm Child Psychol*. 2018;46(5):993-1009. doi:10.1007/s10802-017-0364-8

49. Feldstein Ewing SW, Filbey FM, Chandler LD, Hutchison KE. Exploring the Relationship Between Depressive and Anxiety Symptoms and Neuronal Response to Alcohol Cues. *Alcohol Clin Exp Res*. 2010;34(3):396-403. doi:10.1111/j.1530-0277.2009.01104.x

50. Moran LV, Betts JM, Ongur D, Janes AC. Neural Responses to Smoking Cues in Schizophrenia. *Schizophr Bull*. 2018;44(3):525-534. doi:10.1093/schbul/sbx085

51. Baurley JW, McMahan CS, Ervin CM, Pardamean B, Bergen AW. Biosignature Discovery for Substance Use Disorders using Statistical Learning. *Trends Mol Med*. 2018;24(2):221-235. doi:10.1016/j.molmed.2017.12.008

52. Rawls E, Kummerfeld E, Zilverstand A. An integrated multimodal model of alcohol use disorder generated by data-driven causal discovery analysis. *Communications Biology*. 2021;4(1):1-12. doi:10.1038/s42003-021-01955-z

53. Ray LA, Courtney KE, Hutchison KE, Mackillop J, Galvan A, Ghahremani DG. Initial evidence that OPRM1 genotype moderates ventral and dorsal striatum functional connectivity during alcohol cues. *Alcohol Clin Exp Res*. 2014;38(1):78-89. doi:10.1111/acer.12136

54. Schacht JP, Anton RF, Voronin KE, et al. Interacting effects of naltrexone and OPRM1 and DAT1 variation on the neural response to alcohol cues. *Neuropsychopharmacology*. 2013;38(3):414-422. doi:10.1038/npp.2012.195

55. Ekhtiari H, ACRI Secretariat. A Systematic Review on fMRI Drug Cue Reactivity Studies. Published online May 18, 2020. Accessed November 29, 2020. https://osf.io/eb972/

56. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/M18-0850

57. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*. 2015;4:1. doi:10.1186/2046-4053-4-1

58. Roy Rosenzweig Center for History and New Media. Zotero. www.zotero.org/download

59. R Core Team. R: A language and environment for statistical computing. Published online 2013.

60. Bivona G, Agnello L, Ciaccio M. Biomarkers for Prognosis and Treatment Response in COVID-19 Patients. *Ann Lab Med*. 2021;41(6):540-548. doi:10.3343/alm.2021.41.6.540

61. Ligezka AN, Radenkovic S, Saraswat M, et al. Sorbitol Is a Severity Biomarker for PMM2-CDG with Therapeutic Implications. *Ann Neurol*. 2021;90(6):887-900. doi:10.1002/ana.26245

62. de Mendonça EB, Schmaltz CA, Sant'Anna FM, et al. Anemia in tuberculosis cases: A biomarker of severity? *PLoS One*. 2021;16(2):e0245458. doi:10.1371/journal.pone.0245458