Perspective

Functional neuroimaging as a catalyst for integrated neuroscience

https://doi.org/10.1038/s41586-023-06670-9

Received: 16 March 2023

Accepted: 22 September 2023

Published online: 8 November 2023

Check for updates

Emily S. Finn[1]✉, Russell A. Poldrack[2]✉ & James M. Shine[3]✉

Functional magnetic resonance imaging (fMRI) enables non-invasive access to the awake, behaving human brain. By tracking whole-brain signals across a diverse range of cognitive and behavioural states or mapping differences associated with specific traits or clinical conditions, fMRI has advanced our understanding of brain function and its links to both normal and atypical behaviour. Despite this headway, progress in human cognitive neuroscience that uses fMRI has been relatively isolated from rapid advances in other subdomains of neuroscience, which themselves are also somewhat siloed from one another. In this Perspective, we argue that fMRI is well-placed to integrate the diverse subfields of systems, cognitive, computational and clinical neuroscience. We first summarize the strengths and weaknesses of fMRI as an imaging tool, then highlight examples of studies that have successfully used fMRI in each subdomain of neuroscience. We then provide a roadmap for the future advances that will be needed to realize this integrative vision. In this way, we hope to demonstrate how fMRI can help usher in a new era of interdisciplinary coherence in neuroscience.

Interest and investment in brain science have expanded rapidly in recent decades, rewarding us with exciting discoveries about how nervous systems orchestrate their command over complex adaptive behaviour. Progress is slowed, however, by the fact that many subdomains within neuroscience are fundamentally siloed from one another. There are understandable reasons for this: neuroscience research is conducted in diverse organisms, from humans to rodents to invertebrates; on scales ranging from nanometres (for example, single molecules) to metres (for example, socially interacting humans); using measurement techniques that range from detecting gene expression in single cells to recording electrical activity in groups of neurons and observing complex behaviour (for example, human speech).

This wide range of approaches raises conceptual and empirical barriers that make it inherently difficult to integrate insights or advances across subfields. Yet, ultimately, we will need to bridge across the domains of neuroscience to fully understand how the brain works, as well as to identify and treat its many distinct disorders and syndromes. Without these links, we are tacitly accepting our inability to translate insights from one field (for example, circuit mechanisms for synaptic plasticity following brain stimulation[1,2]) into another (for example, understanding how and where to stimulate the brain of an individual with a particular disorder in order to confer maximal clinical benefit[3]).

In this Perspective, we suggest that the technique of fMRI is well-placed to act as an integrative bridge to connect between different subfields of neuroscience (Fig. 1). fMRI has a number of enviable features, such as non-invasive access to both spatial and temporal information at the whole-brain level[4], the ability to interrogate human-level cognitive, emotional and motoric capacities[5] as well as to acquire commensurate data across species, and a strong grounding in open science practices that are driving heightened validity and reproducibility[6,7]. Along with these strengths, fMRI also has notable weaknesses, including the fact that the blood oxygen level-dependent (BOLD) signal is an indirect measure of neural activity[8]. Crucially, these limitations align with prominent strengths of techniques used in other fields, such as the ability to causally drive neural dynamics in ways that reduce ambiguity about the source of signals being measured[8]. If the unique strengths of fMRI can be leveraged while finding ways to compensate for its weaknesses, network science shows that better integration across a network (that is, increasing the density or strength of connector hubs[9–11]) will help to catalyse robust progress in the field of neuroscience in general.

Strengths and weaknesses of fMRI

Unlike structural MRI, which provides detailed anatomical images of the brain, fMRI detects changes in blood flow dynamics that occur while the brain is active[12]. The most common contrast mechanism, known as the BOLD signal, is based on the principle that there is an increased delivery of oxygenated blood supply to regions of heightened neuronal activity that can be detected using MRI. fMRI is very safe in that it is non-invasive and does not require the use of ionizing radiation. Although fMRI is perhaps most strongly associated with cognitive neuroscience—indeed, one of its core strengths is to provide a window into brain function as people perform specific tasks involving, for example, perception, memory and decision making—its applications are not limited to the cognitive domain. It has also made key contributions to systems, computational and clinical neuroscience, as discussed below.

[1]Department of Psychological and Brain Sciences, Dartmouth College, Dartmouth, NH, USA. [2]Department of Psychology, Stanford University, Stanford, CA, USA. [3]School of Medical Sciences, University of Sydney, Sydney, New South Wales, Australia. ✉e-mail: esfinn@gmail.com; russpold@stanford.edu; mac.shine@sydney.edu.au

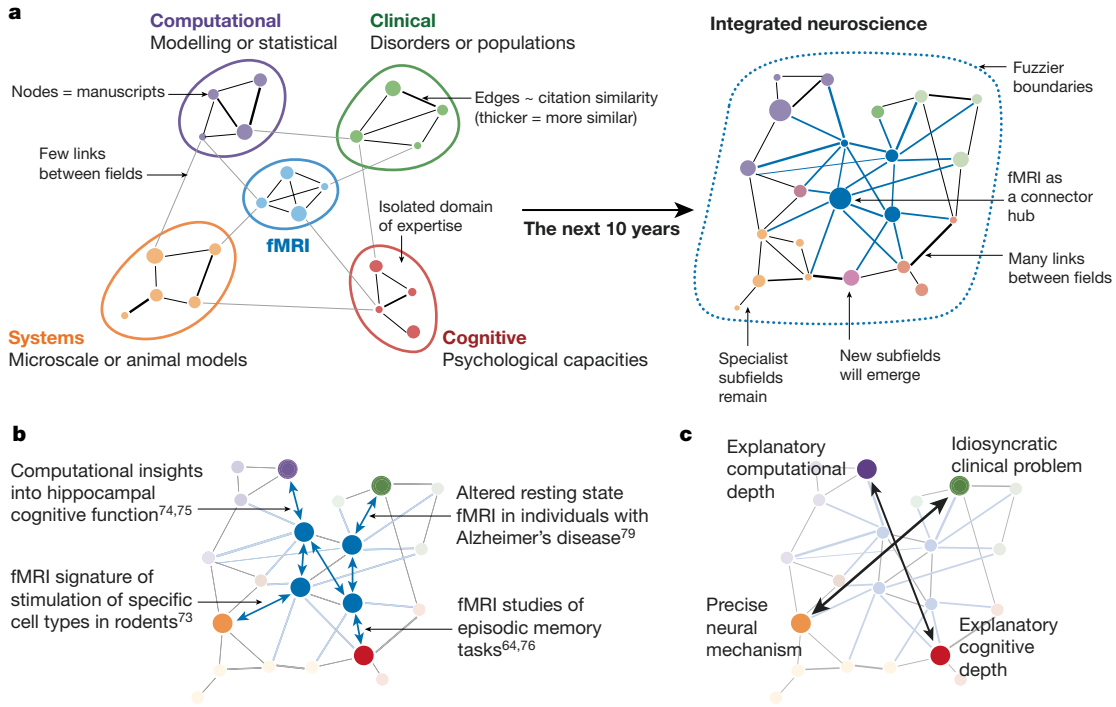

Fig. 1 | fMRI as an integrative catalyst to dissolve modular boundaries in the existing neuroscience network. a, At present, most subdomains in neuroscience are highly insular, with few studies linking across subdomains; a co-citation network of papers published in neuroscience would appear topologically segregated. We foresee an opportunity for fMRI studies to help dissolve these subdomain boundaries, which would integrate the field of neuroscience in ways that will catalyse the formation of new subfields and novel insights into how the brain works. Note that the networks depicted in this figure are entirely conceptual. **b**, Illustrative example of how, by combining studies in each subdomain with fMRI, a richer perspective can emerge linking across scales and species. **c**, The benefit of integrating across subfields is that there will be stronger links between precise neural mechanisms and idiosyncratic clinical problems, which will be enriched with computational and cognitive depth (and may also be important targets of inquiry in their own right).

One key feature of fMRI that makes it well-positioned to serve as a bridge across subfields is that it can be performed in humans as well as all mammalian model systems (for example, rodents and non-human primates), yielding commensurate data that open a direct translational route between species. Another major advantage is its whole-brain access: it can acquire a full three-dimensional brain volume at each readout, with reasonable temporal resolution (on the order of 1 Hz or faster; Box 1 and Fig. 2) and with the highest spatial resolution of any non-invasive imaging technique (on the order of 1 mm or less; Box 1). Although more invasive techniques are preferable for most applications in model systems, even the state of the art in these techniques cannot yet match this whole-brain, simultaneous, depth-agnostic feature of fMRI—for example, in mice, wide-field calcium imaging is typically limited to relatively superficial subsets of the cortical surface, with no ability to simultaneously image subcortical structures. Furthermore, the same safe and non-invasive nature that makes fMRI suitable for humans also makes it more suitable for certain applications in model systems, such as long-term longitudinal designs. Whereas more invasive techniques, such as direct electrophysiological recordings and/or calcium imaging, suffer from problems with localizing the same neurons or neuronal populations across sessions, or even with repeated acquisitions damaging the health of the tissue, fMRI makes it safe and easy to acquire an arbitrary number of repeated measures in the same subject (whether human or animal) and accurately cross-register these measurements with one another. Thus, although fMRI requires substantial technical expertise to acquire and analyse the data, it provides unique benefits that are difficult to obtain with other imaging techniques.

Other imaging approaches used in humans offer complementary strengths to those of fMRI; however, each of these has its own weaknesses that we argue make it less suitable to act as an interdisciplinary bridge. Electroencephalography (EEG) provides a more direct measure of electrical activity with millisecond temporal resolution, but the spatial resolution is poor and largely limited to the cortical surface. This impairs our ability to draw links with systems neuroscience, since many key structures are small and/or subcortical (for example, the hippocampus). Collecting surface EEG in small animals is empirically challenging, making it a poor choice for achieving correspondence in cross-species measurements. Magnetoencephalograhy (MEG) provides excellent temporal resolution and spatial resolution that is better than EEG (though still poor in comparison to fMRI). However, MEG requires specialized equipment that is less widely available than MRI machines, and is also difficult or impossible to perform in most smaller model organisms. Positron emission tomography (PET) offers direct links to molecular and cellular neuroscience, but both the temporal and spatial resolution are substantially poorer than fMRI and it is invasive, prohibitively expensive and limited in its availability, making it challenging to scale. It is also inappropriate for some populations (for example, children or pregnant women) owing to exposure to ionizing radiation. Intracranial EEG, in which electrodes are placed directly on or inside neural tissue, comes closest to a 'gold standard' for measuring brain activity, but its availability in humans is extremely limited (to neurosurgical patients). Furthermore, it does not offer whole-brain coverage, and electrode placement is determined by clinical rather than research goals.

Despite its strengths over other techniques available in humans, fMRI has several weaknesses worth noting (Fig. 2). First, the BOLD signal is an indirect probe of neuronal activity, tracking instead a complex relationship between spiking activity, local field potentials, glial cell function and vascular smooth muscle cells[13–16]. These complex relationships can make it challenging to infer the likely neuronal correlates of localized BOLD changes, particularly as the mechanisms supporting

Box 1

Major advances in functional neuroimaging

There have been a number of major advances in the fMRI field over the last few years, including (but not limited to):

Advances in data acquisition

In recent years, fMRI has seen substantial improvements in both spatial and temporal resolution. Smaller voxel sizes enabled by more powerful magnetic fields (7 Tesla and above) and/or pulse sequences have made it possible to record meaningful signal from distinct cortical layers and columns[134,135], opening the door to understanding cognitive processes in terms of directional circuits (see 'fMRI in cognitive neuroscience'). Although the temporal resolution of fMRI scans that rely on the BOLD contrast is fundamentally limited by the sluggishness of the blood response itself, there is evidence that signatures of much faster processes can be recovered even from what is traditionally considered a slow signal[4]. These improved acquisitions are enabling researchers to make new progress on old questions (for example, visual sequence detection[136] and word-by-word responses in natural spoken language[137]), as well as progress on new questions not previously accessible with fMRI (such as phenomena linked to the human 'slow' oscillation band at around 0.7 Hz including processes of sleep, memory and awareness[41]).

Advances in experimental paradigms

fMRI research has greatly expanded its repertoire of experimental paradigms on two—in some ways opposing—fronts. First, 'resting-state' fMRI, in which participants are imaged in the absence of any explicit task, has exploded in popularity over the past 15 years. The resting-state revolution both spurred the development of several families of techniques for finding structure in these high-dimensional datasets and provided ample opportunity to link across species and populations. Second, the field has also embraced rich experimental paradigms that derive their power by being either more naturalistic (for example, films, natural spoken language, interactive games and virtual reality[138–140]) and/or more clever in their constraints[67,141]. These new-wave paradigms often more closely mimic what the brain does in real-world contexts.

Advances in data analysis

The field has adopted new analytical methods to detect and characterize fine-grained activity patterns with much more precision. These include both techniques that directly model features of the task or paradigm such as encoding[142,143] and decoding approaches[144,145] and representational similarity analysis[146], as well as largely model-free techniques that aim to find structure in patterns of brain activity in a more data-driven way, such as those that leverage signal coherence across different brains[147,148] and/or those aimed at discovering latent states that can then be related to behaviour[149–151]. In fact, data-driven methods applied to complex neuroimaging data can lead to remarkable advances in our textbook understanding of the organization of the nervous system[152]. Combined with the carefully crafted experimental paradigms described above, these techniques are enabling researchers to adjudicate between possible models for how the brain performs complex tasks and test new theories of cognition and behaviour.

the BOLD response can vary across regions. The BOLD response is also relatively slow, and the spatial resolution of the neurovasculature can also make it difficult to resolve small structures, as well as distinct cell types or sub-cellular processes. Finally, cross-species comparisons can be challenging owing to the fact that animals typically require anaesthesia to undergo MRI scanning.

Despite these limitations, there is reason to be optimistic, as novel imaging sequences are continuing to improve spatial signal to noise[17]. Furthermore, even invasive imaging techniques considered to capture neuronal activity more directly cannot fully overcome these complex transfer functions: for example, wide-field calcium imaging, arguably fMRI's main competitor technique for invasive studies in animal models, also requires specialized deconvolution algorithms to separate signal from background[18] and link the recovered signal back to the inferred spike-driven events that give rise to it, and just like with the BOLD signal, the choice of algorithm can substantially impact results[19]. The second main weakness of fMRI is its relatively poor temporal resolution compared to electrophysiological imaging techniques (that is, EEG and MEG). However, especially for cross-species applications, we believe that accurate spatial localization should be prioritized over temporal resolution: when comparing how specific circuits give rise to behaviour across species, the more fundamental task is to establish correspondence in space, to ensure that the signals are coming from the same or homologous regions; only then can the relative timing of events be meaningfully interpreted. Finally, although MRI-based methods are among the safest imaging techniques, they are expensive and can be still inappropriate for some populations, such as those with certain medical implants or who cannot lay prone in the scanner for sufficient periods of time. However, these constraints exclude a much smaller percentage of the population than the constraints of more invasive techniques (for example, PET or intracranial EEG). Therefore, given this landscape, and despite its inherent limitations, we believe that fMRI is the strongest candidate for integrating across subfields of neuroscience.

fMRI in systems neuroscience

There is a historic divide between microscopic, neuronal-level and macroscopic, whole-brain studies in neuroscience[20,21]. At the small scale, precision is retained over individual elements, but how these microscopic details factor into the coherent whole is often overlooked. By contrast, at the whole-brain level, we gain an appreciation for the low-dimensional constraints imposed on a nervous system embedded within a corporeal body and constrained environment, but trade-off this clarity for an incognizance regarding the precise cellular-level composition of the signals that we track. Discovering the identity of the transfer functions that interconnect these different scales of analysis is a crucial missing piece for the ability to understand how the brain works.

Although there is ongoing debate regarding precisely whether pre- or post-synaptic activity in individual neurons[13,16,22,23] or their interactions with astrocytes[24] are responsible for the BOLD signal tracked with fMRI (Box 1), there remains little doubt regarding the fact that there is tremendous spatial and temporal organization in BOLD signals recorded across the brain. For instance, tracking simple patterns of zero-lagged correlation between coarsely mapped spatial locations distributed across the brain identifies a set of robust groupings (often referred to as 'large-scale networks') that are stable across sessions, individuals

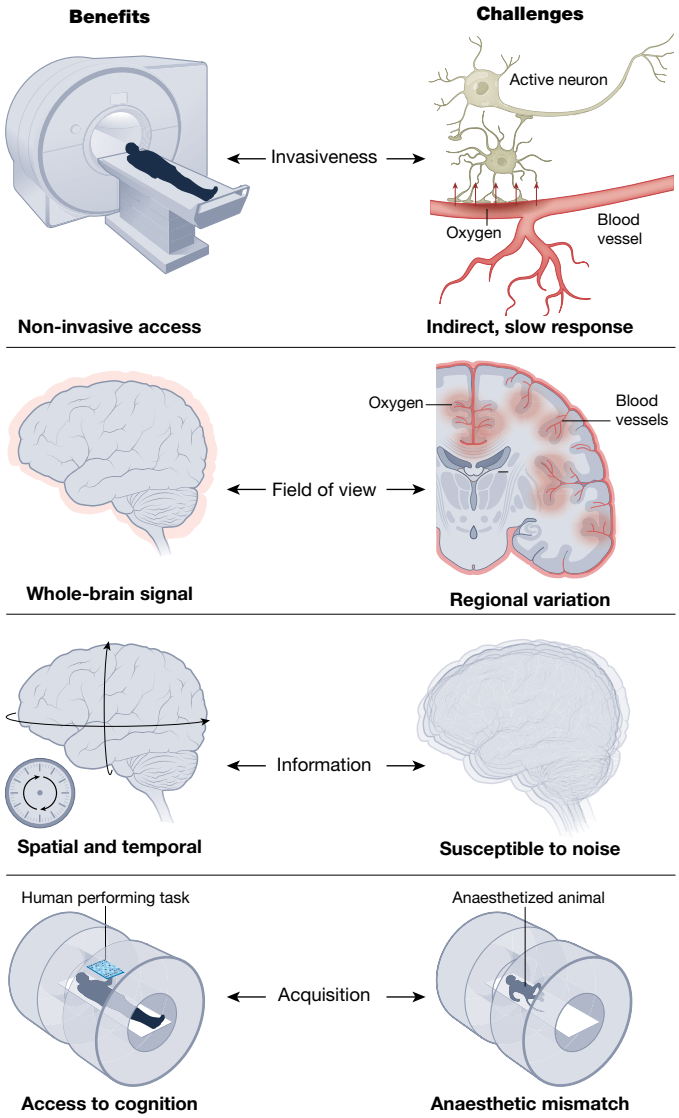

Fig. 2 | Benefits and challenges associated with fMRI. Benefits include non-invasive access to the whole brain, with both spatial and temporal information as well as access to a wide range of cognitive functions in awake human participants. Challenges include the indirect nature of the BOLD response, its sluggish temporal nature, regional variations in the strength and nature of the BOLD signal, its susceptibility to various forms of noise (such as head motion and physiological artefacts), as well as the fact that cross-species comparisons can be challenging owing to the typical need for anaesthesia in animal studies. fMRI is also well suited to within-subject designs, but is also expensive and technically challenging (not shown).

and different MRI scanners[25,26], but also reconfigure as a function of task performance[27], arousal[28], learning[29], across development[30] and as a function of a range of disease processes[31]. Tremendous recent progress has been made in systematically characterizing these patterns, which have been shown to contain robust low-dimensional structure[32], while also covarying with recordings from other neuroimaging modalities, such as the intrinsic time-scale in magnetoencephalography data[33], the approximate myelin content from structural imaging data[34,35] and the differential expression of a wide range of distinct genes based on postmortem studies[36–38]. This suggests a promising link between the recordings made at the whole-brain level and the microscopic elements that comprise the brain, but as yet there are no effective means for translating between these different scales.

Integrating across scales is a difficult problem for a variety of reasons: recording techniques often trade-off micro-level precision for macro-scale coverage, and few techniques are capable of simultaneously resolving signals at multiple unique scales. This is problematic, as the scales at which microscopic-level neuroscience are conducted are orders of magnitude smaller than those that can currently be detected with even the most powerful high-field fMRI techniques. This gap is of course becoming ever smaller with the advent of improved technologies that map thousands of populations of neurons simultaneously in awake animals, typically using invasive calcium sensors[39]. However, an important issue remains: there are no currently accessible technologies that can map these microscale dynamics at the resolution of the whole brain, and especially without doing so invasively. However, although whole-brain calcium imaging in mammalian model organisms or any scale of calcium imaging in humans are not yet possible, it is possible to record whole-brain fMRI in humans and other mammal models. In addition, recent advances have made it possible to scan animals without the need for anaesthesia[40], offering exciting opportunities to compare and contrast whole-brain imaging signatures of behaviour across species. Thus, a profitable middle-ground could therefore be to use fMRI as a translator between scales, particularly given recent advances in layer-resolved fMRI[17] and fast fMRI[41] (Box 1).

The ability to state causal hypotheses about circuit-level function in the nervous system that then make testable predictions about macroscopic patterns recorded at the whole-brain scale in humans is crucial for the advancement of the field. Importantly, simultaneously tracking fMRI with microscopic measures of neural activity is not enough—the notoriously difficult problems of learning the transfer functions between signals must also be solved[14], while also taking seriously important differences in neural circuitry across species[42,43]. There are several recent examples of work in this space that provide hope for the future. By combining causal optogenetic manipulation of precise cellular populations in the brain with recordings of whole-brain BOLD signal in lightly anaesthetized rodents, researchers have been able to link the causal perturbation of neural populations to the large-scale network signatures that can be recorded at the whole-brain level[44,45]. Analogous approaches that electrically perturb the brain of humans with implanted electrodes during fMRI recording provide similar causal access to the brain, albeit at a coarser spatial resolution than those provided by optogenetic methods[46,47]. Optogenetic fMRI techniques have also been combined with gene-expression maps to shed light onto how neuromodulators may regulate aspects of human behaviour via large-scale receptor networks[48]. Other approaches have combined BOLD with other imaging modalities, such as calcium recordings[49] or electrophysiology[50,51], to determine whether the patterns inherent within each technique coincide or differ. Early progress appears promising, although there are many idiosyncratic details inherent to neural circuitry, including sensitivity to arousal states[52,53] and neurochemical alterations across rostrocaudal hierarchies in the brain[54,55], that remain to be effectively dealt with before the field can deliver upon the pressing need for a translational tool that enables mapping across the vastly different spatial scales required to link these different methods.

fMRI in cognitive neuroscience

Once more robust links between the microcircuitry of the brain and the macroscopic measures that we record from fMRI have been created, that microcircuitry can be linked to the emergent capacities of nervous systems by taking advantage of what has been a longstanding advantage of whole-brain imaging in humans: the capacity to shed light on the internal cognitive, emotional and motivational processes that give rise to human behaviour and thought. Developments in signal acquisition and novel analytic strategies (Box 1) have enabled substantial and meaningful changes in the role of fMRI in this process.

On the data acquisition front, by acquiring depth-dependent or 'layer-specific' measurements using high-field fMRI, we can characterize directional circuits via the fact that top-down (feedback) and bottom-up (feed-forward) activity have distinct depth profiles across the cortical column. Critically, these layer-specific signatures of activity in humans can also be compared to observations using depth-dependent probes in animal models[17,56–58], thus affording opportunities to test longstanding theories of predictive coding in human participants[59]. Whereas early layer-specific fMRI work sought to validate these methods by extending well-established findings from animal model systems into humans, more recent studies have begun to make novel discoveries that would be difficult or impossible to test in animal models. These studies are now informing theories of cognition in their own right. For example, theoretical studies comparing perception (actually experiencing an external stimulus) with mental imagery (calling a stimulus to mind or imagining a behaviour without performing it) suggest precise differences in the cortical circuits responsible for these two processes[60]; layer-specific fMRI studies are beginning to tease these phenomena apart in terms of their top-down versus bottom-up profiles[61,62]. Findings from layer-specific fMRI are also informing theories of attention[63–65], the effects of which can markedly change both brain activity and behaviour for otherwise identical stimuli. For example, one study[64] was able to dissociate multisensory interactions from top-down attention into two distinct depth-dependent profiles, suggesting that these are two (at least partially) separable mechanisms by which the brain regulates and prioritizes information flow.

The combination of incisive experimental paradigms with nuanced approaches to data analysis has also led to rapid progress in cognitive neuroscience. For example, careful task designs combined with model-based analysis strategies have revealed that the brain uses a grid-like code to navigate both spatial and nonspatial (that is, conceptual) relationships[66,67] and to organize episodic memories[68], akin to the grid cells that have been observed during spatial navigation in non-human animals. Building on longstanding work in the field of reinforcement learning[69], recent studies have used fMRI to distinguish between possible models of how the brain computes subjective value[70], to delineate mechanisms for learning from direct experience versus social observation[71], and to dissociate the effects of errors of action selection versus execution on learning[72]. Of note, these studies depart somewhat from earlier model-based approaches, which fit models to behavioural data first to choose the winning model, then search for the neural correlates of the relevant cognitive operations[73]. Rather, in these cases, the models under study make similar predictions about ultimate behaviour, meaning that it would it be difficult or impossible to adjudicate between them using behaviour alone. Instead, in these studies, models are fit directly to the neural data to search for evidence of the hypothesized latent operations at intermediate steps in the cognitive process. Although it is possible that the same conclusions could have been reached using EEG, MEG, or another technique, it is unlikely that they could have been reached with the same spatial specificity of fMRI, given that many of the structures in question were relatively deep (for example, striatum, thalamus and medial prefrontal cortex). This spatial specificity in humans is important for links across species, and should be especially prioritized in the fields of reward learning and decision making, where there are exciting opportunities to translate similar tasks and circuits between humans and animal models. These examples thus underscore a critical role for fMRI in deepening our understanding of the mechanisms of cognition in ways that can also meaningfully interface with other subfields of neuroscience.

With this new wave of studies fuelled by methodological advances, the capacity for fMRI to help integrate cognitive neuroscience with the domains of systems, computational and clinical neuroscience offers exciting opportunities for advancement. One example of where these efforts are already bearing fruit is studies of the medial temporal lobe (particularly the hippocampus; Fig. 1b) and cross-species phenomena such as replay and reactivation, processes by which humans consolidate memories and learn new information[74,75]. These processes typically occur 'offline' (while individuals are not focused on a particular task, and often without explicit awareness) and are therefore not accessible to behavioural probes, and also critically involve deep brain structures in the medial temporal lobe, such that fMRI has been a critical tool for characterizing these phenomena in humans. These phenomena are thus a prime example of how fMRI is well-positioned to link across scales, which requires high spatial precision which can be achieved using both non-invasive fMRI (in humans or animal models) and invasive recordings (in animal models[76] or occasionally, humans undergoing intracranial recordings).

To further interface with systems neuroscience, future work could leverage these clear links across species and scales to refine our understanding of the general transfer function(s) that map invasive recordings in animal models to fMRI in humans; these function(s) could in turn be applied to the measurements collected in other experiments to make more precise translational predictions and/or more precisely evaluate the extent to which cross-species findings converge or diverge, despite the substantial differences in recording techniques. On the computational side, a great deal is understood regarding the computational nature of hippocampal circuitry[77,78], and these algorithmic features can be refined through the design of sophisticated cognitive tasks that would be difficult to perform in animals[67,79]. Finally, many clinical disorders, such as Alzheimer's disease[80] and epilepsy[81], depend critically on pathology within the medial temporal lobe, suggesting that advances in our mapping of these disorders using functional imaging[82] will be reciprocally informed by links to clinical neuroscience and data from patient populations (Fig. 1b). Importantly, this is merely one example of a broader capacity inherent within functional neuroimaging to advance our understanding of how brain systems support cognitive functions in a way that meaningfully intersects with systems, computational and clinical neuroscience (Fig. 1c).

fMRI in computational neuroscience

Armed with functional neuroimaging access to cognitive processes, neuroscientists are faced with the novel challenge of how to make sense of the complex patterns inherent within functional neuroimaging data. Although the trade-off between functional localization and integration has long been appreciated[83,84], functional neuroimaging was for many years primarily focussed on 'brain mapping'—that is, developing links between particular regions or networks of the brain and specific mental functions. This approach led to major advances in our appreciation of brain function—for instance, in the characterization of distributed networks that are engaged consistently across different tasks[85] and are predictive of cognitive capacities[86]. However, the approach falls short in explaining precisely how the different specialized regions work together to give rise to higher cognitive functions, such as fluid intelligence and problem-solving[87]. This is, of course, a difficult problem that requires a 'birds-eye' view of the system (that is, precisely the vantage point offered by fMRI); however it is also challenging to discern coherent algorithmic order from a set of relatively heterogeneous studies. What is needed are ways to sift through the complexity of whole-brain imaging to make sense of the principles that link patterns of brain activity to mental functions.

A promising approach to addressing this question has emerged through the use of artificial neural networks. Whereas previous computational approaches to fMRI analysis often used relatively abstract models of a particular computational subcomponent of a task, artificial neural networks have inspired an alternative approach to mapping neural computations to brain activity[88,89]: namely, by training the connections within a structured network of brain-inspired computational units in such a way that the network can perform tasks in an end-to-end

manner (that is, from stimulus to response). These models can now perform as well as (or better) than humans on many complex tasks[90]. Although it is clear that these network are not accurate re-creations of detailed biological networks—for example, the commonly used back-propagation training algorithm may be quite distinct from the learning rules used by biological neural networks[88] and artificial neural networks fail to account for key findings from human psychology—the key issue is that the distributed nature of the networks forces the system to solve the computational problem using some of the same kinds of degrees of freedom available to biological neural networks (that is, as distributed networks of units with varying strengths of connections between them[91]). To the extent that the brain solves computational problems in a similar way to artificial neural networks, trained networks can be thought of as a type of 'computational model organism' that can be interrogated in order to identify the computational components of the model that are most predictive of signals recorded from biological brains, thus providing further insight into the function of the brain.

Perhaps most crucially, the types of tasks performed by human participants can tap into complex, language-based cognitive capacities that are essentially impossible to study using non-human models. When the data from these human experiments is subsequently mapped to computational models, the models can be extended beyond the reaches of the behaviours that can be interrogated in animals, while retaining contact with the many hard-earned facts about the nervous system identified using the tools of systems neuroscience. That is, inspiration can be drawn from computational models that successfully replicate human behaviour to help answer the difficult 'how' question (the algorithmic level, in the language of Marr[92]), while evaluating candidate solutions against measurements from fMRI—and then grounding those results with more precise measurements from systems neuroscience with known links to fMRI—to determine whether a given computational solution has a plausible neural implementation. This iterative process will afford the ability to refine existing cognitive ontologies[93,94] using mechanisms and methods that are ideally suited to unpacking the structure of cognition.

To this end, a promising area that requires further development is determining precisely if and how existing artificial neural network architectures map onto the functions of the human brain. For instance, pioneering work showed that hierarchical convolutional neural networks, which combine many layers and filters to mimic dimensionality collapse and expansion[95], provide a robust means for modelling key features of the mammalian visual system[96,97]. Precisely how these models map onto the complexity of the visual system remains an open question[98]; however, the open release of high-quality datasets designed for just this purpose will no doubt accelerate our progress in this sphere[99]. Recent work has also demonstrated similar isomorphisms between human brain activity related to language understanding (as measured using invasive electrophysiology) and the autoregressive neural network ('transformer') architectures that have become highly successful at many linguistic tasks[100].

This approach holds great promise for mapping more abstract, cognitive functions using the same tools—for example, by training recurrent neural networks to perform simulacra of cognitive tasks[101,102], after which time the function of the networks can be interrogated using the same methods that used to analyse functional neuroimaging data[103]. By comparing signatures of neural networks performing cognitive tasks with fMRI patterns from human participants, the similarities and differences in the computational solutions to cognitive challenges can be identified. Conversely, one can also use fMRI to inform the generation of new task-optimized neural network models[104]. Yet another promising approach has used artificial neural networks to assess the degree to which stimuli are actually diagnostic with respect to specific theoretical debates[105] (in this case regarding the role of perirhinal cortex in memory versus perception), leading to resolution of apparent inconsistencies between human and macaque lesion results[106]. In short, there are of course crucial ways in which human brains differ from artificial neural networks[88,107–109], yet the similarities and the differences between the two provide a number of avenues to potentially inform our understanding of how the brain gives rise to cognition.

fMRI in clinical neuroscience

At the advent of fMRI, many believed that its ultimate promise would be to develop tools for clinical and other real-world settings. Currently, fMRI is used for pre-surgical mapping to ensure that key brain regions are not damaged during resections[110], and more recently, to guide targeted placement of electrodes for deep brain stimulation[111] as well as non-invasive stimulation techniques such as transcranial magnetic stimulation[112]. However, these applications remain relatively niche, and many agree that the full promise of fMRI for diagnosing and/or treating psychiatric and neurological illnesses has not been borne out[113]. One major challenge that has prevented fMRI from making meaningful clinical inroads is that while nearly all the discoveries discussed above were based on aggregating data from multiple individuals, to develop tools with real-world value, we need to identify biomarkers for neurological or psychiatric illnesses that are sensitive and specific at the individual '*n* of 1' level, which is a fundamentally difficult statistical challenge.

The dream of fMRI-based tools that are suitable for widespread, routine clinical use may be far from reality, as proposed imaging-based biomarkers have turned out to be unreliable[114] or too weak to support robust inference at the individual level[115], perhaps because they have been built on insufficient data[116]. This is not to say that fMRI in its current state cannot make meaningful contributions to clinical neuroscience; rather, we suggest that a renewed focus on how basic scientific discoveries occurring in systems, cognitive and computational neuroscience can inform clinical neuroscience, and vice versa.

One challenge for *n* of 1 biomarkers is that behaviour and brain function are marked by considerable variability both within and across individuals. Ultimately, then, understanding links between brain, behaviour, and clinical status will require mapping of the full space of possible values for the measures extracted on a spectrum from health to pathology. Although behavioural work has long recognized within-subject fluctuations in how subjects perform a particular task (for example, trial-to-trial variability), fMRI can reveal the neural basis for why these fluctuations occur—in other words, how endogenous brain states influence if and how incoming sensory information is processed[28,117–119]. Such neural and behavioural fluctuations across a range of different timescales—from seconds to years—have the potential to inform our understanding of symptom trajectories and responses to treatment[120]. They also highlight a fundamental problem in the logic of most current neuroimaging studies of mental health disorders: By sampling individuals at a single point in time, these studies systematically neglect the temporal variability that is known to occur in many such conditions. Longitudinal designs that relate within-subject brain changes to behavioural changes in the name of prognosis may be a more fruitful avenue for clinical applications of fMRI than single-timepoint diagnostic prediction. Similarly, while rich across-subject variability (that is, individual differences) is well established across many domains (personality, affect, cognitive and motor skills, and so on), understanding the brain basis of these differences can reveal how and why different brains perform the same task or react to the same information differently, and how this might relate to disease or risk of disease[121–123]. Again, even in the absence of *n* of 1 biomarkers, basic science regarding neural variability can inform clinical practice in other ways, such as by suggesting novel interventions. For example, knowing which features or events within a complex stimulus, such as a film, trigger differential brain responses across individuals might suggest novel targets for cognitive-behavioural therapy.

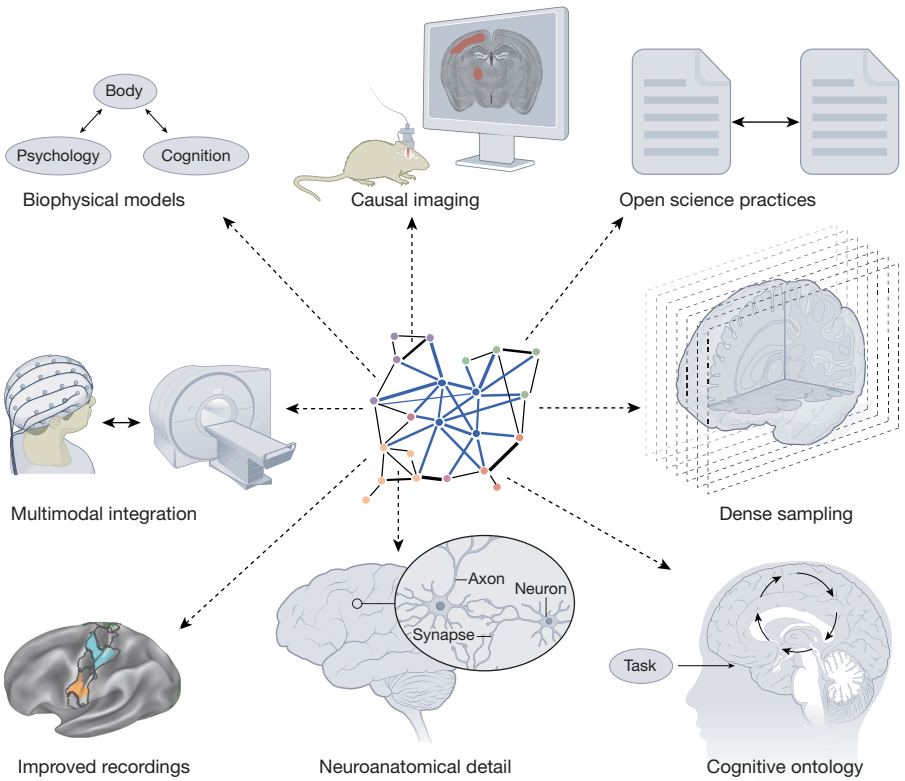

Fig. 3 | Potential paths towards a more integrative neuroscience. These paths include (clockwise, from top-left): using biophysical models to link cognition to psychology; combining precise causal perturbation of the brain (for example, with optogenetics) with whole-brain BOLD imaging to improve our understanding of the signals recorded from purely descriptive experiments; open science practices to ensure that results are robust, reproducible and generalizable; dense-sampling approaches that balance individual precision with robust statistical analyses to provide strong conclusions that retain heightened interpretability; recordings across a wide array of cognitive tasks to leverage a key strength of fMRI and mitigate the limitations of approaches in non-human animals; using circuit-level insights to populate biophysical models of the nervous system to refine our understanding of the neural mechanisms of cognitive function; improved recording techniques, such as layer-resolved fMRI and heightened resolution in the subcortex, to augment our appreciation of patterns of whole-brain coordinated activity; and understanding the links between different imaging modalities.

Another way in which fMRI can inform clinical neuroscience is by contributing to a mechanistic understanding of symptoms. One powerful example is hallucinations, or perceptual experiences in the absence of external stimulation. Hallucinations are a transdiagnostic phenomenon: while auditory (often verbal) hallucinations are a hallmark of schizophrenia, they are also associated with several other conditions and sensory domains (for example, visual hallucinations in Parkinson's disease). Several theories attempt to explain hallucinations in terms of predictive coding and aberrant weighting of perceptual priors in constructing sensory experience[124]. Older neuroimaging studies were concerned with identifying loci of activity associated with hallucinations[125] to understand the extent to which hallucinations resemble true sensory experiences in their functional neuroanatomy, whereas more recent fMRI studies are leveraging methodological advances to build a more mechanistic understanding of when, how and why hallucinations occur. For example, layer-specific fMRI can provide empirical evidence to constrain predictive coding accounts, helping to adjudicate whether aberrant prediction errors result from impaired top-down or bottom-up signals[126], and combining fMRI with computational models of conditioned hallucination tasks can shed light on the neural circuitry underlying latent cognitive processes that give rise to hallucinations[127]. In addition to having the potential to advance clinical science (for example, by suggesting new targets for therapies), both of these directions also strengthen bridges between clinical and systems, cognitive and computational neuroscience.

While much attention is rightfully focused on how discoveries in basic science can inform clinical problems, there is also a richness of possible reciprocal contributions that fMRI can facilitate from clinical neuroscience back to the basic science modules of systems, cognitive and computational neuroscience. Psychoactive medications offer a way to causally manipulate microscale phenomena—for example, by targeting specific neurotransmitter systems—in humans in an admittedly blunt, yet ethically acceptable way. If it is known from animal models how manipulating these microscale phenomena should affect more macro-scale patterns of brain activity, pharmacological fMRI studies can be used to validate that the same principles apply in humans. With respect to computational neuroscience, to the extent that it is possible to model how mental processes work and how they are instantiated in the brain, it will be possible to test whether disruptions at various levels within those models yield data that match neural and/or behavioural deficits seen in disease. As an example, pharmacological agents that perturb dopaminergic function have been used to confirm the association of fMRI signals in the striatum with dopaminergic reward prediction errors[128]. This approach is a core principle of the rapidly growing field of computational psychiatry[129].

A roadmap for the future of fMRI

This Perspective began by highlighting the relatively segregated nature of the current field of neuroscience (Fig. 1). We then argued that, despite its limitations (Fig. 2), fMRI is a tool that has a number

Box 2

The metascience of fMRI

fMRI has seen its share of controversies, including the challenges associated with circularities in data analysis[153], multiple comparisons problems[154,155], variability of results across analysis pipelines[156] and the stability of brain-wide associations with behaviour[116]. Rather than signalling the death knell of fMRI, this checkered history actually reflects a heightened degree of self-criticism that has led to improved research practices. In part as a response to the discovery of these weaknesses, the neuroimaging community is a remarkable example of a field that has effectively adopted open science[157,158] and has been honestly self-reflective and self-corrective.

These reckonings have spurred several best practices that should be celebrated. Data sharing is now becoming standard, both prospectively[159,160] and retrospectively[6]. The effectiveness of data sharing efforts has greatly benefited from efforts to build data standards, most notably the Brain Imaging Data Structure[161]. The field also benefits from a strong culture of code sharing along with the development of robust, standardized, validated and widely accessible open source analysis tools, as is evident from—for example, the rapid adoption of the fMRIPrep pre-processing workflow[162]. Further acceleration will come from the development and adoption of community standards for software development that will enable frictionless sharing of software components between development teams, as is currently being developed by the NMIND (this Neuroimaging Method Is Not Duplicated) Consortium[163]. We are optimistic that these improved and still improving research practices will accelerate discovery for years to come, and hope that the lessons learned and practices adopted in fMRI can be leveraged by other scientific communities.

of positive features that allow it to cut across and integrate these relatively isolated sub-disciplines. Through this integration, we envision a future in which fMRI acts as a connector hub[9], bringing the pieces of our field back together to rediscover communication between otherwise segregated subfields.

Advances on a number of fronts will accelerate this progression (Fig. 3). Key amongst these is a need for more precise appreciation of how the BOLD response aligns with detailed neuronal recordings at the microscopic scale, which we anticipate will emerge from multimodal image integration. By refining the biophysical understanding that enables translation between different imaging modalities, predictions created in one domain (such as the causal role that a specific cortical cell type has in a particular cognitive task[130]) and perhaps more complex cognitive contexts that are difficult or impossible to image in non-human species (such as complex social interactions[131] or highly abstract, rule-based tasks[132]) will be able to be tested in other species. We also anticipate a heightened importance of the neuroanatomical characterization of the cellular components of the human brain, specifically so as to ensure that comparisons across species are conducted with high precision; this aim will be greatly facilitated by the ongoing efforts of the BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative Cell Atlas Network. Combining this cellular-level detail with the causal techniques used in non-human model organisms[133] as well as BOLD imaging will markedly improve our ability to interpret the measures that we collect non-invasively in human subjects. The complex causal networks identified using these approaches will further augment existing biophysical models that expose the mechanisms underlying higher-level brain functions.

Progress in the field will also be catalysed by continued advances in signal acquisition and paradigm design, particularly those that improve our ability to discern different neuroanatomically relevant signals in the brain during the performance of cognitive tasks. By carefully considering the breadth and scope of cognitive tasks used in fMRI studies in both humans and in animal models, the wide range of our cognitive capacities will be more effectively mapped in both healthy and disease states. To this end, studies designed to robustly track individual differences in whole-brain organization will enable us to combine the power of statistical approaches with the precision afforded by idiosyncratic cognitive and affective capacities. Finally, from a meta-scientific perspective, in recent years the fMRI community has been at the forefront of practices such as data and code sharing and efforts to ensure replicability and generalizability that can serve as a model for other scientific communities (Box 2). This, in turn, can reinforce the integrative role of fMRI as well as help ensure overall scientific progress within and across subfields.

The complex, multi-scale organization of the brain does not lend itself to interrogation from a single vantage point. Overall, we maintain that the next wave of breakthroughs in neuroscience will be catalysed by a pluralistic approach that leverages the benefits of particular empirical tools to bolster the weaknesses of others, with fMRI serving as a key connector hub. In sum, there is a timely opportunity to refine our understanding of how the coordinated activity of the brain shapes the intrinsic processes that generate human behaviour and thought.

1. Wagner, T., Rushmore, J., Eden, U. & Valero-Cabre, A. Biophysical foundations underlying TMS: setting the stage for an effective use of neurostimulation in the cognitive neurosciences. *Cortex* **45**, 1025–1034 (2009).
2. Murphy, S. C., Palmer, L. M., Nyffeler, T., Müri, R. M. & Larkum, M. E. Transcranial magnetic stimulation (TMS) inhibits cortical dendrites. *eLife* **5**, e13598 (2016).
3. Fitzgerald, P. B. Targeting repetitive transcranial magnetic stimulation in depression: do we really know what we are stimulating and how best to do it? *Brain Stimul.* **14**, 730–736 (2021).
4. Lewis, L. D., Setsompop, K., Rosen, B. R. & Polimeni, J. R. Fast fMRI can detect oscillatory neural activity in humans. *Proc. Natl Acad. Sci. USA* **113**, E6679–E6685 (2016). **This study demonstrates a means for accelerating acquisition times in fMRI sequences, thus providing researchers access to rapid fluctuations in BOLD signals**.
5. Barch, D. M. et al. Function in the human connectome: task-fMRI and individual differences in behavior. *NeuroImage* **80**, 169–189 (2013).
6. Markiewicz, C. J. et al. The OpenNeuro resource for sharing of neuroscience data. *eLife* **10**, e71774 (2021).
7. Zuo, X.-N. et al. An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci. Data* **1**, 140049 (2014).
8. Logothetis, N. K. et al. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
9. Guimerà, R. & Nunes Amaral, L. A. Functional cartography of complex metabolic networks. *Nature* **433**, 895–900 (2005).
10. Bertolero, M. A., Yeo, B. T. T. & D'Esposito, M. The diverse club. *Nat. Commun.* **8**, 1277 (2017).
11. Shine, J. M. et al. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* **92**, 544–554 (2016).
12. Glover, G. H. Overview of functional magnetic resonance imaging. *Neurosurg. Clin. N. Am.* **22**, 133–139 (2011).
13. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
14. Logothetis, N. K. & Wandell, B. A. Interpreting the BOLD Signal. *Annu. Rev. Physiol.* **66**, 735–769 (2004).
15. MacVicar, B. A. & Newman, E. A. Astrocyte regulation of blood flow in the brain. *Cold Spring Harb. Perspect. Biol.* **7**, a020388 (2015).
16. Kahn, I. et al. Optogenetic drive of neocortical pyramidal neurons generates fMRI signals that are correlated with spiking activity. *Brain Res.* **1511**, 33–45 (2013).
17. Huber, L. et al. Layer-dependent functional connectivity methods. *Prog. Neurobiol.* **207**, 101835 (2020).
18. Zhang, Y. et al. Rapid detection of neurons in widefield calcium imaging datasets after training with synthetic data. *Nat. Methods* **20**, 747–754 (2023).
19. Evans, M. H., Petersen, R. S. & Humphries, M. D. On the use of calcium deconvolution algorithms in practical contexts. Preprint at *bioRxiv* https://doi.org/10.1101/871137 (2019).
20. Park, H.-J. & Friston, K. Structural and functional brain networks: from connections to cognition. *Science* **342**, 1238411 (2013).
21. Betzel, R. F. & Bassett, D. S. Multi-scale brain networks. *NeuroImage* **160**, 73–83 (2017).

22. Ekstrom, A. How and when the fMRI BOLD signal relates to underlying neural activity: the danger in dissociation. *Brain Res. Rev.* **62**, 233–244 (2010).
23. Howarth, C., Mishra, A. & Hall, C. N. More than just summed neuronal activity: how multiple cell types shape the BOLD response. *Phil. Trans. R. Soc. B* **376**, 20190630 (2021).
24. Takata, N. et al. Optogenetic astrocyte activation evokes BOLD fMRI response with oxygen consumption without neuronal activity modulation. *Glia* **66**, 2013–2023 (2018).
25. Damoiseaux, J. S. et al. Consistent resting-state networks across healthy subjects. *Proc. Natl Acad. Sci. USA* **103**, 13848–13853 (2006).
26. Smith, S. M. et al. Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* **17**, 666–682 (2013).
27. Krienen, F. M., Yeo, B. T. T. & Buckner, R. L. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Phil. Trans. R. Soc. B* **369**, 20130526 (2014).
28. Chang, C. et al. Tracking brain arousal fluctuations with fMRI. *Proc. Natl Acad. Sci. USA* **113**, 4518–4523 (2016).
In this paradigm-shifting study, the authors tracked covariance between BOLD signals and a range of physiological measures, highlighting the close inter-relationships between these markers.
29. Bassett, D. S., Yang, M., Wymbs, N. F. & Grafton, S. T. Learning-induced autonomy of sensorimotor systems. *Nat. Neurosci.* **18**, 744–751 (2015).
30. Uddin, L. Q. Typical and atypical development of functional human brain networks: insights from resting-state fMRI. *Front. Syst. Neurosci.* **4**, 21 (2010).
31. Jack, C. R. et al. Magnetic resonance imaging in Alzheimer's Disease Neuroimaging Initiative. *Alzheimers Dement.* **11**, 740–756 (2015).
32. Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl Acad. Sci. USA* **113**, 12574–12579 (2016).
33. Demirtaş, M. et al. Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron* **101**, 1181–1194.e13 (2019).
34. Vázquez-Rodríguez, B. et al. Gradients of structure–function tethering across neocortex. *Proc. Natl Acad. Sci. USA* **116**, 21219–21227 (2019).
35. Paquola, C. et al. Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLoS Biol.* **17**, e3000284 (2019).
36. Fulcher, B. D. & Fornito, A. A transcriptional signature of hub connectivity in the mouse connectome. *Proc. Natl Acad. Sci. USA* **113**, 1435–1440 (2016).
37. Vértes, P. E. et al. Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. *Phil. Trans. R. Soc. B* **371**, 20150362 (2016).
38. Richiardi, J. et al. Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241–1244 (2015).
39. Lin, M. Z. & Schnitzer, M. J. Genetically encoded indicators of neuronal activity. *Nat. Neurosci.* **19**, 1142–1153 (2016).
40. Ferris, C. F. Applications in awake animal magnetic resonance imaging. *Front. Neurosci.* **16**, 854377 (2022).
41. Polimeni, J. R. & Lewis, L. D. Imaging faster neural dynamics with fast fMRI: a need for updated models of the hemodynamic response. *Prog. Neurobiol.* **207**, 102174 (2021).
42. Gutierrez-Barragan, D. et al. Unique spatiotemporal fMRI dynamics in the awake mouse brain. *Curr. Biol.* **32**, 631–644.e6 (2022).
43. Zerbi, V., Grandjean, J., Rudin, M. & Wenderoth, N. Mapping the mouse brain with rs-fMRI: An optimized pipeline for functional network identification. *NeuroImage* **123**, 11–21 (2015).
44. Zerbi, V. et al. Rapid reconfiguration of the functional connectome after chemogenetic locus coeruleus activation. *Neuron* **103**, 702–718.e5 (2019).
This study combined whole-brain BOLD signal detection with triggering of activity in the locus coeruleus of mice, demonstrating macroscopic network reconfigurations similar to those identified in humans during cognitive task performance, conceptually bridging the microscopic and macroscopic scales of analysis.
45. Lee, J. H. et al. Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. *Nature* **465**, 788–792 (2010).
This landmark paper showed that combined optogenetic control of both the cerebral cortex and thalamus in the mouse brain elicits positive BOLD signals.
46. Oya, H. et al. Mapping effective connectivity in the human brain with concurrent intracranial electrical stimulation and BOLD-fMRI. *J. Neurosci. Methods* **277**, 101–112 (2017).
47. Thompson, W. H. et al. A data resource from concurrent intracranial stimulation and functional MRI of the human brain. *Sci. Data* **7**, 258 (2020).
48. Salvan, P. et al. Serotonin regulation of behavior via large-scale neuromodulation of serotonin receptor networks. *Nat. Neurosci.* **26**, 53–63 (2023).
49. Lake, E. M. R. et al. Simultaneous cortex-wide fluorescence Ca^{2+} imaging and whole-brain fMRI. *Nat. Methods* **17**, 1262–1271 (2020).
This article presents a method for concurrent widefield optical imaging and fMRI, enabling cell-type-specific investigations of how different neural populations contribute to the fMRI signal as well as more precise translation between mouse models and human studies.
50. Valdes-Sosa, P. A. et al. Model driven EEG/fMRI fusion of brain oscillations. *Hum. Brain Mapp.* **30**, 2701–2721 (2009).
51. Mele, G. et al. Simultaneous EEG-fMRI for functional neurological assessment. *Front. Neurol.* **10**, 848 (2019).
52. Destexhe, A., Rudolph, M. & Paré, D. The high-conductance state of neocortical neurons in vivo. *Nat. Rev. Neurosci.* **4**, 739–751 (2003).
53. Brown, E. N., Purdon, P. L. & Van Dort, C. J. General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annu. Rev. Neurosci.* **34**, 601–628 (2011).
54. García-Cabezas, M. Á., Zikopoulos, B. & Barbas, H. The Structural Model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. *Brain Struct. Funct.* **72**, 429 (2019).
55. Hansen, J. Y. et al. Mapping neurotransmitter systems to the structural and functional organization of the human neocortex. *Nat. Neurosci.* **25**, 1569–1581 (2022).
This study collated data from multiple different neuroimaging modalities to track the spatial relationships between different classes of neuromodulatory receptors, thus providing a benchmark for future studies to orient their results to a broader neuroanatomical framework.
56. Weiler, N., Wood, L., Yu, J., Solla, S. A. & Shepherd, G. M. G. Top-down laminar organization of the excitatory network in motor cortex. *Nat. Neurosci.* **11**, 360–366 (2008).
57. Finn, E. S., Huber, L., Jangraw, D. C., Molfese, P. J. & Bandettini, P. A. Layer-dependent activity in human prefrontal cortex during working memory. *Nat. Neurosci.* **22**, 1687–1695 (2019).
This article represents the first application of layer-specific fMRI to human association cortex, demonstrating that the technique can go beyond primary sensory cortex to reveal cortical-depth-dependent processes during a higher-order cognitive task.
58. Bastos, A. M., Loonis, R., Kornblith, S., Lundqvist, M. & Miller, E. K. Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proc. Natl Acad. Sci. USA* **115**, 1117–1122 (2018).
59. Stephan, K. E. et al. Laminar fMRI and computational theories of brain function. *NeuroImage* **197**, 699–706 (2019).
60. Kosslyn, S. M., Ganis, G. & Thompson, W. L. Neural foundations of imagery. *Nat. Rev. Neurosci.* **2**, 635–642 (2001).
61. Iamshchinina, P. et al. Perceived and mentally rotated contents are differentially represented in cortical depth of V1. *Commun. Biol.* **4**, 1069 (2021).
62. Persichetti, A. S., Avery, J. A., Huber, L., Merriam, E. P. & Martin, A. Layer-specific contributions to imagined and executed hand movements in human primary motor cortex. *Curr. Biol.* **30**, 1721–1725.e3 (2020).
63. Klein, B. P. et al. Cortical depth dependent population receptive field attraction by spatial attention in human V1. *NeuroImage* **176**, 301–312 (2018).
64. Gau, R., Bazin, P.-L., Trampel, R., Turner, R. & Noppeney, U. Resolving multisensory and attentional influences across cortical depth in sensory cortices. *eLife* **9**, e46856 (2020).
65. Liu, C. et al. Layer-dependent multiplicative effects of spatial attention on contrast responses in human early visual cortex. *Prog. Neurobiol.* **207**, 101897 (2021).
66. Doeller, C. F., Barry, C. & Burgess, N. Evidence for grid cells in a human memory network. *Nature* **463**, 657–661 (2010).
67. Constantinescu, A. O., O'Reilly, J. X. & Behrens, T. E. J. Organizing conceptual knowledge in humans with a gridlike code. *Science* **352**, 1464–1468 (2016).
This article uses model-based fMRI to show that the grid-like organization used by the brain to represent physical space also extends to conceptual knowledge, and is present in the entorhinal cortex, prefrontal cortex and several other brain regions.
68. Deuker, L., Bellmund, J. L., Navarro Schröder, T. & Doeller, C. F. An event map of memory space in the hippocampus. *eLife* **5**, e16534 (2016).
69. Gläscher, J. P. & O'Doherty, J. P. Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data. *WIREs Cogn. Sci.* **1**, 501–510 (2010).
70. Williams, T. B. et al. Testing models at the neural level reveals how the brain computes subjective value. *Proc. Natl Acad. Sci. USA* **118**, e2106237118 (2021).
71. Zhang, L. & Gläscher, J. A brain network supporting social influences in human decision-making. *Sci. Adv.* **6**, eabb4159 (2020).
72. McDougle, S. D. et al. Neural signatures of prediction errors in a decision-making task are modulated by action execution failures. *Curr. Biol.* **29**, 1606–1613.e5 (2019).
73. Lee, S. W., Shimojo, S. & O'Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **81**, 687–699 (2014).
74. Schuck, N. W. & Niv, Y. Sequential replay of nonspatial task states in the human hippocampus. *Science* **364**, eaaw5181 (2019).
75. Kurth-Nelson, Z., Economides, M., Dolan, R. J. & Dayan, P. Fast sequences of non-spatial state representations in humans. *Neuron* **91**, 194–204 (2016).
76. Abe, Y. et al. Opto-fMRI analysis for exploring the neuronal connectivity of the hippocampal formation in rats. *Neurosci. Res.* **74**, 248–255 (2012).
77. Basu, J. & Siegelbaum, S. A. The corticohippocampal circuit, synaptic plasticity, and memory. *Cold Spring Harb. Perspect. Biol.* **7**, a021733 (2015).
78. Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M. & Norman, K. A. Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Phil. Trans. R. Soc. B* **372**, 20160049 (2017).
79. Chanales, A. J. H., Oza, A., Favila, S. E. & Kuhl, B. A. Overlap among spatial memories triggers repulsion of hippocampal representations. *Curr. Biol.* **27**, 2307–2317.e5 (2017).
80. Double, K. L. et al. Topography of brain atrophy during normal aging and Alzheimer's disease. *Neurobiol. Aging* **17**, 513–521 (1996).
81. Engel, J. Mesial temporal lobe epilepsy: what have we learned? *Neuroscientist* **7**, 340–352 (2001).
82. Dennis, E. L. & Thompson, P. M. Functional brain connectivity Using fMRI in aging and Alzheimer's disease. *Neuropsychol. Rev.* **24**, 49–62 (2014).
83. Friston, K. J., Frith, C. D., Liddle, P. F. & Frackowiak, R. S. Functional connectivity: the principal-component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* **13**, 5–14 (1993).
84. McIntosh, A. R. Contexts and catalysts: a resolution of the localization and integration of function in the brain. *Neuroinformatics* **2**, 175–182 (2004).
85. Smith, S. M. et al. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl Acad. Sci. USA* **106**, 13040–13045 (2009).
86. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
87. Shine, J. M. & Poldrack, R. A. Principles of dynamic network reconfiguration across diverse brain states. *NeuroImage* **180**, 396–405 (2017).

88. Richards, B. A. et al. A deep learning framework for neuroscience. *Nat. Neurosci.* **22**, 1761–1770 (2019).
89. Yamins, D. L. K. & DiCarlo, J. J. Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* **19**, 356–365 (2016).
90. Bubeck, S. et al. Sparks of artificial general intelligence: early experiments with GPT-4. Preprint at *arXiv* https://doi.org/10.48550/ARXIV.2303.12712 (2023).
91. Doerig, A. et al. The neuroconnectionist research programme. *Nat. Rev. Neurosci.* **24**, 431–450 (2023).
92. Marr, D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (MIT Press, 2010).
93. Poldrack, R. A. & Yarkoni, T. From brain maps to cognitive ontologies: informatics and the search for mental structure. *Annu. Rev. Psychol.* **67**, 587–612 (2016).
94. Bolt, T., Anderson, M. L. & Uddin, L. Q. Beyond the evoked/intrinsic neural process dichotomy. *Netw. Neurosci.* https://doi.org/10.1162/netn_a_00028 (2018).
95. Lindsay, G. W. Convolutional neural networks as a model of the visual system: past, present, and future. *J. Cogn. Neurosci.* **33**, 2017–2031 (2021).
96. Yamins, D. L. K. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl Acad. Sci. USA* **111**, 8619–8624 (2014).
97. Kriegeskorte, N. Deep neural networks: a new framework for modeling biological vision and brain information processing. *Annu. Rev. Vis. Sci.* **1**, 417–446 (2015).
98. Bowers, J. S. et al. Deep problems with neural network models of human vision. *Behav. Brain Sci.* https://doi.org/10.1017/S0140525X22002813 (2022).
99. Allen, E. J. et al. A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nat. Neurosci.* **25**, 116–126 (2022).
100. Goldstein, A. et al. Shared computational principles for language processing in humans and deep language models. *Nat. Neurosci.* **25**, 369–380 (2022).
101. Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T. & Wang, X.-J. Task representations in neural networks trained to perform many cognitive tasks. *Nat. Neurosci.* **22**, 297–306 (2019).
102. Mastrogiuseppe, F. & Ostojic, S. Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron* **99**, 609–623.e29 (2018).
This landmark study demonstrated how the low-rank yet noisy structure of a connectome can directly shape the low-dimensional manifold that characterizes the emergent behaviour of a recurrent neural network.
103. Thomas, A. W., Ré, C. & Poldrack, R. A. Interpreting mental state decoding with deep learning models. *Trends Cogn. Sci.* **26**, 972–986 (2022).
104. Ito, T., Yang, G. R., Laurent, P., Schultz, D. H. & Cole, M. W. Constructing neural network models from brain data reveals representational transformations linked to adaptive behavior. *Nat. Commun.* **13**, 673 (2022).
105. Bonnen, T., Yamins, D. L. K. & Wagner, A. D. When the ventral visual stream is not enough: a deep learning account of medial temporal lobe involvement in perception. *Neuron* **109**, 2755–2766.e6 (2021).
106. Bonnen, T. & Eldridge, M. A. G. Inconsistencies between human and macaque lesion data can be resolved with a stimulus-computable model of the ventral visual stream. *eLife* **12**, e84357 (2023).
107. Zador, A. M. A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat. Commun.* **10**, 3770 (2019).
108. Hassabis, D., Kumaran, D., Summerfield, C. & Botvinick, M. Neuroscience-inspired artificial intelligence. *Neuron* **95**, 245–258 (2017).
109. Hasson, U., Nastase, S. A. & Goldstein, A. Direct fit to nature: an evolutionary perspective on biological and artificial neural networks. *Neuron* **105**, 416–434 (2020).
110. Silva, M. A., See, A. P., Essayed, W. I., Golby, A. J. & Tie, Y. Challenges and techniques for presurgical brain mapping with functional MRI. *NeuroImage* **17**, 794–803 (2018).
111. Boutet, A. et al. Predicting optimal deep brain stimulation parameters for Parkinson's disease using functional MRI and machine learning. *Nat. Commun.* **12**, 3043 (2021).
112. Cole, E. J. et al. Stanford neuromodulation therapy (SNT): a double-blind randomized controlled trial. *AJP* **179**, 132–141 (2022).
113. Abi-Dargham, A. et al. Candidate biomarkers in psychiatric disorders: state of the field. *World Psychiatry* **22**, 236–262 (2023).
114. Dinga, R. et al. Evaluating the evidence for biotypes of depression: methodological replication and extension of. *NeuroImage* **22**, 101796 (2019).
115. Dukart, J., Weis, S., Genon, S. & Eickhoff, S. B. Towards increasing the clinical applicability of machine learning biomarkers in psychiatry. *Nat. Hum. Behav.* **5**, 431–432 (2021).
116. Marek, S. et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).
This article calls attention to the need for very large sample sizes when performing mass univariate correlations between measures of structural and/or functional brain connectivity and trait-like behaviours across individuals.
117. Rosenberg, M. D. et al. Functional connectivity predicts changes in attention observed across minutes, days, and months. *Proc. Natl Acad. Sci. USA* **117**, 3797–3807 (2020).
This article shows that the same functional brain connections that predict overall sustained attention ability also predict within-subject changes in attention observed over several timescales as well as during pharmacological manipulation, shedding light on the relationship between trait- and state-level factors in a manner with implications for studying symptom fluctuations in clinical conditions.
118. Podvalny, E., Flounders, M. W., King, L. E., Holroyd, T. & He, B. J. A dual role of prestimulus spontaneous neural activity in visual object recognition. *Nat. Commun.* **10**, 3910 (2019).
119. Sadaghiani, S., Poline, J.-B., Kleinschmidt, A. & D'Esposito, M. Ongoing dynamics in large-scale functional connectivity predict perception. *Proc. Natl Acad. Sci. USA* **112**, 8463–8468 (2015).
120. Burciu, R. G. et al. Functional MRI of disease progression in Parkinson disease and atypical parkinsonian syndromes. *Neurology* **87**, 709–717 (2016).
121. Finn, E. S., Corlett, P. R., Chen, G., Bandettini, P. A. & Constable, R. T. Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. *Nat. Commun.* **9**, 2043 (2018).
122. Rikandi, E. et al. Precuneus functioning differentiates first-episode psychosis patients during the fantasy movie Alice in Wonderland. *Psychol. Med.* **47**, 495–506 (2017).
123. Salmi, J. et al. ADHD desynchronizes brain activity during watching a distracted multi-talker conversation. *NeuroImage* **216**, 116352 (2020).
124. Sterzer, P. et al. The predictive coding account of psychosis. *Biol. Psychiatry* **84**, 634–643 (2018).
125. Jardri, R., Pouchet, A., Pins, D. & Thomas, P. Cortical activations during auditory verbal hallucinations in schizophrenia: a coordinate-based meta-analysis. *Am. J. Psychiatry* **168**, 73–81 (2011).
126. Haarsma, J., Kok, P. & Browning, M. The promise of layer-specific neuroimaging for testing predictive coding theories of psychosis. *Schizophr. Res.* **245**, 68–76 (2022).
127. Powers, A. R., Mathys, C. & Corlett, P. R. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science* **357**, 596–600 (2017).
128. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).
129. Huys, Q. J. M., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* **19**, 404–413 (2016).
130. Christensen, A. J., Ott, T. & Kepecs, A. Cognition and the single neuron: How cell types construct the dynamic computations of frontal cortex. *Curr. Opin. Neurobiol.* **77**, 102630 (2022).
131. Misaki, M. et al. Beyond synchrony: the capacity of fMRI hyperscanning for the study of human social interaction. *Soc. Cogn. Affect. Neurosci.* **16**, 84–92 (2021).
132. Shine, J. M. et al. The low-dimensional neural architecture of cognitive complexity is related to activity in medial thalamic nuclei. *Neuron* **104**, 849–855.e3 (2019).
133. Deisseroth, K. Optogenetics: 10 years of microbial opsins in neuroscience. *Nat. Neurosci.* **18**, 1213–1225 (2015).
134. Yacoub, E., Harel, N. & Uğurbil, K. High-field fMRI unveils orientation columns in humans. *Proc. Natl Acad. Sci. USA* **105**, 10607–10612 (2008).
135. De Martino, F. et al. Frequency preference and attention effects across cortical depths in the human primary auditory cortex. *Proc. Natl Acad. Sci. USA* **112**, 16036–16041 (2015).
136. Wittkuhn, L. & Schuck, N. W. Dynamics of fMRI patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nat. Commun.* **12**, 1795 (2021).
137. Rocca, R. et al. Language beyond the language system: Dorsal visuospatial pathways support processing of demonstratives and spatial language during naturalistic fast fMRI. *NeuroImage* **216**, 116128 (2020).
138. Sonkusare, S., Breakspear, M. & Guo, C. Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci.* **23**, 699–714 (2019).
139. Hamilton, L. S. & Huth, A. G. The revolution will not be controlled: natural stimuli in speech neuroscience. *Lang. Cogn. Neurosci.* **35**, 573–582 (2020).
140. Lenormand, D. & Piolino, P. In search of a naturalistic neuroimaging approach: exploration of general feasibility through the case of VR-fMRI and application in the domain of episodic memory. *Neurosci. Biobehav. Rev.* **133**, 104499 (2022).
141. Tomov, M. S., Tsividis, P. A., Pouncy, T., Tenenbaum, J. B. & Gershman, S. J. The neural architecture of theory-based reinforcement learning. *Neuron* **111**, 1331–1344.e8 (2023).
142. Kay, K. N., Naselaris, T., Prenger, R. J. & Gallant, J. L. Identifying natural images from human brain activity. *Nature* **452**, 352–355 (2008).
143. Mitchell, T. M. et al. Predicting human brain activity associated with the meanings of nouns. *Science* **320**, 1191–1195 (2008).
144. Haxby, J. V. et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425–2430 (2001).
145. Polyn, S. M., Natu, V. S., Cohen, J. D. & Norman, K. A. Category-specific cortical activity precedes retrieval during memory search. *Science* **310**, 1963–1966 (2005).
146. Kriegeskorte, N. Representational similarity analysis—connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* https://doi.org/10.3389/neuro.06.004.2008 (2008).
147. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
148. Nastase, S. A., Gazzola, V., Hasson, U. & Keysers, C. Measuring shared responses across subjects using intersubject correlation. *Soc. Cogn. Affect. Neurosci.* **14**, 667–685 (2019).
149. Song, H., Shim, W. M. & Rosenberg, M. D. Large-scale neural dynamics in a shared low-dimensional state space reflect cognitive and attentional dynamics. *eLife* **12**, e85487 (2023).
150. van der Meer, J. N., Breakspear, M., Chang, L. J., Sonkusare, S. & Cocchi, L. Movie viewing elicits rich and reliable brain state dynamics. *Nat. Commun.* **11**, 5004 (2020).
151. Munn, B. R., Müller, E. J., Wainstein, G. & Shine, J. M. The ascending arousal system shapes neural dynamics to mediate awareness of cognitive states. *Nat. Commun.* **12**, 6016 (2021).
152. Gordon, E. M. et al. A somato-cognitive action network alternates with effector regions in motor cortex. *Nature* **617**, 351–359 (2023).
153. Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F. & Baker, C. I. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* **12**, 535–540 (2009).
154. Bennett, C., Miller, M. & Wolford, G. Neural correlates of interspecies perspective taking in the post-mortem Atlantic salmon: an argument for multiple comparisons correction. *NeuroImage* **47**, S125 (2009).
155. Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl Acad. Sci. USA* **113**, 7900–7905 (2016).
156. Botvinik-Nezer, R. et al. Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* **582**, 84–88 (2020).
157. Zuo, X.-N., Xu, T. & Milham, M. P. Harnessing reliability for neuroscience research. *Nat. Hum. Behav.* **3**, 768–771 (2019).
158. Gorgolewski, K. J. & Poldrack, R. A. A practical guide for improving transparency and reproducibility in neuroimaging research. *PLoS Biol.* **14**, e1002506 (2016).
This important benchmark paper highlights goals for improving standard practice in neuroimaging research.

159. Van Essen, D. C. et al. The WU-Minn Human Connectome Project: an overview. *NeuroImage* **80**, 62–79 (2013).
160. Karcher, N. R. & Barch, D. M. The ABCD study: understanding the development of risk for mental and physical health outcomes. *Neuropsychopharmacology* **46**, 131–142 (2021).
161. Gorgolewski, K. J. et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* **3**, 160044 (2016).
162. Esteban, O. et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
163. Kiar, G. et al. Align with the NMIND consortium for better neuroimaging. *Nat. Hum. Behav.* **7**, 1027–1028 (2023).

Acknowledgements The authors thank D. Fair for his engagement on earlier versions of this manuscript and K. Nautiyal for helpful discussions.

Author contributions All authors contributed in the generation, execution and revision of the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to Emily S. Finn, Russell A. Poldrack or James M. Shine.
Peer review information *Nature* thanks Thomas Nickl-Jockschat, Nicholas Turk-Browne, Lucina Uddin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.