HHS Public Access

Author manuscript

Nat Rev Neurosci. Author manuscript; available in PMC 2022 August 18.

Published in final edited form as:

Nat Rev Neurosci. 2022 June ; 23(6): 361–375. doi:10.1038/s41583-022-00583-8.

Causal mapping of human brain function

Shan H. Siddiqi[*,1,2], **Konrad P. Kording**[3], **Josef Parvizi**[4], **Michael D. Fox**[1,5]

[1]Center for Brain Circuit Therapeutics, Brigham & Women's Hospital, Boston, MA, USA

[2]Department of Psychiatry, Harvard Medical School, Boston, MA, USA

[3]Department of Neuroscience, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

[4]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Palo Alto, CA, USA

[5]Department of Neurology, Harvard Medical School, Boston, MA, USA

Abstract

Mapping human brain function is a longstanding goal of neuroscience that promises to inform the development of new treatments for brain disorders. Early maps of human brain function were based on locations of brain damage or brain stimulation that caused a change in function. Over time, this approach was largely replaced by technologies such as functional neuroimaging, which identify brain regions correlated with behaviors or symptoms. Despite their advantages, these technologies reveal correlation, not causation. This creates challenges for interpreting the data generated from these tools and using them to develop treatments for brain disorders. A return to causal mapping of human brain function based on brain lesions and brain stimulation is underway. New approaches can combine these causal sources of information with modern neuroimaging and electrophysiology techniques to gain new insights into the functions of specific brain areas. In this Review, we provide a definition of causality for translational research, propose a continuum along which to assess the relative strength of causal information from human brain mapping studies, and discuss recent advances in causal brain mapping and their relevance for developing treatments.

Introduction

Human brain mapping studies often rely on neurophysiological correlates of symptoms or behavior[1–4]. This approach has yielded a wealth of knowledge about brain organization[5–8], but this knowledge has not routinely translated into therapeutic targets[3,9–11]. The difference between correlation and causation may be an important reason for this lack of translation[12,13]. Correlative methods do not distinguish between brain regions that are causing a symptom, compensating for a symptom, or incidentally related to a symptom. As such, interventions based on neurophysiological correlates could improve a symptom, worsen a symptom, or have no effect.

* shsiddiqi@bwh.harvard.edu .

This "causality gap" is a well-recognized problem in human neuroscience[10,13,14], and echoes a problem that other disciplines have also faced. For instance, the field of econometrics experienced a "credibility revolution"[15] after an influential push to use causal methods in the early 1980s[16]. By emulating this example, the causality gap in neuroscience may be better framed as a causality opportunity. Wider adoption of causal methods could improve our ability to translate brain mapping findings into successful interventions[13,14,17,18]. Causal mapping studies using brain lesions and brain stimulation helped lead to mainstream clinical treatments such as deep brain stimulation (DBS) for Parkinson disease[19–22] and transcranial magnetic stimulation (TMS) for major depression[23,24]. These successes may provide a road map for identifying new therapeutic targets based on causal mapping of human brain function.

In this Review, we provide a clinically-relevant definition of causality, summarize different approaches to causal brain mapping, propose a framework for appraising the strength of causality in brain mapping studies, and describe how these studies can be translated into anatomically-targeted treatments for neuropsychiatric disorders.

Brain mapping and causal inference

A brief history of human brain mapping

Early brain maps were based on lesions or stimulation sites that directly caused a functional change, ranging from influential lesion cases such as Phineas Gage[25] and H.M.[26] to systematic experiments such as Wilder Penfield's intracranial stimulation studies[27–29]. However, in the latter part of the 20th century, the advent of functional neuroimaging[1,2,4,30] was accompanied by a decline in lesion and stimulation studies (Fig. 1)[31]. Neuroimaging allowed for noninvasive mapping of human brain function in patients without lesions, including healthy subjects. Neuroimaging correlates often aligned with prior lesion and stimulation results[32], and provided additional insights about complex functions such as language and memory[33,34].

Nevertheless, this new capability came with a cost, as these correlative tools moved away from the causal inference that was possible with brain lesions and brain stimulation[9,14,17,18,30]. Some saw this as a minor limitation[35], but others criticized functional neuroimaging as "the new phrenology"[36]. More importantly, despite its popularity, functional neuroimaging was not leading to the development of new clinical treatments[14,17,18]. This led to essays highlighting the causality gap in human neuroscience[14,17] and the need to combine functional neuroimaging with causal techniques[17,18].

A brief history of causal inference

The history of causal inference has largely focused on addressing one fundamental problem: to conclude that an intervention caused an outcome, one must determine what would have happened if the intervention had been absent[37,38]. The counterfactual, or what would happen if an activity had been different, is unobservable by definition. Therefore, definitive causal inference is impossible. Instead, researchers make careful assumptions to estimate the

counterfactual by studying the effects of a perturbation, such as a focal disruption of brain activity[10,39].

The value of understanding causality was recognized by philosophers as early as Aristotle, who believed that true knowledge is unattainable without asking "why"[40,41]. Aristotle's tradition inspired medieval theologians such as the Islamic physician–philosopher ibn Sina and the Catholic priest Thomas Aquinas[42]. Both ibn Sina and Aquinas reasoned that any causal chain either can be traced infinitely or should end in a primary cause, which Aquinas labeled as "God"[42]. To explore the concept of free will, ibn Sina also described secondary causes which occur further down the chain set in place by God. As a physician, ibn Sina emphasized the need to modify outcomes by manipulating these secondary causes[43].

Philosophers debated this topic for centuries until David Hume codified eight logical "rules" for causal inference in 1748[44]. This early attempt inspired subsequent systems that were concrete enough to apply scientifically. For instance, Robert Koch's microbiological "postulates" were successfully used to identify the etiology of tuberculosis and cholera[45]. More recently, Austin Bradford-Hill's epidemiological "criteria" revealed many important causal relationships, such as the link between smoking and lung cancer[46–48].

The Bradford-Hill criteria include temporality (the only mandatory criterion), specificity, effect size, reproducibility, dose-response, physiological plausibility, experimental manipulation, analogy, reversibility (an optional criterion), and coherence at different levels. The link between smoking and lung cancer exemplifies how these criteria can be used to establish a causal relationship: smoking consistently precedes lung cancer (temporality); lung cancer increases specifically after smoking, but not after related exposures such as chewing tobacco and alcohol (specificity); smoking has a large effect on the risk of lung cancer (effect size); the relationship between smoking and lung cancer is consistent across different settings (reproducibility); more smoking leads to greater lung cancer risk (dose-response); smoking causes single-strand DNA breaks, which predispose to neoplasia (physiological plausibility); in animal models, smoking increases cancer risk more than control interventions (experimental manipulation); the effects of smoking are similar to those of other known carcinogens (analogy); some smoking-induced changes are reversible upon cessation of smoking (reversibility); and these criteria are consistently satisfied by *in vitro* studies, animal studies, human studies, and epidemiological studies (coherence).

Causal inference in biomedical science

The medical sciences did not explicitly focus on causality until ibn Sina's vision of controlled clinical trials[49] was finally realized in the mid-20th century[50]. Randomized trials changed the face of clinical therapeutics because they established causality by minimizing confounders, as the control group is identical to the intervention group in nearly every way except for the intervention itself. Randomized controlled trials thus became the gold standard for estimating the counterfactual by modeling a situation in which the intervention was absent[50].

However, before designing a clinical trial, one must identify components of the causal chain as potential treatment targets[51] — otherwise, there is a risk of intervening based

on misleading associations[52]. For instance, tuberculosis rates correlate with malaria rates, body temperature, mycobacterial infection, and travel to endemic countries[53]. This might lead to the incorrect conclusion that tuberculosis could be treated with antimalarials or antipyretics. In reality, the correlation with malaria is spurious, as both diseases are endemic to lower-income countries, and the correlation with body temperature is due to reverse causation, as tuberculosis causes fever. By contrast, mycobacterial infection is causal, as the counterfactual has been demonstrated with confidence using Koch's postulates[54]. Travel to endemic countries is also in the causal chain of events, but only because it increases the risk of mycobacterial infection. Thus, before the first randomized clinical trial evaluated antimycobacterial therapy for tuberculosis, the likely outcome was already well-understood[50].

When experimental manipulation in a randomized controlled trial is impractical or unethical, causality can still be assessed observationally using systematic approaches such as the Bradford-Hill criteria[55]. Importantly, these observational approaches do not necessarily rule out all unmeasured confounders. For instance, it is plausible that the link between smoking and lung cancer is explained by an unobserved genetic factor that independently increases the risk for both. Econometricians often overcome this limitation by identifying natural experiments, in which an intervention is applied in a manner that is practically random. For instance, if lung cancer rates decline consistently in different districts that raise tobacco taxes, this effect is unlikely to be mediated by an unobserved genetic factor. Accordingly, natural experiments can be used to infer causality[56,57].

Causal inference in human brain mapping

Human brain mapping researchers often attempt causal inference to achieve one of two objectives. This Review focuses on the objective of symptom localization, which aims to identify causal links between symptoms and neuroanatomy[58]. When a symptom is successfully localized, it may potentially be treated by modulating the corresponding neuroanatomy. The second objective, mapping the direction of information flow, aims to understand how one brain region can causally influence another. These experiments attempt to estimate the direction of information flow between two or more nodes in the brain using various measures of effective connectivity[10]. Different methods for symptom localization are summarized in Box 1, while different methods for measuring effective connectivity are summarized in Box 2.

For any of these methods to provide causal knowledge, the study must be designed to estimate what would have happened if the intervention had been different[40]. When direct experimental manipulation is not possible, this may still be accomplished by regressing out potential confounders, fitting progressively larger models that contain the interactions, or estimating how different factors influence the probability of an outcome[59]. These approaches differ in how they estimate the counterfactual, but can be combined into a single axiomatic definition for causality: an event is causal if, with all else being equal, its presence or absence affects the probability of an outcome[10,37,60]. Even if the cause is not necessary or sufficient to induce the effect, this probabilistic definition can still reveal treatment targets. For instance, smoking does not always cause lung cancer and lung cancer is not always

caused by smoking, but smoking singlehandedly increases the risk of lung cancer. Following this definition, we can reduce the risk of lung cancer with interventions that reduce smoking.

Although the fundamental goal of this approach is to estimate the counterfactual, this is sometimes not enough to clearly demonstrate causality. For instance, task-based fMRI experiments often compare the effects of a task to control conditions that estimate what would have happened if the task were not conducted. However, for the reasons described in Box 1, these experiments are rarely used to argue for causality.

In addition to estimating the counterfactual, causal inference in brain mapping may be assessed by borrowing criteria from Bradford-Hill's framework[61,62] (Fig. 2). Of these criteria, six are particularly useful for assessing the strength of causal inference across different human brain mapping studies: specificity, experimental manipulation, dose-response, coherence/convergence, reversibility, and temporality. Other criteria including effect size, reproducibility, plausibility, and analogy can be studied using any method, and thus are less useful for comparing the strength of causal inference across brain mapping studies.

When such criteria are applied rigorously, causal inference from observational data can approach the strength of causal inference from randomized controlled trials[57]. However, there is no established system for estimating the strength of causal inference from human brain mapping data. Building on Hume's framework and the many systematic approaches that it inspired, we suggest a 'causality continuum' for human brain mapping (Fig. 2b). While it is impossible to conclusively determine the causal role of an intervention, this continuum may be applied to score the relative strength of causality for different brain mapping techniques.

Conceptual causal mapping frameworks

Among different human brain mapping approaches, brain lesions and brain stimulation are the furthest along the causality continuum, assuming that a study using one of these approaches is rigorously designed to estimate the counterfactual and to demonstrate specificity[10,18]. However, the number of possible lesions or stimulation sites is essentially infinite, as is the number of possible behavioral outcomes, making it impractical to prospectively disentangle the behavioral effects of every single brain region from all other brain regions[30]. This practical challenge can be overcome by identifying natural experiments with incidental or near-random variance in the location of the lesion or the stimulation site. Using large datasets, it is possible to test for anatomical specificity by comparing hundreds of incidental lesions or stimulation sites across the brain[63]. This can yield a large-scale map of brain regions associated with certain lesion-induced symptoms[64,65] or stimulation-induced responses[66]. This is commonly achieved by comparing the locations of lesion or stimulation locations that modify a symptom, by comparing the connectivity of these lesions or stimulation sites, or by modifying symptoms in real-time with direct stimulation and recording.

Mapping near-random brain lesions

Most early lesion and stimulation studies were based on case reports or small case series in which a lesion or a stimulation was sufficient to cause a certain outcome. These studies provided useful knowledge about the neuroanatomy of clearly measurable functions, such as sensation, movement, and vision. However, some higher-order functions are difficult to study using individual cases, as stimulation of higher association areas does not always induce a measurable response[67]. This limitation can be mitigated using a probabilistic approach to causality, in which it is determined how a lesion or a stimulation modifies the probability of the outcome[63]. The probabilistic approach still lends itself to identifying therapeutic targets, as a treatment that reduces the probability of a symptom can still be clinically useful[37].

In a probabilistic study, large datasets are used to compare the effects of multiple lesions or stimulation sites. This was first illustrated with voxel lesion symptom mapping (VLSM), which compares symptom-causing lesions with other lesions[63]. For example, in Vietnam Veterans with penetrating brain injury, lesions to the amygdala and anterior temporal lobe reduced the risk of posttraumatic stress disorder[68], while lesions to the posterior temporal lobe increased the risk of anxiety[69]. The probabilistic approach used in VLSM has also been expanded to brain stimulation studies — for instance, in patients receiving prefrontal TMS for major depression, relatively anterior and lateral stimulation sites are more effective than relatively posterior and medial sites[70–72]. However, the statistical power of VLSM is limited because most lesions and stimulation sites do not overlap with one another (Fig. 3a)[58], so large sample sizes are required to detect statistical associations at any given brain location. VLSM also does not directly account for the role of diaschisis and inter-connected brain networks[30,58], although network-level relationships may become apparent with sufficient spatial coverage and sample size.

Localizing symptoms to brain circuits

When similar symptoms are induced by non-overlapping lesions in different patients, it can be challenging to localize the symptom to specific neuroanatomy (Fig. 3a). These individual differences may be explained by damage to a common network with different components that have been affected across individuals[67,73–84] — in other words, lesions in different locations caused the same symptom because they intersected the same brain circuit[58].

To test for localization to brain circuits, lesion and stimulation data can be combined with circuit maps derived from brain imaging or electrophysiology data (Box 3). This approach has been used widely in studies that employ lesion network mapping (LNM) and related techniques, which use correlative neuroimaging to explain heterogeneous causal findings[58]. Rather than simply comparing lesion locations, LNM first identifies the circuit connected to each lesion and then compares the circuits (Fig. 3b), as lesions in different brain regions may cause a similar symptom if they damage the same brain circuit[58]. In this manner, correlative neuroimaging can reveal a mechanistic connection between seemingly discordant causal results.

Because LNM maps the whole-brain connectivity of each lesion, every patient has data for every voxel in the brain, providing LNM with more statistical power than VLSM. LNM also quantifies the degree to which a lesion or stimulation site overlaps with different circuits, thus uncovering dose-response relationships[85]. Many recent studies have illustrated the potential of using normative brain connectivity to detect a causal link between lesions or stimulation sites and different clinical outcomes, even when VLSM shows no significant overlap. For instance, brain lesions are more likely to cause amnesia if they are connected to the subiculum[85]. DBS sites are more likely to relieve motor symptoms of Parkinson's disease if they are connected to the supplementary motor area[86]. TMS sites are more likely to relieve depression if they are functionally anti-correlated to the subgenual cingulate[72]. Similar approaches have now been used to map the causal neuroanatomy of movement disorders[86–91], mood disorders[71,90,92–95], anxiety-related disorders[96–98], psychotic disorders[99–101], disorders of consciousness[102–104], and various other neuropsychiatric phenomena[105–109].

Of note, most LNM studies define the circuit using a normative connectome database, which acts as a 'wiring diagram' that depicts the expected connectivity profile of any lesion location. This wiring diagram is based on resting-state functional connectivity or diffusion tractography data from a large cohort of participants without brain lesions or stimulation. Ongoing work is underway to compare wiring diagrams generated by different methods, including structural versus functional connectivity[110,111] and normative versus individualized connectivity[75,76], each of which may independently add value[75,76,111].

There are also some disadvantages to studying lesions and stimulation sites at the circuit level. These approaches can increase the risk of erroneous causal reasoning by adding an extra layer of complexity. If a brain lesion causes a particular symptom, it is reasonable to infer that the lesion location is causally involved in the symptom. If different lesion locations causing the same symptom are connected to one common hub, it is tempting to infer a causal relationship between the symptom and the hub. However, the hub simply defines the connectivity network that encompasses the lesion locations that cause a symptom. In other words, the causal inference is for the network, not the hub that defines the network. The hub may represent an important treatment target, but it is one step removed from the causal inference applied to the lesion locations themselves. One would need to independently study the effect of lesions or stimulation to the hub region to infer a causal link between the hub and the symptoms.

Real-time causal mapping

Another approach to causal inference addresses the temporality, experimental manipulation, and reversibility criteria by using real-time stimulation experiments in which different brain regions are focally stimulated to induce transient effects.

The most precise tool for real-time mapping of human brain function is intracranial electrical stimulation (iES) in conscious individuals[66], although recent studies suggest that thermal stimulation may achieve similar effects[112,113]. These studies are usually conducted in patients who are implanted with multifocal electrode arrays for seizure monitoring before surgery for epilepsy. While these electrodes are implanted, investigators can use them to

deliver focal stimulation and measure changes in a participant's subjective experiential state or experimental task performance[66,114]. This method also enables simultaneous intracranial EEG (iEEG) recording during and after stimulation, yielding measures of effective connectivity. While iES does not provide whole-brain coverage in the same manner as LNM and other circuit-level analyses, one can still estimate the counterfactual by comparing clinical outcomes from thousands of possible stimulation sites across multiple patients[67].

Although iES dates back to Penfield's work in the mid-20th century, a recent renaissance in iES research has been accelerated by the development of neuroimaging and circuit-based techniques[66]. High resolution imaging methods can localize the stimulated brain areas with millimeter precision to calculate the extent of the brain targeted by a given electrical dose[115].

Some recent studies provide examples of how iES-induced effects can be used to validate anatomical boundaries defined by other techniques such as functional neuroimaging or even animal models. For example, visual perception of faces is modified by stimulation of a patch of the posterior fusiform face area defined by task-based fMRI[80]. A will to persevere can be induced by stimulation of a hub in the anterior cingulate cortex defined by resting-state fMRI[79]. An altered sense of self can be induced by stimulation of part of the posteromedial cortex implicated in self-referential thought from task-based fMRI[116]. Furthermore, the probability of observing any behavioral response to stimulation aligns with brain network boundaries defined by resting state fMRI[67]. Finally, a dissociative state can be induced by stimulating the posterior cingulate cortex, a target identified based on an optogenetic mouse model of dissociation[117]. These studies illustrate how iES can add value to results from correlative neuroimaging experiments or studies in animal models, potentially facilitating their translation into clinical treatments.

Clinical causal brain mapping

Optimizing brain stimulation targets

Causal research is uniquely valuable in medicine because of its potential to identify targets for clinical intervention. If a particular factor is causal, it affects the probability of an outcome with all else being equal. Therefore, modifying that factor should modify the probability of that outcome. In clinical neuroscience, this principle can be tested using existing data in patients who received therapeutic TMS and DBS. If a particular brain region is causally implicated in a symptom, then stimulating that region should modify the probability or severity of that symptom. For instance, TMS can be used to localize the language centres in individual patients by studying the effects of targeted inhibition on speech function[118,119]. This approach can be used by neurosurgeons to avoid damaging language function during surgery[120].

Furthermore, circuit mapping studies have used connectivity to gain insight into the most effective TMS and DBS targets. In major depression, the most effective TMS sites are functionally anti-correlated to the subgenual cingulate cortex[71,72,94], while the most effective DBS sites overlap with tracts that connect the subgenual cingulate to the orbitofrontal cortex and the limbic system[121,122]. In Parkinson disease, the most effective

DBS sites for motor symptoms are structurally connected to the supplementary motor area and functionally anti-correlated to the primary motor cortex[86]. In obsessive-compulsive disorder, the most effective DBS sites overlap with a tract that connects the subthalamic nucleus to the prefrontal cortex[96,97] and are functionally connected to the anterior cingulate, insula, and precuneus[123]. In each case, stimulation sites might be refined to better target the therapeutic connections or to potentially avoid connections associated with DBS side effects[90]. Whether this improves clinical outcomes remains to be validated prospectively.

Beyond optimizing treatment targets for any given disorder, causal circuit mapping approaches have been used to disentangle which symptom clusters respond to different stimulation sites. In two cohorts of patients with major depression, TMS sites connected to a particular circuit were more likely to improve “dysphoric” symptoms such as sadness and suicidality, while TMS sites connected to a different circuit were more likely to improve “anxiosomatic” symptoms such as insomnia and sexual dysfunction[98]. Disentangling symptom clusters in this fashion lays the foundation for personalized target selection[124] and highlights the need for deeper phenotyping, which may reveal more distinct treatment targets for distinct symptoms.

Identifying brain stimulation targets

This framework may be extended to identify new TMS and DBS targets based on lesions that modify different symptoms. If damage to a particular circuit can cause a symptom, it is reasonable to hypothesize that stimulation of the same circuit may relieve that symptom. In fact, a similar approach was used to identify some of the neuromodulation targets currently in clinical use. Early studies on incidental stroke lesions led to identification of the dorsolateral prefrontal cortex as an effective TMS target for depression[24,125,126], while early findings on therapeutic lesioning led to identification of the ventral intermediate nucleus, subthalamic nucleus, and globus pallidus interna as effective DBS targets for Parkinson disease[19,20,22,127].

Broadly, these successes suggest that brain lesions can inform treatment targeting with brain stimulation. The causal circuit mapping approach also illustrates coherence across multiple methods, a rarely met criterion on the causality continuum. For example, while LNM can help us detect a causal link between lesions and outcomes based on overlap with a network, the causal agent is still the lesion location, not the network hub (Fig. 3). To determine if the network hubs are causally implicated in the behavior, targeted stimulation could then be used, thereby meeting the coherence criterion. This postulate is illustrated by two recent examples of coherence between LNM and iES results. First, lesions connected to the dorsal cingulate can disrupt volitional will[105], while iES to the dorsal cingulate can induce a will to persevere (Fig. 4a)[79]. Second, lesions connected to the right posterior fusiform gyrus can disrupt facial recognition[107], while iES of the right posterior, but not the left or anterior, fusiform gyrus can distort face perception[80,128,129] (Fig. 4b). These findings illustrate how stimulation and lesions can provide complementary information about the function of a network hub.

A similar postulate suggests that different stimulation modalities can also provide complementary causal information. If a circuit is causally implicated in a symptom, then

the symptom should be affected by different types of stimulation to that circuit. A review of published TMS and DBS targets explored this topic across 14 different disorders, finding that published TMS targets for various disorders were connected to the same brain regions as published DBS targets for the same disorders[109]. The authors thus hypothesized that better DBS targets may be identified based on the connectivity of effective TMS targets, and vice versa. If TMS and DBS sites converge on the same circuit, the strength of the causal inference linking this circuit to the symptom increases.

A recent study systematically tested this hypothesis by examining depression severity in 461 patients with brain lesions (five datasets), 151 patients receiving therapeutic TMS (four datasets), and 101 patients receiving therapeutic DBS (five datasets)[95]. Across all fourteen datasets, this analysis revealed a common circuit that was connected to the lesions that caused depression, the TMS sites that improved depression, and the DBS sites that worsened or improved depression (Fig. 5). Connectivity to this circuit also predicted the clinical efficacy of TMS and DBS sites in a leave-one-dataset-out cross-validation. Similar concordance was also observed for motor symptoms of Parkinson disease — lesions causing parkinsonism were connected to the same circuit as DBS sites that relieve parkinsonism, and both were connected to the primary motor cortex, the most common TMS target for parkinsonism[95]. Thus, the connectivity of brain lesions causing a symptom may reveal effective stimulation targets for the same symptom[95]. This provides a clear model for how convergent results can strengthen causal inference — if a circuit is causally linked to a symptom, then different manipulations of the circuit should modify the probability or severity of that symptom. Subsequent studies have expanded this principle to other symptoms and disorders, finding that lesion locations can predict which DBS targets are most likely to modify tics in Tourette disorder[130] and cognitive symptoms in Parkinson disease[131].

Roadmap for future research

Integrating correlative and causal maps

Beyond LNM, which directly combines correlative and causal techniques into a single analysis, there are several ways to test for coherence (or lack thereof) between correlative and causal techniques.

This principle was nicely illustrated in two recent studies, in which task-based iEEG (causality score = 2) was used to identify a medial-lateral dissociation between simple tone perception and complex speech perception. These correlative results were then used to guide a causal manipulation (iES, causality score = 5.5), revealing that medial sites induced hallucinations of simple sounds, while lateral sites interrupted complex word perception[132,133]. Thus, by combining advanced analysis of correlative data with causal stimulation, these studies discovered a novel medial-lateral gradient in the auditory cortex.

Causal techniques can also be used to compare different correlative brain mapping results. For instance, a recent study used task-based fMRI (causality score = 2) to compare different approaches for parcellating the brain based on task-free neuroimaging (causality score = 0)[134]. Another recent study used incidental lesions (causality score = 5.5) to compare

different approaches for defining ‘cortical hubs’ based on resting-state fMRI (causality score = 0)[135]. In both cases, the more causal technique showed that some of the correlative approaches were more functionally relevant than the others. Importantly, the strength of causal inference in such studies is defined by the more causal technique, not the less causal technique. In the future, techniques further along the causality continuum should be used to test results from less causal methods whenever feasible.

By contrast, correlative imaging can also be used as a biomarker when comparing the effects of different causal techniques, especially techniques that do not induce a measurable clinical effect. For instance, functional neuroimaging can be used to estimate the effect of TMS to different targets in the prefrontal cortex, even when these targets do not induce clear behavioral changes[136–138]. In addition, by using correlative imaging to measure the physiological effects of an intervention before the clinical effects become apparent, it may be possible to derive useful therapeutic biomarkers[139].

Finally, multifocal stimulation protocols may be designed to target different parts of a circuit defined by correlative techniques. The effect of stimulating different nodes of the same circuit could differ, and stimulating one site may modify excitability of a secondary site[132]. Various studies have attempted multifocal TMS and some have suggested that it may be safe to combine TMS with DBS applied to the same circuit, but more work is needed to identify optimal multifocal stimulation approaches[140–143].

Best practices for causal brain mapping

Different human brain mapping techniques can yield insights at different levels of causal inference. It is important to recognize which level of causal inference applies to a given study and to form appropriate conclusions as a result, especially when the goal is uncovering a therapeutic target. Here, we provide some guidelines that may be helpful in this regard.

First, purely correlative studies should avoid making claims of causality and be cautious in suggesting therapeutic relevance. While correlative studies can plausibly identify biomarkers for diagnosis and monitoring[139], there is a fundamental gap between correlative results and identifying treatment targets — while some correlations may reflect a causal relationship, others could be spurious or reverse-causal.

Second, the strength of causal inference should be explicitly assessed in studies that make any causal conclusions, including limitations. Before proposing a treatment target, brain mapping results should be explicitly compared with results from studies using approaches further along the causality continuum. For instance, atrophy-derived brain maps are more likely to be causal if they agree with lesion-derived maps of the same behavior[144]. If there is concordance across different causal mapping approaches, and consistency as one moves higher up on the causality continuum, this increases the probability that a correlative finding provides a useful mechanistic insight towards a treatment target.

Third, lesion and stimulation studies should explicitly consider the degree of concordance between modalities. This may become increasingly common as more lesion and stimulation datasets become available. We provided several examples of concordance between lesion

and brain stimulation data (Fig. 4 and 5)[95,130,131] which increases confidence in a causal relationship.

Fourth, to identify treatment targets, it is useful to focus not only on the strength of causality but also on its clinical relevance. For instance, if stimulating a specific region or circuit modifies the probability of specific symptoms, this is clinically meaningful and may represent a therapeutic target. By contrast, if stimulating a specific region or circuit induces an electrophysiological or hemodynamic response in a distant region, this information may be less clinically relevant. This knowledge may be useful for mechanistic and methodological purposes, but it does not imply that the distant region should be used as a treatment target.

Fifth, there is a critical difference between the strength of causality and the overall quality of a study. Our proposed causality continuum applies to studies that are well-designed and well-executed with rigorous methods and appropriate controls. If a study uses unjustified methods or lacks an appropriate control group, then the strength of causal inference cannot be reliably assessed, regardless of which methods were used. Thus, the causality continuum is only one of many dimensions that should be assessed when appraising a human brain mapping study.

Last, perhaps most importantly, prospective randomized clinical trials are still needed to translate causal circuit mapping results into practical treatments. Most existing clinical trials compare the effect of treatment with a sham treatment or placebo. This approach is valuable to determine the effect of the treatment modality, but it does not clarify the effect of the targeted circuit. To address this, prospective trials could isolate the effect of the targeted circuit by comparing the clinical effects of stimulating different brain circuits. If different targets can modulate different symptoms, these trials would be a powerful tool for testing circuit-based hypotheses and translating these findings into treatments.

Limitations and open questions

The mechanisms by which lesions or stimulations change an outcome remain unclear. It is relatively straightforward to understand why brain damage from a structural lesion will cause a loss of function, but it is unclear how some lesions can lead to positive outcomes[145,146]. Lesions are generally presumed to be inhibitory, but homeostatic compensation and diaschisis may lead to excitatory effects elsewhere the brain[147,148]. The effects of focal brain damage may extend beyond the lesion itself and these effects may differ for gray matter versus white matter lesions[148,149]. A better understanding of the neural mechanisms of diaschisis may shed light on these open questions.

The mechanisms of transcranial and intracranial stimulation can add complexity to the interpretation of causal brain mapping studies, as the roles of different stimulation parameters remain unclear. What are the effects of different stimulation frequencies and intensities[150,151]? How are different cortical layers affected[152]? How are outcomes related to cognitive states or tasks[153]? These and other factors may influence which neurons are affected by a stimulation[154,155]. Subtle variations in the stimulated field can lead

to measurable differences in behavioral outcomes[156], which may explain why some of Penfield's experiments have failed to replicate[157].

Results have been particularly heterogenous in the TMS literature. This is partly related to dosing – over the last two decades, TMS studies have used progressively higher doses and observed progressively larger effects[158]. However, individual differences and mechanisms remain poorly-understood. For example, high-frequency TMS is conventionally thought to increase cortical excitability, but up to 30% of individuals may show no change in or even a decrease in excitability[159]. Similarly, TMS has been hypothesized to function by modulating NMDA receptor-mediated plasticity, which can be modified by various drugs[160], but these drugs can have minimal effect on TMS clinical outcomes[161]. As such, even a rigorously-designed TMS study may produce limited effect sizes, potentially underestimating the true causal relationship between the targeted brain region and behavioral effect.

Finally, when attempting to map causal information to brain circuits, it's important to note that brain circuits can be defined in different ways[5,162–165]. Human brain circuits can be defined using thousands of individual voxels[63,100,166], hundreds of distinct parcels[5,162,164,165], or a handful of functionally independent networks[6,163], and can be defined at the group or individual level[167–170]. This heterogeneity leads to high researcher degrees of freedom and may contribute to poor reproducibility of neuroimaging results[171]. This fragmentation was illustrated by a recent study in which 70 different research teams were asked to test the same hypotheses in the same neuroimaging dataset, and no two teams chose the same analytical procedure[172]. While many causal mapping studies have been robust to subtle methodological variations[93,98], others have yielded different results when using different approaches[173]. Methodological standardization may improve reproducibility (one of Bradford-Hill's criteria), and thus help move the entire field closer to causal inference.

Conclusions

Owing to decades of research in clinical and systems neuroscience, the field has amassed a wealth of knowledge about the organization of different brain regions. The recent re-emergence of lesion and stimulation-based studies has also led to a rapid growth in our ability to infer the causal role of these different regions, potentially enabling more informed clinical trials of targeted neuromodulation. To achieve this, causality should be interpreted as 'stronger' or 'weaker' along a continuous spectrum, and purely correlative findings should be explicitly acknowledged as such. By using tools and approaches that are further along the causality continuum, we can come closer to translating brain maps into clinical treatment targets for neuropsychiatric disease. Following the example set by the field of econometrics[15], this may help human brain mapping achieve its own credibility revolution.

Acknowledgments

The authors thank the National Institute of Mental Health, the Brain & Behavior Research Foundation, Brigham & Women's Hospital, and Stanford University BioX for funding the present work.

Competing interests

SHS serves as a consultant for Kaizen Brain Center, Acacia Mental Health, and Magnus Medical. SHS owns stock in Brainsway Inc (publicly traded) and Magnus Medical (not publicly traded). SHS and MDF have jointly received investigator-initiated research support from Neuronetics Inc. None of these organizations were involved in the present work. SHS and MDF each own independent intellectual property on the use of brain network mapping to target neuromodulation. The present work did not utilize any of this intellectual property. The authors report no other conflicts of interest related to the present work.

References

1. Fox MD & Raichle ME Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nature reviews neuroscience 8, 700–711 (2007). [PubMed: 17704812]
2. Raichle ME A brief history of human brain mapping. Trends in neurosciences 32, 118–126 (2009). [PubMed: 19110322]
3. Genon S, Reid A, Langner R, Amunts K & Eickhoff SB How to Characterize the Function of a Brain Region. Trends in Cognitive Sciences 22, 350–364, doi:10.1016/j.tics.2018.01.010 (2018). [PubMed: 29501326]
4. Gusnard DA & Raichle ME Searching for a baseline: functional imaging and the resting human brain. Nature reviews neuroscience 2, 685–694 (2001). [PubMed: 11584306]
5. Glasser MF et al. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178, doi:10.1038/nature18933 (2016). [PubMed: 27437579]
6. Power JD et al. Functional network organization of the human brain. Neuron 72, 665–678 (2011). [PubMed: 22099467]
7. Bullmore E & Sporns O The economy of brain network organization. Nature Reviews Neuroscience 13, 336–349 (2012). [PubMed: 22498897]
8. Logothetis NK What we can do and what we cannot do with fMRI. Nature 453, 869–878, doi:10.1038/nature06976 (2008). [PubMed: 18548064]
9. Jonas E & Kording KP Could a Neuroscientist Understand a Microprocessor? PLoS computational biology 13, e1005268, doi:10.1371/journal.pcbi.1005268 (2017). [PubMed: 28081141]
10. Reid AT et al. Advancing functional connectivity research from association to causation. Nature Neuroscience 22, 1751–1760, doi:10.1038/s41593-019-0510-4 (2019). [PubMed: 31611705]
11. Laumann TO & Snyder AZ Brain activity is not only for thinking. Current Opinion in Behavioral Sciences 40, 130–136, doi:10.1016/j.cobeha.2021.04.002 (2021).
12. Etkin A A Reckoning and Research Agenda for Neuroimaging in Psychiatry. The American journal of psychiatry 176, 507–511, doi:10.1176/appi.ajp.2019.19050521 (2019). [PubMed: 31256624]
13. Weichwald S & Peters J Causality in Cognitive Neuroscience: Concepts, Challenges, and Distributional Robustness. Journal of Cognitive Neuroscience 33, 226–247, doi:10.1162/jocn_a_01623 (2021). [PubMed: 32812827]
14. Etkin A Addressing the Causality Gap in Human Psychiatric Neuroscience. JAMA Psychiatry 75, 3–4, doi:10.1001/jamapsychiatry.2017.3610 (2018). [PubMed: 29167887]
15. Angrist JD & Pischke J-S The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. Journal of economic perspectives 24, 3–30 (2010).
16. Leamer EE Let's Take the Con Out of Econometrics. The American Economic Review 73, 31–43 (1983).
17. Etkin A Mapping Causal Circuitry in Human Depression. Biol Psychiatry 86, 732–733, doi:10.1016/j.biopsych.2019.09.009 (2019). [PubMed: 31648679]
18. Bergmann TO & Hartwigsen G Inferring Causality from Noninvasive Brain Stimulation in Cognitive Neuroscience. Journal of Cognitive Neuroscience 33, 195–225, doi:10.1162/jocn_a_01591 (2021). [PubMed: 32530381]
19. Pycroft L, Stein J & Aziz T Deep brain stimulation: An overview of history, methods, and future developments. Brain and Neuroscience Advances 2, 2398212818816017, doi:10.1177/2398212818816017 (2018).

20. Lozano AM et al. Effect of GPi pallidotomy on motor function in Parkinson's disease. The Lancet 346, 1383–1387, doi:10.1016/S0140-6736(95)92404-3 (1995).
21. Groiss SJ, Wojtecki L, Sudmeyer M & Schnitzler A Deep brain stimulation in Parkinson's disease. Ther Adv Neurol Disord 2, 20–28, doi:10.1177/1756285609339382 (2009). [PubMed: 21180627]
22. Benabid AL et al. Long-term suppression of tremor by chronic stimulation of the ventral intermediate thalamic nucleus. The Lancet 337, 403–406 (1991).
23. George MS et al. A controlled trial of daily left prefrontal cortex TMS for treating depression. Biol Psychiatry 48, 962–970, doi:10.1016/s0006-3223(00)01048-9 (2000). [PubMed: 11082469]
24. Pascual-Leone A, Rubio B, Pallardó F & Catalá MD Rapid-rate transcranial magnetic stimulation of left dorsolateral prefrontal cortex in drug-resistant depression. Lancet (London, England) 348, 233–237, doi:10.1016/s0140-6736(96)01219-6 (1996).
25. Harlow JM Passage of an iron rod through the head. The Boston Medical and Surgical Journal (1828–1851) 39, 0_1 (1848).
26. Scoville WB & Milner B Loss of recent memory after bilateral hippocampal lesions. Journal of neurology, neurosurgery, and psychiatry 20, 11 (1957). [PubMed: 13406589]
27. Penfield W & Jasper H Epilepsy and the functional anatomy of the human brain. (1954).
28. Penfield W & Perot P The brain's record of auditory and visual experience: a final summary and discussion. Brain 86, 595–696 (1963). [PubMed: 14090522]
29. Penfield W & Boldrey E Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain 60, 389–443 (1937).
30. Rorden C & Karnath H-O Using human brain lesions to infer function: a relic from a past era in the fMRI age? Nature Reviews Neuroscience 5, 812–819, doi:10.1038/nrn1521 (2004).
31. Michel J-B et al. Quantitative analysis of culture using millions of digitized books. Science 331, 176–182, doi:10.1126/science.1199644 (2011). [PubMed: 21163965]
32. Gaillard R et al. Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. Neuron 50, 191–204 (2006). [PubMed: 16630832]
33. Petersen SE, Fox PT, Posner MI, Mintun M & Raichle ME Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 331, 585–589, doi:10.1038/331585a0 (1988). [PubMed: 3277066]
34. Dolan RJ & Fletcher P Dissociating prefrontal and hippocampal function in episodic memory encoding. Nature 388, 582–585 (1997). [PubMed: 9252188]
35. Zalesky A, Fornito A & Bullmore E On the use of correlation as a measure of network connectivity. NeuroImage 60, 2096–2106, doi:10.1016/j.neuroimage.2012.02.001 (2012). [PubMed: 22343126]
36. Uttal WR The new phrenology: The limits of localizing cognitive processes in the brain. (The MIT press, 2001).
37. Pearl J Causality. (Cambridge University Press, 2009).
38. Angrist JD & Pischke J-S Mostly harmless econometrics: An empiricist's companion. (Princeton university press, 2008).
39. Holland PW Statistics and causal inference. Journal of the American statistical Association 81, 945–960 (1986).
40. Pearl J & Mackenzie D The book of why: the new science of cause and effect. (Basic books, 2018).
41. Gotthelf A Aristotle's conception of final causality. The Review of Metaphysics, 226–254 (1976).
42. Black DL Mental Existence in Thomas Aquinas and Avicenna. Mediaeval Studies 61, 45–79 (1999).
43. Taylor R Primary and Secondary Causality. (The Routledge Companion to Islamic Philosophy, 2012).
44. Smith NK The philosophy of David Hume: a critical study of its origins and central doctrines. (1941).
45. Koch R in 10th International Medical Congress.
46. Bradford Hill A The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine 58, 295–300, doi:10.1177/003591576505800503 (1965). [PubMed: 14283879]

47. Fedak KM, Bernal A, Capshaw ZA & Gross S Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. Emerg Themes Epidemiol 12, 14, doi:10.1186/s12982-015-0037-4 (2015). [PubMed: 26425136]
48. Araújo L, Dalgalarrondo P & Banzato C On the notion of causality in medicine: Addressing Austin Bradford Hill and John L. Mackie. Archives of Clinical Psychiatry 41, 56–61, doi:10.1590/0101-60830000000010 (2014).
49. Sajadi MM, Mansouri D & Sajadi MR Ibn Sina and the clinical trial. Annals of internal medicine 150, 640–643, doi:10.7326/0003-4819-150-9-200905050-00011 (2009). [PubMed: 19414844]
50. Bradford Hill A Memories of the British streptomycin trial in tuberculosis: the first randomized clinical trial. Controlled Clinical Trials 11, 77–79 (1990). [PubMed: 2161313]
51. Gillies D Causality, Probability, and Medicine. (Taylor & Francis, 2018).
52. Kaplan RM & Irvin VL Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. PLOS ONE 10, e0132382, doi:10.1371/journal.pone.0132382 (2015). [PubMed: 26244868]
53. Baluku JB et al. Prevalence of Malaria and TB Coinfection at a National Tuberculosis Treatment Centre in Uganda. Journal of Tropical Medicine 2019, 3741294, doi:10.1155/2019/3741294 (2019). [PubMed: 31428162]
54. Koch R The etiology of tuberculosis. Mittheilungen aus dem Kaiserlichen Gesundheitsamte 2, 1–88 (1884).
55. Hernán MA & Robins JM Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. American Journal of Epidemiology 183, 758–764, doi:10.1093/aje/kwv254 (2016). [PubMed: 26994063]
56. Marinescu IE, Lawlor PN & Kording KP Quasi-experimental causality in neuroscience and behavioural research. Nature human behaviour 2, 891–898, doi:10.1038/s41562-018-0466-5 (2018).
57. Cook TD, Shadish WR & Wong VC Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. Journal of Policy Analysis and Management 27, 724–750, doi:10.1002/pam.20375 (2008).
58. Fox MD Mapping Symptoms to Brain Networks with the Human Connectome. N Engl J Med 379, 2237–2245, doi:10.1056/NEJMra1706158 (2018). [PubMed: 30575457]
59. Williamson J Establishing Causal Claims in Medicine. International Studies in the Philosophy of Science 32, 33–61, doi:10.1080/02698595.2019.1630927 (2019).
60. Russo F & Williamson J Interpreting causality in the health sciences. International studies in the philosophy of science 21, 157–170 (2007).
61. Evans AS Causation and disease: the Henle-Koch postulates revisited. Yale J Biol Med 49, 175–195 (1976). [PubMed: 782050]
62. Cassidy JM, Mark JI & Cramer SC Functional connectivity drives stroke recovery: shifting the paradigm from correlation to causation. Brain, doi:10.1093/brain/awab469 (2021).
63. Bates E et al. Voxel-based lesion–symptom mapping. Nature Neuroscience 6, 448–450, doi:10.1038/nn1050 (2003). [PubMed: 12704393]
64. Wu O et al. Role of Acute Lesion Topography in Initial Ischemic Stroke Severity and Long-Term Functional Outcomes. Stroke 46, 2438–2444, doi:doi:10.1161/STROKEAHA.115.009643 (2015). [PubMed: 26199314]
65. Corbetta M et al. Common behavioral clusters and subcortical anatomy in stroke. Neuron 85, 927–941, doi:10.1016/j.neuron.2015.02.027 (2015). [PubMed: 25741721]
66. Parvizi J & Kastner S Promises and limitations of human intracranial electroencephalography. Nature Neuroscience 21, 474–483, doi:10.1038/s41593-018-0108-2 (2018). [PubMed: 29507407]
67. Fox KCR et al. Intrinsic network architecture predicts the effects elicited by intracranial electrical stimulation of the human brain. Nature human behaviour 4, 1039–1052, doi:10.1038/s41562-020-0910-1 (2020).
68. Koenigs M et al. Focal brain damage protects against post-traumatic stress disorder in combat veterans. Nat Neurosci 11, 232–237, doi:10.1038/nn2032 (2008). [PubMed: 18157125]

69. Knutson KM et al. Injured brain regions associated with anxiety in Vietnam veterans. Neuropsychologia 51, 686–694, doi:10.1016/j.neuropsychologia.2013.01.003 (2013). [PubMed: 23328629]
70. Herbsman T et al. More lateral and anterior prefrontal coil location is associated with better repetitive transcranial magnetic stimulation antidepressant response. Biol Psychiatry 66, 509–515, doi:10.1016/j.biopsych.2009.04.034 (2009). [PubMed: 19545855]
71. Fox MD, Buckner RL, White MP, Greicius MD & Pascual-Leone A Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. Biol Psychiatry 72, 595–603, doi:10.1016/j.biopsych.2012.04.028 10.1016/j.biopsych.2012.04.028. Epub 2012 Jun 1. (2012). [PubMed: 22658708]
72. Weigand A et al. Prospective Validation That Subgenual Connectivity Predicts Antidepressant Efficacy of Transcranial Magnetic Stimulation Sites. Biol Psychiatry 84, 28–37, doi:10.1016/j.biopsych.2017.10.028 (2018). [PubMed: 29274805]
73. Gordon EM et al. Individual-specific features of brain systems identified with resting state functional correlations. Neuroimage 146, 918–939, doi:10.1016/j.neuroimage.2016.08.032 (2017). [PubMed: 27640749]
74. Gordon EM, Laumann TO, Adeyemo B & Petersen SE Individual Variability of the System-Level Organization of the Human Brain. Cereb Cortex 27, 386–399, doi:10.1093/cercor/bhv239 (2017). [PubMed: 26464473]
75. Siddiqi SH, Weigand A, Pascual-Leone A & Fox MD Identification of personalized TMS targets based on subgenual cingulate connectivity: an independent replication. Biol Psychiatry (2021 (In Press)).
76. Cash RFH, Cocchi L, Lv J, Fitzgerald PB & Zalesky A Functional Magnetic Resonance Imaging–Guided Personalization of Transcranial Magnetic Stimulation Treatment for Depression. JAMA Psychiatry, doi:10.1001/jamapsychiatry.2020.3794 (2020).
77. Howell B et al. Quantifying the axonal pathways directly stimulated in therapeutic subcallosal cingulate deep brain stimulation. Human brain mapping 40, 889–903, doi:10.1002/hbm.24419 (2019). [PubMed: 30311317]
78. Siddiqi SH et al. Repetitive Transcranial Magnetic Stimulation with Resting-State Network Targeting for Treatment-Resistant Depression in Traumatic Brain Injury: A Randomized, Controlled, Double-Blinded Pilot Study. J Neurotrauma 36, 1361–1374, doi:10.1089/neu.2018.5889 (2019). [PubMed: 30381997]
79. Parvizi J, Rangarajan V, Shirer WR, Desai N & Greicius Michael D. The Will to Persevere Induced by Electrical Stimulation of the Human Cingulate Gyrus. Neuron 80, 1359–1367, doi:10.1016/j.neuron.2013.10.057 (2013). [PubMed: 24316296]
80. Schrouff J et al. Fast temporal dynamics and causal relevance of face processing in the human temporal cortex. Nature Communications 11, 656, doi:10.1038/s41467-020-14432-8 (2020).
81. Yih J, Beam DE, Fox KCR & Parvizi J Intensity of affective experience is modulated by magnitude of intracranial electrical stimulation in human orbitofrontal, cingulate and insular cortices. Social cognitive and affective neuroscience 14, 339–351, doi:10.1093/scan/nsz015 (2019). [PubMed: 30843590]
82. Downar J et al. Anhedonia and reward-circuit connectivity distinguish nonresponders from responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. Biol Psychiatry 76, 176–185, doi:10.1016/j.biopsych.2013.10.026 (2014). [PubMed: 24388670]
83. Drysdale AT et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med, doi:10.1038/nm.4246 10.1038/nm.4246. (2016).
84. Cash RFH et al. Using Brain Imaging to Improve Spatial Targeting of Transcranial Magnetic Stimulation for Depression. Biol Psychiatry, doi:10.1016/j.biopsych.2020.05.033 (2020).
85. Ferguson MA et al. A human memory circuit derived from brain lesions causing amnesia. Nature Communications 10, 3497, doi:10.1038/s41467-019-11353-z (2019).
86. Horn A et al. Connectivity Predicts deep brain stimulation outcome in Parkinson disease. Ann Neurol 82, 67–78, doi:10.1002/ana.24974 (2017). [PubMed: 28586141]

87. Joutsa J, Horn A, Hsu J & Fox MD Localizing parkinsonism based on focal brain lesions. Brain 141, 2445–2456, doi:10.1093/brain/awy161 (2018). [PubMed: 29982424]

88. Corp DT et al. Network localization of cervical dystonia based on causal brain lesions. Brain 142, 1660–1674, doi:10.1093/brain/awz112 (2019). [PubMed: 31099831]

89. Joutsa J et al. Identifying therapeutic targets from spontaneous beneficial brain lesions. Ann Neurol 84, 153–157, doi:10.1002/ana.25285 (2018). [PubMed: 30014594]

90. Irmen F et al. Left prefrontal impact links subthalamic stimulation with depressive symptoms. Ann Neurol, doi:10.1002/ana.25734 (2020).

91. Laganiere S, Boes AD & Fox MD Network localization of hemichorea-hemiballismus. Neurology 86, 2187–2195, doi:10.1212/wnl.0000000000002741 (2016). [PubMed: 27170566]

92. Padmanabhan JL et al. A human depression circuit derived from focal brain lesions. Biol Psychiatry, doi:10.1016/j.biopsych.2019.07.023 (2019).

93. Cotovio G et al. Mapping mania symptoms based on focal brain damage. The Journal of Clinical Investigation 130, 5209–5222, doi:10.1172/JCI136096 (2020). [PubMed: 32831292]

94. Cash RFH et al. Subgenual Functional Connectivity Predicts Antidepressant Treatment Response to Transcranial Magnetic Stimulation: Independent Validation and Evaluation of Personalization. Biol Psychiatry 86, e5–e7, doi:10.1016/j.biopsych.2018.12.002 (2019). [PubMed: 30670304]

95. Siddiqi SH et al. Brain stimulation and brain lesions converge on common causal circuits in neuropsychiatric disease. Nature human behaviour (2021).

96. Baldermann JC et al. Connectivity Profile Predictive of Effective Deep Brain Stimulation in Obsessive-Compulsive Disorder. Biol Psychiatry 85, 735–743, doi:10.1016/j.biopsych.2018.12.019 (2019). [PubMed: 30777287]

97. Li N et al. A unified connectomic target for deep brain stimulation in obsessive-compulsive disorder. Nature Communications 11, 3364, doi:10.1038/s41467-020-16734-3 (2020).

98. Siddiqi SH et al. Distinct Symptom-Specific Treatment Targets for Circuit-Based Neuromodulation. The American journal of psychiatry 177, 435–446, doi:10.1176/appi.ajp.2019.19090915 (2020). [PubMed: 32160765]

99. Kim NY et al. Lesions causing hallucinations localize to one common brain network. Molecular Psychiatry 26, 1299–1309, doi:10.1038/s41380-019-0565-3 (2021). [PubMed: 31659272]

100. Boes AD et al. Network localization of neurological symptoms from focal brain lesions. Brain : a journal of neurology 138, 3061–3075, doi:10.1093/brain/awv228 (2015). [PubMed: 26264514]

101. Darby RR, Laganiere S, Pascual-Leone A, Prasad S & Fox MD Finding the imposter: brain connectivity of lesions causing delusional misidentifications. Brain : a journal of neurology 140, 497–507, doi:10.1093/brain/aww288 (2017). [PubMed: 28082298]

102. Snider SB et al. Cortical lesions causing loss of consciousness are anticorrelated with the dorsal brainstem. Human brain mapping 41, 1520–1531, doi:10.1002/hbm.24892 (2020). [PubMed: 31904898]

103. Fischer DB et al. A human brain network derived from coma-causing brainstem lesions. Neurology 87, 2427–2434, doi:10.1212/wnl.0000000000003404 (2016). [PubMed: 27815400]

104. Joutsa J, Shih LC & Fox MD Mapping holmes tremor circuit using the human brain connectome. Ann Neurol 86, 812–820, doi:10.1002/ana.25618 (2019). [PubMed: 31614012]

105. Darby RR, Joutsa J, Burke MJ & Fox MD Lesion network localization of free will. Proc Natl Acad Sci U S A 115, 10792–10797, doi:10.1073/pnas.1814117115 (2018). [PubMed: 30275309]

106. Darby RR, Horn A, Cushman F & Fox MD Lesion network localization of criminal behavior. Proc Natl Acad Sci U S A 115, 601–606, doi:10.1073/pnas.1706587115 (2018). [PubMed: 29255017]

107. Cohen AL et al. Looking beyond the face area: lesion network mapping of prosopagnosia. Brain 142, 3975–3990, doi:10.1093/brain/awz332 (2019). [PubMed: 31740940]

108. Cohen AL et al. Tuber Locations Associated with Infantile Spasms Map to a Common Brain Network. Ann Neurol 89, 726–739, doi:10.1002/ana.26015 (2021). [PubMed: 33410532]

109. Fox MD et al. Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. Proc Natl Acad Sci U S A 111, E4367–4375, doi:10.1073/pnas.1405003111 (2014). [PubMed: 25267639]

110. Horn A & Fox MD Opportunities of connectomic neuromodulation. NeuroImage 221, 117180, doi:10.1016/j.neuroimage.2020.117180 (2020). [PubMed: 32702488]
111. Bowren M et al. Post-Stroke Cognitive and Motor Outcomes Predicted from Lesion Location and Lesion Network Mapping. Brain (2022 (In Press)).
112. Long Michael A. et al. Functional Segregation of Cortical Regions Underlying Speech Timing and Articulation. Neuron 89, 1187–1193, doi:10.1016/j.neuron.2016.01.032 (2016). [PubMed: 26924439]
113. Ibayashi K et al. Focal Cortical Surface Cooling is a Novel and Safe Method for Intraoperative Functional Brain Mapping. World Neurosurgery 147, e118–e129, doi:10.1016/j.wneu.2020.11.164 (2021). [PubMed: 33307258]
114. Borchers S, Himmelbach M, Logothetis N & Karnath HO Direct electrical stimulation of human cortex - the gold standard for mapping brain functions? Nature reviews. Neuroscience 13, 63–70, doi:10.1038/nrn3140 (2011). [PubMed: 22127300]
115. Winawer J & Parvizi J Linking Electrical Stimulation of Human Primary Visual Cortex, Size of Affected Cortical Area, Neuronal Responses, and Subjective Experience. Neuron 92, 1213–1219, doi:10.1016/j.neuron.2016.11.008 (2016). [PubMed: 27939584]
116. Parvizi J et al. Altered sense of self during seizures in the posteromedial cortex. Proceedings of the National Academy of Sciences 118, e2100522118, doi:10.1073/pnas.2100522118 (2021).
117. Vesuna S et al. Deep posteromedial cortical rhythm in dissociation. Nature 586, 87–94, doi:10.1038/s41586-020-2731-9 (2020). [PubMed: 32939091]
118. Knecht S et al. Degree of language lateralization determines susceptibility to unilateral brain lesions. Nature neuroscience 5, 695–699 (2002). [PubMed: 12055632]
119. Pascual-Leone A, Gates JR & Dhuna A Induction of speech arrest and counting errors with rapid-rate transcranial magnetic stimulation. Neurology 41, 697–702 (1991). [PubMed: 2027485]
120. Lefaucheur JP & Picht T The value of preoperative functional cortical mapping using navigated TMS. Neurophysiologie clinique = Clinical neurophysiology 46, 125–133, doi:10.1016/j.neucli.2016.05.001 (2016). [PubMed: 27229765]
121. Riva-Posse P et al. Defining critical white matter pathways mediating successful subcallosal cingulate deep brain stimulation for treatment-resistant depression. Biol Psychiatry 76, 963–969, doi:10.1016/j.biopsych.2014.03.029 (2014). [PubMed: 24832866]
122. Riva-Posse P et al. A connectomic approach for subcallosal cingulate deep brain stimulation surgery: prospective targeting in treatment-resistant depression. Molecular psychiatry 23, 843–849, doi:10.1038/mp.2017.59 (2018). [PubMed: 28397839]
123. Li N et al. A Unified Functional Network Target for Deep Brain Stimulation in Obsessive-Compulsive Disorder. Biol Psychiatry, doi:10.1016/j.biopsych.2021.04.006 (2021).
124. Nestor Sean M., M.D., Ph.D., & Blumberger Daniel M., M.D., M.Sc. Mapping Symptom Clusters to Circuits: Toward Personalizing TMS Targets to Improve Treatment Outcomes in Depression. American Journal of Psychiatry 177, 373–375, doi:10.1176/appi.ajp.2020.20030271 (2020). [PubMed: 32354264]
125. Robinson RG, Kubos KL, Starr LB, Rao K & Price TR Mood disorders in stroke patients. Importance of location of lesion. Brain 107 (Pt 1), 81–93, doi:10.1093/brain/107.1.81 (1984). [PubMed: 6697163]
126. Robinson RG & Szetela B Mood change following left hemispheric brain injury. Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society 9, 447–453 (1981).
127. Bergman H, Wichmann T & DeLong MR Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. Science 249, 1436–1438, doi:10.1126/science.2402638 (1990). [PubMed: 2402638]
128. Parvizi J et al. Electrical Stimulation of Human Fusiform Face-Selective Regions Distorts Face Perception. The Journal of Neuroscience 32, 14915–14920, doi:10.1523/jneurosci.2609-12.2012 (2012). [PubMed: 23100414]
129. Rangarajan V et al. Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. The Journal of neuroscience : the official journal

of the Society for Neuroscience 34, 12828–12836, doi:10.1523/jneurosci.0527-14.2014 (2014). [PubMed: 25232118]

130. Ganos C et al. A neural network for tics: insights from causal brain lesions and deep brain stimulation. Brain, doi:10.1093/brain/awac009 (2022).
131. Reich MM et al. A brain network for deep brain stimulation induced cognitive decline in Parkinson's disease. Brain, doi:10.1093/brain/awac012 (2022).
132. Hamilton LS, Oganian Y, Hall J & Chang EF Parallel and distributed encoding of speech across human auditory cortex. Cell 184, 4626–4639.e4613, doi:10.1016/j.cell.2021.07.019 (2021). [PubMed: 34411517]
133. Hamilton LS, Edwards E & Chang EF A spatial map of onset and sustained responses to speech in the human superior temporal gyrus. Current Biology 28, 1860–1871. e1864 (2018). [PubMed: 29861132]
134. Zhi D, King M & Diedrichsen J Evaluating brain parcellations using the distance controlled boundary coefficient. bioRxiv, 2021.2005.2011.443151, doi:10.1101/2021.05.11.443151 (2021).
135. Warren DE et al. Network measures predict neuropsychological outcome after brain injury. Proceedings of the National Academy of Sciences 111, 14247–14252, doi:10.1073/pnas.1322173111 (2014).
136. Eldaief MC, Halko MA, Buckner RL & Pascual-Leone A Transcranial magnetic stimulation modulates the brain's intrinsic activity in a frequency-dependent manner. Proceedings of the National Academy of Sciences 108, 21229–21234, doi:10.1073/pnas.1113103109 (2011).
137. Chen AC et al. Causal interactions between fronto-parietal central executive and default-mode networks in humans. Proceedings of the National Academy of Sciences 110, 19944–19949, doi:10.1073/pnas.1311772110 (2013).
138. Ozdemir RA et al. Individualized perturbation of the human connectome reveals reproducible biomarkers of network dynamics relevant to cognition. Proceedings of the National Academy of Sciences 117, 8115–8125, doi:10.1073/pnas.1911240117 (2020).
139. Gallen CL & D'Esposito M Brain Modularity: A Biomarker of Intervention-related Plasticity. Trends Cogn Sci 23, 293–304, doi:10.1016/j.tics.2019.01.014 (2019). [PubMed: 30827796]
140. Chiu D et al. Multifocal transcranial stimulation in chronic ischemic stroke: A phase 1/2a randomized trial. Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association 29, 104816, doi:10.1016/j.jstrokecerebrovasdis.2020.104816 (2020). [PubMed: 32321651]
141. Brys M et al. Multifocal repetitive TMS for motor and mood symptoms of Parkinson disease: A randomized trial. Neurology 87, 1907–1915, doi:10.1212/WNL.0000000000003279 (2016). [PubMed: 27708129]
142. Magsood H, Syeda F, Holloway K, Carmona IC & Hadimani RL Safety Study of Combination Treatment: Deep Brain Stimulation and Transcranial Magnetic Stimulation. Frontiers in Human Neuroscience 14, doi:10.3389/fnhum.2020.00123 (2020).
143. Deng Z, Lisanby SH & Peterchev AV in 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. 6821–6824.
144. Tetreault AM et al. Network localization of clinical, cognitive, and neuropsychiatric symptoms in Alzheimer's disease. Brain 143, 1249–1260, doi:10.1093/brain/awaa058 (2020). [PubMed: 32176777]
145. Joutsa J et al. Identifying therapeutic targets from spontaneous beneficial brain lesions. Annals of Neurology 84, 153–157, doi:10.1002/ana.25285 (2018). [PubMed: 30014594]
146. Kapur N Paradoxical functional facilitation in brain-behaviour research. A critical review. Brain 119 (Pt 5), 1775–1790, doi:10.1093/brain/119.5.1775 (1996). [PubMed: 8931597]
147. Carrera E & Tononi G Diaschisis: past, present, future. Brain 137, 2408–2422, doi:10.1093/brain/awu101 (2014). [PubMed: 24871646]
148. Sperber C Rethinking causality and data complexity in brain lesion-behaviour inference and its implications for lesion-behaviour modelling. Cortex; a journal devoted to the study of the nervous system and behavior 126, 49–62, doi:10.1016/j.cortex.2020.01.004 (2020). [PubMed: 32062142]

149. Griffis JC, Metcalf NV, Corbetta M & Shulman GL Structural Disconnections Explain Brain Network Dysfunction after Stroke. Cell reports 28, 2527–2540.e2529, doi:10.1016/j.celrep.2019.07.100 (2019). [PubMed: 31484066]

150. Klomjai W, Katz R & Lackmy-Vallee A Basic principles of transcranial magnetic stimulation (TMS) and repetitive TMS (rTMS). Annals of physical and rehabilitation medicine 58, 208–213, doi:10.1016/j.rehab.2015.05.005 10.1016/j.rehab.2015.05.005. Epub 2015 Aug 28. (2015). [PubMed: 26319963]

151. Su D et al. Frequency-dependent effects of subthalamic deep brain stimulation on motor symptoms in Parkinson's disease: a meta-analysis of controlled trials. Scientific Reports 8, 14456, doi:10.1038/s41598-018-32161-3 (2018). [PubMed: 30262859]

152. Radman T, Ramos RL, Brumberg JC & Bikson M Role of cortical cell type and morphology in subthreshold and suprathreshold uniform electric field stimulation in vitro. Brain Stimul 2, 215–228.e213, doi:10.1016/j.brs.2009.03.007 (2009). [PubMed: 20161507]

153. Feredoes E, Heinen K, Weiskopf N, Ruff C & Driver J Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. Proc Natl Acad Sci U S A 108, 17510–17515, doi:10.1073/pnas.1106439108 (2011). [PubMed: 21987824]

154. Duffley G, Anderson DN, Vorwerk J, Dorval AD & Butson CR Evaluation of methodologies for computing the deep brain stimulation volume of tissue activated. Journal of neural engineering 16, 066024, doi:10.1088/1741-2552/ab3c95 (2019). [PubMed: 31426036]

155. Thielscher A, Antunes A & Saturnino GB Field modeling for transcranial magnetic stimulation: A useful tool to understand the physiological effects of TMS? Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2015, 222–225, doi:10.1109/embc.2015.7318340 (2015).

156. Raij T et al. Prefrontal Cortex Stimulation Enhances Fear Extinction Memory in Humans. Biol Psychiatry 84, 129–137, doi:10.1016/j.biopsych.2017.10.022 (2018). [PubMed: 29246436]

157. Curot J et al. Awake Craniotomy and Memory Induction Through Electrical Stimulation: Why Are Penfield's Findings Not Replicated in the Modern Era? Neurosurgery 87, E130–e137, doi:10.1093/neuros/nyz553 (2020). [PubMed: 31914177]

158. Cole EJ et al. Stanford Accelerated Intelligent Neuromodulation Therapy for Treatment-Resistant Depression. The American journal of psychiatry 177, 716–726, doi:10.1176/appi.ajp.2019.19070720 (2020). [PubMed: 32252538]

159. Maeda F, Keenan JP, Tormos JM, Topka H & Pascual-Leone A Interindividual variability of the modulatory effects of repetitive transcranial magnetic stimulation on cortical excitability. Experimental brain research 133, 425–430, doi:10.1007/s002210000432 (2000). [PubMed: 10985677]

160. Brown JC et al. NMDA receptor partial agonist, d-cycloserine, enhances 10 Hz rTMS-induced motor plasticity, suggesting long-term potentiation (LTP) as underlying mechanism. Brain Stimul 13, 530–532, doi:10.1016/j.brs.2020.01.005 (2020). [PubMed: 32289670]

161. Minzenberg MJ & Leuchter AF The effect of psychotropic drugs on cortical excitability and plasticity measured with transcranial magnetic stimulation: Implications for psychiatric treatment. Journal of Affective Disorders 253, 126–140, doi:10.1016/j.jad.2019.04.067 (2019). [PubMed: 31035213]

162. Schaefer A et al. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex 28, 3095–3114, doi:10.1093/cercor/bhx179 (2018). [PubMed: 28981612]

163. Yeo BT et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106, 1125–1165, doi:10.1152/jn.00338.2011 10.1152/jn.00338.2011. Epub 2011 Jun 8. (2011). [PubMed: 21653723]

164. Gordon EM et al. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cerebral cortex 26, 288–303, doi:10.1093/cercor/bhu239 (2016). [PubMed: 25316338]

165. Destrieux C, Fischl B, Dale A & Halgren E Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53, 1–15 (2010). [PubMed: 20547229]

166. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC & Wager TD Large-scale automated synthesis of human functional neuroimaging data. Nature Methods 8, 665, doi:10.1038/nmeth.1635 https://www.nature.com/articles/nmeth.1635#supplementary-information (2011). [PubMed: 21706013]

167. Laumann TO et al. Functional System and Areal Organization of a Highly Sampled Individual Human Brain. Neuron 87, 657–670, doi:10.1016/j.neuron.2015.06.037 (2015). [PubMed: 26212711]

168. Gordon EM et al. Precision Functional Mapping of Individual Human Brains. Neuron 95, 791–807 e797, doi:10.1016/j.neuron.2017.07.011 (2017). [PubMed: 28757305]

169. Wang D et al. Parcellating cortical functional networks in individuals. Nat Neurosci 18, 1853–1860, doi:10.1038/nn.4164 (2015). [PubMed: 26551545]

170. Hacker CD et al. Resting state network estimation in individual subjects. Neuroimage 82, 616–633, doi:10.1016/j.neuroimage.2013.05.108 (2013). [PubMed: 23735260]

171. Cohen AL & Fox MD Reply: The influence of sample size and arbitrary statistical thresholds in lesion-network mapping. Brain 143, e41–e41, doi:10.1093/brain/awaa095 (2020). [PubMed: 32365379]

172. Botvinik-Nezer R et al. Variability in the analysis of a single neuroimaging dataset by many teams. Nature 582, 84–88, doi:10.1038/s41586-020-2314-9 (2020). [PubMed: 32483374]

173. Salvalaggio A et al. Reply: Lesion network mapping predicts post-stroke behavioural deficits and improves localization. Brain 144, e36–e36, doi:10.1093/brain/awab004 (2021). [PubMed: 33948628]

174. Bergmann TO et al. Concurrent TMS-fMRI for causal network perturbation and proof of target engagement. NeuroImage 237, 118093, doi:10.1016/j.neuroimage.2021.118093 (2021). [PubMed: 33940146]

175. Golay X, Hendrikse J & Lim TC Perfusion imaging using arterial spin labeling. Topics in Magnetic Resonance Imaging 15, 10–27 (2004). [PubMed: 15057170]

176. Fischl B FreeSurfer. Neuroimage 62, 774–781 (2012). [PubMed: 22248573]

177. MacInnes JJ, Dickerson KC, Chen NK & Adcock RA Cognitive Neurostimulation: Learning to Volitionally Sustain Ventral Tegmental Area Activation. Neuron 89, 1331–1342, doi:10.1016/j.neuron.2016.02.002 (2016). [PubMed: 26948894]

178. Bauer CCC et al. Real-time fMRI neurofeedback reduces auditory hallucinations and modulates resting state connectivity of involved brain regions: Part 2: Default mode network -preliminary evidence. Psychiatry Res 284, 112770, doi:10.1016/j.psychres.2020.112770 (2020). [PubMed: 32004893]

179. Tarakad A & Jankovic J Anosmia and Ageusia in Parkinson's Disease. International review of neurobiology 133, 541–556, doi:10.1016/bs.irn.2017.05.028 (2017). [PubMed: 28802932]

180. Williams ZM, Bush G, Rauch SL, Cosgrove GR & Eskandar EN Human anterior cingulate neurons and the integration of monetary reward with motor responses. Nature Neuroscience 7, 1370–1375, doi:10.1038/nn1354 (2004). [PubMed: 15558064]

181. Sheth SA et al. Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. Nature 488, 218–221, doi:10.1038/nature11239 (2012). [PubMed: 22722841]

182. Shine JM et al. Distinct patterns of temporal and directional connectivity among intrinsic networks in the human brain. The Journal of Neuroscience, 1574–1517, doi:10.1523/jneurosci.1574-17.2017 (2017).

183. Rafiei F, Safrin M, Wokke ME, Lau H & Rahnev D Transcranial magnetic stimulation alters multivoxel patterns in the absence of overall activity changes. Human Brain Mapping n/a, doi:10.1002/hbm.25466.

184. Buzsáki G, Anastassiou CA & Koch C The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. Nature Reviews Neuroscience 13, 407–420, doi:10.1038/nrn3241 (2012). [PubMed: 22595786]

185. Granger CWJ & Newbold P Forecasting economic time series. (Academic Press, 1977).

186. Friston KJ, Harrison L & Penny W Dynamic causal modelling. Neuroimage 19, 1273–1302, doi:10.1016/s1053-8119(03)00202-7 (2003). [PubMed: 12948688]

187. Barnett L, Barrett AB & Seth AK Misunderstandings regarding the application of Granger causality in neuroscience. Proc Natl Acad Sci U S A 115, E6676–e6677, doi:10.1073/pnas.1714497115 (2018). [PubMed: 29991604]

188. Friston K Causal modelling and brain connectivity in functional magnetic resonance imaging. PLoS biology 7, e33, doi:10.1371/journal.pbio.1000033 (2009). [PubMed: 19226186]

189. Webb JT, Ferguson MA, Nielsen JA & Anderson JS BOLD Granger Causality Reflects Vascular Anatomy. PLOS ONE 8, e84279, doi:10.1371/journal.pone.0084279 (2013). [PubMed: 24349569]

190. Granger CW Time series analysis, cointegration, and applications. American Economic Review 94, 421–425 (2004).

191. Mehler D & Kording K The lure of causal statements: Rampant mis-inference of causality in estimated connectivity. (2018).

Box 1 |

Neuroimaging techniques for mapping symptoms or behaviors

Some symptom localization approaches can yield stronger causal inference than others (Fig. 2). Of note, many correlative techniques can provide useful complementary information when combined with causal techniques[174].

Task-free neuroimaging

Task-free neuroimaging measures the brain at rest. Blood flow can be measured using positron emission tomography[2] or arterial spin labeling[175], connectivity can be estimated using resting-state functional MRI (fMRI) and diffusion tensor imaging[2], structure can be measured using volumetric MRI[57] and electrical activity can be measured using electroencephalography (EEG) or magnetoencephalography[176]. These measurements can be compared with clinical outcomes to identify neuroanatomical correlates of a behavior. Although task-free neuroimaging does not satisfy any of the criteria for causal inference[29,45], it can be combined with causal techniques to better understand the circuit-based mechanisms of a causal relationship[55]. Task-free neuroimaging can also be used to reliably map brain organization, potentially revealing targets for causal manipulation[174].

Task-based neuroimaging

Task-based neuroimaging measures the functional neuroanatomical correlates of a behavioral task. Studies using this approach estimate the counterfactual by comparing an experimental task to carefully designed control tasks. However, task-based imaging studies are rarely used to argue for causality owing to the temporal direction of the experiment — rather than manipulating brain activity and observing the resulting behavior, these studies manipulate the behavior and observe the resulting brain activity[3].

Neurofeedback

Neurofeedback is a variant of task-based neuroimaging in which a participant consciously manipulates their own task-based brain activity. As such, this technique modifies brain activity and measures a change in behavior, which allows for stronger causal inference than traditional task-based neuroimaging, which modifies behavior and measures changes in brain activity[177,178].

Incidental atrophy

Incidental atrophy patterns can be compared to behavioral outcomes, potentially illustrating specificity (assuming that the atrophy occurs in a near-random pattern) and dose-response (if greater atrophy leads to greater symptom severity). However, temporality is difficult to assess because atrophy progresses slowly. It can be unclear whether the atrophy caused the symptom, the symptom caused the atrophy, or both were caused by an unmeasured variable. For instance, substantia nigra atrophy is associated with anosmia because both phenomena can be caused by Parkinson disease[179].

Incidental lesions

Incidental lesions can illustrate specificity and dose-response in the same way as incidental atrophy. Furthermore, because lesions occur rapidly, their temporal effects are usually clear. Some lesions are even reversible, as in multiple sclerosis. Thus, lesion studies can provide reasonably strong causal inference[3,10,30,58,63]. However, because lesions occur in an uncontrolled setting, it is difficult to apply experimental manipulations or estimate the effect of unmeasured pre-lesion factors[3,17,29]. Lesions can also be challenging to interpret because of diaschisis or downstream effects on distantly connected brain regions[147].

Incidental stimulation sites

Incidental stimulation sites exist because therapeutic brain stimulation is often applied imprecisely to approximate targets. For example, TMS is usually targeted using scalp landmarks, leading to incidental variability in the stimulation site on the brain. When the actual stimulation sites are localized retrospectively, near-random variability is revealed in the precise region that is stimulated in any given patient[84,95]. Thus, the counterfactual can be estimated by comparing the post-treatment state to the pre-treatment state, while specificity can be demonstrated by comparing effective to ineffective stimulation sites. However, this approach still relies on incidental variance (a natural experiment) rather than prospective experimental manipulation, and thus may be subject to unmeasured confounders.

Targeted lesions

Targeted lesions are often created surgically as a therapeutic intervention. In addition to the insights gleaned from incidental lesions, prospectively targeted lesions can be experimentally manipulated, although this is rare given that the procedure is invasive and generally irreversible[180,181]. This enables strong causal inference, but is currently limited by the inability to assess reversibility.

Targeted stimulation

Targeted stimulation has many of the same advantages as targeted lesions, with the added benefit of reversibility. It also has the same advantages as incidental stimulation sites, with the added benefit of enabling precise experimental manipulation. Thus, targeted stimulation provides data about nearly all of the criteria on the causality continuum[10,14,17,18]. However, there are often assumptions about the effect of stimulation on neural function that may not be accurate (for example, excitation versus inhibition). Furthermore, to illustrate coherence, targeted stimulation must be complemented with other techniques.

Convergent causal mapping

Convergent causal mapping incorporates multiple causal approaches into a single analysis. For instance, if a symptom is caused by lesioning a specific circuit and the same symptom is relieved by stimulating the same circuit, this cross-modal coherence enables stronger causal inference than either modality alone.

Box 2 |

Mapping causal interactions between brain regions

Measuring effective connectivity using brain stimulation

Rather than assessing how certain brain regions causally influence specific behaviors, some studies assess how certain brain regions causally influence other brain regions. Causality can be inferred by combining focal stimulation with distant real-time physiological recordings[174]. For instance, TMS or DBS can be paired with immediate fMRI or EEG. These combinations are high on the causality continuum, as they employ targeted stimulation.

These real-time experiments investigate effective connectivity, which measures the direction of information flow in a brain circuit. If the experiments are rigorously controlled, then alternative and sham stimulation conditions can be used to estimate the counterfactual[174]. For instance, applying excitatory TMS versus inhibitory TMS to the default mode network (DMN) leads to dissociable effects on DMN connectivity[136,143]. Similarly, applying excitatory TMS to the frontoparietal control/central executive network (FPCN) can induce immediate changes in connectivity to the default mode network (DMN), while inhibiting the FPCN or stimulating other networks does not induce the same changes[137]. This capability is consistent across modalities, as intracranial electrical stimulation elicits distinct electrophysiological responses when stimulating the DMN versus the FPCN or the salience network[182]. This principle of using alternate stimulation targets to estimate the counterfactual also extends to other networks, as TMS to the dorsal attention network versus the DMN can induce network-specific EEG changes[138]. Collectively, these findings also suggest that TMS may be able to indirectly modify deeper brain regions that cannot be accessed directly.

Effective connectivity can provide mechanistic insights about direct or indirect relationships between sites of interest, and possible directionality and/or timing of signal flow between them. However, it does not offer causal inference about the importance of any of the sites for specific behaviors or functional outcomes[183].

Indirectly estimating connectivity

Effective connectivity may be estimated indirectly based on temporal relationships between activity in different brain regions[10]. These temporal forecasting methods are particularly useful for data with high temporal precision, such as intracranial measurements of local field potentials at the neuronal level[184]. The most common temporal forecasting approaches are Granger causality and dynamic causal modeling (DCM). Granger causality, which was originally designed for economic time series analysis, quantifies the directional relationship between two variables based on a temporal delay in their activity[185]. DCM uses a similar principle, but also incorporates multiple regions into a probabilistic graphical model allowing for different hypotheses to be specified by varying its nodes and directed edges. The downstream effects of an incidental fluctuation or behavioral task can then be estimated by selecting the simplest explanatory model[186].

These indirect approaches provide weaker causal inference than brain stimulation experiments, as temporal forecasting is not immune to spurious relationships[10,187,188]. Countless unmeasured confounders could cause a temporal delay[189], particularly for techniques with lower temporal resolution such as functional neuroimaging[10]. Nevertheless, Granger causality and DCM are often misunderstood as tools with strong causal inference, possibly due to misleading nomenclature[190,191].

The relationship between brain regions can also be estimated using task-free neuroimaging techniques such as resting-state functional connectivity and diffusion tractography. These approaches are low on the causality continuum, and provide no information about the causal direction of information flow.

Box 3 |

Circuit mapping approaches that may be combined with causal techniques

Several imaging techniques can provide reliable information about the organization of brain circuits. These approaches can be combined with causal information from lesions or brain stimulation sites to map effects to brain circuits. Each brain mapping technique has different strengths and limitations:

Structural connectivity

Diffusion MRI can measure the locations of white matter tracts, enabling us to study the effects of stimulating different parts of the same tract[86,90,97,110]. This provides high spatial resolution, but is unable to map polysynaptic relationships or crossing white matter fibers.

Functional connectivity

This approach estimates connectivity by measuring correlations between spontaneous fluctuations of distant brain regions, usually measured with fMRI. Unlike diffusion MRI, this technique can estimate the effects of a stimulation site on whole-brain polysynaptic networks[75,76]. However, it is limited by the assumption that temporal correlation implies a connection between different brain regions.

Normative connectivity

Rather than measuring structural or functional connectivity in individuals, it is possible to use a large connectome database to estimate the whole-brain connectivity of a particular brain region. This technique provides greater reliability at the expense of mapping inter-individual variability in brain architecture[58].

Surface electrophysiology

Electroencephalography (EEG) or magnetoencephalography (MEG) can be used to measure the immediate electrical effects of stimulating distant brain regions, providing high temporal precision[14]. These techniques enable the study of the rapid propagation of the stimulation to the rest of the brain, but are limited by poor spatial resolution.

Intracranial electrophysiology

Owing to greater anatomical precision than surface EEG, intracranial EEG (iEEG) enables precise mapping of specific behaviors. However, these tools are invasive and are usually only practical in specialized clinical centres where patients are evaluated for surgical treatment of refractory epilepsy[66].

Other approaches

In theory, it may be possible to define a priori circuits using any neuroimaging approach, including task fMRI, effective connectivity, regional cerebral blood flow, or volumetric imaging. These approaches have not yet been widely used to map the circuits connected to specific lesions or stimulation sites, so methodological validation is still needed.

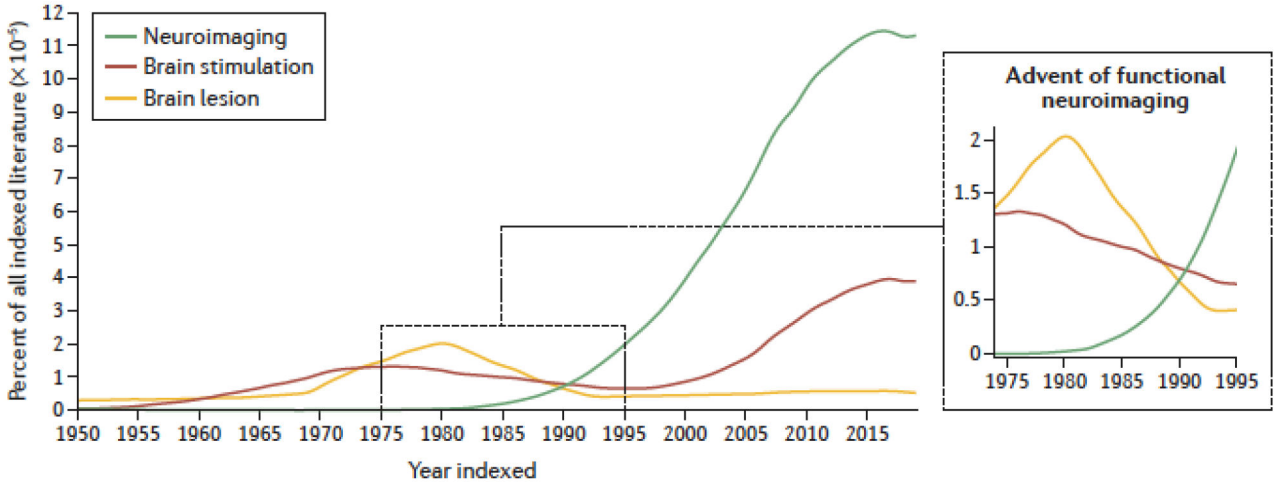

Fig. 1 |. Evolution of brain mapping literature since 1950.

In the late 20th century, the growth of neuroimaging literature coincided with a decline in brain lesion and brain stimulation literature. The data plotted on the figure were obtained by conducting a Google Ngram search using the terms 'neuroimaging,' 'brain stimulation,' and 'brain lesion'[31], with the data binned to show between 1950–2019 and the smoothing set to 5.

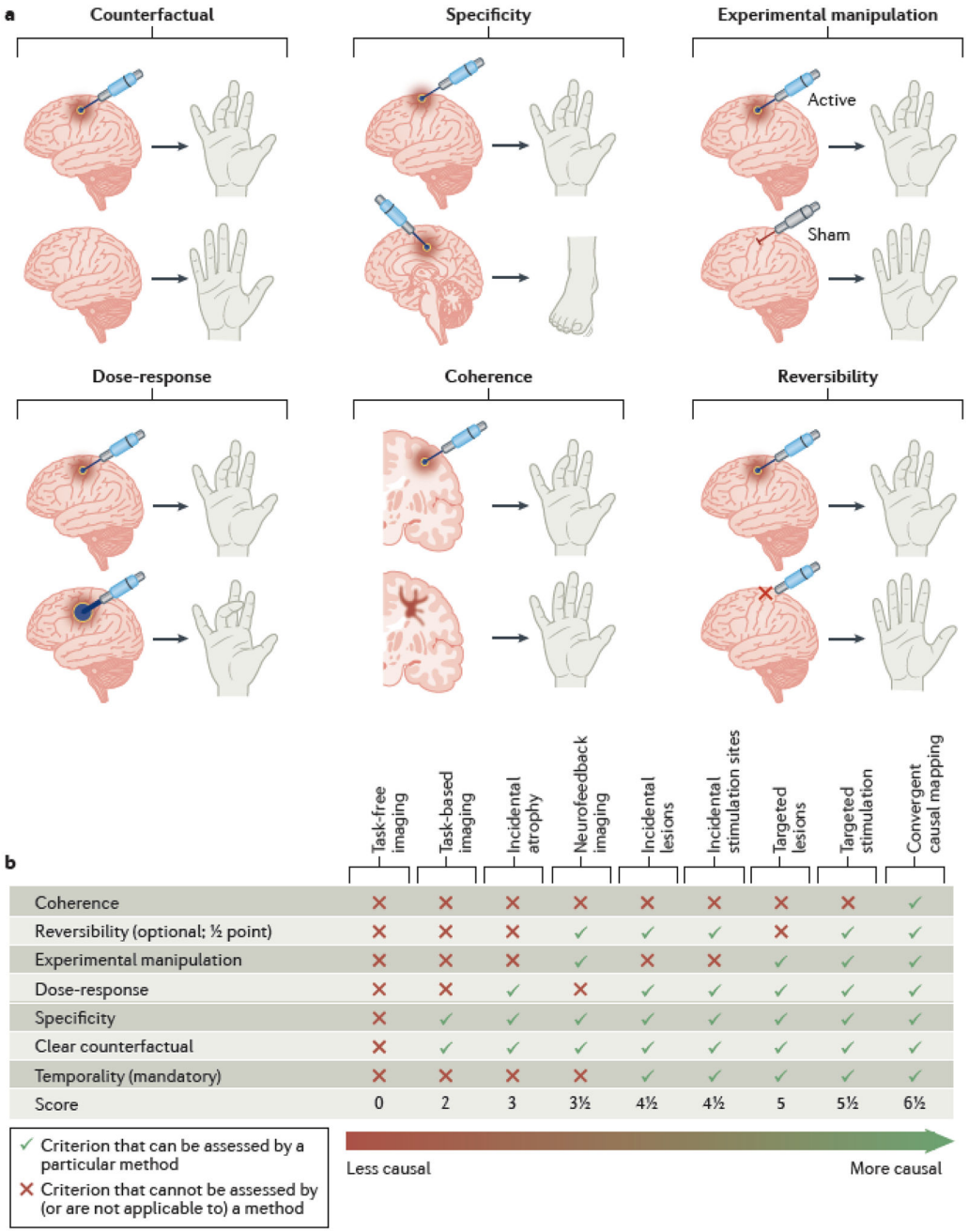

Fig. 2 |. Appraising causality in human brain mapping studies

a | Six criteria for appraising causality in human brain mapping studies, adapted from the Bradford-Hill criteria. Intracranial electrical stimulation is used to illustrate each of these criteria. A counterfactual can be estimated by modeling a hypothetical situation in which brain activity was not modulated or was modulated differently. Specificity can be demonstrated when slightly different interventions induce measurably different outcomes. For instance, stimulating different parts of the motor cortex leads to target-specific effects on motor function. An experimental manipulation can be used to selectively modulate activity in a specific brain region compared to a control intervention (in this case a sham). A

dose-response relationship is evident when higher-intensity stimulation induces a higher intensity outcome, in this case a more intense muscle contraction. Coherence can be demonstrated if different approaches converge on similar results. For instance, if stimulating a location induces a finger movement, then a lesion at the same location should induce finger weakness. Reversibility can be illustrated by reversal of the behavior when the modification of brain function is discontinued. **b** | Our proposed causality continuum in human brain mapping. An example of how different brain mapping techniques may be appraised using criteria adapted from the Bradford-Hill criteria, ordered from least causal to most causal along a causality continuum. Targeted brain lesions and brain stimulation satisfy more causality criteria than other brain mapping techniques (described in Box 1 and Box 2).

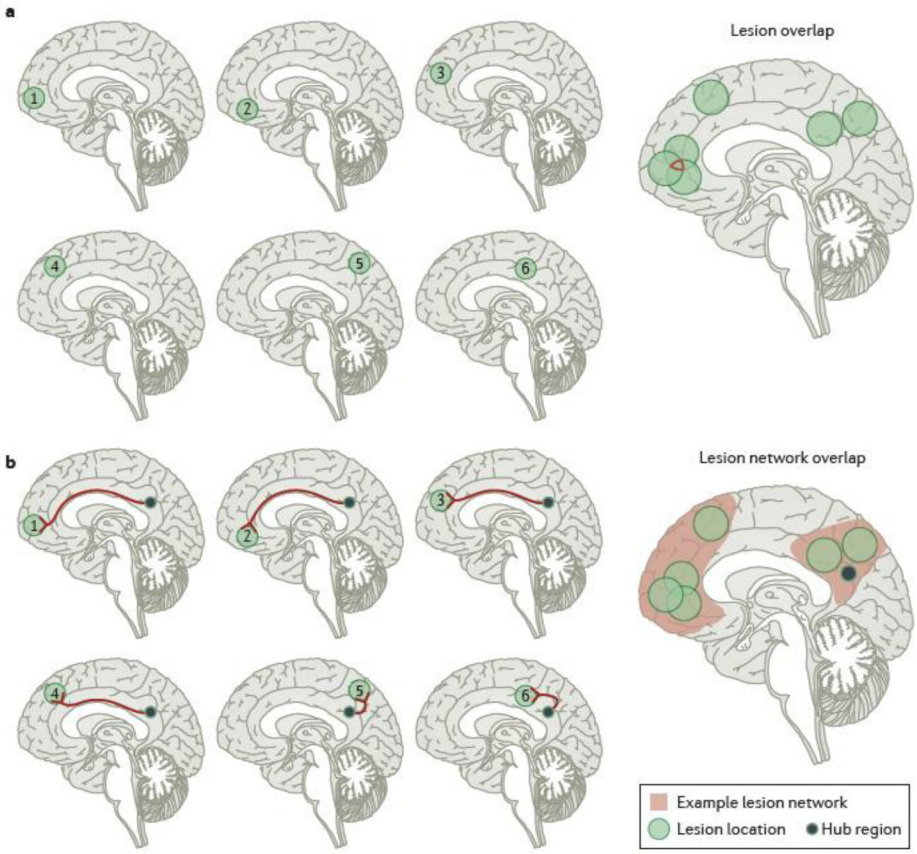

Fig. 3 |. Heteogeneous lesions causing the same symptom can complicate causal inference.
a | If lesions causing the same symptom (but not other lesions) intersect a common brain region (area of overlap for lesions 1–3; outlined in red), one can infer a causal role of this region in symptom generation. If other lesions causing the same symptom fail to intersect this brain region (lesions 4–6), this causal inference becomes weaker. **b** | If lesion locations causing the same symptom (but not any other lesions) are connected to a common hub region (lesions 1–6), then the connectivity with this hub region defines a brain network (orange) that encompasses all lesion locations causing the symptom. One can then infer a causal role of this brain network in symptom generation. Note that the hub region is not necessarily causally linked to the symptom, but rather defines a network that is causally linked to the symptom.

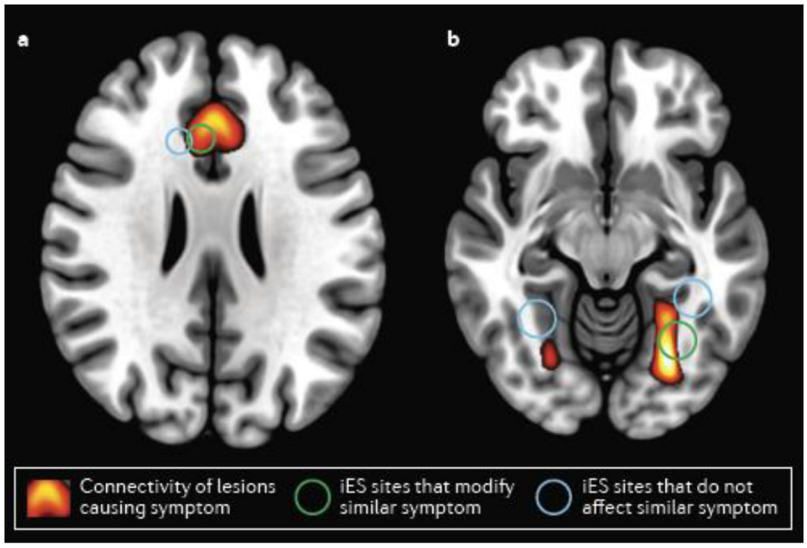

Fig. 4 |. Coherence between LNM and iES studies.

a | Lesion network mapping (LNM) revealed that lesions causing decreased will to perform actions are connected to the anterior cingulate (warm colors). **Intracranial electrical stimulation** (iES) sites overlapping with this network (green circle) can induce a will to persevere, unlike other nearby iES sites (white). **b | LNM revealed that** lesions causing disordered face perception are connected to the right posterior fusiform gyrus (FG, warm colors). iES to the right posterior FG (green circle), but not the right anterior FG or the left FG (white circles), causes distorted face perception.

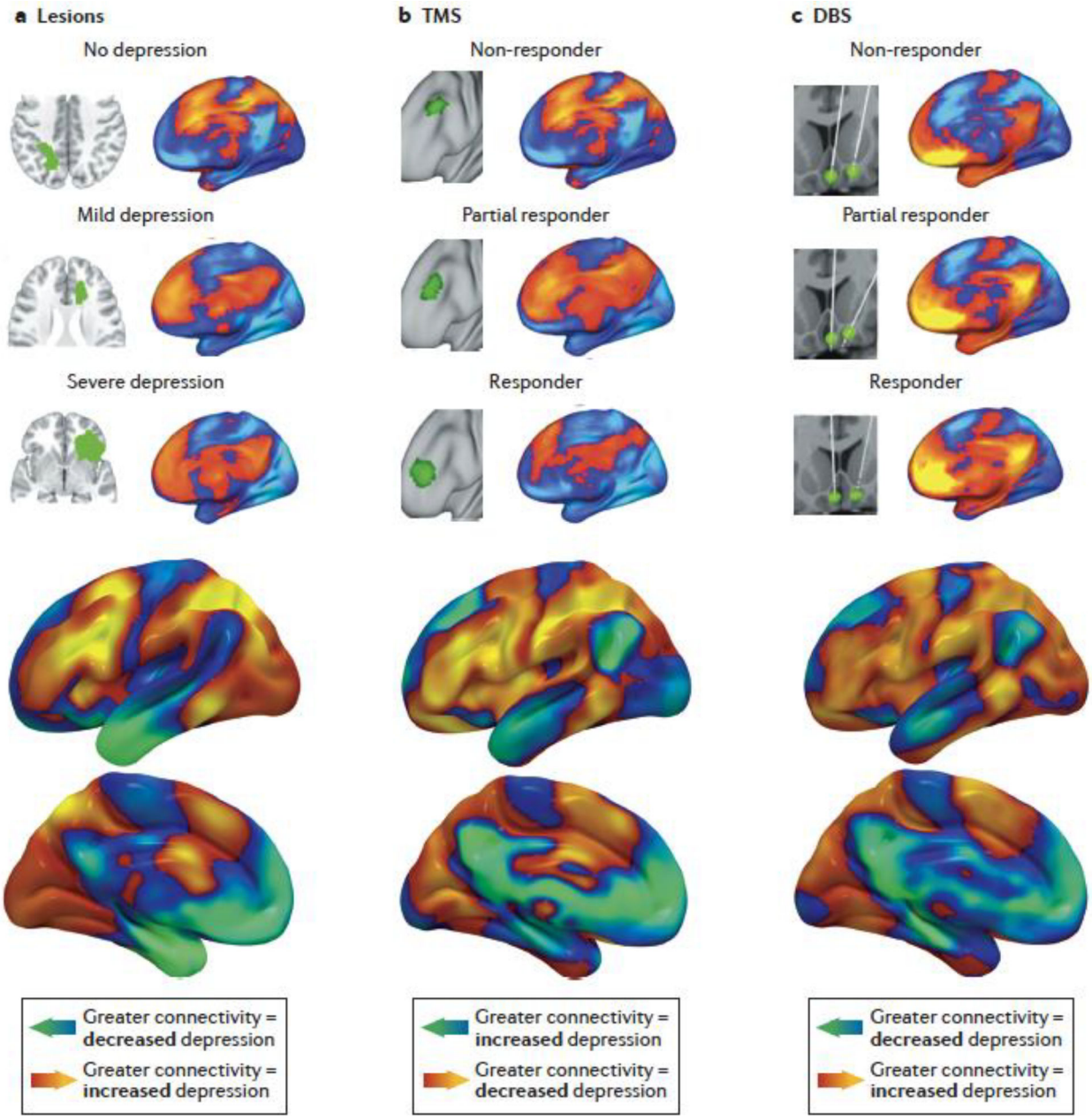

Fig. 5 |. Coherence between TMS, DBS, and brain lesions.

A common brain circuit was connected to brain lesions that cause depression, TMS sites that relieve depression, and DBS sites that modify depression. **a** | For 461 incidental brain lesions (top left), whole-brain connectivity was estimated using a normative connectome database (top right). This connectivity map was compared with depression severity, yielding a map of the whole-brain circuit connected to lesions that increase depression severity (bottom). **b** | The same procedure was conducted for 151 incidentally-variable TMS sites, yielding a map of the whole-brain circuit connected to TMS sites that improve depression. The color scale was inverted for this map because TMS sites that improve depression were expected to be anti-correlated to lesion locations associated with lower depression severity[109]. **c** | The same

procedure was also conducted for 101 DBS sites, yielding a map of the whole-brain circuit connected to DBS sites that modify depression. These three maps were significantly more similar than expected by chance. Adapted with permission from REF. 95.