Whether magnetic resonance imaging studies can provide useful information to clinicians who treat people with major depressive disorder remains to be established. There are, however, several recent findings that suggest that likelihood of response may be predicted by imaging findings. For example, morphometric studies have examined whether hippocampus volume is associated with clinically meaningful outcomes such as response to treatment. In general, patients who remit have larger pretreatment hippocampus volumes bilaterally compared with those who do not remit. There are similar preliminary findings for the anterior cingulate cortex. There are also a number of functional imaging studies that have identified different activity patterns in those who are likely to respond to treatment compared with those who are not. Using positron emission tomography, investigators have reported different patterns of response to treatment in those treated with medication compared with those treated with psychotherapy. Some of the potential barriers to the routine use of imaging in psychiatric practice are reviewed briefly.

Introduction

Magnetic resonance imaging (MRI) is included in clinical trials investigating a variety of neurologic disorders, including Alzheimer disease, multiple sclerosis and stroke. Ciumas and colleagues1 recently argued that MRI-derived parameters have advantages, including allowing for inferences regarding progression of the underlying disease, having better test–retest reliability than clinical outcome measures and, consequently, reducing the sample sizes required for clinical trials. End points based on MRI have been used to support the approval of medications such as glatiramer acetate and interferon-β for use in people with multiple sclerosis.2 To date, most MRI-based clinical trials have relied on morphometric assessment of regional brain volumes, but functional MRI (fMRI) provides an opportunity to examine the impact of interventions on the functional capacities of the brain.

Major depressive disorder (MDD) is a leading cause of disability worldwide, but it remains under-recognized and under-treated. There are many unmet needs in depression, but one is a need to develop the capacity to translate our current understanding of the brain dysfunction underlying MDD into clinically relevant information, as is becoming common with neurologic disorders. One way to link the neuropsychobiology of depression with clinical findings is through brain imaging studies that examine regional structure, regional function or connectivity to improve diagnostic or prognostic acumen. Although clinicians may diagnose MDD with good reliability, and it may be difficult to imagine imaging data supplanting clinical diagnosis, our ability to use clinical information to predict who will respond to a specific treatment in a timely way is more limited. If structural or fMRI studies could provide an accurate probability of a patient’s chances of responding to a specific treatment modality such as antidepressant or psychotherapeutic treatment, then there could be clinically important utility to the information obtained from an MRI scan. The objectives of this review were to discuss studies that examined the relations between imaging findings and clinical outcomes in patients with MDD and to consider whether it is likely that there will someday be a routine role for imaging data in the assessment or monitoring of patients with MDD.

Brain regions of interest in understanding depression

Abnormalities in corticolimbic structure and function are apparent in patients with MDD. These include dysfunction in...
regions within a neural system central to emotional processing comprising the amygdala, ventral striatum, hippocampus, anterior cingulate cortex (ACC) and ventromedial prefrontal cortex (VMPFC). Treatment with antidepressant medications may normalize abnormalities in corticolimbic function. Preclinical studies provide further evidence of structural changes in corticolimbic structures with stress that may be normalized with treatments for MDD. There is a paucity of studies, however, using within-subject prospective designs and sensitive methods of MRI analyses to detect regional brain changes in association with treatment for depression. Studies that have examined associations between regional brain volume or brain activity and clinical outcome are reviewed in subsequent sections.

Using regional brain volumes to predict clinical response

Cross-sectional and longitudinal approaches are now used to evaluate whether there are associations between regional brain volumes and clinical response in patients with conditions such as Alzheimer disease. Cross-sectional studies usually involve obtaining baseline brain volumes and following patients through treatment to determine whether baseline volumes can predict good or poor response to treatment. Longitudinal studies involve the repeated assessment of regional brain volumes over the course of treatment. Strictly speaking, proof of causality cannot be extrapolated from either cross-sectional or longitudinal studies because of the possibility of unknown factors influencing the measures. However, such studies can generate useful hypotheses regarding the relations between underlying brain substrate, disease progression and treatment effect, and longitudinal studies do provide information that can only be proposed by cross-sectional studies. A well-known example is the observed correlation between London taxi-driving and hippocampus volume in cross-sectional studies that led to the suggestion that driving around London, England, resulted in larger regional brain volumes. An alternative explanation, testable with a longitudinal approach, is that drivers with larger hippocampus volumes become successful and remain on the job longer than those with small hippocampus volumes and minimal propensity for spatial navigation. Longitudinal imaging studies are also more powerful for studies that have the goal of developing biomarkers that are relevant for early detection of disease, prediction of disease progression or development of treatment strategies.

More than 30 cross-sectional MRI studies have examined hippocampus volumes in patients with MDD, and several meta-analyses have confirmed that in the aggregate, people with major depression have hippocampus volumes that are about 5%–8% smaller than healthy controls. Cross-sectional studies have reported that small hippocampus volumes are associated with depression severity, age at onset, nonresponsiveness to treatment, untreated days of illness, illness burden, history of childhood abuse, level of anxiety and certain genetic polymorphisms.

The hippocampus is also small in patients with a variety of other neuropsychiatric conditions, including psychotic disorders, dementia and posttraumatic stress disorder. It is important to recognize, therefore, that whereas there is an association between hippocampus volume and recurrent depression, this association is not specific to depression. Whether the pathological mechanisms that result in small hippocampus volumes in patients with mood or anxiety disorders, psychosis or dementia are unique for each condition or represent shared pathophysiological processes is not confirmed. Bipolar disorder represents an interesting situation in which it has been very difficult to determine whether there are reliable differences between patients and healthy controls. Treatment effects, including an effect of lithium, may be partially responsible for the heterogeneity in reported findings, while comorbidity, course of illness and variability in subtypes of patients scanned may also contribute to the lack of consistency in studies of bipolar disorder.

Several prospective cross-sectional studies have examined whether hippocampus volume is associated with response to treatment in patients with depression (Table 1). A study of patients with geriatric depression reported that those with hippocampus volumes in the lowest quartile of the sample were less likely to achieve remission than those with volumes in the highest quartile, and another study reported that women who responded to 8 weeks of fluoxetine had larger right hippocampus volumes than nonresponders. Frodl and colleagues reported that depressed patients who were not remitted from an episode of depression 1 year after discharge had smaller left and right hippocampus volumes at baseline than those who were remitted. At 3 years of follow-up, patients with small hippocampus volumes and previous depressive episodes had a worse clinical outcome than patients.

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of patients</th>
<th>Outcome assessed</th>
<th>Sex effect</th>
<th>Laterality effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vakili et al.</td>
<td>38</td>
<td>Response after 8 wk of fluoxetine</td>
<td>Effect apparent only in women</td>
<td>Association apparent in right hippocampus</td>
</tr>
<tr>
<td>MacQueen et al.</td>
<td>46</td>
<td>Remission after 8 wk of antidepressant treatment</td>
<td>Not apparent</td>
<td>Association apparent bilaterally</td>
</tr>
<tr>
<td>Hsieh et al.</td>
<td>60*</td>
<td>Remission after 12 wk of antidepressant treatment</td>
<td>Not apparent</td>
<td>Association stronger in right hippocampus</td>
</tr>
<tr>
<td>Frodl et al.</td>
<td>30</td>
<td>In remission at 1 yr</td>
<td>Not reported</td>
<td>Association stronger in right hippocampus</td>
</tr>
<tr>
<td>Kronmüller et al.</td>
<td>57</td>
<td>Sustained recovery for 2 yr</td>
<td>Effect apparent in men</td>
<td>Association apparent bilaterally</td>
</tr>
<tr>
<td>Frodl et al.</td>
<td>30</td>
<td>Remission for 3 yr</td>
<td>Not reported</td>
<td>Association apparent bilaterally</td>
</tr>
</tbody>
</table>
with large hippocampus volumes. These investigators suggested that small hippocampus volumes may represent a vulnerability factor for poor treatment outcomes in patients with MDD. However, none of these studies examined variation in subregions of the hippocampus, and none examined patients with no prior treatment to control for the possibility that past treatment responsiveness resulted in, rather than from, small hippocampus volumes. Table 1 also summarizes the studies examining whether hippocampus volumes are associated with clinical outcome.

We recently examined hippocampus volumes in a group of 63 participants who had baseline MRI scans and then completed at least 8 weeks of first treatment for depressive symptoms. We compared anterior and posterior hippocampus volumes of patients who met criteria for clinical remission with those of patients who did not meet criteria for remission to determine whether there was an association with regional hippocampus volumes and clinical response. On average, patients who entered remission had larger posterior hippocampus volumes than those who did not enter remission. Because these patients had never been treated before the study, we excluded the possibility that past responsiveness to medication had resulted in both larger hippocampus volumes and good outcome for the index episode.

A large preclinical literature implicates the hippocampus as a key target of antidepressant medication, supporting the hypothesis that the integrity of the hippocampus may have predictive value as a marker of treatment responsiveness. Indeed, Frodl and colleagues found that hippocampus volumes increased over 3 years among patients who took antidepressant medication during that time. In addition to MRI studies of patients with MDD, positron emission tomography (PET) studies have been used extensively to evaluate the metabolic activity of discrete brain regions in depressed patients compared with healthy controls by observing the increased glucose uptake (metabolism) and regional blood flow resulting from metabolically active tissues. There is an extensive literature of PET studies in depression that is beyond the scope of this review, but it is worth briefly considering the literature reporting on blood-flow changes in the hippocampus in depression to determine whether there is any support for linking the hippocampus with clinical responsiveness. Some studies of regional cerebral glucose metabolism do not report differences in activity in the hippocampus of depressed patients compared with controls, whereas others report increased blood flow, and therefore increased regional metabolism, to the hippocampus.

Another provocative line of research has focused on the blood flow to the hippocampus in depression using a different methodological approach. Relying on a relation between angiogenesis and neurogenesis, Pereira and colleagues used MRI measurements of cerebral blood volume as an imaging correlate of neurogenesis. Healthy controls were scanned before and after 3 months of aerobic exercise; cerebral blood flow in the dentate gyrus increased over time. These results may reflect exercise-induced neurogenesis in this key region of the hippocampus. The investigators argue that "the imaging tools presented here are uniquely suited to investigate potential pharmacological modulators of neurogenesis," but whether such an approach can be operationalized routinely to investigate neurogenesis in the human hippocampus is a matter of ongoing debate. The results do highlight that the adult hippocampus appears to be a highly plastic structure, at least in healthy individuals. Whether such plasticity is necessary for good clinical response and whether patients can lose this plasticity over the course of illness is unknown.

Most structural studies evaluating the relation between regional brain volume and treatment responsiveness have focused on the hippocampus. Chen and colleagues have also identified regions of the ACC that correlate with the rate of antidepressant response. Others have reported an increase in lateral prefrontal cortical grey matter volume in patients with chronic fatigue syndrome after treatment with cognitive behaviour therapy (CBT). The investigators suggest that the result provides an example of macroscopic cortical plasticity in the adult human brain. Whether similar changes would be observed with CBT in other patient groups is unknown.

Diffusion tensor imaging with voxel-based analysis of fractional anisotropy has also been used recently to examine elderly patients with nonpsychotic depression. Participants who did not reach remission had lower fractional anisotropy in multiple frontal limbic brain areas, including the rostral and dorsal ACC, dorsolateral prefrontal cortex, genu of the corpus callosum, white matter adjacent to the hippocampus, multiple posterior cingulate cortex regions and insular white matter, compared with patients who did enter remission. The investigators speculated that there may be a disconnection syndrome associated with poor antidepressant response, occurring as a result of microstructural abnormalities in cortico–striato–limbic systems, although the etiology of these microanatomical abnormalities is unknown. The finding that remission is associated with fractional anisotropy is an interesting one, but it remains to be determined whether similar results might be apparent for younger people with depression as, to date, these findings are restricted to elderly patients.

Future studies are required to confirm whether volumetric assessment of key regions in frontolimbic networks may have utility for identifying patients who are particularly likely or unlikely to respond to pharmacotherapy and whether similar relations hold for other therapeutic modalities. The capacity to identify patients who are likely to respond to various treatment modalities could eventually have relevance to clinical practice; in the shorter term, such methods might be more readily applied to reducing the sample sizes required in clinical trials by identifying and excluding patients who are unlikely to respond to certain treatment modalities.

Using fMRI to predict clinical response

Resting state studies

Greicius has recently argued that fMRI has largely failed to fulfill its promise in the clinical realm, in part because of limitations of fMRI when used in a standard task-activation paradigm. He has further argued, however, that resting-state
functional connectivity in which participants do not have to perform a task may overcome some of these limitations. A large-scale functional network exhibiting increased activity during the resting state has been described.3 This “default-mode” network includes the medial, lateral–frontal and lateral–parietal regions; the precuneus/posterior cingulate gyrus and the hippocampus, and in the absence of any stimulus the network exhibits temporally coherent low-frequency fluctuations of the blood oxygen–level dependent signal.10,11

Functional connectivity is used to examine relations between networked regions and can be described as the temporal correlation of a neurophysiological index measured in spatially discrete brain areas.12 Independent component analysis and region-of-interest (ROI) analyses are the main approaches used to characterize functional connectivity in resting state networks, and the relative merits of each approach have been enumerated.11 With either an independent component analysis or an ROI approach, the patient group connectivity map is generally compared with the control group connectivity map for conclusions about between-group differences. In these approaches, the resulting components must be evaluated for goodness of fit. A study investigating inter-rater (human–human) and intermethod (human–machine) reliability for determining default mode network activation in healthy individuals reported very high inter-rater reliability but only moderate intermethod reliability.39

The first resting state studies of depression were reported by Anand and colleagues.39 They used an ROI approach to examine resting-state connectivity between the dorsal cingulate cortex (Brodmann area [BA] 24) and the medial thalamus, amygdala and the pallidostriatum, reporting that the dorsal cingulate cortex had reduced connectivity with the 3 other regions in patients compared with controls. The investigators further noted an increased limbic activation in the depressed group to negative pictures, suggesting that reduced connectivity between the cingulate and limbic regions resulted in a loss of cortical control over limbic responsivity. Cingulate connectivity increased in the depressed group following treatment with sertraline.40

Greicius and colleagues41 subsequently used an independent component analysis approach to examine connectivity in depressed patients, reporting that patients had increased default mode network connectivity in several regions, including the thalamus and subgenual cingulate cortex (BA 25). Notably, there was a correlation with the duration of the current depressive episode, such that the longer someone was ill, the greater the subgenual connectivity within the default mode network. It remains to be determined whether resting state information can be used to distinguish patients from controls or treatment responders from nonresponders, but the changes observed within participants from pre- to posttreatment suggest that these issues should be further explored.

Activation studies

To have clinical utility as a diagnostic tool, the pattern of brain activity obtained in an imaging paradigm should distinguish patients from healthy controls such that the brain scan from a novel participant will be accurately ascribed to either the patient or healthy group at the individual level. Fu and colleagues42 recently used an fMRI paradigm that involved the incidental affective processing of sad facial stimuli with modulation of the intensity of the emotional expression (low, medium and high intensity). They examined the fMRI data at each level of affective intensity with a support vector machine pattern classification method. Using this approach, they correctly classified patients and controls with a sensitivity of 84% and a specificity of 89%, which corresponded to an overall accuracy rate of 86%. They also had a trend toward significance when they attempted to classify patients’ later clinical response from scan data obtained before the initiation of treatment; they correctly classified 75% of patients who subsequently showed a partial clinical response and 62% of those who showed a full response following treatment. Notably, they achieved this degree of accuracy despite a relatively small sample of patients (n = 19); further work is necessary to determine whether this method can be used to identify those patients who are likely to respond to treatment in a larger sample.

This same group of investigators also used support vector machine methods to examine the utility of activity patterns generated during verbal working memory tasks as a diagnostic method for depression.43 The functional neuroanatomy of verbal working memory was not judged likely to have clinically important utility. However, classification of clinical response did achieve a sensitivity of 85%, supporting the notion that further investigation of this approach is warranted. A few studies have examined whether patients with other psychiatric disorders can be identified reliably based on imaging data. Davatzikos and colleagues44 reported an accuracy rate of 81% in identifying patients with schizophrenia based on structural data. Using a monetary reward task, Zhang and colleagues45 identified patients with substance abuse with an accuracy rate of up to 83%.

Other studies have also focused on whether neural activity during an acute depressive episode predicts subsequent clinical response. From a variety of imaging modalities we know that activity in the ACC is predictive of clinical response to antidepressant medication.4 There is evidence that greater amygdala activation to emotional facial expressions among depressed patients predicts symptom resolution.4 A smaller number of studies have focused on predicting clinical response to CBT.46,47 Baseline dorsal ACC activity in patients treated with CBT for depression was predictive of subsequent clinical response.47 In another study, depressed patients with low subgenual cingulate cortex (BA 25) sustained reactivity to emotional stimuli displayed the strongest improvement with CBT.48 Although both studies identified the importance of the ACC activity in prediction of subsequent clinical response to CBT, task differences may account for variability in the observed patterns of association between ACC activity and response probability. Another study recently examined whether response to 8 sessions of CBT in patients with posttraumatic stress syndrome could be predicted by pretreatment activity observed in response to fearful and neutral facial expressions.49 Lack of improvement after treatment was...
predicted by greater levels of amygdala and ventral ACC activation.

In one recent study of adolescents with depression, the investigators examined blood oxygen–level dependent response to anticipation and outcome of reward. They reported that during the anticipation and receipt of monetary rewards, young people with MDD had reduced response in dorsal striatal reward areas. Furthermore, activation in caudate regions, in which the group with depression had reduced response relative to the comparison group, was correlated with positive affect reported in natural settings. Notably, a reasonable amount of the variance in positive affect was accounted for by reward processing when age and sex were also considered. Although the group did not specifically link activity in reward circuitry to outcome following an intervention, the relation between activity and subjective reports of positive affect suggests that response to anticipation of reward would be interesting to study in intervention trials.

A comprehensive review has recently discussed the information from imaging studies supporting the notion that cognitive therapy and medication have shared and unique mechanisms associated with response. DeRubeis and colleagues suggest that depression is associated with decreased prefrontal function, possibly arising in part from increased amygdala reactivity. They propose that cognitive behaviour therapy primarily enhances prefrontal function, whereas antidepressant medications have a direct impact on the amygdala. Positron emission tomography studies have examined differential brain responses to treatment with either pharmacotherapy or psychotherapy, with interesting implications about the possible shared and unique contributions of various treatment modalities. Functional MRI studies that examine patients randomly assigned to one treatment modality or another are lacking.

In addition to studies examining whether brain activity patterns can predict response in patients with MDD, other investigators have examined the association between baseline activity patterns and treatment response in selected other conditions. Whalen and colleagues, for example, examined whether pretreatment amygdala and rostral ACC reactivity to facial expressions could predict outcome following 8 weeks of antidepressant treatment in women with generalized anxiety disorder. The magnitude of the treatment response was predicted by greater reactivity to fearful faces in the rostral ACC (rACC) and less reactivity in the amygdala, although the overall magnitude of pretreatment rACC and amygdala reactivity did not differ between patients and controls. The investigators suggested that such a pattern of rACC–amygdala activity could have utility as a marker of responsiveness to medication in patients with generalized anxiety. This group has also shown that anticipatory activity in the ACC is associated with clinical outcome following 8 weeks of treatment for generalized anxiety disorders.

Activation studies to date have relied on a retrospective analysis of regions or voxels of interest when linking the activity patterns to clinical outcome; identifying a priori the ROIs in predicting outcome represents a different and distinct challenge. If studies consistently identify regions such as areas of the cingulate where greater activity at baseline is associated with treatment outcome, it may be possible to confirm these associations in prospective trials.

Use of peripheral markers to predict clinical response

In addition to recent studies examining whether imaging data may be used to classify patients or predict response, other studies have examined peripheral markers for possible predictive links with treatment responsiveness. Peripheral brain derived neurotrophic factor (BDNF) levels may represent a viable biomarker of antidepressant response. Brain derived neurotrophic factor is expressed at high levels in the brain, but it is also expressed in peripheral tissues and is found at high levels in serum. A series of reports now suggest that serum and plasma BDNF levels are decreased in patients with MDD and there are also data suggesting that low levels of BDNF are normalized by treatment with antidepressant medications. Finally, degree of improvement in depressive symptoms, as measured by the Hamilton Depression Rating Scale, is associated with change in BDNF levels. However, the sensitivity and specificity of BDNF as a marker for MDD have not been well evaluated. Arguably, its utility could be greatest not in identifying patients who are depressed, but in distinguishing patients who will or will not have a satisfactory clinical response to a course of treatment.

Whether early change in BDNF levels following initiation of treatment can predict later clinical response remains to be determined. Studies that use both imaging and peripheral marker measurements are lacking, therefore there is little information on whether peripheral BDNF levels can be linked with altered brain activity levels. In studies of patients with recent onset of schizophrenia, however, carriers of the BDNF methionine (Met) polymorphism at codon 66 had greater reductions in frontal grey matter volume than valine (Val) homozygous patients over 3 years of follow-up. The investigators did not measure the hippocampus specifically in this investigation. The Met variant is associated with inefficient BDNF trafficking and reduced activity-dependent BDNF release, and this longitudinal study suggests not only that the Met variant may be important during development, but also that it may continue to exert an influence on neuroplasticity in young adults. Notably, changes in cognition or symptoms over the follow-up period were not predicted by genotype in this study. The imaging findings are consistent with previous cross-sectional studies reporting that Met allele carriers with schizophrenia have smaller frontal and temporal grey matter volumes than Val homozygote carriers. Associations have been reported between Met carrier status, risk for depression and anxiety and small hippocampus volumes. In one recent study, the combination of Met carrier status and exposure to early-life stress was associated with small grey matter volume in the hippocampus and high depression scores. To date, however, findings similar to those of Ho and colleagues showing relations between Met carrier status and changes in regional brain volumes following illness onset are lacking for depression. Furthermore, the association...
between the val/met BDNF polymorphism and depression has been inconsistently reported. As the recent study by Gatt and colleagues supports, it may be necessary to examine early-life stress in conjunction with genotype in order for such relations to become apparent.

There is substantial interest in the role that other growth factors such as vascular endothelial growth factor and insulin growth factor have in the pathophysiology of MDD. The extensive preclinical and postmortem literature that links BDNF to MDD is lacking for other growth factors, and at this time there is little evidence that other growth factors or peripheral markers will be used to screen for or monitor depressive symptoms in the near term.

**Summary**

Clinicians are appropriately wary of any method that claims to be able to diagnose or predict the course of MDD. It may be that no test will ever supplant the joint exercise in which a patient and clinician engage when they assign a diagnostic label to a set of symptoms and experiences. On the other hand, there are instances where clinical ambiguity is substantial, and it is reasonable to consider where there is a role for noninvasive measures in the approach to understanding the complex or ambiguous collections of symptoms, signs and motivations with which patients can present.

More compelling perhaps, is the hope that noninvasive technologies will enhance our clinical capacity to predict which patient will do well with which treatment within a reasonable period of time. The argument is frequently made that while we need new treatments for MDD, we also need better ways of matching patients to optimal treatment. Measurement of peripheral markers, pharmacogenetics and even metabolomics are all approaches that have promise in the emerging world of personalized medicine. Magnetic resonance imaging investigations of the brain that are brief, noninvasive and yield reliable information about whether a patient is likely to respond to treatment represent another modality that may one day have clinical utility. The routine use of new technologies will come at a price. Scanning patients or drawing blood before treatment may be cost-effective if it enhances our ability to choose the best treatment for a specific patient. Studies that measure the economic costs and benefits of using such technologies in patients with psychiatric illnesses are not apparent on the horizon.

The sensitivity and awareness of the art and humanism of medicine is prominent in psychiatry and some may view technologies such as imaging or genomics as diminishing the importance of the relationships between health care provider and recipient. It is possible therefore that most clinicians will not soon accept the notion that technology can inform and enhance patient outcome without feeling that the essence of clinical practice is threatened. On the other hand, clinicians and patients may embrace new technologies with enthusiasm as a means to validate the severity of psychiatric illness and to put to rest the lingering notions that these illnesses are somehow less “real” than other medical conditions. We may suspect that we have moved beyond mind–brain dualism in psychiatry, but Mirescu and Kirmayer provided empirical data that this is not yet so. They reported that mental health professionals still tended to assign psychological versus neurobiological causes to behavioral problems. Further, if the cause of the problem is viewed as psychological, patients are considered more responsible and worthy of blame for the symptoms. Data from neuroimaging studies make it harder to sustain that dualistic approach to the interpretation of behavior. Whatever the ultimate role of imaging and other technologies in routine clinical practice, it seems likely that neurobiological and imaging studies will require that we continue to be mindful of the shifting (or not) nature of psychiatric practice.

**Competing interests:** None declared.

**References**
