Voxel-based morphometry: current perspectives

Abstract: Voxel-based morphometry (VBM) is a widely used automated technique for the analysis of neuroanatomical images. This work, which reviews important VBM clinical findings of the last few years, is divided into two main sections. After briefly introducing VBM methodology, in the first part, VBM findings on neurological and psychiatric diseases have been discussed separately. The reported studies were divided into studies that examine the diagnostic value of VBM results and their usefulness for the differential diagnosis between two disorders; studies investigating the potential of VBM for the diagnosis in an early stage of the illness, and studies that examine the utility of VBM for predicting the transition from prodromal phase to full-blown disease. In the second part, this review focuses on the most recent findings on the single-case approach. This analysis is useful for promoting the translational impact of VBM results in clinical practice, where clinicians need to make inferences at the level of the individual patient. Finally, within the single-case approach, a paragraph is dedicated to the potential forensic applications of VBM. Indeed, if used to support and integrate results obtained with classical forensic evaluations, VBM may provide objective data that could be used to reduce controversy in forensic psychiatric evaluations of mental insanity. In this second part, particular emphasis is given to the problem of results interpretation, which should be based mainly on the presence of anatomoclinical correlation. The review finishes with a provocative note reporting an interesting result of a single-case VBM analysis that highlights the risk to fall into the “reverse inference” reasoning. This example has been chosen because it effectively highlights the risks that are encountered when this technique is inappropriately used.

Keywords: neurological disorders, psychiatric disorders, clinical translation, forensic application, false positives, single case analysis

Introduction

Over the past two decades, hundreds of studies have shed light on the neuroanatomical structural correlates of neurological1–3 and psychiatric4–7 disorders. Many of these studies were performed using voxel-based morphometry (VBM), a whole-brain technique for characterizing between groups’ regional volume and tissue concentration differences from structural magnetic resonance imaging (MRI) scans.8–11 Before the advent of VBM, the manual delineation of region of interest was the gold standard for measuring the volume of brain structures. However, compared to the region of interest approach, VBM presents a large number of advantages that explain its wide popularity within the neuroimaging community. Indeed, it is an automated and relatively easy-to-use, time-efficient, whole-brain tool that could detect the focal microstructural differences in...
brain anatomy in vivo between groups of individuals without requiring any a priori decision concerning which structure to evaluate. Moreover, VBM exhibits comparable accuracy to manual volumetry;12 indeed, several studies have shown good correspondence between the two techniques,13–15 providing confidence in the biological validity of the VBM approach.

The classical VBM analysis works on T1-weighted volumetric MRI and relies on parametric statistics (two-sample t-test), which requires the data to be sampled from normally distributed populations. Before performing the statistical analysis allowing group-level inferences, the raw MRI images need to be preprocessed in four main steps.11 First, the images need to be segmented into their main tissue components (gray matter [GM], white matter, and cerebrospinal fluid) based on voxel intensities. Second, the images need to be spatially normalized to the same stereotactic space in order to correct for individual differences in brain shape and dimension to allow different brains to be compared directly. Third, the images may be modulated (optional) in order to compensate for volume changes during the spatial normalization such that modulated data measure the absolute volume of GM, while unmodulated data measure the relative concentration of GM. Fourth, the images should be smoothed (ie, after smoothing, the value of each voxel represents the average between its own original value and the value of surrounding voxels), in order to ascertain normal distribution of the data (an important assumption of VBM) and to reduce intersubject variability. Each of these preprocessing steps could potentially introduce bias (ie, misalignment of brain structures, misclassification of tissue type), which might lead to errors in the statistical analysis and to false-positive results. In order to prevent these potential errors affecting the results, the results should be corrected for multiple comparisons (the family-wise error – correction is implemented in VBM). The outcome of these steps is a statistical parametric map, highlighting all voxels of the brain where intensities (volume or GM concentration depending on whether the modulation step has been applied or not) in a group images are significantly lower/higher than those in the other group under investigation. More specific technical and methodological information about the VBM steps has been described in other reviews.11,16 This approach (ie, comparing groups with parametric statistics with family-wise error correction) was proven to be reliable and to have a low susceptibility to false-positive results when comparing groups of subjects.17

In this review, we first summarize findings from the more recent and important studies in which VBM has been applied to investigate neurological and psychiatric pathologies. Although the distinction between these two types of disorders has been called into question on the basis of different scientific data,18 a recent meta-analysis showed that in a neuroimaging perspective, neurological and psychiatric disorders represent two distinct classes of disorders characterized by different neural correlates.19 Consequently, in what follows, we treated the two categories of diseases separately. The reported studies were divided into 1) studies that examine the diagnostic value of VBM and studies that examine the potential of VBM for the differential diagnosis between syndromes that present similar symptoms; 2) studies that examine the potential of VBM for the early detection of illness; and 3) studies that examine the potential of VBM for predicting the transition from prodromal phase to full-blown disease.

We then discuss the potential impact of these findings on clinical practice, concluding with a brief description of the forensic perspective. In this last part regarding the analysis on single subjects, particular emphasis is given to the problem of the results interpretation.

VBM in neurological disorders

Diagnosis and differential diagnosis

Diagnosis

Since its introduction, VBM has proven to be a very useful tool for investigating possible neuroanatomical biomarkers of neurological disorders. Within this class of pathologies, VBM has been mostly used for studying patterns of brain atrophy in the context of several neurodegenerative pathologies, including Alzheimer’s disease (AD). Clinically characterized by a progressive impairment of cognitive functions and behavior, AD is the most common cause of dementia in the general population and commonly occurs in patients older than 60 years of age. Because the diagnosis of AD has lower sensitivity (~70%) and specificity (<70%) than the gold standard neuropathological diagnosis,20 the characterization of the prototypical neural substrates of AD may be helpful for its diagnosis. As a result, a number of VBM studies have been conducted to investigate brain volume abnormalities associated with the diagnosis of AD in comparison to healthy controls (HCs), reporting remarkable results and suggesting that VBM can be useful to improve diagnostic accuracy and for monitoring disease progression in AD.

Although published results vary significantly in terms of the site of GM volume differences, a GM reduction in
the medial temporal lobe (MTL), including hippocampus, entorhinal cortex, and parahippocampal gyrus, appears to be a consistent and robust finding that is reliably demonstrated until the early stage of AD. Indeed, core pathological changes observed in AD take place in these regions and are initiated years before the first clinical manifestations of dementia. This result is consistent with histopathological studies showing that MTLs, and particularly hippocampus and parahippocampus, are the first regions to be affected by neurofibrillary tangles and amyloid plaques and exhibit the greatest loss of neurons in AD. Furthermore, as confirmed by longitudinal studies, the hippocampus and entorhinal cortex have been identified to suffer from the most severe atrophy (up to 40%), consistent with the memory deficits experienced by all AD patients.

Significant differences in GM volume in AD patients have also been identified in other brain areas, supporting the hypothesis that the initial pathological changes in AD occur in the MTL and then spread to other regions. Specifically, regions such as the temporal neocortex, parietal neocortex, insula, precuneus, anterior cingulated cortex, posterior cingulated cortex, frontal cortex, thalamus, and caudate nucleus have been frequently reported to be reduced in volume in AD compared to normal controls. Conversely, other brain regions have been identified only occasionally, such as the putamen, basal forebrain, and hypothalamus.

Recently, two meta-analyses in the attempt to draw conclusions from a synthesis of previous studies, have confirmed the presence of atrophic patterns predominantly and early involving the MTL structures that spread throughout parietal and temporal regions and posterior parietal cortex as the hallmark of AD pathology compared to normal aging since the early stage of dementia.

Differential diagnosis

After AD, frontotemporal lobar degeneration (FTLD) is the most frequent cause of dementia among individuals under the age of 65 years. FTLD is a heterogeneous neurodegenerative syndrome caused by progressive degeneration of the frontal lobes, anterior temporal lobes, or both and consisting of three clinical subtypes: behavioral variant frontotemporal dementia, semantic dementia, and progressive nonfluent aphasia. Although AD and FTLD are recognized as distinct disorders and clinical criteria have been defined for both, the frequent overlap of the clinical symptoms may pose serious problems in the differential diagnosis between the dementias, particularly in the early stage of the disease. As a result, there has been increased interest in using MRI to better understand the distinctive impact of these pathologies on the gross and fine morphology of the brain and hence to aid in differential diagnosis.

Compared to other structural MRI techniques, such as hippocampal volumetry and tensor-based morphometry, VBM has been demonstrated to show higher accuracy in its ability to differentiate between FTLD and AD. Several VBM studies have been conducted in order to compare patterns of brain atrophy in FTLD and AD and have identified various discriminatory regions between these disorders that vary depending on the specific FTLD subtype. For instance, compared to AD, semantic dementia patients manifested a greater volume loss throughout the temporal lobe (anterior, medial, and lateral areas), mirroring the speech and language deficits that are the most salient features of this pathology. On the other hand, AD is associated with significantly more cinguloparietal volume loss, mainly reflecting the problems in orientation and memory experienced by AD patients. Differently, behavioral variant frontotemporal dementia subtype (in which behavioral and personality abnormalities are the most salient features) is characterized by a more extensive GM atrophy in orbitofrontal and inferior frontal areas, anterior cingulate and insula, than AD patients, once again mirroring the specific clinical manifestation of the disorder.

Prodromal stage

Characterized by impairment in cognition extending beyond what one would expect based on age and education alone that does not interfere notably with activities of daily life, mild cognitive impairment (MCI) is thought to be an intermediate stage between normal aging and dementia. According to the profile of cognitive impairment, MCI can be further classified into two major subtypes: amnesic MCI (aMCI), characterized by episodic memory complaints without significant impairment in global cognition, and nonamnesic MCI (naMCI), dominated by symptoms other than memory impairment. Because the two different MCI subtypes have been clinically proved to be at higher risk to develop to AD or other type of dementia, respectively, the development of objective neuroimaging markers to differentiate MCI from healthy elderly subjects constitutes a major research goal. Indeed, emerging interventions are most likely to be successful prior to generalized cognitive decline.

A number of VBM studies have been conducted to investigate brain volume abnormalities associated with the diagnosis of MCI. Overall, they have documented smaller
GM volumes in patients relative to controls mainly affecting the MTL and cingulated gyri and extending into the temporal neocortex.\textsuperscript{46–48} Consistent with these data, a recent meta-analysis\textsuperscript{39} of 34 neuroimaging studies using VBM highlighted GM volume reduction in the MTL, temporal, thalamus, and cingulate cortex in MCI compared to HCs. Furthermore, contrasting AD and MCI, they found similar neural networks of GM abnormalities between these pathologies but that was more severe in AD than in MCI.

Interestingly, VBM studies have also assessed the presence of differences in the distribution of GM loss between subclassifications of MCI patients, providing evidence of distinct brain structural abnormalities in different samples of MCI\textsuperscript{49–51} according to their main clinical symptoms. For example, compared to normal controls, MCI with a focal memory deficit had more involvement of the mesial temporal structures,\textsuperscript{49–51} naMCI subjects with language impairment had loss in the left anterior inferior temporal lobe,\textsuperscript{50} while naMCI with attention/executive deficits showed loss in the orbitofrontal cortex, hypothalamus, and basal ganglia.\textsuperscript{50,51} There is also some suggestion that the patterns of atrophy vary depending on the number of cognitive domains impaired, with MCI that shows deficits in multiple domains showing a more widespread pattern of atrophy.\textsuperscript{49} These studies also provided evidence for the presence of anatomoclinical correlations even in the early stage of the disease.

Transition from prodromal stage to full-blown disease

Although previous studies have shown an increased rate of conversion to dementia in MCI, some of the MCI subjects will remain stable or even improve during follow-up.\textsuperscript{52} Currently, no specific methods are available to predict who will develop dementia, and hence biomarkers that help in predicting MCI subjects likely to convert to dementia are under extensive investigation. Consequently, one important challenge for neuroimaging is discriminating MCI patients progressing later to AD (converter-MCI) from MCI patients who will not (nonconverter-MCI). In this regard, VBM allows to longitudinally assess brain changes associated with the progression from MCI to dementia, and therefore, it is considered a promising tool in the research of biomarkers for the early detection of disease.

As a result, several longitudinal VBM studies have explored the spread of GM loss over time in MCI.\textsuperscript{23,49,50,53–57} Overall, these findings show that subjects with MCI who go on to develop AD have significantly lower baseline GM volume in MTLs, posterior cingulated, lateral temporal, and parietal cortices compared with HCs or stable MCI, although they can have heterogeneous results in terms of laterality and atrophy extent. For example, Bozzali et al\textsuperscript{23} showed that, in comparison to normal controls, MCI-converters exhibited more widespread areas of reduced GM density than nonconverters, with a pattern of abnormalities similar to that seen in patients with AD. In addition, the converting-MCI sample displayed significant GM decreases in the inferior frontal gyrus bilaterally, the left supramarginal gyrus, and the right hippocampus relative to nonconverting-MCI group. A further study\textsuperscript{50} also reported a pattern of bilateral GM loss affecting the MTL, inferior temporal lobe, temporoparietal association cortex, and frontal lobe in aMCI-converters compared to elderly controls. When compared to MCI who remained clinically stable for 3 years, MCI-converter subjects presented greater GM loss in the MTL, inferior temporal lobe, temporoparietal neocortex, anterior and posterior cingulate gyri, precuneus, and frontal lobe.\textsuperscript{50}

In order to highlight a consistent pattern of GM abnormalities across studies, meta-analyses of VBM studies have been conducted. Two recent meta-analyses\textsuperscript{2,39} demonstrated that the left MTL is the most affected region in MCI patients who will later develop AD, especially the hippocampus and the parahippocampal gyrus, and that this structure may be an anatomical marker to supervise disease progression from MCI to AD. Again, it is important to underline the strong anatomoclinical correlation: the majority of MCI patients complained about memory difficulties, which is mirrored by GM atrophy in structures known to be involved in memory.

VBM in psychiatric disorders

I didn’t have the luxury of taking reality for granted. And I wouldn’t say I hated people who did […] They didn’t live in my world. But that never stopped me from wishing I lived in theirs. Francesca Zappia, Made You Up

Diagnosis and differential diagnosis

Diagnosis

Although the neuroanatomical basis of most neurological diseases is relatively well understood, this is not true for psychiatric illnesses, which for a long time were considered “functional” disorders without a reliable neuroanatomical basis.\textsuperscript{58} The development of structural neuroimaging techniques such as VBM have provided substantial evidence that psychiatric diseases are associated with abnormalities in brain structure. This has brought about significant breakthroughs in our understanding of the neurobiology of such
illnesses. The peculiarity of psychiatric disorders being brain disorders without a clear brain correlate makes them highly interesting. The psychiatric diseases are very heterogeneous and the differential diagnosis is often made difficult by the heterogeneity and inconsistency of symptoms and by the possible effect driven by drug therapies. Within these diseases, VBM has been extensively used for studying patterns of GM abnormalities in the context of psychosis, in particular schizophrenia (SCZ), a psychiatric disorder which involves distortions in thought and perception, blunted affect, and behavioral disturbances. SCZ has always been considered the most fascinating of all the psychoses and has inspired books (eg, “Spider” by Patrick McGrath), movies (“A Beautiful Mind”, based on the life of the Nobel Laureate John Nash), and art (the painter Salvador Dali has been quoted as saying that the creative mind should mimic SCZ).

A number of VBM studies have been conducted to investigate brain volume abnormalities associated with the diagnosis of SCZ. However, these studies often reported inconsistent results that reflect the clinical heterogeneity of the included patients. In order to highlight consistent pattern of GM abnormalities across the studies, meta-analyses of VBM studies have been conducted. An early meta-analysis found that the most consistent GM differences between patients and controls are expressed as a smaller GM volume in the superior and medial temporal gyri in SCZ. A subsequent meta-analysis further expanded these results as it revealed GM structural findings in SCZ, characterized by bilateral anterior cortical (anterior cingulate gyrus/medial frontal cortex) and limbic (thalamus and left amygdale) GM abnormalities. Male sex, chronic illness, and negative symptoms were associated with more severe GM abnormalities. Moreover, compared to HCs, the patient group showed not only decreased GM, in particular in the right cuneus, the right superior frontal gyrus, and the right insula, but also increased GM in the bilateral putamen, the right parahippocampal gyrus, the left precentral gyrus, and the left inferior temporal gyrus.

The heterogeneity of clinical manifestation of SCZ may be responsible for the inconsistency of the results in the literature. One possible way to overcome this limitation could be the investigation of brain alterations associated with specific symptoms. For example, Modinos et al meta-analyzed VBM studies examining the neural basis of auditory verbal hallucinations in SCZ patients and revealed a close association between the severity of hallucinations and GM volume reductions bilaterally in the superior temporal gyrus. The neural basis of a symptom rather than of a disease is still under investigation in psychiatric research, with few studies emerging in the last few years, but it could help to disentangle the conundrum of the anatomoclinical correlation of psychiatric symptoms in the near future.

Differential diagnosis
Together with SCZ, bipolar disorder (BD) is the most frequent psychiatric disorder. BD is a devastating mental disorder, affecting ~1% of the population, characterized by remitting and relapsing episodes of depression and (hypo)mania and can also include psychotic symptoms.

Although SCZ and BD are recognized as distinct disorders and clinical criteria have been defined for both, there is a frequent overlap of the clinical symptoms, especially in the early stage of the disease where the initial presentation of BD overlapped with either the depressive features of major depressive disorder or the manic psychotic features of SCZ. This can pose serious problems in the differential diagnosis between the two psychiatric illnesses. As a result, there has been increasing interest in using VBM to better understand the distinctive impact of these pathologies on the morphology of the brain and hence to aid in differential diagnosis.

However, SCZ and BD not only share some clinical and behavioral symptoms but also some GM alterations. Interestingly, a recent meta-analysis of VBM studies investigated the presence of common areas of GM volume increase or decrease across Axis I diagnosis. GM volume in the anterior cingulated cortex and in the bilateral insulae has been found as consistently depleted in all the considered diagnoses, included SCZ and BD. Therefore, these brain regions are suggested as a transdiagnostic neural signature of psychiatric illness.

Given that brain structural changes are a putative feature of SCZ and possibly also BD, recent studies focused their attention on the different patterns of reduced GM. Some authors compared both SCZ and BD to controls and found that, despite similarities of the involved brain regions, SCZ manifest more extensive prefrontal, thalamic, and hippocampal deficits than BD. In addition to hippocampus atrophy, a study also found smaller GM volume in the putamen and amygdale in SCZ compared to BD. Other authors directly compared GM volume between the two psychiatric disorders and highlighted a pattern of brain alterations that is suggestive of a more widespread GM alteration in chronic SCZ compared to chronic BD. In particular, both cortical (dorsolateral and medial prefrontal and precentral) and subcortical regions (putamen and thalamus) were found to be mainly involved in SCZ.
Overall, these findings suggest greater brain alteration in SCZ compared with BD, with significant overlap between the two disorders. These impressions have been confirmed by a meta-analysis of VBM studies on SCZ and BD, which highlight that SCZ is characterized by extensive GM deficit in frontal, temporal, cingulated, and insular cortices and thalamus, coupled with an increased GM in the basal ganglia. Similarly, BD is characterized by GM reduction in the anterior cingulated and bilateral insula only.

**Prodromal stage**

The first episode of a psychotic disorder is usually preceded by a prodromal period characterized by a progressive decline in functioning and the emergence of attenuated psychotic symptoms. Individuals with these clinical features are considered UHR as well. Studies using VBM have examined whether UHR subjects are affected by neuroanatomical abnormalities. These have revealed that, relative to HCs, UHR subjects have reduced GM volume in frontal, lateral, and medial temporal regions. When dividing UHR groups into individuals with a genetic risk and those with attenuated psychotic symptoms and comparing them to a HC group, the former showed left anterior frontal, right caudate, a smaller right hippocampus, and an amygdala reduction, while the latter subgroup showed right middle temporal cortical reductions. Despite the variability of results reported in different studies, consistent results may be highlighted using meta-analysis. For example, a meta-analysis on 896 UHR revealed that individuals at UHR for psychosis consistently showed reduced GM volume compared with controls in the right superior temporal gyrus, left precuneus, left medial frontal gyrus, right middle frontal gyrus, bilateral parahippocampal/hippocampal regions, and bilateral anterior cingulated cortex.

**Transition from prodromal stage to full-blown disease**

Despite GM abnormalities being found in UHR, not all people with these features progress to develop a full-blown psychotic disorder. Indeed, within 24 months, 20%–50% develop SCZ, but the remaining do not. Currently, no methods are available to identify patients who will develop a full-blown SCZ before the transition occurs, and no biomarkers exist for psychiatric disorders. Indeed, one goal for neuroimaging is to provide evidence to support the existence of neuroanatomical differences between UHR patients who develop psychosis and those who do not. One pioneering study in this field, which has the advantage of being multicentric, revealed that reduced left parahippocampal volume was specifically associated with the later onset of psychosis in UHR individuals. Therefore, alterations in this region may be crucial to the expression of illness. This result is consistent with previous literature; reductions in parahippocampal volume have been described in high-risk individuals with transient or isolated psychotic symptoms, and longitudinal reductions have been described in high-risk individuals who developed psychosis. However, this has not been consistently replicated.

While the earlier reported studies might be intrinsically biased by the presence of subclinical SCZ individuals who have not yet experienced the first psychotic symptoms in the nontransition group, the neuroimaging studies in early psychosis show promise in the ability to identify core neurobiological alterations at the onset of the disorder. The study of the neuroanatomical alteration of patients who manifested the first episode of psychosis is advantageous since it allows the identification of brain alterations that are unrelated to treatment with antipsychotic medication. To date, hundreds of studies have been conducted on the first episode of psychosis, and data provided consistent evidence supporting a close relationship between structural and functional brain alterations in the medial frontal/anterior cingulate cortices and in the bilateral insulae, where patients showed a decrease in GM volume as well as abnormal functional response.

Within those regions, the right insula has recently attracted a great deal of interest since its alterations are associated with a higher risk for transition to both psychosis and the first episode of psychosis. Remarkably, a smaller GM volume in this region has been found to be a consistent datum not only across different psychiatric diagnoses, as reported in the previous paragraph, but also across different ethnic groups. This is in line with the notion of a neuroanatomical signature of early SCZ above and beyond ethnic differences in the manifestation of the disease. Moreover, the GM volume in the right anterior insula seems to be inversely related to both negative and positive symptomatology, such that the insula becomes smaller as symptoms become more severe. In conclusion, the right anterior insular alteration could reflect cognitive impairment and/or emotional dysfunction – two deficits have been associated with anterior insular alteration and that tend to co-occur across developmental and adult disorders.

In conclusion, the research on neuroanatomical correlates of psychiatric disorders is made difficult by 1) the heterogeneity of clinical manifestation of the disease across...
subjects; 2) possible and still underinvestigated effect of drug therapy on brain structure; and 3) still weak knowledge of anatomoclinical correlation in psychiatry.

Future research should be focused on the investigation of the neural basis of symptoms, rather than of disorder, in order to reduce variability across studies.

**Translation to clinical practice**

I’d looked at hundreds of psychopaths’ brains in my career, but the consistency of their brain abnormalities never ceased to amaze me. Kent Kiehl, The Psychopath Whisperer

The majority of research cited in the previous paragraphs reported group-level results, that is, results obtained by comparing two groups of individuals. Indeed, a classical VBM study compares a group of patients with a group of HCs and tests for neuroanatomical differences between the two groups relying on parametric statistics, which required the data to be normally distributed. However, the results of these studies have had limited translational impact in everyday clinical practice, where a clinician needs to make inferences at the level of the individual patient. In recent years, an increasing number of research groups have attempted to overcome this by performing single-case studies in which an individual patient is compared against a group of HCs. Table 1, an update by performing single-case studies in which an individual patient is compared against a group of HCs, summarizes the existing studies using single-case VBM.

Although the standard VBM technique is suitable for comparing groups of individuals, the assumptions of group-level parametric statistics are no longer met when an individual subject is compared against a group of subjects, resulting in very high false-positive rates, despite the use of the appropriate statistical corrections (Figure 1, top panel). In other words, in single-case VBM studies, disease-free individuals’ brains are often identified as abnormal. However, as reported in the previous paragraph, single-case VBM is widely used to investigate neuroanatomical abnormalities in individual subjects. Because of this, there is an urgent need for alternative analytical approaches that could be used to examine neuroanatomical abnormalities in individual subjects without suffering from high false-positive rates. Since the high rates of false positives arise from the statistic on which the classical VBM relies (because the data cannot be normally distributed by definition in single-case analysis), one possible solution seems to be the use of nonparametric statistics in VBM. This allows the examination of neuroanatomical abnormalities in individual subjects without the very high false-positive rates that have been reported with standard single-case VBM (Figure 1, bottom panel). However, this alternative approach solves a problem but creates another: when using nonparametric statistics, a control group of at least 100 individuals is required in order to accurately compute the P-value. The requirement for such a large control group makes the use of nonparametric statistics impractical in a clinical setting, where it is usually difficult to have such a large control group.

Besides the statistical issue, the interpretation of the results is also tricky. Indeed, it is worth noting that, when using the classical VBM approach that relies on parametric statistics, the false positives appeared to be mainly located in the frontal and temporal lobes. Ironically, these are also the areas of the brain that are typically affected in most psychiatric and neurological disorders, further complicating the interpretation of the results of single-case studies in these clinical populations. However, it is important not to fall into the base rate fallacy error – the tendency to ignore the existence of some a priori data. In this specific context, it is important not to ignore the fact that different brain regions are more likely to be affected in each specific disease. For example, a recent meta-analysis of VBM studies found that group-level anatomical differences between schizophrenic patients and controls are more likely to be expressed in the anterior cingulated, medial, inferior frontal, and temporal cortices. Thus, if a single-case VBM analysis highlights lower frontal GM values in a patient who show clinical symptoms suggestive of SCZ, this result is more likely to reflect a true positive than a false positive, even if this specific brain region suffers from high false-positive rate. This is because the a priori probability for SCZ to be expressed as decreased GM within the frontal lobe is high, based on previous literature. Contrarily, if a significant finding in the frontal lobe will be found in a comparison between a patient with cortical posterior atrophy (which involves the posterior part of the brain by definition) and a control group, this is more likely to reflect a false positive.

The single-subject analysis using VBM could in the future have transdiagnostic applications across psychiatry and neurology. However, the research on single-case VBM analysis is still in its infancy. To date, it is clear that VBM might be helpful in the diagnostic process but it should be used with caution, always remembering the error rate and understanding that a statistically significant result could be considered indicative of a brain disease only if a correspondence between the brain region emerged from the analysis and the clinical symptoms manifested by the patients (the so-called anatomoclinical correlation; Table 1) is highlighted. However, VBM usefulness to predict the patients’ decline or to predict the success of therapies (pharmacological or not)
<table>
<thead>
<tr>
<th>Study</th>
<th>Subject Disease</th>
<th>Number of control subjects</th>
<th>GM increase or decrease</th>
<th>Brain regions</th>
<th>Anatomoclinical correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taubert et al, 2015</td>
<td>Top athlete</td>
<td>27♂ 28♂ 30♂</td>
<td>Increase</td>
<td>Putamen + thalamus, Putamen + caudatus, Putamen + caudatus</td>
<td>UK</td>
</tr>
<tr>
<td>Martinaud et al, 2015</td>
<td>Developmental visual impairment</td>
<td>14♀</td>
<td>Decrease</td>
<td>R ventral occipitotemporal region, bilateral cingulate sulcus, paracentral lobule, SMA</td>
<td>Yes</td>
</tr>
<tr>
<td>Luedecke et al, 2014</td>
<td>Presenilin I mutation</td>
<td>50♀</td>
<td>Decrease</td>
<td>Precuneus, parietal lobe, the R hippocampus</td>
<td>Yes</td>
</tr>
<tr>
<td>Kitamura et al, 2014</td>
<td>REM sleep behavior disorder</td>
<td>72♀</td>
<td>Decrease</td>
<td>Tegmental portion of the pons, anterior lobes of the cerebellum, hippocampus</td>
<td>UK</td>
</tr>
<tr>
<td>Serino et al, 2014</td>
<td>Electrocution-induced heart attack</td>
<td>44♀</td>
<td>Decrease</td>
<td>R precuneus</td>
<td>UK</td>
</tr>
<tr>
<td>Yasuno et al, 2014</td>
<td>TBI</td>
<td>30♀</td>
<td>Decrease</td>
<td>Bilateral insular cortex, posterior cingulate cortex, hypothalamus, R parietal cortex, medial occipital cortex</td>
<td>Yes</td>
</tr>
<tr>
<td>Yoo et al, 2013</td>
<td>Isolated congenital anosmia</td>
<td>14♀</td>
<td>Decrease</td>
<td>L OFC, precuneus, gyrus rectus, middle cingulate cortex, thalamus, subcallosal gyrus, fusiform gyrus</td>
<td>Yes</td>
</tr>
<tr>
<td>Migliaccio et al, 2012</td>
<td>PCA 6 subjects</td>
<td>29</td>
<td>Decrease</td>
<td>Bilateral ventral occipital and temporal (three of seven patients), R supramarginal gyrus (two of seven patients), bilateral inferior parietal lobule (seven of seven patients)</td>
<td>Yes</td>
</tr>
<tr>
<td>Beeson et al, 2011</td>
<td>PPA</td>
<td>67♀</td>
<td>Decrease</td>
<td>Bilateral posterior perisylvian regions, bilateral mild lateral temporal lobes</td>
<td>Yes</td>
</tr>
<tr>
<td>Klingner et al, 2011</td>
<td>Bell palsy</td>
<td>23♀</td>
<td>Decrease</td>
<td>L M1, bilateral SMA, L insula, bilateral cerebellum</td>
<td>Yes</td>
</tr>
<tr>
<td>Migliaccio et al, 2011</td>
<td>PCA</td>
<td>58♀</td>
<td>Decrease</td>
<td>Bilateral superior occipital gyrus, cuneus, bilateral inferior occipital gyrus, bilateral lingual gyrus, bilateral superior parietal lobule, bilateral thalamus</td>
<td>Yes</td>
</tr>
<tr>
<td>Sehm et al, 2011</td>
<td>Focal retrograde amnesia</td>
<td>24♀</td>
<td>Decrease</td>
<td>L temporopolar cortex, R lingual cortex</td>
<td>Yes</td>
</tr>
<tr>
<td>Valdes-Sosa et al, 2011</td>
<td>Prosopagnosia</td>
<td>73♀</td>
<td>Decrease</td>
<td>L ventral and occipital temporal cortex, L fusiform gyrus</td>
<td>Yes</td>
</tr>
<tr>
<td>Bianchini et al, 2010</td>
<td>Topographical disorientation</td>
<td>22♀</td>
<td>Decrease</td>
<td></td>
<td>--</td>
</tr>
<tr>
<td>Erguchi et al, 2010</td>
<td>Citrullinemia epilepsy</td>
<td>31♀</td>
<td>Decrease</td>
<td>L hippocampus</td>
<td>UK</td>
</tr>
<tr>
<td>Freudermann et al, 2010</td>
<td>Vascular encephalopathy</td>
<td>72♀</td>
<td>Decrease</td>
<td>R postcentral, R IPL, bilateral putamen, L cingulated cortex</td>
<td>UK</td>
</tr>
<tr>
<td>Maguire et al, 2010</td>
<td>SD</td>
<td>70♀</td>
<td>Decrease</td>
<td>L striatum, LSTG, bilateral hippocampus, L IFG, R temporal pole, R cerebellum, R ITG</td>
<td>No</td>
</tr>
<tr>
<td>Nanri et al, 2010</td>
<td>Autoimmune ataxia + Basedow disease</td>
<td>84♀</td>
<td>Decrease</td>
<td>R cerebellar cortex</td>
<td>Yes</td>
</tr>
<tr>
<td>Riddoch et al, 2010</td>
<td>Stroke</td>
<td>74♂ 70♂ 65♂</td>
<td>Decrease</td>
<td>R IPL, R IFG, R angular, R supramarginal gyrus, R parietooccipital regions, L parietotemporal regions, R IFG, R STG, R supramarginal, R angular gyrus</td>
<td>Yes</td>
</tr>
<tr>
<td>Rigoni et al, 2010</td>
<td>Psychopathy</td>
<td>24♀</td>
<td>Decrease</td>
<td>L MFG, L SFG, L lateral temporal cortex, L superior occipital cortex, L SFG, L MFG</td>
<td>Yes</td>
</tr>
<tr>
<td>Muhlau et al, 2009</td>
<td>HD</td>
<td>22 subjects (12%)</td>
<td>Decrease</td>
<td>Bilateral head of caudate, R insula, R temporal pole</td>
<td>UK</td>
</tr>
<tr>
<td>Narvid et al, 2009</td>
<td>FTD + HME</td>
<td>66♂</td>
<td>Decrease</td>
<td>R ventral and medial frontal cortex, R insular cortex</td>
<td>Yes</td>
</tr>
<tr>
<td>Study</td>
<td>Subject</td>
<td>Disease</td>
<td>Number of control subjects</td>
<td>GM increase or decrease</td>
<td>Brain regions</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>------------------------</td>
<td>----------------------------</td>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Tramoni et al, 2009&lt;sup&gt;134&lt;/sup&gt;</td>
<td>34♂</td>
<td>Functional amnesia</td>
<td>25 subjects – (10♀)</td>
<td>–</td>
<td>Bilateral pre- and post-central gyrus, R middle frontal gyrus, L middle frontal gyrus, R SFG, R IFG, R Punctamen, R lenticular nucleus, R parietal lobe</td>
</tr>
<tr>
<td>Bozzalli et al, 2008&lt;sup&gt;135&lt;/sup&gt;</td>
<td>77♀</td>
<td>CBD</td>
<td>8</td>
<td>Decrease</td>
<td>Bilateral frontal and ACC, L parietal gyrus</td>
</tr>
<tr>
<td>Epelbaum et al, 2008&lt;sup&gt;136&lt;/sup&gt;</td>
<td>46♂</td>
<td>Pure alexia, epilepsy</td>
<td>17</td>
<td>Decrease</td>
<td>Bilateral head of body of hippocampus and entorhinal cortex</td>
</tr>
<tr>
<td>Feldman et al, 2008&lt;sup&gt;137&lt;/sup&gt;</td>
<td>54♀</td>
<td>PCA</td>
<td>20</td>
<td>Decrease</td>
<td>R occipital mesial and inferior, R fusiform, R ACC, R SFG, R precuneus, L occipital inferior gyrus</td>
</tr>
<tr>
<td>Adlam et al, 2006&lt;sup&gt;138&lt;/sup&gt;</td>
<td>Six subjects</td>
<td>PPA</td>
<td>12</td>
<td>Decrease</td>
<td>L ventral temporal lobe (three of six subjects), L dorsoventral temporal lobe (two of six subjects), R temporal pole (one of six subjects), R rostroventral temporal lobe (one of six subjects)</td>
</tr>
<tr>
<td>Brazdil et al, 2006&lt;sup&gt;139&lt;/sup&gt;</td>
<td>25♀</td>
<td>Epilepsy</td>
<td>–</td>
<td>Increase</td>
<td>Anterior rim of the L central sulcus</td>
</tr>
<tr>
<td>Cipollotti et al, 2006&lt;sup&gt;140&lt;/sup&gt;</td>
<td>74♂</td>
<td>Hippocampal amnesia</td>
<td>15</td>
<td>Decrease</td>
<td>Bilateral head of body of hippocampus and entorhinal cortex</td>
</tr>
<tr>
<td>Colliot et al, 2006&lt;sup&gt;141&lt;/sup&gt;</td>
<td>27♂</td>
<td>FCD</td>
<td>39</td>
<td>Increase</td>
<td>Regions with FCD lesions (21/27 patients), frontal and temporal extralesional regions (16/27 patients), frontal and temporal extralesional regions (eight of 27 patients)</td>
</tr>
<tr>
<td>Joubert et al, 2006&lt;sup&gt;142&lt;/sup&gt;</td>
<td>61♀</td>
<td>Right FTD</td>
<td>28</td>
<td>Decrease</td>
<td>R inferior temporal gyrus, R hippocampus, R middle temporal gyrus, R caudate, L STG</td>
</tr>
<tr>
<td>Suzuki et al, 2005&lt;sup&gt;143&lt;/sup&gt;</td>
<td>43♀</td>
<td>Schizophrenia</td>
<td>20 (10♀)</td>
<td>Decrease</td>
<td>L IFG and L frontal pole (four patients only)</td>
</tr>
<tr>
<td>Zahn et al, 2005&lt;sup&gt;144&lt;/sup&gt;</td>
<td>Ten subjects (6♀)</td>
<td>AD</td>
<td>10 (5♀)</td>
<td>Decrease</td>
<td>Bilateral anterior lateral temporal (eight of ten AD and five of five PPA); L medial temporal (nine AD and one PPA); R posterior parietal (nine AD and five PPA); R posterior cingulate and R precuneus (nine AD and zero PPA)</td>
</tr>
<tr>
<td>Gorno-Tempini et al, 2004a&lt;sup&gt;145&lt;/sup&gt;</td>
<td>56♀</td>
<td>PPA + CBD</td>
<td>16</td>
<td>Decrease</td>
<td>L IFG, L anterior insula, L SMA, bilateral caudate</td>
</tr>
<tr>
<td>Gorno-Tempini et al, 2004b&lt;sup&gt;146&lt;/sup&gt;</td>
<td>67♀</td>
<td>Right SD</td>
<td>35</td>
<td>Decrease</td>
<td>R anterior temporal lobe, R amygdala/anterior hippocampus, R collateral sulcus, R fusiform gyrus, R posterior insula/STG, L amygdala/anterior hippocampus, L collateral sulcus/fusiform gyrus</td>
</tr>
<tr>
<td>Thompson et al, 2004&lt;sup&gt;147&lt;/sup&gt;</td>
<td>63♀</td>
<td>SD</td>
<td>17</td>
<td>Decrease</td>
<td>L infero lateral temporal lobe, L ACC, L insula, R temporal pole</td>
</tr>
<tr>
<td>Salmond et al, 2003&lt;sup&gt;148&lt;/sup&gt;</td>
<td>14♀</td>
<td>Autism</td>
<td>18 (12♀)</td>
<td>Decrease</td>
<td>Bilateral hippocampus (in seven subjects), amygdala, OFC (in 13 subjects), STG, cerebellum (in eleven subjects)</td>
</tr>
<tr>
<td>Rosen et al, 2002&lt;sup&gt;149&lt;/sup&gt;</td>
<td>54♀</td>
<td>PPA</td>
<td>20 for each Decrease subject</td>
<td>Decrease</td>
<td>L DLPFC, L IFG, bilateral temporoparietal cortex, bilateral IPL, L anterior temporal</td>
</tr>
<tr>
<td>Gitelman et al, 2001&lt;sup&gt;150&lt;/sup&gt;</td>
<td>33♀</td>
<td>Herpes simplex</td>
<td>10 for each Decrease subject</td>
<td>Decrease</td>
<td>Bilateral amygdala, hippocampus, entorhinal cortex, insula, gyrus rectus, nucleus accumbens</td>
</tr>
</tbody>
</table>
Figure 1 Distribution of false positives in parametric and nonparametric single-case VBM.

Notes: In the top panel, the figure represents the percentage of false positives found in each brain regions when using parametric statistics with \( P = 0.05 \) FWE correction: colors reflected the percentage of false positives (adapted from Neuroimage, 70, Scarpazza C, Sartori G, De Simone M, Mechelli A, When the single matters more than the group: very high false positive rates in single case voxel based morphometry, 175–188, 2013 Copyright ©2013, with permission from Elsevier.). In the bottom panel, the figure represents the localization of false positives based on peak coordinates with \( P = 0.05 \) FWE correction (adapted from Scarpazza C, Nichols T, Seramondi D, Maumet C, Sartori G, Mechelli A. When the single matters more than the group (II): addressing the problem of high false positive rates in single case voxel based morphometry using non-parametric statistics. Front Neurosci. 2016;10:6.). colors reflected the different brain regions in which the false positives were found to be located. In both panels, (A) GM increases in the single subject compared with the control group at \( P<0.05 \) corrected and (B) GM decreases in the single subject compared with the control group at \( P<0.05 \) corrected.

Abbreviations: VBM, voxel-based morphometry; FWE, family-wise error; GM, gray matter.
or to longitudinally monitor the clinical course of patients is still unknown.

Forensic perspective

DB life history was full of facts that might indicate brain abnormalities. […] The legal team wanted to make the argument that psychopathy constituted a developmental disorder of emotion, and the new neuroscience might help make their point. Kent Kiehl, The Psychopath Whisperer

VBM has been widely used in order to investigate whether structural brain abnormalities exist within those who are criminals compared to those who are not. For example, Puri et al94 detected significant smaller GM volume in schizophrenic patients who had a history of serious and violent offending, compared with schizophrenic patients without such a history. The differences between the groups were located in the cerebellum and in the supramarginal gyrus, brain regions involved in the verbal working memory.94 Moreover, significant GM reductions in the frontal and temporal brain regions of violent psychopaths compared with nonviolent psychopaths have been found.95 Interestingly, Schiffer et al96 found smaller GM volume in the ventral striatum and in the orbitofrontal cortex of the brains of those with pedophilic tendencies when compared with those of homosexual and heterosexual subjects. Importantly, this latter study has been cited in order to point out that it does not investigate a criminal behavior, rather the sexual orientation of subjects. Indeed, individuals with pedophilic tendencies do not necessarily engage in sexually deviant or criminal acts.

Although there are hundreds of studies investigating the neural basis of criminal and deviant behaviors by means of VBM, the majority are based on group analysis. As for the clinical setting, the translation from group to individual analysis is problematic. In the previous paragraph, we described the statistical issue causing high false-positive rates and the problem of the interpretation of the results. These problems are also present in forensic settings where the most critical issue is how results are interpreted.

In the following paragraphs, we are describing how single-case VBM and the error probability map should be used at individual level when applied to the forensic setting. As a general premise, we would like to emphasize that the aim of this work is not to enter into a forensic debate but only to describe a possible use of VBM in this field. It is important also to underline that we are perfectly aware that VBM results alone, that is, without a clear clinical correlate, do not have any forensic meaning since incompetency or diminished responsibility is considered to be proven if evidence is given of a causal link between a pathological mental state and the criminal behavior. Thus, the criteria to determine responsibility are, and should always be, behavioral. VBM is only helpful in providing objective support to the behavioral findings. Moreover, the use of VBM as a supporting evidence in the courtroom does not presuppose the belief in the existence of a one-to-one correspondence between brain states and behavior. It is a well-established fact in contemporary neuropsychology that brain alterations are mere risk factors for the occurrence of abnormal behaviors.97 For a critical discussion on the use of neuroscientific findings into forensic setting, refer Sartori et al.98

The results of our previous investigation91 could be particularly useful for the assessment of mental insanity in a forensic setting. Such assessment is made complicated by the fact that psychiatric symptoms can be easily faked or exaggerated and that most defendants assessed for mental insanity do not have a previous psychiatric history. Indeed, in psychiatry, the diagnosis is usually possible only by basing it on the symptoms complained of by the patients—and this is a critical difference between psychiatric and neurological disorders. For instance, if a patient complains of memory deficits, neuropsychological tests are available to test the genuineness of the symptom.99 On the other hand, if a patient complains of hallucinations, clinicians have no objective measures to explore whether the patient is actually experiencing hallucinations or if he is simply feigning them. While in the clinical setting usually there is no reason to suspect a malingering of the symptoms, obviously the forensic context is profoundly different. One strategy for dealing with malingering psychiatric disorders consists of validating the reported symptoms with an anatomoclinical correlation.98,100 Importantly, the possibility to quantify the error in single-case VBM makes this technique suitable to be used as an instrument in a forensic setting, although with extreme caution. Indeed, single-case VBM now reflects the Daubert criteria, accepted worldwide for scientific evidence to be admissible in court101 (ie, whether the method is verifiable, whether the method can be falsified, the awareness of the error rate, and acceptance by the scientific experts after a scientific peer review).

The use of VBM in the forensic setting might be traced back to 2009, when the brain of the serial killer DB has been compared with the brain of HCs. In this specific case, the legal team asked an expert to determine whether DB’s brain could be considered compatible with the established clinical diagnosis of psychopathy. Results revealed a pattern of atrophy in DB brain similar to those observed in psychopathic individuals,102–105 that is, the paralimbic system appeared to be smaller in BD compared with HCs. Note that DB scored
in the 99th percentile on the Psychopathy Checklist. Thus, the VBM results supported the clinical findings.\textsuperscript{106}

To date, only one criminal case in which VBM has been used in forensic setting has been published. The defendant was a young woman convicted for a second-degree murder.\textsuperscript{107} The defendant underwent a forensic psychiatric evaluation in order to explore her psychopathological profile. In addition, she also underwent a neuropsychological examination in order to determine if she was able to inhibit her impulses. Her psychopathological profile revealed impulsivity, lack of sensitivity, rejection of conventional standards, and inability to consider alternative solutions to the problems. Moreover, the neuropsychological tests unveiled her inability to inhibit the prominent response, to attribute emotional states to others, and to identify the violation of social norms. Interestingly, both the psychopathological and neuropsychological evaluations provided convergent results suggestive of frontal lobe dysfunc-
tion.\textsuperscript{108–110} Critically, the VBM results revealed two clusters of decreased GM volume in the defendant compared with the control groups, both of them located in the frontal lobe, specifically in the left middle and superior frontal lobe. Let us try to anachronistically interpret these results on the light of the findings of Scarpazza et al\textsuperscript{91} (the 2013 paper publishing the error probability and distribution map was not available in 2010): when investigating the likelihood of a decreased GM volume in the brain of a single subject with the brain of
a control group, the probability to find a false positive in the left middle frontal gyrus ranges from 0.5% to 5%, while in the left superior frontal gyrus ranges from 5.5% to 10%. In Figure 2, to visually compare the results of the murderer with the probability distribution map of false positives has been shown. Although the detected brain abnormalities in the murderer are located in brain regions suffering from a relatively high degree of false positives, the reported results closely mirrored the psychopathological and neuropsychological results, thus decreasing the possibility to be false positives. This case has been reported here in order to highlight which should be the usefulness of VBM in forensic setting: although the use of VBM in court cannot change the rationale underlying the determination of criminal liability, which must be based on a causal association between a mental disorder and a crime, VBM may provide objective, biological data that can be used to reduce controversy in forensic psychiatric evaluations of mental insanity.

As a final note, it is important to underline that it is not possible to reason backward – inferring the presence of a neuropsychological or behavioral deficit from the presence of significant results at VBM: this is called “reverse inference” and is not deductively valid.

Conclusion
This review has highlighted the utility of VBM in the diagnosis (differential and early diagnosis) of both neurological and psychiatric disorders. Moreover, it has focused on the potential clinical translation of this technique and its forensic application. In the last two paragraphs, we have tried to convey the message that VBM results are meaningless if they are not supported by the clinical symptoms complained of by the patient. Thus, the message that should be remembered when applying VBM to the single individual is the following: when a statistically significant result is detected, it should be interpreted in the light of 1) its coherence with the patient’s symptoms and 2) its coherence with the results of previous literature.

As a provocative and final note, we would like to present the VBM results of a single-case comparison: one of the authors’ brain has been compared with the brain of other healthy volunteers. Figure 3 shows GM increases in the author’s brain compared to the control group. In particular, the GM increases are located in the dorsolateral prefrontal cortex.

What does this mean? Using existing literature to interpret the results: I am a binge drinker then! This is a wonderful example showing the risk of “reverse inference” since VBM results alone (ie, not coupled with clinical symptoms) are meaningless. The truth is that I am teetotal. The other author is still laughing.

Disclosure
The authors report no conflicts of interest in this work.

References


