

# Uncovering Capgras delusion using a large-scale medical records database

Vaughan Bell, Caryl Marshall, Zara Kanji, Sam Wilkinson, Peter Halligan and Quinton Deeley

## Background

Capgras delusion is scientifically important but most commonly reported as single case studies. Studies analysing large clinical records databases focus on common disorders but none have investigated rare syndromes.

## Aims

Identify cases of Capgras delusion and associated psychopathology, demographics, cognitive function and neuropathology in light of existing models.

## Method

Combined computational data extraction and qualitative classification using 250 000 case records from South London and Maudsley Clinical Record Interactive Search (CRIS) database.

## Results

We identified 84 individuals and extracted diagnosis-matched comparison groups. Capgras was not ‘monothematic’ in the majority of cases. Most cases involved misidentified family members or close partners but others were misidentified in 25% of cases, contrary to dual-route face recognition models.

Neuroimaging provided no evidence for predominantly right hemisphere damage. Individuals were ethnically diverse with a range of psychosis spectrum diagnoses.

## Conclusions

Capgras is more diverse than current models assume. Identification of rare syndromes complements existing ‘big data’ approaches in psychiatry.

## Declaration of interests

V.B. is supported by a Wellcome Trust Seed Award in Science (200589/Z/16/Z) and the UCLH NIHR Biomedical Research Centre. S.W. is supported by a Wellcome Trust Strategic Award (WT098455MA). Q.D. has received a grant from King’s Health Partners.

## Copyright and usage

© The Royal College of Psychiatrists 2017. This is an open access article distributed under the terms of the Creative Commons Non-Commercial, No Derivatives (CC BY-NC-ND) license.

Capgras delusion is the delusional belief that someone, often a family member or close relative, has been replaced by similar-looking impostor or double.<sup>1–3</sup> Although considered rare, cases of Capgras delusion have played an important role in the history and development of the science of psychopathology. It was one of the founding areas of research in cognitive neuropsychiatry<sup>4</sup> and continues to be conceptually important in the development of cognitive theories of delusions.<sup>5–7</sup> Anonymised data from electronic health records have been used to conduct large-scale psychiatric studies on common disorders, allowing associations to be observed in very large data-sets. However, these data-sets would also allow for a different approach: the identification of relatively large series of very rare disorders, along with a rich set of clinical data to elucidate psychopathology and its neuropsychological and neuropathological associations. The use of ‘big data’ to identify rare cases has been discussed in cancer genomics,<sup>8</sup> but it has not yet been applied to psychiatry.

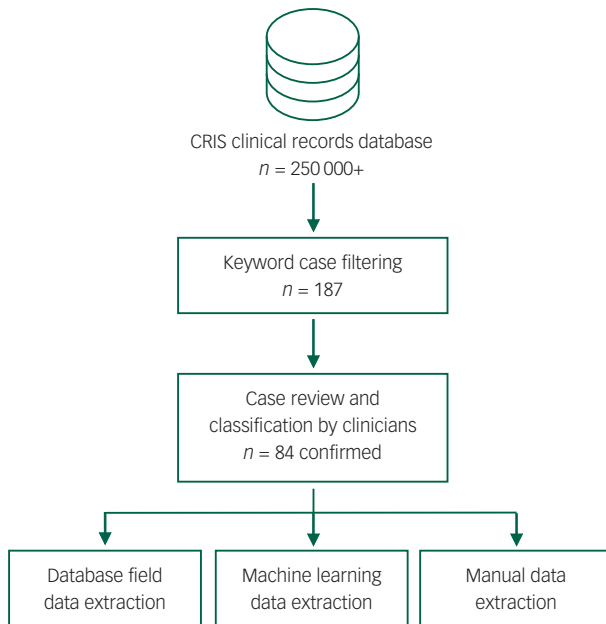
One of the challenges of using medical record databases for cognitive neuropsychiatry is that many rare syndromes, like Capgras, do not have a specific diagnostic code and are usually documented in narrative written notes. Therefore, a combined approach is needed that involves manual classification combined with methods of computational data extraction. This is the approach taken by this study, where we used a high-specificity search strategy to identify cases in a large anonymised database of clinical records, which were then analysed to examine demographics, psychopathology, cognitive function and neuropathology associated with the presence of Capgras delusion. Our aim was to examine assumptions about the condition that have largely been based on single-case studies and explore their implications for theories of delusion formation, both as an investigation of the delusion itself and as a test case for the use of large clinical data-sets to identify rare psychiatric syndromes.

## Method

The study used the Clinical Record Interactive Search (CRIS) system, an electronic research database of anonymised clinical records from the South London and Maudsley NHS Foundation Trust (SLaM) – the public service provider for secondary and tertiary mental health services for four London boroughs (Lambeth, Southwark, Lewisham and Croydon) with a total population of about 1.2 million residents. The CRIS system contains anonymised health records for over 250 000 people starting in 2008. Over this time period, SLaM has provided almost the entire range of public mental health services for the population across the lifespan and across presentation types as well as a range of local and national specialist services. Technical details of the CRIS system are reported in Stewart *et al.*<sup>9</sup> Details of the CRIS patient cohort used for this study, including statistics for demographics and diagnoses in the sample, are reported in Perera *et al.*<sup>10</sup> Ethical approval as an anonymised database for secondary analysis was originally granted in 2008, and renewed for a further 5 years in 2013 (Oxford C Research Ethics Committee, reference 08/H0606/71+5). Data extraction procedures and stages are illustrated in Fig. 1.

## Case selection

Preliminary use of the keyword ‘Capgras’ indicated it had sufficient specificity to identify candidate records for manual classification based on an informal inspection (‘eyeballing’) of a selection of retrieved case records. The keyword ‘Capgras’ was then used in a formal search of free-text record fields, retrieving records from 187 patients. Previous case-identification studies have tended to select cases based on informal labels used by consulting clinicians,<sup>11</sup> or single-rater review of case notes using a brief description of the delusion (see Tamam *et al.*<sup>12</sup>). In this study, the 187 anonymised records were subsequently independently



**Fig. 1** Capgras case data extraction procedure. CRIS, Clinical Records Interactive Search.

rated by a psychiatrist and clinical psychologist for the likely presence of Capgras delusion using the structured classification system described in Fig. 2. The features of Capgras delusion were based on standard definitions in the literature<sup>1-3</sup> and were drawn up to avoid misclassification of a closely related delusional misidentification syndrome (Fregoli delusion<sup>13</sup>) to ensure Capgras would be specifically identified.

Cohen's kappa was used to assess the level of independent agreement between raters, with  $\kappa=0.749$  indicating an acceptable level of agreement. After independent rating, disagreements were resolved by discussion between raters. Consequently, 20 cases were classified as not Capgras, 20 cases as past Capgras only, 63 cases as Capgras possibly present and 84 cases were classified as strongly indicating the presence of Capgras delusion. Only the last cases were retained for subsequent analysis. The date of the record from which raters first identified strong evidence of Capgras for each case was noted (henceforth referred to as the date of 'Case ID') and used for subsequent analyses. Data extraction for comparison groups is described in the relevant sections below.

### Data extraction

Data for demographics and primary diagnosis were extracted directly from specific database fields. Symptom phenomenology and service context were collaboratively extracted by the two raters from written records. Data for neuroimaging assessment

results (computed tomography (CT), magnetic resonance imaging (MRI) and electroencephalography (EEG)) were extracted from reports arising from the assessments (original images of scans and/or EEG charts were not available). Data for antipsychotic medication and Mini-Mental State Examination (MMSE) score for identified cases were extracted through the use of specific CRIS text-mining applications built using the General Architecture for Text Engineering machine-learning framework based on natural language-processing techniques, a system described in Perera *et al.*<sup>10</sup> At the time of data extraction, the medication extraction application had a precision (equivalent to specificity) of 0.94 and a recall (sensitivity) of 0.57, and the MMSE application had a precision of 0.91 and a recall of 0.88 (M. Broadbent, 2016, personal communication). Comparison samples for MMSE and neuroimaging results were derived *post hoc* to match the Capgras sample and are detailed in the relevant sections below. Statistics were calculated using R version 3.3.2.

## Results

### Demographics

Cases consisted of 56 females and 28 males. The mean age at date of Case ID date was 44.6 years (s.d.=19.4, range 14.7–94.8).

The sample was ethnically diverse. Ethnicity was classified as British ( $n=36$ ), African ( $n=16$ ), any other White background ( $n=8$ ), any other Black background ( $n=6$ ), Caribbean ( $n=6$ ), any other ethnic group ( $n=5$ ), any other Asian background ( $n=2$ ), Pakistani ( $n=2$ ), any other mixed background ( $n=1$ ), Chinese ( $n=1$ ) and Indian ( $n=1$ ). Out of the total cases, 50 did not have information for country of origin but for the 34 that did they were classified as the UK ( $n=23$ ), Nigeria ( $n=3$ ), Cyprus ( $n=2$ ), Ghana ( $n=2$ ), Italy ( $n=2$ ), Jamaica ( $n=2$ ) and Yemen ( $n=2$ ). One case from each of the following countries of origin was recorded: Brazil, Congo, Cote d'Ivoire, Croatia, England, France, Hong Kong, India, Pakistan, Portugal, Romania, Sierra Leone, Somalia and Sri Lanka.

### Service context

Out of the 84 cases, 16 individuals were presenting to mental health services for the first time at the date of Case ID, 60 had a history of contact with mental health services and in the remaining 8 cases past contact was not clear from the retrieved information. Case ID was based on a record where a mental state examination was reported by a psychiatrist or trainee psychiatrist in 54 cases.

### Diagnosis

Of the 84 cases, 69 individuals had existing diagnoses at the time of Case ID (Table 1). Considering previously reported associations between Capgras delusion and dementia,<sup>14</sup> we examined the number of patients with Capgras at Case ID converted to dementia from other diagnoses. Over the total recorded contact with mental

<b>Strongly:</b> Capgras is mentioned as a present delusion plus evidence of present or recent relevant delusional misidentification of people is described (mention of people being replaced, or impostors, or lookalikes, or identical looking people, or clones, or robots etc).
<b>Possibly:</b> Capgras is mentioned as a present delusion but no additional description of delusion content is given, or delusional nature is questioned, or the description is clearly not person misidentification.
<b>Not present:</b> Capgras is excluded, or mentioned erroneously, or conflicts with the description of the delusion (clearly not misidentification).
<b>Only past:</b> Capgras delusion is only mentioned as previously present with no evidence of current misidentification, or is described as fully resolved.

**Fig. 2** Categories and definitions for case note classification used by independent raters.

**Table 1** Diagnoses at date of case identification

ICD-10 code	Diagnosis	n
F20	Schizophrenia or paranoid schizophrenia	23
F28/F29	Other or unspecified nonorganic psychotic disorders	8
F99	Mental disorder, not otherwise specified	7
F32	Depressive episode	6
F00	Dementia in Alzheimer's disease	4
F23	Acute and transient psychotic disorders	4
F31	Bipolar affective disorder	4
F22	Delusional disorder	3
F25	Schizoaffective disorders	3
Z71.1	Person with feared complaint in whom no diagnosis is made	2
F05.0	Delirium not superimposed on dementia, so described	1
F11.1	Mental and behavioural disorders because of use of opioids; opioid abuse	1
F19.8	Mental and behavioural disorders because of multiple-drug use and use of other psychoactive substances: Other mental and behavioural disorders	1
F41.1	Generalised anxiety disorder	1
F43.1	Post-traumatic stress disorder	1
-	No diagnosis	15

health services, nine individuals received a diagnosis of dementia, four held a dementia diagnosis at time of Case ID, and five converted to a primary diagnosis of dementia from a primary diagnosis of depressive disorder ( $n=2$ ), schizophrenia ( $n=1$ ), delirium not superimposed on dementia ( $n=1$ ) and mental disorders not otherwise specified ( $n=1$ ).

In light of similar debates about the relationship between Capgras and the diagnoses of schizophrenia, delusional disorder and organic psychosis,<sup>15</sup> we examined the transition to these diagnoses in the sample.

Although only 23 patients had a diagnosis of schizophrenia or paranoid schizophrenia at the time of Case ID, 39 received a schizophrenia or paranoid schizophrenia diagnosis during their total recorded contact with mental health services. Last recorded diagnosis for this cohort of 39 included schizophrenia or paranoid schizophrenia for 31 individuals, schizoaffective disorder in 3 cases, Alzheimer's disease in 1 case, 'unspecified mental disorder because of brain damage and dysfunction and to physical disease' in 1 case, 'mental disorder not otherwise specified' in 1 case, and 'person with feared complaint in whom no diagnosis' made in 1 case.

In total, six individuals had a diagnosis of delusional disorder during their total recorded contact with mental health services (three at date of Case ID). Final diagnoses for this cohort included delusional disorder ( $n=4$ ) and schizophrenia/paranoid schizophrenia ( $n=2$ ).

There were no individuals with a primary diagnosis of any ICD-10 F06 category organic mental disorder at date of Case ID and only one that had such a diagnosis during their total recorded contact with mental health services. This patient's last recorded primary diagnosis was a moderate depressive episode.

### Symptom phenomenology

Details of the Capgras delusion were extracted from the text of the records and were categorised collaboratively by two raters. Totals sum to more than the total number of cases because of some patients having more than one delusion or a belief that could be classified in various ways.

The individual of the delusional replacement was a family member or close partner in 71 out of 84 cases. Non-family or non-partner misidentification was reported in 21 cases and included

a care professional or professionals (doctor, care staff) in 9 cases, a friend or friends in 6 cases, an acquaintance or strangers in 3 cases, the police in 2 cases, a solicitor in 1 case, and unknown or not specified in 4 cases (the individual was referred to but not specified in the text or the individual was referred to in general terms – e.g. 'people'). Twelve cases had non-family and non-partner misidentification only and nine had a combination of family and non-family or partner misidentification. In addition, three of the identified cases involved beliefs that the patients themselves had personally been replaced. The delusional belief was described as involving the replacement of more than one person in 33 cases.

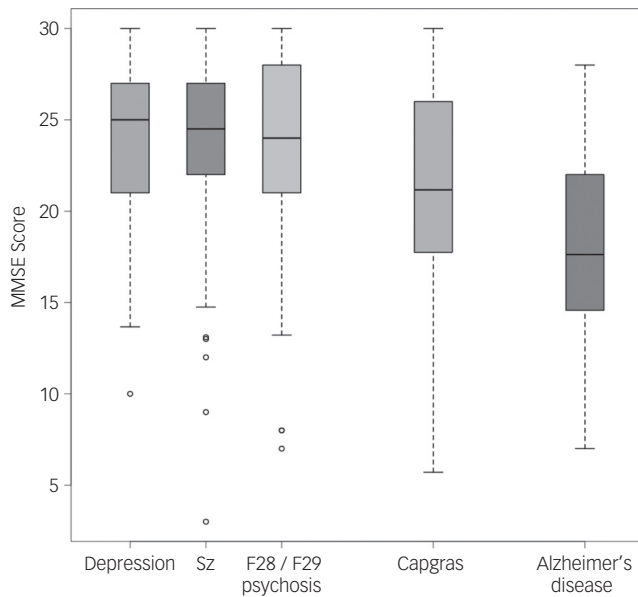
The identity of the 'replacer' was described in terms of a specific agent or agents in 10 cases (e.g. a 'swami', 'someone called (male name)', 'police officers'). Family members featured as the described 'replacer' in three cases ('father and father's girlfriend', 'uncle', 'brother') and a supernatural agent featured as a replacer in one case ('jinn').

In 59 cases, there was no record of patient-reported justification for the belief. In 10 cases, physical appearance was cited and these ranged from seeing an 'identical-looking person in an unexpected location', 'appearance had changed (eyes, improved skin, body shape)', a lack of resemblance to a spouse despite taking his or her place in the home, looking similar but being of smaller body type, the patient's children looking older than expected, and an externally visible eye problem in the individual that was perceived as being on the 'wrong side' of the face. In seven cases, what were likely to be delusional explanations were reported (e.g. another person inside the patient informed them of the impostors, the patient met her dead sister in a local shop, police are being covered by a membrane that disguises them). In six cases, altered behaviour or personality was reported as the justification (e.g. being told to 'do this do that' which is out of character for the person who was 'replaced', 'replaced' person had a 'cold attitude', frequent arguing with the 'replaced' person). In four cases, alterations in the sound of the individual's voice were reported as justification (e.g. voice sounded fake on the phone, voice sounded different). In two cases, likely auditory hallucinations or thought insertion were cited (receiving messages from Interpol informing the patient of the impostors, information from external auditory hallucinations). In one case, a 'sense' that the person was different was reported, in one case 'hearsay' was given as justification, and in one case a likely visual hallucination of seeing the person's face in other faces was reported.

Of the 84 cases of likely Capgras delusion, 61 were reported as having other non-Capgras-related delusions whereas the remaining 23 had only Capgras-themed delusions reported. Twenty-four out of the 84 cases were reported as experiencing auditory hallucinations at time of Case ID.

### Cognition

Scores from the MMSE<sup>16</sup> were available for 20 cases with a total of 70 data points. Mean MMSE scores were calculated for each patient by creating a patient average if there was more than one MMSE score on record, and then calculating the mean for the sample. Mean MMSE score for Capgras cases was 21.28 (s.d.=6.12). As a comparison, 50 patients with MMSE scores were randomly selected from the CRIS database for each of the following diagnoses: depression (MMSE mean 23.87; s.d.=4.59), schizophrenia (MMSE mean 23.09; s.d.=5.75), F28/F29 other or unspecified nonorganic psychosis (MMSE mean 23.48; s.d.=5.73), dementia in Alzheimer's disease (MMSE mean 18.95; s.d.=6.62). Score distributions for the Capgras cases and comparison samples are presented in Fig. 3.



**Fig. 3** MMSE score box plot for Capgras cases and comparison samples. Sz, schizophrenia; F28/F29 psychosis; other/unspecified nonorganic psychotic disorders.

### Neuroimaging assessment results

Of the 84 Capgras cases, 40 had results from neuroimaging assessments available. Records for a comparison sample ( $n=84$ ) were extracted on the basis of primary diagnosis matching where an equivalent number of cases for each diagnosis in the Capgras sample were randomly matched from cases in the clinical records database with the same primary diagnosis. Fifteen Capgras cases had no recorded diagnosis at the time of Case ID. We matched these 'no diagnosis' patients with additional randomly selected cases of people diagnosed with schizophrenia rather than other 'no diagnosis' patients to ensure consistent matching with a psychotic disorder diagnosis.

Of the 40 cases in the Capgras sample with any available neuroimaging assessment results, 26 had CT scans, 16 had MRI scans, and 8 had EEGs. Regardless of the availability of assessment results, 54 had a neuroimaging assessment recommended. For the comparison sample, 22 cases had any neuroimaging results available (12 CT, 7 MRI, 3 EEG) and 37 had a neuroimaging assessment recommended. When tested with  $\chi^2$  using Monte Carlo permutation, Capgras individuals were significantly more likely to have neuroimaging recommended ( $\chi^2=6.929$ ,  $P=0.0138$ ) and more likely to have neuroimaging results available ( $\chi^2=10.37$ ,  $P=0.0031$ ). We compared the frequency of abnormal results in Capgras and comparison samples for each imaging type: CT (Capgras=11, comparison=4), MRI (Capgras=2, comparison=2) and EEG (Capgras=2, comparison=0). When tested with  $\chi^2$  using Monte Carlo permutation over the whole sample, abnormal results were not significantly more common in the Capgras sample than the comparison sample for any imaging modality. Of the 14 Capgras cases with abnormal neuroimaging assessments, right-hemisphere pathology was the least common: 9 patients had diffuse pathological changes, 7 had left-sided changes, 3 had right-sided changes and 1 had a reported result with unclear lateralisation. Two patients had ambiguously reported results. Tables detailing the abnormal findings in the Capgras and comparison samples are presented in the Data supplement (Tables DS1 and DS2).

### Antipsychotic prescribing

Out of the 84 patients, 66 had a recorded antipsychotic prescription history. The most commonly prescribed antipsychotic medications were olanzapine (40 cases), risperidone (37 cases), aripiprazole (24 cases), quetiapine (15 cases), haloperidol (13 cases), zuclopenthixol (9 cases), amisulpride (7 cases), paliperidone (7 cases), flupenthixol (6 cases), clozapine (5 cases), fluphenazine (4 cases), pipotiazine (4 cases), trifluoperazine (4 cases), chlorpromazine (3 cases) and sulpiride (2 cases).

### Discussion

Using a combination of computational data extraction and qualitative classification, we searched over 250 000 case records and identified 84 cases of Capgras delusion and extracted delusion characteristics, demographics, MMSE results and neuropathological findings, finding that Capgras cases as described in clinical records are more diverse than models of the Capgras delusion assume.

### Implications for cognitive models and syndromal definition

The most common subject of the misidentification was a family member or partner but in a significant minority of cases the subjects included friends, acquaintances, care professionals or strangers. There are no diagnostic criteria or official definitions of Capgras delusion, but it is typically described as involving a belief that a close relative or family member has been replaced by an identical or near-identical-looking impostor.<sup>1-3</sup>

From the perspective of cognitive models of Capgras that explain the delusion in terms of an inference based on the lack of an expected emotional response from the face of a familiar or close individual ('she looks like my mother, but doesn't feel like her')<sup>17-19</sup> these misidentifications involving mere acquaintances or strangers are hard to explain. This suggests two alternatives: either that 'familiar-person Capgras' and 'non-familiar person Capgras' are distinct syndromes that might need to be explained via neuropsychologically distinct causal pathways, or that cognitive models that don't rely on the familiar or non-familiar distinction may more accurately capture the full range of clinical presentations (see Wilkinson<sup>20,21</sup>). The extent to which different presentations should equally be considered 'Capgras' given its traditional definition is perhaps an argument over semantics. However, we note in this case that the debate could be informed by neuropsychological testing: a putative 'non-familiar person Capgras' might be a stronger candidate as a distinct syndrome if affected people show an unconscious recognition response for familiar faces that is apparently missing in traditionally defined Capgras, assuming that the original studies<sup>17-19</sup> are sound.

Notably, Capgras was most commonly present in the context of other delusions, contrary to its typical characterisation as a 'monothematic delusion'.<sup>22</sup> It was also identified in people with the full range of psychotic disorder diagnoses as well as dementia diagnoses and patients who had no pre-existing diagnoses. In terms of demographics, the typical profile was middle-aged, previous contact with mental health services and more likely to be female by a ratio of 2:1. Patients gave a range of self-report justifications for their delusions. The marked ethnic and cultural diversity in the sample reflects the high levels of ethnic diversity in the catchment area but also raises the question as to whether Capgras delusion is a fundamental aspect of psychosis shared widely across cultures. This is not to say that Capgras delusion is 'culturally impenetrable'. Analysis of the symptom phenomenology showed clear cultural influence (with *jinn* and a religious



leader featuring). However, the evidence here tentatively suggests that the core features of misidentification may well be best accounted for in terms of universal cognitive mechanisms.

### Neurocognition

Cognition in the Capgras group fell in the mid-point between the depression, schizophrenia and unspecified non-organic psychoses comparison samples, and the Alzheimer's comparison sample, although there was a marked overlap in performance between all groups and no specific conclusions can be drawn from these data. Additionally, the MMSE is a relatively simple 'bedside' cognitive screening test and was only available for a limited number of cases.

In terms of neuroimaging, clinically abnormal results were more common in Capgras cases than the diagnosis-matched comparison sample, although this was non-significant when tested statistically. Furthermore, the type of abnormalities reported in the Capgras cases were diagnosable but typically minor or likely-incidental changes which are typical for clinical assessments in psychosis<sup>23</sup> with only a small number of discrete lesions. Clinical definitions of Capgras often include right-hemisphere pathology<sup>1,2</sup> or suggest it is strongly linked or even indicative of right-hemisphere neuropathology,<sup>24,25</sup> although this presentation was actually the least prevalent in the sample reported here.

In interpreting the significance of assessment results taken from clinical records, one concern is confounding by indication where, in the case of assessments, those who present with risk factors for a pathology are more likely to receive assessments confirming the pathology, potentially biasing prevalence estimates. In this instance, individuals who present with indicators of neuropathology might be more likely to receive neuroimaging assessments. Indeed, we confirmed that the presence of Capgras delusion predicted both recommendations for neuroimaging and the presence of neuroimaging results when compared with a diagnosis-matched control group, suggesting that Capgras is being considered as a potential indicator of neuropathology by clinicians, in line with implications in the literature. However, confounding by indication should lead to an over-detection of neuropathology and yet there was still no reliable evidence for higher rates of neuropathology in the Capgras group compared with a diagnosis-matched control group, and no evidence for a higher frequency of right-hemisphere lesions.

Nevertheless, in a recent review of 61 cases of lesion-related Capgras, right-lateralised neuropathology predominated.<sup>26</sup> The distinction between lesion and non-lesion-related Capgras is not consistently made in the literature and the results reported here suggest that for individuals who present to psychiatric services, the delusion is unlikely to be a reliable indicator of gross neuropathology on its own. It is worth noting that existing research has still yet to adequately compare lesion and non-lesion-related Capgras in terms of presenting features, so it is not clear to what extent they represent different aetiologies underlying a similar psychopathology, or to what extent Capgras delusion in the context of clear neuropathology has a distinct phenomenological presentation.

From a clinical perspective, it is notable that, at least within the sample presented here, neuroimaging seems to be over-recommended whereas the MMSE, a brief basic cognitive test typically used in screening for dementia, was frequently used. It is possible that the motivation for the use of this form of cognitive screening is the belief that Capgras delusion 'necessarily indicates an organic syndrome' (for which we find no evidence here), or perhaps is motivated by knowledge of prior evidence that misidentification delusions correlate with MMSE score in Alzheimer's disease.<sup>27</sup> However, there is no evidence from our sample that MMSE score differentiates between Capgras and

comparison groups, although this can only be adequately tested through a formal study where collection of cognitive data is not based on clinical judgement.

### Methodological implications

Methodologically, this study illustrates the potential contribution of 'mining' electronic health records for the identification of large case series of presumably rare conditions, their associated clinical features and factors relevant to causal models. A previous review of lesion-related Capgras delusion reported a modal number of cases per publication of 1,<sup>26</sup> and we assume from other published cases that the same is true for the Capgras literature as a whole. Considering that research on rare conditions has been influential in cognitive neuropsychiatry, which aims to better understand typical neuropsychological processes through studying altered functioning, we suggest that the use of large medical records databases may be an important complement to existing neuropsychiatry research.

The approach taken here – first filtering by keyword, then submitting those records to case identification by clinicians using structured criteria, and then using database field and machine-learning data extraction on identified cases – aimed for high-specificity case identification. This allowed the efficient identification of a large case series when compared to the existing scientific literature but did not control false negative rate in terms of the medical records database being studied, and therefore cannot give an estimate of prevalence. For example, there may have been many more cases of Capgras delusion that could be clearly identified based on case description but which were not recorded using the word 'Capgras'. Indeed, high rates of false negatives in this study are suggested by previous studies on delusions in Alzheimer's disease that have reported relatively high rates of misidentification delusions (34.2%<sup>28</sup>) and Capgras delusion (10%<sup>14</sup>), suggesting many more cases are likely to be present than have been detected by our methods, particularly given the fact that over 2000 dementia patients have been reported in previous studies using the same clinical records database that was used in this study.<sup>29</sup>

Our method for identifying cases took advantage of the fact that Capgras is relatively well known, largely because of having an impact on the science of psychopathology beyond its typical prevalence in the clinic, meaning it is more likely to be identified and recorded as 'Capgras' in the medical notes by clinicians. However, this approach may not be as effective for other syndromes that may be scientifically interesting but less likely to be described by clinicians using specific searchable text.

One potential solution to the problems of estimating prevalence and identifying lesser-known rare psychiatric syndromes would be to identify cases using clinician-assisted case identification to (i) create a training set of genuine cases and then (ii) train machine-learning text classification algorithms to identify further cases in the wider database. Supervised learning algorithms such as support vector machines have already proved efficient in identifying further cases when trained on pre-classified text from written notes<sup>30</sup> and biomedical data<sup>31</sup> and these techniques have the potential to form the basis of epidemiological studies or case identification for further investigation.

### Limitations

We are aware that the data included in this study were obtained from clinical notes and rely on recording by clinicians. Here, simple error or more frequently clinician's bias may play an important role in terms of what is recorded, or indeed, asked about in clinical interviews, and we are aware that there may be additional aspects of the experience, phenomenology or medical history that were simply not recorded. With regard to computational data extraction, the natural language processing algorithm

for extracting MMSE data had a high precision and recall but the algorithm for extracting antipsychotic prescription had poor recall, meaning antipsychotic prescription was likely underestimated. As noted previously, results from cognitive and neuroimaging assessments may have been subject to confounding by indication. This is less a concern in reference to hypotheses that Capgras is indicative of gross right-hemisphere pathology, where we are looking to disconfirm the hypothesis and confounding by indication would lead to over-detection, but it does mean that broader generalisations from the neuroimaging data outside this specific hypothesis must involve significant caveats.

It is also worth noting that CT scans were more common than MRI scans in both the case and comparison samples. Although CT scans have a high sensitivity for space-occupying lesions, bleeds, strokes and malformations, they do not, for example, have the same sensitivity as MRI scans for contusions, shearing injuries and subdural or epidural haematomas,<sup>32</sup> meaning some lesions may have been missed by CT assessment. Nevertheless, we have no reason to suspect that these differences biased the results in terms of the detection of lateralised neuropathology. However, neither scan type was able to detect functional brain differences above and beyond specific lesions, which have been reported in previous Capgras studies that used functional magnetic resonance imaging (fMRI)<sup>33</sup> or single-photon emission computed tomography (SPECT).<sup>34</sup> EEG was the only imaging type reported here that could potentially detect functional alterations and was the least common in our sample.

In sum, we report the use of large medical records database for rare syndrome identification in psychiatry and applied this to Capgras delusion, showing clinical features to be more diverse than has previously been reported in the literature, suggesting that existing cognitive models, important for wider theories of delusion formation, must account for a wider range of presentation and causal pathways. It also seems that portrayals of Capgras as a syndrome of 'right-hemisphere organic dysfunction' in psychiatry are inaccurate given that this presentation was rare, and additional clarity and detail is needed in scientific descriptions of the syndrome to account for this diversity.

One additional implication of this study is that descriptive psychopathology and carefully recorded phenomenological interviewing beyond diagnosis is more, not less relevant, in the age of 'big data' analytics and electronic health records, as it has the potential to identify rare and important psychiatric syndromes – either as they are currently defined or as they could be discovered amid the data.

**Vaughan Bell**, PhD DClinPsy, Division of Psychiatry, University College London, London, UK; South London and Maudsley NHS Foundation Trust, London, UK; **Caryl Marshall**, MBBS MRCPsych, Lewisham Mental Health Learning Disabilities Team, Behavioural & Developmental, Psychiatry Clinical Academic Group, South London and Maudsley NHS Foundation Trust, London, UK; **Zara Kanji**, MSc, Psychological Interventions Clinic for Outpatients with Psychosis, Maudsley Psychology Centre, Maudsley Hospital, London, UK; **Sam Wilkinson**, PhD, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK; **Peter Halligan**, PhD DSc, School of Psychology, Cardiff University, Cardiff, UK; **Quinton Deeley**, PhD, MRCPsych, Cultural and Social Neuroscience Research Group, Institute of Psychiatry, Psychology and Neuroscience, London, UK

**Correspondence:** Vaughan Bell, Division of Psychiatry, University College London, 6th Floor, Maple House, 149 Tottenham Court Road, London W1T 7NF, UK. E-mail: Vaughan.Bell@ucl.ac.uk

First received 3 May 2017, final revision 26 Jun 2017, accepted 27 Jun 2017

## References

- Christodoulou GN, Margariti M, Kontaxakis VP, Christodoulou NG. The delusional misidentification syndromes: strange, fascinating, and instructive. *Curr Psychiatry Rep* 2009; **11**: 185–9.
- Edelstyn NMJ, Oyebo F. A review of the phenomenology and cognitive neuropsychological origins of the Capgras syndrome. *Int J Geriatr Psychiatry* 1999; **14**: 48–59.
- Enoch D, Ball H. Capgras delusion. In *Uncommon Psychiatric Syndromes* (4th ed.): 1–19. Arnold, 2001.
- Halligan PW, David AS. Cognitive neuropsychiatry: towards a scientific psychopathology. *Nat Rev Neurosci* 2001; **2**: 209–15.
- Fineberg SK, Corlett PR. The doxastic shear pin: delusions as errors of learning and memory. *Cogn Neuropsychiatry* 2016; **21**: 73–89.
- Bortolotti L. In defence of modest doxasticism about delusions. *Neuroethics* 2012; **5**: 39–53.
- Coltheart M, Langdon R, McKay R. Schizophrenia and monothematic delusions. *Schizophr Bull* 2007; **33**: 642–7.
- Bernstein R. Shifting ground for big data researchers. *Cell* 2014; **157**: 283–4.
- Stewart R, Soremekun M, Perera G, Broadbent M, Callard F, Denis M, et al. The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLAM BRC) case register: development and descriptive data. *BMC Psychiatry* 2009; **9**: 51.
- Perera G, Broadbent M, Callard F, Chang CK, Downs J, Dutta R, et al. Cohort profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLAM BRC) Case Register: current status and recent enhancement of an Electronic Mental Health Record-derived data resource. *BMJ Open* 2016; **6**: e008721.
- Joseph AB. Observations on the epidemiology of the delusional misidentification syndromes in the Boston metropolitan area: April 1983–June 1984. *Psychopathology* 1994; **27**: 150–3.
- Tamam L, Karatas G, Zeren T, Ozpoyraz N. The prevalence of Capgras syndrome in a university hospital setting. *Acta Neuropsychiatr* 2003; **15**: 290–5.
- Langdon R, Connaughton E, Coltheart M. The Fregoli delusion: a disorder of person identification and tracking. *Top Cogn Sci* 2014; **6**: 615–31.
- Harwood DG, Barker WW, Ownby RL, Duara R. Prevalence and correlates of Capgras syndrome in Alzheimer's disease. *Int J Geriatr Psychiatry* 1999; **14**: 415–20.
- Atta K, Forlenza N, Gujski M, Hashmi S, Isaac G. Delusional misidentification syndromes: separate disorders or unusual presentations of existing DSM-IV categories? *Psychiatry (Edgmont)* 2006; **3**: 56–61.
- Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; **12**: 189–98.
- Ellis HD, Young AW, Quayle AH, Pauw KWD. Reduced autonomic responses to faces in Capgras delusion. *Proc Biol Sci* 1997; **264**: 1085–92.
- Hirstein W, Ramachandran VS. Capgras syndrome: a novel probe for understanding the neural representation of the identity and familiarity of persons. *Proc Biol Sci* 1997; **264**: 437–44.
- Brighetti G, Bonifacci P, Borlimi R, Ottaviani C. 'Far from the heart far from the eye': evidence from the Capgras delusion. *Cogn Neuropsychiatry* 2007; **12**: 189–97.
- Wilkinson S. Delusions, dreams and the nature of identification. *Philos Psychol* 2015; **28**: 203–26.
- Wilkinson S. A mental files approach to delusional misidentification. *Rev Philos Psychol* 2016; **7**: 389–404.
- Davies M, Coltheart M, Langdon R, Breen N. Monothematic delusions: toward a two-factor account. *Philos Psychiatry Psychol* 2001; **8**: 133–58.
- Goulet K, Deschamps B, Evoy F, Trudel JF. Use of brain imaging (computed tomography and magnetic resonance imaging) in first-episode psychosis: review and retrospective study. *Can J Psychiatry* 2009; **54**: 493–501.
- Alexander MP, Stuss DT, Benson DF. Capgras syndrome: a reduplicative phenomenon. *Neurology* 1979; **29**: 334–9.
- Feinberg TE. Neuropathologies of the self: clinical and anatomical features. *Conscious Cogn* 2011; **20**: 75–81.
- Darby R, Prasad S. Lesion-related delusional misidentification syndromes: a comprehensive review of reported cases. *J Neuropsychiatry Clin Neurosci* 2016; **28**: 217–22.
- Wilkosz PA, Miyahara S, Lopez OL, Dekosky ST, Sweet RA. Prediction of psychosis onset in Alzheimer disease: the role of cognitive impairment, depressive symptoms, and further evidence for psychosis subtypes. *Am J Geriatr Psychiatry* 2006; **14**: 352–60.

## Acknowledgements

The authors thank Megan Pritchard for her advice and help with the CRIS system during this study.

- 28** Perini G, Carlini A, Pomati S, Alberoni M, Mariani C, Nemni R, Farina E. Misidentification delusions: prevalence in different types of dementia and validation of a structured questionnaire. *Alzheimer Dis Assoc Disord* 2016; **30**: 331–7.
- 29** Perera G, Khondoker M, Broadbent M, Breen G, Stewart R. Factors associated with response to acetylcholinesterase inhibition in dementia: a cohort study from a secondary mental health care case register in London. *PLoS One* 2014; **9**: e109484.
- 30** Wright A, McCoy AB, Henkin S, Kale A, Sittig DF. Use of a support vector machine for categorizing free-text notes: assessment of accuracy across two institutions. *J Am Med Inform Assoc* 2013; **20**: 887–90.
- 31** Khondoker M, Dobson R, Skirrow C, Simmons A, Stahl D. A comparison of machine learning methods for classification using simulation with multiple real data examples from mental health studies. *Stat Methods Med Res* 2016; **25**: 1804–23.
- 32** Orrison WW, Gentry LR, Stimac GK, Tarrel RM, Espinosa MC, Cobb LC. Blinded comparison of cranial CT and MR in closed head injury evaluation. *AJNR Am J Neuroradiol* 1994; **15**: 351–6.
- 33** Thiel CM, Studte S, Hildebrandt H, Huster R, Weerda R. When a loved one feels unfamiliar: a case study on the neural basis of Capgras delusion. *Cortex* 2014; **52**: 75–85.
- 34** Horikawa H, Monji A, Sasaki M, Maekawa T, Onitsuka T, Nitazaka Y, et al. Different SPECT findings before and after Capgras' syndrome in interictal psychosis. *Epilepsy Behav* 2006; **9**: 189–92.

