This is a repository copy of An Interactional Profile to Assist the Differential Diagnosis of Neurodegenerative and Functional Memory Disorders.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/126490/

Version: Accepted Version

Article:
Reuber, M., Blackburn, D.J., Elsey, C. et al. (7 more authors) (2018) An Interactional Profile to Assist the Differential Diagnosis of Neurodegenerative and Functional Memory Disorders. Alzheimer Disease & Associated Disorders. ISSN 0893-0341

https://doi.org/10.1097/WAD.0000000000000231

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
An interactional profile to assist the differential diagnosis of
neurodegenerative and functional memory disorders

Revised Manuscript

Markus Reuber**,1, Daniel Blackburn**,1, Chris Elsey2, Sarah Wakefield1, Kerry Ardern1, Kirsty Harkness3, Annalena Venneri1, Danielle Jones4, Chloe Shaw5, Paul Drew6

*: Corresponding author (Daniel Blackburn)
**: Markus Reuber and Daniel Blackburn contributed in equal measure to this study.

1) Department of Neuroscience, University of Sheffield, Sheffield, UK.
2) School of Allied Health Sciences, De Montfort University, Leicester, UK.
3) Department of Neurology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK.
4) School of Dementia Studies, University of Bradford UK.
5) UCL Institute for Women’s Health, London, UK.
6) Department of Social Sciences, Loughborough University, Loughborough, UK.

Corresponding Author:

* Daniel Blackburn, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ. Telephone: 0114 222 2267 Email: d.blackburn@sheffield.ac.uk

Other authors:

Markus Reuber, Academic Neurology Unit, University of Sheffield, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF. Telephone: 0114 226 8688 Email: markus.reuber@sth.nhs.uk

Chris Elsey, School of Allied Health Sciences, De Montfort University, The Gateway, Leicester, LE1 9BH. Email: christopher.elsey@dmu.ac.uk

Sarah Wakefield, Department of Neuroscience, Medical School, University of Sheffield, Sheffield, S10 2TN. Email: s.wakefield@sheffield.ac.uk

Kirsty Harkness, Department of Neurology, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF. Email: kirsty.harkness@sth.nhs.uk

Kerry Ardern, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ. kaardern1@sheffield.ac.uk
Annalena Venneri, Department of Neuroscience, Medical School, University of Sheffield, Beech Hill Road, Sheffield, S10 2RX. Telephone: 0114 2713430. Email: a.venneri@sheffield.ac.uk

Danielle Jones, School of Dementia Studies, University of Bradford, Bradford, BD7 1DP. Telephone: 01274 236469 Email: d.k.jones1@bradford.ac.uk

Chloe Shaw, Institute for Women’s Health, University College London, 89-96 Chenies Mews, London WC1 6HX. Telephone: 020 7679 6036 Email: chloe.shaw@ucl.ac.uk

Paul Drew, Department of Social Sciences, Loughborough University, Epinal Way, Leicestershire, LE11 3TU. Telephone: 01509 223365 Email: p.drew@lboro.ac.uk
Objective: Specialist services for dementia are seeing an increasing number of patients. We investigated whether interactional and linguistic features in the communication behaviour of patients with memory problems could help distinguish between those with problems secondary to neurological disorders (ND) and those with Functional Memory Disorder (FMD).

Methods: In Part 1 of this study, a Diagnostic Scoring Aid (DSA) was developed encouraging linguists to provide quantitative ratings for 14 interactional features. An optimal cut-off differentiating ND and FMD was established by applying the DSA to 30 initial patient–doctor memory clinic encounters. In Part 2, the DSA was tested prospectively in ten additional cases analysed independently by two Conversation Analysts blinded to medical information.

Results: In part one, the median score of the DSA was +5 in ND and -5 in FMD (p<0.001). The optimal numeric DSA cut off (+1) identified patients with ND with a sensitivity of 86.7% and a specificity of 100%. In part two, DSA scores of rater one correctly predicted 10/10 and those of rater two 9/10 diagnoses.

Conclusions This study indicates that interactional and linguistic features can help distinguish between patients developing dementia and those with FMD and could aid the stratification of patients with memory problems.
INTRODUCTION

Demographic changes have increased pressure on specialist services for patients with dementia, causing healthcare professionals and service commissioners in many countries to focus on improvements to diagnostic pathways for people with memory complaints. In the UK, the National Dementia Strategy identified the closure of the ‘dementia gap’ (the difference between the predicted number of people with dementia versus those diagnosed as having dementia) as an area of particular concern. An audit by the Royal College of Psychiatrists found that the number of people assessed in specialist clinics with memory concerns increased fourfold between 2010 and 2013. A further 31% increase was seen from 2013 to 2014. While the ‘dementia gap’ has narrowed with this increase in activity, it has not been reduced at the same rate at which the number of memory clinic referrals has risen. One reason for this is that the proportion of patients with functional memory disorder (FMD) or other non-progressive memory disorders has also increased. The referral of patients to memory clinics is costly and can cause avoidable distress. These observations – combined with studies showing that current screening procedures lack sensitivity – suggest that case selection for referral to specialist clinics is suboptimal.

It is well recognised that patients exhibit linguistic impairments and deficits in spontaneous speech even in the earlier stages of dementia. Language impoverishment, through grammatical simplification, loss of vocabulary, semantic paraphasias, and overuse of semantically empty words, becomes progressively evident in dementia, as do impaired semantic processing and classification errors. While the analysis of patients’ language may therefore contribute towards identifying those at risk of developing dementia, the detection of such language impoverishment i) requires complex linguistic analysis, ii) may be diagnostically ambiguous, and iii) does not take account of more directly observable conversational or interactional features of language. Automated analysis of spontaneous speech could address some of the practical problems with the assessment of language in routine practice. However, to date, it remains uncertain how well this method would perform.
as a screening procedure in clinical situations. What is more, previous approaches have focused on vocal/aural features of speech, abstracted from their communicative context. Thus, they would not capture impairments in specifically interactional capabilities, that can be detected with methods focusing on the co-construction of conversation and which may well be particularly early and specific indicators of cognitive complaints secondary to neurological disorders (ND).

We aimed to use problems with communication between patients, doctors and third parties in medical consultations as a diagnostic tool. Our project was inspired by studies demonstrating the potential of Conversation Analysis (CA)-derived interactional and linguistic observations in the differentiation of epilepsy and (non-epileptic) dissociative seizures (DS) \(^ {18}\). Using previously described conversational profiles of patients with seizures, CA experts were able to predict the medical “gold standard” diagnosis of epilepsy or DS with a sensitivity and specificity of around 85\% \(^ {19}\).

Mirroring the study design pursued in seizure clinics, Conversation Analysts have previously described a number of interactional and linguistic features, which appeared to distinguish patients with ND from those with FMD in two qualitative studies \(^ {20, 21}\). The current paper describes the initial validation and assessment of a quantitative Diagnostic Scoring Aid (DSA) guiding analysts to rate each doctor-patient encounter on a number of the interactional, topical and linguistic features described in these qualitative studies. Part 1 of the present study was designed to establish an optimal discriminatory DSA cut-off for the distinction of ND and FMD. In Part 2 of this study this numeric cut-off was applied prospectively to DSA ratings of clinic interactions with newly recruited patients.

**METHODS**

**Participant recruitment and assessment**
Participant recruitment and assessment have been described in a previous article exploring the scope of automated analysis of conversational interaction\textsuperscript{22}. Briefly, all participants had been referred to the neurology-led memory clinic in Sheffield, UK. Patients are routinely encouraged to bring someone along to their memory clinic appointment if possible (accompanying person, AP). A member of the study team obtained written informed consent prior to the encounter with a neurologist. Participants and AP were only consented if they had capacity to make their own decision about participation and used English as their first language. Participants whose diagnosis remained uncertain and those whose cognitive problems were considered to be due to other causes than ND or FMD were excluded.

Participants were investigated and followed up by Consultant Neurologists specialising in memory disorders according to clinical need. Participants were referred for detailed neuropsychological testing and MRI brain imaging. The neuropsychological battery (see Wakefield et al 2014\textsuperscript{23} for details) included the Mini Mental State Examination\textsuperscript{24}, tests of short and long term memory (verbal and non-verbal)\textsuperscript{25}, abstract reasoning\textsuperscript{26, 27}, attention and executive function\textsuperscript{28}, language comprehension, naming by confrontation, category and letter fluency\textsuperscript{29}.

All participants were recruited before their first ever appointment in the memory clinic. Most patients with ND were in the early disease stages, but some already had moderately severe dementia. Diagnoses were reached by multidisciplinary consensus; taking into account clinical history, neurological examination, neuropsychological scores and neuro-radiological findings. Alzheimer’s disease was diagnosed according to the NINCDS-ADRDA criteria\textsuperscript{30}. A diagnosis of mixed dementia (AD plus vascular cognitive impairment) was made if moderate to severe small vessel ischaemic changes or cortical infarctions were present on MRI brain imaging. Vascular Cognitive Impairment not demented (VCIND) was used to label those with extensive radiological evidence of vascular impairment who, however, did not reach the threshold of dementia- i.e. the deterioration in function sufficient to impair activities of daily living; these patients were not included in the ND group. The
diagnosis of behavioural variant Frontotemporal Dementia (bvFTD) was made according to the Rascovsky criteria. Mild Cognitive Impairment (MCI) was diagnosed according to the Petersen criteria. We did not use biomarkers for amyloid or neurodegeneration (tau or FDG PET) because these tests are currently not available at our institution for routine assessment and because they are currently not widely used for clinical decision-making in the NHS. Initial diagnoses in the ND group were, however, confirmed by clinical follow-up.

The diagnosis of FMD was based on the criteria proposed by Schmidtke et al. 2008 with the exception of the age cut-off of <70 years. We considered this criterion overly restrictive because there have been previous reports of cases of ‘functional’ memory problems in people aged over 70. All participants with a Patient Health Questionnaire-9 (PHQ9) score of >15 (indicative of current depression) or symptoms of active moderate depression as judged by the clinician (and whose memory symptoms may have been due to depressive pseudodementia, DPD), were excluded from further analysis. Participants were also screened for Generalised Anxiety Disorder using the GAD7. However, in keeping with the criteria for the diagnosis of FMD proposed by Schmidtke et al., they were not excluded from the study on the basis of GAD7 scores. This means that some patients with FMD included in this study will have had significant problems with anxiety symptoms other than anxiety about memory.

Like those with DPD, participants in whom memory problems were found to be due to vascular cognitive impairment (VCI) not related to dementia and those for whom the diagnosis remained uncertain were not further analysed. Although VCI and DPD are important diagnostic categories, at this early stage of development we were keen to test the DSA methodology in only two homogeneous diagnostic groups.

Part 1 of this study is based on analyses of the first 15 patients with ND and the first 15 patients with FMD whose conversational data were analysable. Part 2 is based on the blinded analysis of the consecutive five next patients
with ND and next five patients with FMD who agreed to participate in this study after the recruitment for Part 1 of this study had been completed.

Data preparation for Conversation Analysis (CA)
Video or audio recordings of the history-taking phase (from a patient’s entry into the neurologist’s office, to the start of cognitive testing) were transcribed in detail, using a transcription system capturing the timing of speech (e.g. overlaps between speakers, pauses both within and between speaker turns), certain intonation and prosodic features of speech (falling/rising intonation, loudness, emphasis, sound stretching) 36.

Initial CA approach
Interactions with patients with ND or FMD were subjected to examination using the perspective and methods of CA. CA is a micro-analytic, qualitative method, but it is well-suited to combination with statistical measures. It has been applied widely to medical interactions in exploratory research37,38, in research aimed at improving the effectiveness of doctor-patient communication,39,40 and in assisting the diagnostic process in other conditions 19,38. It is particularly useful for the identification of detailed aspects of language use and communicative practices (e.g. the ways in which patients with epilepsy ‘normalise’ their seizure experiences, whilst patients with dissociative (nonepileptic) seizures ‘catastrophise’ their seizure descriptions)19. In their analysis, the CA experts involved in this project initially identified the methods individual patients (and accompanying others if present) used to describe memory problems to the doctor (see 20,21 for details).

Part 1: Quantitative examination of qualitative findings
Based on the findings of the initial qualitative analysis 20,21, we developed a diagnostic scoring aid (DSA) to provide a guide for the rating of potentially diagnostic features of communication and in order to transform qualitative observations into a numeric score (see Table 1). The DSA encourages analysts to comment and rate nine separate items. An additional five items
focus on triadic features that can only be observed if patients are accompanined during their memory clinic appointment. The DSA describes findings for each item more in keeping with ND (associated with a score of +1) or observations more in keeping with FMD (given a numeric score of -1).

Items can also be judged as un-ratable and would be given a score of 0. Items could be considered un-ratable because the interactional behaviour did not take place (for instance because the neurologist had not asked the specific question (e.g. “who is more concerned about the memory problems?”), the performance provided mixed evidence, or was neither typical of that expected from patients with ND nor that of patients with FMD. Free text fields are provided for each item allowing the analyst to describe the reasoning for their categorical judgement. Finally, the DSA asks the analyst to make a qualitative judgement taking account of the whole conversation profile.

The applicability of the DSA was initially tested by one CA expert using the DSA on the recordings and transcripts of the 30 cases previously analysed in a purely qualitative way. The analyst categorised each item as more in keeping with ND, more in keeping with FMD or un-rateable. This was translated into a numeric score for each item. Finally, item scores were added up to produce a total score for each patient. An AUROC statistic was carried out based on these numeric ratings to identify an optimal diagnostic cut off score.

Part 2: Blinded analysis using a Diagnostic Scoring Aid

In order to test the discriminatory potential of the DSA, two CA experts independently rated an additional ten doctor-patient encounters (five consecutive cases in each group with ND and FMD). These cases had not been included in the initial qualitative or retrospective quantitative analyses. The Conversation Analysts were expected to predict the neurological diagnosis in these new cases on the basis of their qualitative analysis of the video recordings of the doctor-patient encounter and transcripts of these encounters, guided by the DSA. They were also asked to use the DSA to produce a numeric assessment score for each participant having provided a
score for each DSA feature. The analysts were blinded to all additional medical or demographic information about these patients. Analysts were not aware of the numeric cut-off calculated by the AUROC statistic at the time of this analysis and were encouraged not simply to base their diagnostic prediction on patients’ numeric score. In their overall qualitative judgement, this enabled them to place more diagnostic emphasis on particularly outstanding features. However, in addition to the number of cases correctly diagnosed by their qualitative judgement, we also report the number of cases correctly categorised on the basis of the DSA scores using the diagnostic cut-off calculated in Part 1 of this study.

Statistical analysis
Routinely collected clinical data on consenting and non-consenting patients approached about this study were compared using t-tests to ascertain the representativeness of the patient group included.

In Part 1 of this study, the diagnostic potential of individual DSA items was examined using Fisher’s Exact tests. The significance of differences in median scores of the ND and FMD groups was determined using the Mann-Whitney U test. An Area Under the Receiver Operated Characteristic Curve (AUROC) statistic was used to identify an optimal numeric cut-off for the differentiation of ND and FMD. In Part 2 of this study we report Kappa scores as a measure of the inter-rater reliability of focussed CA using the DSA.

Ethics
The study was approved by the NHS Research Ethics Committee (NRES Committee Yorkshire & The Humber - South Yorkshire). Ref 12/YH/0205.

RESULTS
Part 1
Of 353 patients referred to the specialist memory clinic during the recruitment period and of 148 eligible to take part in this study, 36 declined to participate and 112 were enrolled (see Figure 1 Consort diagram). Three withdrew their consent subsequently, leaving 109 who completed the study. There were
no significant differences in terms of age, gender, anxiety, depression or ACE-R scores between those who consented to take part and those who did not, suggesting that the participants were representative of the wider population served by this memory clinic. There was also no significant difference in terms of diagnostic mix between people who consented and people who did not consent (see supplementary Table 1). The ND patient group (n=20) comprised of eight patients with AD, four with amnestic MCI, two with vascular dementia, two with frontotemporal dementia, three mixed AD and vascular and one unspecified dementia (without detailed neuropsychology). Figure 1 also provides more information about participants who were excluded from the study.

Clinical details of the patients included in the ND and FMD groups described here are provided in Table 2. Two participants out of the twenty with FMD had MRI brain scans reported as possible atrophy. They were both followed up; One followed up at 24 months was aware that they had been working in a very stressful job at time of the first consultation. At follow-up they had changed jobs and now had no memory complaints. The second person was followed up at 18 months. The Montreal Cognitive Assessment (MoCA) was 26/30 at follow-up (prior ACE 85 and MoCA 22). They were functioning normally in a busy job. Two participants out of twenty cases with ND had normal structural scans but one of these had abnormal Single-Photon Emission Computed Tomography (SPECT). The other was seen for follow-up at 12 months and ACE-R had decreased from 87 to 82, the clinical picture at this stage being consistent with AD.

In Part 2 of the study, three out of five FMD cases did not attend for neuropsychology testing, hence the missing MMSE scores in table 2. However, their ACE-R scores were 87, 96 and 97. Two had entirely normal neuroimaging and were discharged. One had an old caudate head infarct and on follow-up one year later was still working, managing a team. Repeat ACE-R was 96 (97 one year earlier). Also in Part 2, one person with ND did not have detailed neuropsychological testing. Neuroimaging showed atrophy. On follow-up this patient showed significant cognitive impairment.
There were no significant differences in demographics, depression, anxiety or ACE-R scores between ND participants in Parts 1 and 2 or between FMD patients in the two parts of this study. 20 of the 30 participants included in Part 1, and seven of ten included in Part 2 of this study were accompanied. Feature 11 (patient’s head turn encouraging accompanying person to answer a question directed at the patient) could not be rated in two of the accompanied encounters because participants had only consented to audio recording the interaction.

Table 3 shows the analyst’s DSA ratings of the interactions included in Part 1 of this study. A more detailed description of the individual items can be found on the DSA form (additional web content). The median total score of the first nine items of the DSA was +5 in the ND (range +8 to -3) and -5 in the FMD group (range 0 to -9, difference p<0.001).

The median total of the five additional items to be rated in accompanied encounters was 2 (range +5 to -3) in the ND group and -1 (range 1 to -5) in the FMD group (difference p=0.003). The fact that only one of the additional items to be rated in accompanied encounters individually yielded a statistically significant between-group difference may (at least in part) be explained by the relatively small number of accompanied interactions available for analysis.

In view of the fact that the additional item scores for accompanied interactions were only available for a subset of the encounters, only the first nine items were used for the AUROC analysis and the estimation of a quantitative diagnostic threshold. The area under the ROC curve was 0.98 (see Figure 2). At the optimal DSA score for the distinction of patients with ND from those with FMD of +1 (with DSA score above this threshold suggesting a diagnosis of ND) the DSA-derived total score identified patients with ND with a sensitivity of 86.7% and a specificity of 100%.

3.4 RESULTS – PART 2
3.4.1 Quantitative scoring using the DSA (blinded results)

Rater 1 was accurate in 10/10 cases, whilst Rater 2 correctly predicted 9/10 diagnoses on the basis of DSA-guided qualitative analysis (Rater 1 was more experienced because of his involvement in Part 1 of the project). The results were identical when the linguistic diagnostic prediction was based on the numeric DSA scores. The case misdiagnosed by Rater 2 as FMD (when the ultimate medical diagnosis was ND) attracted the lowest score Rater 1 gave to any of the patients assessed as having ND (+4) and the highest score Rater 2 gave to any patients thought to have FMD (-1). This suggests that the patient misdiagnosed by Rater 2 had an objectively ambiguous conversational profile, posing a particular discriminatory challenge. The differences in the two raters’ diagnostic prediction was based on a single completely discordant judgment of DSA item 4 (ratings 1 vs. -1) and on non-concordant decisions (0 vs. 1 or 0 vs. -1) on DSA items 3, 7, 9, 13 and 14 (see Table 4 for further DSA scoring details).

3.4.2 Inter-rater reliability of the DSA

In terms of the final diagnosis (either based on the two raters’ qualitative judgements or the quantitative procedure using the diagnostic cut-off derived from the AUROC analysis), the raters agreed in 9/10 cases. The Kappa value for the DSA procedure as a whole was therefore 0.8 (SE of Kappa = 0.19, 95% confidence interval 0.44 to 1.0) suggesting ‘very good’ inter-rater reliability. We also looked at the inter-rater reliability of the 123 individual +1, 0 or -1 ratings from Part 2 of this study. Both ratings were fully concordant for 87 numeric scores (the scores from raters 1 and 2 were 1/1, 0/0 or -1/-1), non-concordant for 30 and discordant for 6. This means that both raters agreed on 70.7% of the observations when agreement on 33.8% of the ratings would have been expected by chance. The Kappa value for all 123 DSA-based ratings combined was 0.56 (SE of Kappa = 0.06, 95% confidence interval 0.44 to 0.68), consistent with ‘moderate’ inter-rater agreement. The two raters’
scores for each item assessed in part 2 of the study and the Kappa-values of each individual item are shown in table 3.

Discussion

Our previous qualitative work has demonstrated that it is possible to describe characteristic conversational profiles of patients describing cognitive problems due to ND or FMD based on their interactional and linguistic contributions to initial encounters in a memory clinic\textsuperscript{20,21}. However, in these descriptive studies, the conversation analysts who analysed video- and audio recordings of memory clinic encounters between neurologists, patients and (sometimes) accompanying persons were always aware of the patients’ medical diagnoses during the analytic process. The present study is the first to demonstrate that these linguistic and interactional features can be used diagnostically to predict diagnoses of ND or FMD made on the basis of standard medical criteria. What is more, we show that qualitative assessments can be structured and likely medical diagnoses formulated using a Diagnostic Scoring Aid with a numeric diagnostic cut-off. The fact that the linguistic raters involved in this study had no expertise in the medical assessment of patients presenting with memory problems, together with the relatively high level of agreement between the two raters, suggests that the raters did not base their diagnostic predictions on an ill-defined hunch but on robust and objectifiable interactional observations.

The correct classification of 9/10 by one rater and 10/10 by a second independent rater, and the very good inter-rater reliability of the DSA-guided procedure as a whole, suggest that the addition of the structured observation of interactional features can make a significant contribution to screening processes for ND. Importantly, this is one of the few studies of cognitive screening ‘tools’ to include participants with FMD; most previous studies compared patients with memory impairment with healthy controls. The inclusion of a group of patients with memory complaints but no neurological disorder adds ecological validity to our findings. Compared to other studies set in clinical situations (such as a study exploring the screening potential of
the 6CIT brief cognitive test in primary care) our approach appears to have
greater reliability and validity.\footnote{9}

Interactional and linguistic observations may increase the confidence of non-
expert clinicians to diagnose clear cases of FMD, enabling them to treat
patients in primary care or to refer them on to services providing appropriate
treatment for functional neurological disorders. Importantly, the interactional
and linguistic observations contributing to the diagnosis of FMD may allow
clinicians to provide more effective reassurance by allowing them to
demonstrate to patients that they are displaying good memory function in
interaction. It should be possible for clinicians to pick up these features during
routine clinic encounters. Previous studies in patients presenting with epileptic
or dissociative (non-epileptic) seizures have demonstrated that doctors can
learn to change their history-taking style to optimise patients' opportunities to
demonstrate particular conversational behaviours and to make diagnostically
useful interactional observations as they take a patient's history.\footnote{41} The DSA
developed here could be modified and used in similar studies investigating
whether clinicians can be trained to identify features, which have diagnostic
value while talking to patients with memory problems.

This study has a number of limitations. First and foremost, we were only able
to explore the potential of CA as a tool capable of predicting medical
diagnoses in the setting of initial memory clinic encounters in a modest
number of patients. Whilst the consecutive recruitment and the levels of
statistical significance in between-group tests on a range of separate
conversational features observed even in such a small patient group make it
unlikely that our findings are spurious, it would be desirable to replicate our
findings in a larger and more diverse group of patients. One particular
limitation of our findings in this regard is that we excluded patients with
depression and those with VCI from this first quantitative study of our method.
These are important differential diagnoses, which will need to be picked up by
screening procedures. Future larger studies will need to demonstrate that the
inclusion of interactional and linguistic observations can contribute to
screening or stratification procedures in which patients with these problems
are allocated to the correct management pathways. We also recognised that
the ND and FMD groups in this study were not age-matched since those with
ND were significantly older. This is not surprising as the biggest risk factor for
ND is increasing age. However as younger patients with memory concerns
are increasingly referred to specialist memory clinics it is important to include
and compare all age groups in studies of this nature. Furthermore the mean
MMSE score of the ND group was lower than that of the FMD group (20.4
versus 28.2), reflecting the relatively late stage in the development of
cognitive disorders at which patients are currently first referred to specialist
services. Only four of the patients in the ND groups had MCI. An optimal
screening tool for the earliest stages of ND would need to be capable of
picking up patients with MCI and near normal MMSE scores). This means that
confirmatory studies capturing more patients at an earlier stage of ND will be
required before the method described here can be embedded in screening
procedures. **The MMSE scores of the MCI patients included in this study were
just above the standard cut-off score of 23 for this test, indicating that,
although unimpaired in their activities of daily living, they already had
extensive global cognitive impairment.** We do not know how effective this tool
will be at distinguishing MCI from FMD. Because a clinical diagnosis of MCI
refers to a very heterogeneous symptom profile, the distinction between MCI
and FMD might be difficult and will require larger number of participants,
along with prospective follow-up to investigate whether it can predict those
who are at high risk of developing AD or other dementias.

We have only studied native English speakers; findings may have been
different in patients speaking other languages or those using English as a
second language. Although several different doctors were involved in the
clinic conversations studied here, it would also be important to test this
procedure in different clinical settings (for instance in community-based
clinics, during home visits and in elderly-care settings). The fact that the ND
group included patients with memory problems of different aetiologies should
not be considered a weakness of this study. Although it is likely the method
employed in this study could also be deployed to identify interactional
differences between different ND (such as Alzheimer's disease or frontotemporal dementia) the fact that we were able to distinguish clearly between patients with a range of ND and those with FMD demonstrates its potential for screening or stratifying patient management.

We did not have access to investigations confirming clinical diagnosis with tests documenting the presence of amyloid or tau (PET or cerebrospinal studies), but this reflects current NICE guidelines [http://www.nice.org.uk/guidance/CG42](http://www.nice.org.uk/guidance/CG42) and our ‘medical’ diagnoses were based on multidisciplinary assessment by experts including detailed neuropsychological testing and structural brain imaging as well as clinical follow up.

Future studies will need to demonstrate how much diagnostic value the observation of interactional features can add to conventional brief cognitive screening tools. A combined approach with an automated low cost, high-speed system to analyse speech will require the use of technology rather than Conversation Analysts. One way in which the research described here can be taken forward involves the computerised analysis of speech which has shown some promise in distinguishing AD, MCI and healthy controls \(^{16}\). Early indications are that computerised speech analysis and machine learning algorithms can also be used to produce an automated system to pick up and evaluate the sort of interactional observations described here and can discriminate between ND and FMD \(^{22}\).

More immediately, the findings described here can be used in the training of clinicians working with patients with memory problems.

**Acknowledgements:** This article presents independent research funded by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number PB-PG-0211-24079). This is a summary of independent research carried out at the NIHR Sheffield Biomedical Research Centre (Translational Neuroscience). The
views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

We would like to thank all participants and care partners who took part. We would also like to thank all participating clinicians working in the memory clinic (Aijaz Khan, Alice Brockington, Dilraj Sokhi and Muhammad Rafiq). We also acknowledge the support of the NIHR Sheffield Clinical Research Facility.

References


2. Audit RMC. 
   http://www.rcpsych.ac.uk/workinpsychiatry/qualityimprovement/nationalclinicalaudit
   /memoryservicesaudit.aspx, 2015


30. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-


Figure 1. CONSORT diagram showing recruitment to study. FMD- Functional Memory Disorder; DPD- Depressive Pseudo Dementia; ND- memory problems secondary to neurological disorders (ND) – this includes neurodegenerative dementias and mild cognitive impairment due to likely underlying neurodegenerative aetiology.

Figure 2: Area under the Receiver Operator Characteristic Curve
A ROC curve was constructed for the sample of 15 ND and 15 FMD cases.

Table 1 Diagnostic Scoring Aid (DSA)
For each interactional feature scores range between 1 and -1 (1: in favour of ND; 0: undecided or unable to rate; -1: in favour of FMD). There are 9/14 (unaccompanied/accompanied) items to score, so the maximum score is 9/14 (unaccompanied/accompanied) and minimum score -9/-14 (unaccompanied/accompanied). A high score corresponds to a stereotypical ND description; a low score to a stereotypical FMD description.

Supplementary Table 1 Comparison of patients eligible for participation in the study who did and did not consent to have their clinic interaction recorded and analysed.

There is no significant difference in demographics or test scores between the people who consented and the people who did not consent. Addenbrooke’s Cognitive Examination; Patient Health Questionnaire-9 PHQ9- Generalised Anxiety Disorder scale 7 GAD7

Table 2 Demographic and neuropsychological results
ACE-R Addenbrooke’s Cognitive Examination. MMSE Mini Mental State Examination. PHQ9- Patient Health Questionnaire 9 item depression scale. GAD7General Anxiety Disorder 7-item Scale CF- Confrontational Naming. VPA- Verbal Paired Associates, P&PT-Pyramid & Palm Trees, Rey’s CF-Rey’s Complex Figure, SF- Semantic Fluency, PF - Phonemic Fluency, DS - Digit Span, VCA- Visuoconstructive Apraxia, TT-Token task, PM - Prose Memory. * 3 missing scores. + 1 missing score. #
Three missing scores from ND group (2 due to different protocol and one participants from part 2). Three missing scores from FMD group due to not attending appointments. Twenty participants with ND; comprised eight with AD, four with amnestic MCI, two 2 with vascular dementia, two with fronto temporal dementia, three with mixed AD and vascular and one unspecified dementia (without detailed neuropsychology).

Table 3: Diagnostic Scoring Aid (DSA) results
Profiles of 30 patients attending a specialist clinic with memory secondary to neurological disorders (ND) – this includes neurodegenerative dementias and mild cognitive impairment due to likely underlying neurodegenerative aetiology or functional memory disorder (FMD). 3a: Items rated in all
encounters / 3b: Items only rated in encounters also involving an accompanying person (AP).
For a full description of the items see supplementary appendix. Some items were unratable because a particular question was not asked (eg. “who is most concerned?”).

Table 4 Diagnostic Scoring Aid (DSA) results of blinded analysis.
Two independent linguistic raters (L1 and L2) of interactions with five patients with a medical diagnosis of FMD (5a) and five patients with medical diagnosis of ND (5b).
An interactional profile to assist the differential diagnosis of neurodegenerative and functional memory disorders

Revised Manuscript

Markus Reuber**,1, Daniel Blackburn**,1, Chris Elsey2, Sarah Wakefield1, Kerry Ardern1, Kirsty Harkness3, Annalena Venneri1, Danielle Jones4, Chloe Shaw5, Paul Drew6

*: Corresponding author (Daniel Blackburn)
**: Markus Reuber and Daniel Blackburn contributed in equal measure to this study.

1) Department of Neuroscience, University of Sheffield, Sheffield, UK.
2) School of Allied Health Sciences, De Montfort University, Leicester, UK.
3) Department of Neurology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK.
4) School of Dementia Studies, University of Bradford UK.
5) UCL Institute for Women’s Health, London, UK.
6) Department of Social Sciences, Loughborough University, Loughborough, UK.

Corresponding Author:

* Daniel Blackburn, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ. Telephone: 0114 222 2267 Email: d.blackburn@sheffield.ac.uk

Other authors:

Markus Reuber, Academic Neurology Unit, University of Sheffield, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF. Telephone: 0114 226 8688 Email: markus.reuber@sth.nhs.uk

Chris Elsey, School of Allied Health Sciences, De Montfort University, The Gateway, Leicester, LE1 9BH. Email: christopher.elsey@dmu.ac.uk

Sarah Wakefield, Department of Neuroscience, Medical School, University of Sheffield, Sheffield, S10 2TN. Email: s.wakefield@sheffield.ac.uk

Kirsty Harkness, Department of Neurology, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF. Email: kirsty.harkness@sth.nhs.uk

Kerry Ardern, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ. kaardern1@sheffield.ac.uk
Annalena Venneri, Department of Neuroscience, Medical School, University of Sheffield, Beech Hill Road, Sheffield, S10 2RX. Telephone: 0114 2713430. Email: a.venneri@sheffield.ac.uk

Danielle Jones, School of Dementia Studies, University of Bradford, Bradford, BD7 1DP. Telephone: 01274 236469 Email: d.k.jones1@bradford.ac.uk

Chloe Shaw, Institute for Women’s Health, University College London, 89-96 Chenies Mews, London WC1 6HX. Telephone: 020 7679 6036 Email: chloe.shaw@ucl.ac.uk

Paul Drew, Department of Social Sciences, Loughborough University, Epinal Way, Leicestershire, LE11 3TU. Telephone: 01509 223365 Email: p.drew@lboro.ac.uk
**Objective:** Specialist services for dementia are seeing an increasing number of patients. We investigated whether interactional and linguistic features in the communication behaviour of patients with memory problems could help distinguish between those with problems secondary to neurological disorders (ND) and those with Functional Memory Disorder (FMD).

**Methods:** In Part 1 of this study, a Diagnostic Scoring Aid (DSA) was developed encouraging linguists to provide quantitative ratings for 14 interactional features. An optimal cut-off differentiating ND and FMD was established by applying the DSA to 30 initial patient–doctor memory clinic encounters. In Part 2, the DSA was tested prospectively in ten additional cases analysed independently by two Conversation Analysts blinded to medical information.

**Results:** In part one, the median score of the DSA was +5 in ND and -5 in FMD (p<0.001). The optimal numeric DSA cut off (+1) identified patients with ND with a sensitivity of 86.7% and a specificity of 100%. In part two, DSA scores of rater one correctly predicted 10/10 and those of rater two 9/10 diagnoses.

**Conclusions** This study indicates that interactional and linguistic features can help distinguish between patients developing dementia and those with FMD and could aid the stratification of patients with memory problems.
INTRODUCTION

Demographic changes have increased pressure on specialist services for patients with dementia, causing healthcare professionals and service commissioners in many countries to focus on improvements to diagnostic pathways for people with memory complaints. In the UK, the National Dementia Strategy identified the closure of the ‘dementia gap’ (the difference between the predicted number of people with dementia versus those diagnosed as having dementia) as an area of particular concern. An audit by the Royal College of Psychiatrists found that the number of people assessed in specialist clinics with memory concerns increased fourfold between 2010 and 2013. A further 31% increase was seen from 2013 to 2014. While the ‘dementia gap’ has narrowed with this increase in activity, it has not been reduced at the same rate at which the number of memory clinic referrals has risen. One reason for this is that the proportion of patients with functional memory disorder (FMD) or other non-progressive memory disorders has also increased. The referral of patients to memory clinics is costly and can cause avoidable distress. These observations—combined with studies showing that current screening procedures lack sensitivity—suggest that case selection for referral to specialist clinics is suboptimal.

It is well recognised that patients exhibit linguistic impairments and deficits in spontaneous speech even in the earlier stages of dementia. Language impoverishment, through grammatical simplification, loss of vocabulary, semantic paraphasias, and overuse of semantically empty words, becomes progressively evident in dementia, as do impaired semantic processing and classification errors. While the analysis of patients’ language may therefore contribute towards identifying those at risk of developing dementia, the detection of such language impoverishment i) requires complex linguistic analysis, ii) may be diagnostically ambiguous, and iii) does not take account of more directly observable conversational or interactional features of language. Automated analysis of spontaneous speech could address some of the practical problems with the assessment of language in routine practice. However, to date, it remains uncertain how well this method would perform.
as a screening procedure in clinical situations. What is more, previous approaches have focused on vocal/aural features of speech, abstracted from their communicative context. Thus, they would not capture impairments in specifically interactional capabilities, that can be detected with methods focusing on the co-construction of conversation and which may well be particularly early and specific indicators of cognitive complaints secondary to neurological disorders (ND).

We aimed to use problems with communication between patients, doctors and third parties in medical consultations as a diagnostic tool. Our project was inspired by studies demonstrating the potential of Conversation Analysis (CA)-derived interactional and linguistic observations in the differentiation of epilepsy and (non-epileptic) dissociative seizures (DS) \(^{18}\). Using previously described conversational profiles of patients with seizures, CA experts were able to predict the medical “gold standard” diagnosis of epilepsy or DS with a sensitivity and specificity of around 85\% \(^{19}\).

Mirroring the study design pursued in seizure clinics, Conversation Analysts have previously described a number of interactional and linguistic features, which appeared to distinguish patients with ND from those with FMD in two qualitative studies \(^{20,21}\). The current paper describes the initial validation and assessment of a quantitative Diagnostic Scoring Aid (DSA) guiding analysts to rate each doctor-patient encounter on a number of the interactional, topical and linguistic features described in these qualitative studies. Part 1 of the present study was designed to establish an optimal discriminatory DSA cut-off for the distinction of ND and FMD. In Part 2 of this study this numeric cut-off was applied prospectively to DSA ratings of clinic interactions with newly recruited patients.

**METHODS**

**Participant recruitment and assessment**
Participant recruitment and assessment have been described in a previous article exploring the scope of automated analysis of conversational interaction. Briefly, all participants had been referred to the neurology-led memory clinic in Sheffield, UK. Patients are routinely encouraged to bring someone along to their memory clinic appointment if possible (accompanying person, AP). A member of the study team obtained written informed consent prior to the encounter with a neurologist. Participants and AP were only consented if they had capacity to make their own decision about participation and used English as their first language. Participants whose diagnosis remained uncertain and those whose cognitive problems were considered to be due to other causes than ND or FMD were excluded.

Participants were investigated and followed up by Consultant Neurologists specialising in memory disorders according to clinical need. Participants were referred for detailed neuropsychological testing and MRI brain imaging. The neuropsychological battery (see Wakefield et al 2014 for details) included the Mini Mental State Examination, tests of short and long term memory (verbal and non-verbal), abstract reasoning, attention and executive function, language comprehension, naming by confrontation, category and letter fluency.

All participants were recruited before their first ever appointment in the memory clinic. Most patients with ND were in the early disease stages, but some already had moderately severe dementia. Diagnoses were reached by multidisciplinary consensus; taking into account clinical history, neurological examination, neuropsychological scores and neuro-radiological findings. Alzheimer’s disease was diagnosed according to the NINCDS-ADRDA criteria. A diagnosis of mixed dementia (AD plus vascular cognitive impairment) was made if moderate to severe small vessel ischaemic changes or cortical infarctions were present on MRI brain imaging. Vascular Cognitive Impairment not demented (VCIND) was used to label those with extensive radiological evidence of vascular impairment who, however, did not reach the threshold of dementia; i.e. the deterioration in function sufficient to impair activities of daily living; these patients were not included in the ND group. The
diagnosis of behavioural variant Frontotemporal Dementia (bvFTD) was made according to the Rascovsky criteria. Mild Cognitive Impairment (MCI) was diagnosed according to the Petersen criteria. We did not use biomarkers for amyloid or neurodegeneration (tau or FDG PET) because these tests are currently not available at our institution for routine assessment and because they are currently not widely used for clinical decision-making in the NHS. Initial diagnoses in the ND group were, however, confirmed by clinical follow-up.

The diagnosis of FMD was based on the criteria proposed by Schmidtke et al. 2008 with the exception of the age cut-off of <70 years. We considered this criterion overly restrictive because there have been previous reports of cases of ‘functional’ memory problems in people aged over 70. All participants with a Patient Health Questionnaire-9 (PHQ9) score of >15 (indicative of current depression) or symptoms of active moderate depression as judged by the clinician (and whose memory symptoms may have been due to depressive pseudodementia, DPD), were excluded from further analysis. Participants were also screened for Generalised Anxiety Disorder using the GAD7. However, in keeping with the criteria for the diagnosis of FMD proposed by Schmidtke et al., they were not excluded from the study on the basis of GAD7 scores. This means that some patients with FMD included in this study will have had significant problems with anxiety symptoms other than anxiety about memory.

Like those with DPD, participants in whom memory problems were found to be due to vascular cognitive impairment (VCI) not related to dementia and those for whom the diagnosis remained uncertain were not further analysed. Although VCI and DPD are important diagnostic categories, at this early stage of development we were keen to test the DSA methodology in only two homogeneous diagnostic groups.

Part 1 of this study is based on analyses of the first 15 patients with ND and the first 15 patients with FMD whose conversational data were analysable. Part 2 is based on the blinded analysis of the consecutive five next patients.
with ND and next five patients with FMD who agreed to participate in this study after the recruitment for Part 1 of this study had been completed.

**Data preparation for Conversation Analysis (CA)**

Video or audio recordings of the history-taking phase (from a patient's entry into the neurologist's office, to the start of cognitive testing) were transcribed in detail, using a transcription system capturing the timing of speech (e.g. overlaps between speakers, pauses both within and between speaker turns), certain intonation and prosodic features of speech (falling/rising intonation, loudness, emphasis, sound stretching) 36.

**Initial CA approach**

Interactions with patients with ND or FMD were subjected to examination using the perspective and methods of CA. CA is a micro-analytic, qualitative method, but it is well-suited to combination with statistical measures. It has been applied widely to medical interactions in exploratory research 37, 38, in research aimed at improving the effectiveness of doctor-patient communication, 39, 40 and in assisting the diagnostic process in other conditions 19, 38. It is particularly useful for the identification of detailed aspects of language use and communicative practices (e.g. the ways in which patients with epilepsy ‘normalise’ their seizure experiences, whilst patients with dissociative (nonepileptic) seizures ‘catastrophise’ their seizure descriptions) 19. In their analysis, the CA experts involved in this project initially identified the methods individual patients (and accompanying others if present) used to describe memory problems to the doctor (see 20, 21 for details).

**Part 1: Quantitative examination of qualitative findings**

Based on the findings of the initial qualitative analysis 20, 21, we developed a diagnostic scoring aid (DSA) to provide a guide for the rating of potentially diagnostic features of communication and in order to transform qualitative observations into a numeric score (see Table 1). The DSA encourages analysts to comment and rate nine separate items. An additional five items
focus on triadic features that can only be observed if patients are
accompanied during their memory clinic appointment. The DSA describes
findings for each item more in keeping with ND (associated with a score of +1)
or observations more in keeping with FMD (given a numeric score of -1).
Items can also be judged as un-ratable and would be given a score of 0.
Items could be considered un-ratable because the interactional behaviour did
not take place (for instance because the neurologist had not asked the
specific question (e.g. “who is more concerned about the memory
problems?”), the performance provided mixed evidence, or was neither typical
of that expected from patients with ND nor that of patients with FMD. Free text
fields are provided for each item allowing the analyst to describe the
reasoning for their categorical judgement. Finally, the DSA asks the analyst to
make a qualitative judgement taking account of the whole conversation
profile.

The applicability of the DSA was initially tested by one CA expert using the
DSA on the recordings and transcripts of the 30 cases previously analysed in
a purely qualitative way. The analyst categorised each item as more in
keeping with ND, more in keeping with FMD or un-rateable. This was
translated into a numeric score for each item. Finally, item scores were added
up to produce a total score for each patient. An AUROC statistic was carried
out based on these numeric ratings to identify an optimal diagnostic cut off
score.

**Part 2: Blinded analysis using a Diagnostic Scoring Aid**

In order to test the discriminatory potential of the DSA, two CA experts
independently rated an additional ten doctor-patient encounters (five
consecutive cases in each group with ND and FMD). These cases had not
been included in the initial qualitative or retrospective quantitative analyses.
The Conversation Analysts were expected to predict the neurological
diagnosis in these new cases on the basis of their qualitative analysis of the
video recordings of the doctor-patient encounter and transcripts of these
encounters, guided by the DSA. They were also asked to use the DSA to
produce a numeric assessment score for each participant having provided a
score for each DSA feature. The analysts were blinded to all additional medical or demographic information about these patients. Analysts were not aware of the numeric cut-off calculated by the AUROC statistic at the time of this analysis and were encouraged not simply to base their diagnostic prediction on patients’ numeric score. In their overall qualitative judgement, this enabled them to place more diagnostic emphasis on particularly outstanding features. However, in addition to the number of cases correctly diagnosed by their qualitative judgement, we also report the number of cases correctly categorised on the basis of the DSA scores using the diagnostic cut-off calculated in Part 1 of this study.

**Statistical analysis**

Routinely collected clinical data on consenting and non-consenting patients approached about this study were compared using t-tests to ascertain the representativeness of the patient group included.

In Part 1 of this study, the diagnostic potential of individual DSA items was examined using Fisher’s Exact tests. The significance of differences in median scores of the ND and FMD groups was determined using the Mann-Whitney U test. An Area Under the Receiver Operated Characteristic Curve (AUROC) statistic was used to identify an optimal numeric cut-off for the differentiation of ND and FMD. In Part 2 of this study we report Kappa scores as a measure of the inter-rater reliability of focussed CA using the DSA.

**Ethics**

The study was approved by the NHS Research Ethics Committee (NRES Committee Yorkshire & The Humber - South Yorkshire). Ref 12/YH/0205.

**RESULTS**

**Part 1**

Of 353 patients referred to the specialist memory clinic during the recruitment period and of 148 eligible to take part in this study, 36 declined to participate and 112 were enrolled (see Figure 1 CONSORT diagram). Three withdrew their consent subsequently, leaving 109 who completed the study. There were
no significant differences in terms of age, gender, anxiety, depression or ACE-R scores between those who consented to take part and those who did not, suggesting that the participants were representative of the wider population served by this memory clinic. There was also no significant difference in terms of diagnostic mix between people who consented and people who did not consent (see supplementary Table 1). The ND patient group (n=20) comprised of eight patients with AD, four with amnestic MCI, two with vascular dementia, two with frontotemporal dementia, three mixed AD and vascular and one unspecified dementia (without detailed neuropsychology). Figure 1 also provides more information about participants who were excluded from the study.

Clinical details of the patients included in the ND and FMD groups described here are provided in Table 2. Two participants out of the twenty with FMD had MRI brain scans reported as possible atrophy. They were both followed up; One followed up at 24 months was aware that they had been working in a very stressful job at time of the first consultation. At follow-up they had changed jobs and now had no memory complaints. The second person was followed up at 18 months. The Montreal Cognitive Assessment (MoCA) was 26/30 at follow-up (prior ACE 85 and MoCA 22). They were functioning normally in a busy job. Two participants out of twenty cases with ND had normal structural scans but one of these had abnormal Single-Photon Emission Computed Tomography (SPECT). The other was seen for follow-up at 12 months and ACE-R had decreased from 87 to 82, the clinical picture at this stage being consistent with AD.

In Part 2 of the study, three out of five FMD cases did not attend for neuropsychology testing, hence the missing MMSE scores in table 2. However, their ACE-R scores were 87, 96 and 97. Two had entirely normal neuroimaging and were discharged. One had an old caudate head infarct and on follow-up one year later was still working, managing a team. Repeat ACE-R was 96 (97 one year earlier). Also in Part 2, one person with ND did not have detailed neuropsychological testing. Neuroimaging showed atrophy. On follow-up this patient showed significant cognitive impairment.
There were no significant differences in demographics, depression, anxiety or ACE-R scores between ND participants in Parts 1 and 2 or between FMD patients in the two parts of this study. 20 of the 30 participants included in Part 1, and seven of ten included in Part 2 of this study were accompanied. Feature 11 (patient’s head turn encouraging accompanying person to answer a question directed at the patient) could not be rated in two of the accompanied encounters because participants had only consented to audio recording the interaction.

Table 3 shows the analyst’s DSA ratings of the interactions included in Part 1 of this study. A more detailed description of the individual items can be found on the DSA form (additional web content). The median total score of the first nine items of the DSA was +5 in the ND (range +8 to -3) and -5 in the FMD group (range 0 to -9, difference p<0.001).

The median total of the five additional items to be rated in accompanied encounters was 2 (range +5 to -3) in the ND group and -1 (range 1 to -5) in the FMD group (difference p=0.003). The fact that only one of the additional items to be rated in accompanied encounters individually yielded a statistically significant between-group difference may (at least in part) be explained by the relatively small number of accompanied interactions available for analysis.

In view of the fact that the additional item scores for accompanied interactions were only available for a subset of the encounters, only the first nine items were used for the AUROC analysis and the estimation of a quantitative diagnostic threshold. The area under the ROC curve was 0.98 (see Figure 2). At the optimal DSA score for the distinction of patients with ND from those with FMD of +1 (with DSA score above this threshold suggesting a diagnosis of ND) the DSA-derived total score identified patients with ND with a sensitivity of 86.7% and a specificity of 100%.

3.4 RESULTS – PART 2
3.4.1 Quantitative scoring using the DSA (blinded results)

Rater 1 was accurate in 10/10 cases, whilst Rater 2 correctly predicted 9/10 diagnoses on the basis of DSA-guided qualitative analysis (Rater 1 was more experienced because of his involvement in Part 1 of the project). The results were identical when the linguistic diagnostic prediction was based on the numeric DSA scores. The case misdiagnosed by Rater 2 as FMD (when the ultimate medical diagnosis was ND) attracted the lowest score Rater 1 gave to any of the patients assessed as having ND (+4) and the highest score Rater 2 gave to any patients thought to have FMD (-1). This suggests that the patient misdiagnosed by Rater 2 had an objectively ambiguous conversational profile, posing a particular discriminatory challenge. The differences in the two raters’ diagnostic prediction was based on a single completely discordant judgment of DSA item 4 (ratings 1 vs. -1) and on non-concordant decisions (0 vs. 1 or 0 vs. -1) on DSA items 3, 7, 9, 13 and 14 (see Table 4 for further DSA scoring details).

3.4.2 Inter-rater reliability of the DSA

In terms of the final diagnosis (either based on the two raters' qualitative judgements or the quantitative procedure using the diagnostic cut-off derived from the AUROC analysis), the raters agreed in 9/10 cases. The Kappa value for the DSA procedure as a whole was therefore 0.8 (SE of Kappa = 0.19, 95% confidence interval 0.44 to 1.0) suggesting 'very good' inter-rater reliability. We also looked at the inter-rater reliability of the 123 individual +1, 0 or -1 ratings from Part 2 of this study. Both ratings were fully concordant for 87 numeric scores (the scores from raters 1 and 2 were 1/1, 0/0 or -1/-1), non-concordant for 30 and discordant for 6. This means that both raters agreed on 70.7% of the observations when agreement on 33.8% of the ratings would have been expected by chance. The Kappa value for all 123 DSA-based ratings combined was 0.56 (SE of Kappa = 0.06, 95% confidence interval 0.44 to 0.68), consistent with 'moderate' inter-rater agreement. The two raters’
scores for each item assessed in part 2 of the study and the Kappa-values of each individual item are shown in table 3.

Discussion

Our previous qualitative work has demonstrated that it is possible to describe characteristic conversational profiles of patients describing cognitive problems due to ND or FMD based on their interactional and linguistic contributions to initial encounters in a memory clinic\(^20,21\). However, in these descriptive studies, the conversation analysts who analysed video- and audio recordings of memory clinic encounters between neurologists, patients and (sometimes) accompanying persons were always aware of the patients’ medical diagnoses during the analytic process. The present study is the first to demonstrate that these linguistic and interactional features can be used diagnostically to predict diagnoses of ND or FMD made on the basis of standard medical criteria.

What is more, we show that qualitative assessments can be structured and likely medical diagnoses formulated using a Diagnostic Scoring Aid with a numeric diagnostic cut-off. The fact that the linguistic raters involved in this study had no expertise in the medical assessment of patients presenting with memory problems, together with the relatively high level of agreement between the two raters, suggests that the raters did not base their diagnostic predictions on an ill-defined hunch but on robust and objectifiable interactional observations.

The correct classification of 9/10 by one rater and 10/10 by a second independent rater, and the very good inter-rater reliability of the DSA-guided procedure as a whole, suggest that the addition of the structured observation of interactional features can make a significant contribution to screening processes for ND. Importantly, this is one of the few studies of cognitive screening ‘tools’ to include participants with FMD; most previous studies compared patients with memory impairment with healthy controls. The inclusion of a group of patients with memory complaints but no neurological disorder adds ecological validity to our findings. Compared to other studies set in clinical situations (such as a study exploring the screening potential of
the 6CIT brief cognitive test in primary care) our approach appears to have greater reliability and validity.

Interactional and linguistic observations may increase the confidence of non-expert clinicians to diagnose clear cases of FMD, enabling them to treat patients in primary care or to refer them on to services providing appropriate treatment for functional neurological disorders. Importantly, the interactional and linguistic observations contributing to the diagnosis of FMD may allow clinicians to provide more effective reassurance by allowing them to demonstrate to patients that they are displaying good memory function in interaction. It should be possible for clinicians to pick up these features during routine clinic encounters. Previous studies in patients presenting with epileptic or dissociative (non-epileptic) seizures have demonstrated that doctors can learn to change their history-taking style to optimise patients’ opportunities to demonstrate particular conversational behaviours and to make diagnostically useful interactional observations as they take a patient’s history. The DSA developed here could be modified and used in similar studies investigating whether clinicians can be trained to identify features, which have diagnostic value while talking to patients with memory problems.

This study has a number of limitations. First and foremost, we were only able to explore the potential of CA as a tool capable of predicting medical diagnoses in the setting of initial memory clinic encounters in a modest number of patients. Whilst the consecutive recruitment and the levels of statistical significance in between-group tests on a range of separate conversational features observed even in such a small patient group make it unlikely that our findings are spurious, it would be desirable to replicate our findings in a larger and more diverse group of patients. One particular limitation of our findings in this regard is that we excluded patients with depression and those with VCI from this first quantitative study of our method. These are important differential diagnoses, which will need to be picked up by screening procedures. Future larger studies will need to demonstrate that the inclusion of interactional and linguistic observations can contribute to screening or stratification procedures in which patients with these problems...
are allocated to the correct management pathways. We also recognised that the ND and FMD groups in this study were not age-matched since those with ND were significantly older. This is not surprising as the biggest risk factor for ND is increasing age. However as younger patients with memory concerns are increasingly referred to specialist memory clinics it is important to include and compare all age groups in studies of this nature. Furthermore the mean MMSE score of the ND group was lower than that of the FMD group (20.4 versus 28.2), reflecting the relatively late stage in the development of cognitive disorders at which patients are currently first referred to specialist services. Only four of the patients in the ND groups had MCI. An optimal screening tool for the earliest stages of ND would need to be capable of picking up patients with MCI and near normal MMSE scores). This means that confirmatory studies capturing more patients at an earlier stage of ND will be required before the method described here can be embedded in screening procedures. The MMSE scores of the MCI patients included in this study were just above the standard cut-off score of 23 for this test, indicating that, although unimpaired in their activities of daily living, they already had extensive global cognitive impairment. We do not know how effective this tool will be at distinguishing MCI from FMD. Because a clinical diagnosis of MCI refers to a very heterogeneous symptom profile, the distinction between MCI and FMD might be difficult and will require larger number of participants, along with prospective follow-up to investigate whether it can predict those who are at high risk of developing AD or other dementias.

We have only studied native English speakers; findings may have been different in patients speaking other languages or those using English as a second language. Although several different doctors were involved in the clinic conversations studied here, it would also be important to test this procedure in different clinical settings (for instance in community-based clinics, during home visits and in elderly-care settings). The fact that the ND group included patients with memory problems of different aetiologies should not be considered a weakness of this study. Although it is likely the method employed in this study could also be deployed to identify interactional
differences between different ND (such as Alzheimer’s disease or frontotemporal dementia) the fact that we were able to distinguish clearly between patients with a range of ND and those with FMD demonstrates its potential for screening or stratifying patient management.

We did not have access to investigations confirming clinical diagnosis with tests documenting the presence of amyloid or tau (PET or cerebrospinal studies), but this reflects current NICE guidelines[^1] and our ‘medical’ diagnoses were based on multidisciplinary assessment by experts including detailed neuropsychological testing and structural brain imaging as well as clinical follow up.

Future studies will need to demonstrate how much diagnostic value the observation of interactional features can add to conventional brief cognitive screening tools. A combined approach with an automated low cost, high-speed system to analyse speech will require the use of technology rather than Conversation Analysts. One way in which the research described here can be taken forward involves the computerised analysis of speech which has shown some promise in distinguishing AD, MCI and healthy controls[^16]. Early indications are that computerised speech analysis and machine learning algorithms can also be used to produce an automated system to pick up and evaluate the sort of interactional observations described here and can discriminate between ND and FMD[^22].

More immediately, the findings described here can be used in the training of clinicians working with patients with memory problems.

**Acknowledgements:** This article presents independent research funded by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (Grant Reference Number PB-PG-0211-24079). This is a summary of independent research carried out at the NIHR Sheffield Biomedical Research Centre (Translational Neuroscience). The

[^1]: [http://www.nice.org.uk/guidance/CG42](http://www.nice.org.uk/guidance/CG42)
views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

We would like to thank all participants and care partners who took part. We would also like to thank all participating clinicians working in the memory clinic (Aijaz Khan, Alice Brockington, Dilraj Sokhi and Muhammad Rafiq). We also acknowledge the support of the NIHR Sheffield Clinical Research Facility.

References

30. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-


Figure 1. CONSORT diagram showing recruitment to study. FMD- Functional Memory Disorder; DPD- Depressive Pseudo Dementia; ND- memory problems secondary to neurological disorders (ND) – this includes neurodegenerative dementias and mild cognitive impairment due to likely underlying neurodegenerative aetiology.

Figure 2: Area under the Receiver Operator Characteristic Curve
A ROC curve was constructed for the sample of 15 ND and 15 FMD cases.

Table 1 Diagnostic Scoring Aid (DSA)
For each interactional feature scores range between 1 and -1 (1: in favour of ND; 0: undecided or unable to rate; -1: in favour of FMD). There are 9/14 (unaccompanied/accompained) items to score, so the maximum score is 9/14 (unaccompanied/accompained) and minimum score -9/-14 (unaccompanied/accompained). A high score corresponds to a stereotypical ND description; a low score to a stereotypical FMD description.

Supplementary Table 1 Comparison of patients eligible for participation in the study who did and did not consent to have their clinic interaction recorded and analysed.

There is no significant difference in demographics or test scores between the people who consented and the people who did not consent. Addenbrooke’s Cognitive Examination; Patient Health Questionnaire-9 PHQ9- Generalised Anxiety Disorder scale 7 GAD7

Table 2 Demographic and neuropsychological results
ACE-R Addenbrooke’s Cognitive Examination. MMSE Mini Mental State Examination. PHQ9- Patient Health Questionnaire 9 item depression scale. GAD7General Anxiety Disorder 7-item Scale CF- Confrontational Naming. VPA- Verbal Paired Associates, P&PT-Pyramid & Palm Trees, Rey’s CF-Rey’s Complex Figure, SF- Semantic Fluency, PF - Phonemic Fluency, DS - Digit Span, VCA- Visuoconstructive Apraxia, TT-Token task, PM - Prose Memory. * 3 missing scores. + 1 missing score. #
Three missing scores from ND group (2 due to different protocol and one participant from part 2). Three missing scores from FMD group due to not attending appointments. Twenty participants with ND; comprised eight with AD, four with amnestic MCI, two 2 with vascular dementia, two with fronto temporal dementia, three with mixed AD and vascular and one unspecified dementia (without detailed neuropsychology).

Table 3: Diagnostic Scoring Aid (DSA) results
Profiles of 30 patients attending a specialist clinic with memory secondary to neurological disorders (ND) – this includes neurodegenerative dementias and mild cognitive impairment due to likely underlying neurodegenerative aetiology or functional memory disorder (FMD). 3a: Items rated in all
encounters / 3b: Items only rated in encounters also involving an 
accompanying person (AP).
For a full description of the items see supplementary appendix. Some items 
were unratable because a particular question was not asked (eg. “who is most 
concerned?”).

**Table 4 Diagnostic Scoring Aid (DSA) results of blinded analysis.**
Two independent linguistic raters (L1 and L2) of interactions with five patients 
with a medical diagnosis of FMD (5a) and five patients with medical diagnosis 
of ND (5b).
353 patients referred to the memory clinic were sent study information

148 eligible

112 consented

205 not eligible

36 did not consent

109 completed study

3 withdrew

Participated in the study

72 Non-dementia (e.g. stroke)
49 DNA/cancelled appointment
44 PIS not read/received
15 English not first language
12 Complex mental health problems
6 Did not have capacity
4 Alcohol/drug related problems
3 Follow up patient

30 FMD
26 DPD
22 Uncertain
19 ND
12 Vascular

16 ND
8 DPD
6 FMD
3 Uncertain
3 Vascular

1 FMD
1 Vascular
1 Uncertain
Differential Diagnosis Scoring Table

This scoring table was developed to assist in the differential diagnosis of patients with neurodegenerative diseases (ND) (e.g. Alzheimer's disease, other type of dementia and mild cognitive impairment) and those with functional memory disorder (FMD) (i.e. memory problems with a non-organic cause). Many of the features included in this scoring chart have been documented in the following sources:


Scores range between 1 and -1 (1: in favour of ND; 0: undecided or unable to rate; -1: in favour of FMD). There are 9/14 (unaccompanied/accompanied) items to score, so the maximum score is 9/14 (unaccompanied/accompanied) and minimum score -9/-14 (unaccompanied/accompanied). A high score corresponds to a stereotypical ND description; a low score to a stereotypical FMD description.
1. Configuration of interaction

Neurologist-patient

Neurologist-patient-accompanying persons

1. Is the patient accompanied?

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>No</td>
<td>-1</td>
</tr>
</tbody>
</table>

Instructions

This item simply records whether the patient is accompanied for their visit to the memory clinic. This can include family members, friends, carers and the like.

Patients with ND are more likely to be accompanied to the consultation, whereas those with FMD are generally more independent and likely to come alone.
2. Responding to neurologists' specific questions about memory problems

2. Specific question - "Who is most concerned about the memory problems?"

The neurologists in the study were instructed to ask a selection of pre-designed questions in order to aid comparison of the patients’ responses. The purpose of this question is to gauge who is aware of and concerned about the reported memory complaints.

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
<th></th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP states they are most concerned, patient no reply or &quot;I don’t know&quot;.</td>
<td>1</td>
<td>Score</td>
</tr>
<tr>
<td>Question (or equivalent) not asked, or mixed answer</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>&quot;It’s me&quot; “I am”</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

Instructions

When responding to this question the FMD tended to state that they were the person most concerned about the memory issues ("[It's] me"). For the ND a different set of responses occurred. Either, the AP offers they own opinion claiming they are most concerned (responsible for appointments, awareness issues); or the patient struggles to respond to this question sometimes failing to answer at all or saying "I don't know".
3. Specific question - "Can you give me an example of the last time your memory let you down?"

The purpose of this question is to gauge whether patients can offer specific and detailed accounts of a recent memory trouble. This question has a few variations or equivalents in which the neurologist might ask for the "a/most recent" example or provide a time frame "last week/month")

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient provides no response, a partial/incomplete answer or offer a general/routine problem (e.g. daily&quot;)</td>
<td>1</td>
</tr>
<tr>
<td>Patient provides a detailed and specific response about a recent occurrence</td>
<td>-1</td>
</tr>
</tbody>
</table>

**Instructions**

If the patient successfully provides a relevant and detailed example of a particular recent event, it is more likely that they have FMD.

In contrast, ND patients will either make no response, or hesitate over the beginnings of a response (e.g., "um" or "er"), or declare that they were unable to remember a specific occasion. These patients might offer an 'example', however, it will consist of a routine or common problem, rather than a specific incident (e.g., "happens all the time" or "it's daily").
3. Working and episodic memory exhibited within the present consultation

4. Ability to recall to recent episodic memory during interaction

Patient displays ability to recall previously mentioned or talked-about information from within the consultation itself.

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes (e.g. &quot;Like I said&quot;, &quot;as I said&quot;)</td>
</tr>
</tbody>
</table>

Instructions

Patients with FMD not only display intact episodic memory in relation to what the other (neurologist or AP) has said in the consultation, but also when recalling and repeating information they have previously voiced themselves (e.g. "Like I said", "as I said"). In contrast those with ND are often unable to display memory in this way in their consultations. They are often unable to retain information about what has been said even a few seconds earlier in the interaction, either by themselves or by the neurologist.

Note for this purpose compound questions are excluded, as they form a separate interactional feature.
5. Responding to compound questions

Exploring how patients respond to question with multiple-parts (e.g. Do you know the reasons why you've been referred to this clinic and, and who's more concerned?").

**Description of relevant evidence**

<table>
<thead>
<tr>
<th>Unable to attend to different parts of compound questions</th>
<th>1</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Can attend to different parts of compound questions</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

**Instructions**

FMD patients are more likely to be able to respond attend to multiple parts of a question and return to them even after extended turns at talk. In contrast, ND patients experience more difficulties responding to such question formulations, frequently replying to single components of the compound questions; they are less likely to be able to recall and respond to other aspects of the original question, resulting in the neurologist having to repeat the omitted parts of the question.
4. How patients respond to neurologists' questions and communication difficulties exhibited

6. Prevalence of "I don't know" responses (excluding head-turning sign)

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall based problems</td>
<td>1</td>
</tr>
<tr>
<td>New questions, not previously considered</td>
<td>-1</td>
</tr>
</tbody>
</table>

Instructions

FMD patients generally respond verbally with "I don't know" only rarely and these are linked to questions that elicit their "expectations" for the visit. These patients exhibit uncertainty that suggests that they have not previously considered the matter asked about and were unsure of the answer. In short the problem is not suggestive of a recall issue.

In contrast for ND patients these utterances occur throughout the consultation. These utterances may take a verbal form of "I don't know" or equivalent phrases.
7. Patients' elaborations and length of turns at talk

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short, 'literal' answers</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Long responses, that provide extra detail</td>
<td>-1</td>
</tr>
</tbody>
</table>

**Instructions**

FMD patients are more likely to volunteer unsolicited details (going beyond the parameter of the original question) that give additional, appropriate, relevant and related information – in short, they elaborate their responses. ND patients tend to offer answers that are brief, and restricted to the parameters of original questions.
8. Repetition

**Description of relevant evidence**

<table>
<thead>
<tr>
<th>More</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Less</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>-1</td>
</tr>
</tbody>
</table>

**Instructions**

ND more prone to repetition of own words and those of others without marking it as such.
9. Production of talk

Focusing particularly on whether a patient is able to reply promptly to neurologists’ questions.

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Struggle to reply to questions, communication difficulties</td>
<td>1</td>
</tr>
<tr>
<td>Able to provide answers when asked,</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>-1</td>
</tr>
</tbody>
</table>

Instructions

In general the speech production of FMD patients is close to that seen in ordinary conversation. These patients display fewer speech production difficulties and are also more likely to give extended/expanded responses. Disfluent talk characterised by long pauses, delayed responses, hesitation, self-interrupts (staccato-like), repetition (of word from self and others), incomplete, aborted or unfinished utterances.
5. Triadic features

These should only be scored for consultations in which an AP is present throughout

10. If the patient is accompanied, what is the main interactional contribution/role of the accompanying persons (AP)?

### Description of relevant evidence

<table>
<thead>
<tr>
<th>Role Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP acts as patient's representative or spokesperson</td>
<td>1</td>
</tr>
<tr>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>AP role limited to confirmation checks and second opinions</td>
<td>-1</td>
</tr>
</tbody>
</table>

### Instructions

This item relates to the amount and type of contributions offered by AP (if they are present). If the AP speaks throughout the consultation, frequently augmenting the patient's responses and providing specific/detailed examples of problems the patient is more likely to have an ND diagnosis. In contrast, for FMD the AP often has a more restricted role in the interaction, which tends to be limited to providing 'confirmation checks' to a patient's answers or at a designated point of the consultation the neurologist might specifically engage the AP to seek a second opinion about the memory problems raised.
11. Presence of head-turning sign (excluding verbal "I don't know" replies)

Description of relevant evidence

<table>
<thead>
<tr>
<th>Patient defers answering to AP by turning to them</th>
<th>1</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Able to answer most questions, turn for confirmation only</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

Instructions

When a patient with ND is struggling to answer they might physically turn to an AP (if present). This non-verbal/embodied display defers or passes the responsibility of responding to the AP (implicitly suggesting an "I don't know" response). This is called a 'head-turning sign'. In contrast patients with FMD are less likely to respond to questions in this way as fewer of the neurologists questions cause such difficulties (if an AP is present).

This feature excludes verbal "I don't know" replies (see feature 6 above).
12. Disagreements and discrepancy between patient responses and AP

<table>
<thead>
<tr>
<th>Description of relevant evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yes</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

**Instructions**

If an AP is present is there evidence of disagreement sequences? For comparative purposes these sequences should take the following format:

- Question by neurologist which seeks a yes/no response (e.g. do you cook at home)
- Answer from the patient matching the yes/no requirements of the question (e.g. "yes"). Often the patient's response indicates the content of the question is "no problem" for them.
- The third turn in the sequence will involve the AP contradicting or correcting the patient's reply ("No, you used to/you don't do that anymore" etc)
- The patient then adjusts their original response (e.g. "no" or "that's true" etc)

In terms of identifying this pattern it is more common in interactions involving patients with ND diagnoses, than those with FMD.
13. Word searches

Description of relevant evidence

<table>
<thead>
<tr>
<th>Display 'word search' difficulties <em>during</em> consultation, AP provide 'missing' information</th>
<th>1</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report 'word search' difficulties <em>in the past</em></td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

Instructions

The item refers to a particular type of memory complaint. This feature is often mentioned as a sign of the memory problems the patient with FMD has noticed previously (reported memory complaints), whereas the ND patients 'display' this type of memory problem during the course of the consultation itself and might seek the assistance of an AP to fill in the blank.
14. Responding to personal questions

This refers to patients' abilities to respond to non-test questions in which the answer is unknown or generally unavailable to the neurologist (including age, education, work details, marital history etc).

### Description of relevant evidence

<table>
<thead>
<tr>
<th>Evidence of difficulties answering these questions</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can answer these questions relatively easily</td>
<td>-1</td>
</tr>
</tbody>
</table>

### Instructions

These questions are more likely to cause difficulties for ND patients, who might be unable to recall the salient details required. As such they might require the help of AP or be corrected by them. FMD find these questions more straightforward.
<table>
<thead>
<tr>
<th>Diagnostic feature</th>
<th>More suggestive of Dementia</th>
<th>More suggestive of functional memory problems</th>
<th>Score (0: if uncertain)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Who attends the memory clinic?</strong></td>
<td></td>
<td></td>
<td>1, 0 or -1</td>
</tr>
<tr>
<td>1. Is the patient accompanied?</td>
<td>Yes (Accompanied persons (AP) include family or friends)</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>2. &quot;Who is most concerned about the memory problems?&quot;</td>
<td>No reply from patient, &quot;I don’t know&quot; or AP states they are most concerned</td>
<td>The patient.</td>
<td></td>
</tr>
<tr>
<td>3. &quot;Can you give me an example of the last time your memory let you down?&quot;</td>
<td>No response, partial or incomplete answer, or offer a routine common problem (it's daily)</td>
<td>Provides detailed specific example</td>
<td></td>
</tr>
<tr>
<td><strong>Working and episodic memory exhibited within the present consultation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Ability to recall to recent episodic memory during interaction</td>
<td>Not demonstrated</td>
<td>Repetitions marked by phrases such as &quot;Like I said&quot; or &quot;as I said&quot;</td>
<td></td>
</tr>
<tr>
<td>5. Responding to compound questions</td>
<td>Unable to attend to different parts of compound questions</td>
<td>Can attend to different parts of compound questions</td>
<td></td>
</tr>
<tr>
<td><strong>How patients respond to neurologists’ questions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Prevalence of verbal &quot;I don’t know&quot; responses</td>
<td>Frequent</td>
<td>Infrequent, relate to new issues not previously considered</td>
<td></td>
</tr>
<tr>
<td>7. Patients’ elaborations and length of turns at talk</td>
<td>Short, literal answers</td>
<td>Long responses, sharing of additional, unsolicited details</td>
<td></td>
</tr>
<tr>
<td>8. Repetition</td>
<td>More frequent repetition of own and others' utterances</td>
<td>Less frequent, marked as repetitions</td>
<td></td>
</tr>
<tr>
<td>9. Production of talk</td>
<td>Struggle to reply to questions, communication difficulties</td>
<td>Able to reply when questioned</td>
<td></td>
</tr>
<tr>
<td><strong>Features to rate if patient accompanied</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. What is the main role of AP?</td>
<td>AP acts as patient's representative or AP role limited to confirming information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>role of the accompanying persons (AP)?</td>
<td>spokesperson</td>
<td>accurate and offering second opinions</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------</td>
<td>---------------------------------------</td>
<td></td>
</tr>
<tr>
<td>11. Presence of head-turning sign (embodied “I don’t know”)</td>
<td>When struggling to answer question patient turns to AP and defers the answer to them</td>
<td>Can answer most questions by themselves</td>
<td></td>
</tr>
<tr>
<td>12. Disagreements</td>
<td>AP’s interject to disconfirm and correct the accuracy of the patient’s responses. Frequent.</td>
<td>Limited evidence of explicit disagreements between patient and AP</td>
<td></td>
</tr>
<tr>
<td>13. Word searches</td>
<td>Display ‘word search’ difficulties during consultation. AP provide ‘missing’ information</td>
<td>Report ‘word search’ difficulties in the past</td>
<td></td>
</tr>
<tr>
<td>14. Responding to personal questions</td>
<td>Evidence of difficulties answering these questions, requiring help from AP to fill in blanks.</td>
<td>Can answer these questions relatively easily and with little hesitation</td>
<td></td>
</tr>
</tbody>
</table>

**Total score**
<table>
<thead>
<tr>
<th></th>
<th>FMD (n=20)</th>
<th>ND (n=20)</th>
<th>Normative Mean</th>
<th>Cut Off</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>57.25 (+/-1.82)</td>
<td>64.2 (+/- 2.13)</td>
<td></td>
<td></td>
<td>p=0.0018</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td>60%</td>
<td>60%</td>
<td>n/s</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACE-R</strong></td>
<td>92.47 (+/-1.18)</td>
<td>65.53 (+/- 4.78)</td>
<td>88</td>
<td>p&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td><strong>MMSE</strong></td>
<td>28.88 (+/-0.19)*</td>
<td>20.44 (+/- 1.71)+</td>
<td>28.88 (1.28)</td>
<td>26.32</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td><strong>PHQ9</strong></td>
<td>5.68 (0.96)</td>
<td>4.47 (+/- 1.67)</td>
<td>5</td>
<td></td>
<td>p=0.51</td>
</tr>
<tr>
<td><strong>GAD7</strong></td>
<td>4.84 (+/- 1.04)</td>
<td>4.47 (+/- 1.23)</td>
<td>5</td>
<td></td>
<td>p=0.82</td>
</tr>
<tr>
<td><strong>CF</strong></td>
<td>19.82 (+/- 0.1)*</td>
<td>17.59 (+/- 0.73)#</td>
<td>19.65 (0.63)</td>
<td>18.39</td>
<td>p=0.0049</td>
</tr>
<tr>
<td><strong>VPA</strong></td>
<td>16.76 (+/- 0.66)#</td>
<td>6.88 (+/- 1.14.)*</td>
<td>14.81 (3.76)</td>
<td>7.29</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td><strong>P&amp;PT</strong></td>
<td>51.18 (+/- 0.18)#</td>
<td>45.81 (+/- 1.94)#</td>
<td>51.23 (0.82)</td>
<td>49.59</td>
<td>p=0.0793</td>
</tr>
<tr>
<td><strong>Rey's CF</strong></td>
<td>33.165 (+/-0.45)*</td>
<td>23.75 (+/- 2.52)#</td>
<td>33.70 (2.30)</td>
<td>29.1</td>
<td>p&lt;0.0004</td>
</tr>
<tr>
<td><strong>SF</strong></td>
<td>50.94 (+/- 3.21)#</td>
<td>28.82 (+/- 4.09)#</td>
<td>59.81 (13.17)</td>
<td>33.47</td>
<td>p&lt;0.0002</td>
</tr>
<tr>
<td><strong>PF</strong></td>
<td>40.47 (+/- 2.79)#</td>
<td>22.25 (+/- 3.86)#</td>
<td>45.58 (12.05)</td>
<td>21.48</td>
<td>p&lt;0.0005</td>
</tr>
<tr>
<td><strong>Digit Span</strong></td>
<td>6.65 (+/- 0.3)*</td>
<td>5.06 (+/- 0.47)#</td>
<td>6.76 (1.48)</td>
<td>3.8</td>
<td>p=0.0071</td>
</tr>
<tr>
<td><strong>VCA</strong></td>
<td>13.24 (+/- 0.18)#</td>
<td>10.65 (+/- 0.79)#</td>
<td>13.77 (0.51)</td>
<td>12.75</td>
<td>p=0.0032</td>
</tr>
<tr>
<td><strong>TT</strong></td>
<td>34.82 (+/- 0.28)*</td>
<td>28.0 (+/- 1.62)#</td>
<td>34.67</td>
<td>1.03</td>
<td>p&lt;0.0002</td>
</tr>
<tr>
<td><strong>PM</strong></td>
<td>14.65 (+/- 0.87)*</td>
<td>5.5 (+/- 0.9)*</td>
<td></td>
<td></td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>Item</td>
<td>Description*</td>
<td>A: Typical of NDD</td>
<td>B: Typical of FMD</td>
<td>Number of NDD cases categorized A/B** (n=15)</td>
<td>Number of FMD cases categorized A/B** (n=15)</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Is the patient accompanied</td>
<td>Yes</td>
<td>No</td>
<td>14/1</td>
<td>6/9</td>
</tr>
<tr>
<td>2</td>
<td>Who is most concerned?</td>
<td>Others</td>
<td>Patient themselves</td>
<td>6/1</td>
<td>0/9</td>
</tr>
<tr>
<td>3</td>
<td>Specific example of memory failure</td>
<td>No or partial / incomplete answer or offers a general / routine problem</td>
<td>Detailed and specific response about a recent occurrence</td>
<td>11/0</td>
<td>1/11</td>
</tr>
<tr>
<td>4</td>
<td>Ability to recall recent episodic memory during interaction memory in interaction</td>
<td>Patient unable to recall earlier talk</td>
<td>Patient able to recall earlier talk (&quot;Like I said&quot;)</td>
<td>11/3</td>
<td>0/8</td>
</tr>
<tr>
<td>5</td>
<td>Responding to compound questions</td>
<td>Unable to attend to different parts of compound questions</td>
<td>Can attend to different parts of compound questions</td>
<td>7/1</td>
<td>3/7</td>
</tr>
<tr>
<td>6</td>
<td>Prevalence of &quot;I don't know&quot; verbal responses</td>
<td>Indicates recall-based problems</td>
<td>Response to unexpected questions</td>
<td>11/1</td>
<td>1/14</td>
</tr>
<tr>
<td>7</td>
<td>Patients’ elaborations and length of turns</td>
<td>Short, 'literal' answers</td>
<td>Long responses, that provide extra detail</td>
<td>9/6</td>
<td>0/11</td>
</tr>
<tr>
<td>8</td>
<td>Repetition</td>
<td>More frequent</td>
<td>Less frequent</td>
<td>10/3</td>
<td>1/11</td>
</tr>
<tr>
<td>9</td>
<td>Production of talk</td>
<td>Struggle to reply to questions, communication difficulties</td>
<td>Able to provide answers when asked</td>
<td>7/2</td>
<td>1/13</td>
</tr>
<tr>
<td>Item</td>
<td>Description*</td>
<td>A: Typical of NDD</td>
<td>B: Typical of FMD</td>
<td>Number of NDD cases categorized A/B** (n=14)</td>
<td>Number of FMD cases categorized A/B** (n=6)</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>10</td>
<td>Main interactional contribution/role of the AP</td>
<td>AP acts as patient's representative or spokesperson</td>
<td>AP's role limited to confirmation checks and second opinions</td>
<td>9/1</td>
<td>1/5 n = 6</td>
</tr>
<tr>
<td>11</td>
<td>Presence of head-turning sign (excluding verbal &quot;I don't know&quot; replies)</td>
<td>Patient defers answering to AP by turning to them</td>
<td>Able to answer most questions, turn for confirmation only</td>
<td>10/4</td>
<td>3/3</td>
</tr>
<tr>
<td>12</td>
<td>Disagreements between patient and AP</td>
<td>Present</td>
<td>Not present</td>
<td>13/1</td>
<td>2/4</td>
</tr>
<tr>
<td>13</td>
<td>Word searches</td>
<td>Displays 'word search' difficulties during consultation, AP provides 'missing' information</td>
<td>Report 'word search' difficulties in the past but does not display 'word search'</td>
<td>3/1</td>
<td>3/3</td>
</tr>
<tr>
<td>14</td>
<td>Responding to personal questions</td>
<td>Evidence of difficulties answering these questions</td>
<td>Can answer these questions relatively easily</td>
<td>6/5</td>
<td>0/6</td>
</tr>
</tbody>
</table>
### 4a: Patients with medical diagnosis of FMD

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DSA Rating score</strong></td>
<td>L1</td>
<td>L2</td>
<td>L1</td>
<td>L2</td>
<td>L1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>8</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>10</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>-5</td>
<td>-8</td>
<td>-5</td>
<td>-6</td>
<td>-5</td>
</tr>
<tr>
<td><strong>Predicted Dx</strong></td>
<td>FMD</td>
<td>FMD</td>
<td>FMD</td>
<td>FMD</td>
<td>FMD</td>
</tr>
<tr>
<td><strong>Actual DX</strong></td>
<td>FMD</td>
<td>FMD</td>
<td>FMD</td>
<td>FMD</td>
<td>FMD</td>
</tr>
</tbody>
</table>

### 4b: Patients with a medical diagnosis of ND

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DSA Rating score</strong></td>
<td>L1</td>
<td>L2</td>
<td>L1</td>
<td>L2</td>
<td>L1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>11</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4</td>
<td>-1</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td><strong>Predicted Dx</strong></td>
<td>ND</td>
<td><strong>FMD</strong></td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td><strong>Actual DX</strong></td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>