It has long been recognised that there is a trade-off between speed and accuracy in the performance of voluntary movements, such that more accurate movements are performed more slowly.\(^1\) This speed/accuracy trade-off may perhaps apply to any task, and since speed is inversely proportional to time it may also be formulated as a time/accuracy trade-off, longer times being required for greater accuracy.

The utility of cognitive screening instruments (CSIs) for the diagnosis of dementia and lesser degrees of cognitive impairment may be indicated by a number of summary parameters, probably the most familiar of which are sensitivity and specificity. It is well recognised that there is always a balance or trade-off to be struck between test sensitivity and specificity, with the chosen test cut-off being determined by the needs of the particular clinical situation. Dr Larner investigates whether there is a trade-off between CSI diagnostic accuracy and administration time, or, in other words, whether shorter CSIs are less accurate than longer ones which may sample more cognitive domains and/or in greater depth.

When using cognitive screening instruments (CSIs) there is always a balance to be struck between test sensitivity and specificity, with the chosen test cut-off being determined by the needs of the particular clinical situation. Dr Larner investigates whether there is a trade-off between CSI diagnostic accuracy and administration time, or, in other words, whether shorter CSIs are less accurate than longer ones which may sample more cognitive domains and/or in greater depth.

<table>
<thead>
<tr>
<th>CSI</th>
<th>Setting</th>
<th>No.</th>
<th>Dementia prevalence</th>
<th>M:F (% male)</th>
<th>Age range (years)</th>
<th>Ref no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE, MMP</td>
<td>Cognitive function clinic</td>
<td>225</td>
<td>21%</td>
<td>130:95 (58)</td>
<td>20–86 (median 62)</td>
<td>14</td>
</tr>
<tr>
<td>ACE</td>
<td>Cognitive function clinic</td>
<td>285</td>
<td>49%</td>
<td>147:138 (52)</td>
<td>N/A</td>
<td>15</td>
</tr>
<tr>
<td>ACE-R</td>
<td>Cognitive function clinic</td>
<td>243</td>
<td>35%</td>
<td>135:108 (56)</td>
<td>24—85 (mean 60)</td>
<td>16</td>
</tr>
<tr>
<td>6CIT</td>
<td>Cognitive function clinic</td>
<td>245</td>
<td>20%</td>
<td>124:121 (51)</td>
<td>16–94 (median 63)</td>
<td>17</td>
</tr>
<tr>
<td>DemTect</td>
<td>Cognitive function clinic</td>
<td>111</td>
<td>52%</td>
<td>52:59 (47)</td>
<td>23–86 (median 59)</td>
<td>18</td>
</tr>
<tr>
<td>MoCA</td>
<td>Cognitive function clinic</td>
<td>150</td>
<td>24%</td>
<td>93:57 (62)</td>
<td>20–87 (median 61)</td>
<td>19</td>
</tr>
<tr>
<td>TYM</td>
<td>Cognitive function clinic and old age psychiatry memory clinic</td>
<td>224</td>
<td>35%</td>
<td>130:94 (58)</td>
<td>20–90 (mean 63)</td>
<td>20</td>
</tr>
</tbody>
</table>

CSI = cognitive screening instrument; MMSE = Mini-Mental State Examination; MMP = Mini-Mental Parkinson; ACE = Addenbrooke’s Cognitive Examination; ACE-R = Addenbrooke’s Cognitive Examination-Revised; 6CIT = Six-Item Cognitive Impairment Test; MoCA = Montreal Cognitive Assessment; TYM = Test Your Memory test.

Table 1. Study demographics

which also need to be factored into the decision as to which CSI(s) to use. For example, in day-to-day clinical practice in the secondary care setting, instruments such as the ADAS-Cog,\(^2\) the Mattis Dementia Rating Scale,\(^3\) and the Neuropsychological Assessment Battery\(^4\) are generally considered too long for routine use, being largely reserved for research settings. In primary care settings, instruments such as the Mini-Mental State Examination
(MMSE) may be considered too long, and briefer instruments have been recommended for this setting such as the Six-item Cognitive Impairment Test (6CIT), the Memory Impairment Screen, and the Mini-Cog.5 Patient self-administered tests such as the Test Your Memory (TYM) test may be of particular value in situations where clinician time is limited, precluding performance of clinician-administered tests.9

A question arising from these considerations is whether there is a trade-off between CSI diagnostic accuracy and administration time, or, in other words, whether shorter CSIs are less accurate than longer ones which may sample more cognitive domains (eg, attention, executive function, visuo-spatial function) and/or in greater depth. Conversely, another possibility is that longer tests might be less accurate than shorter ones, if patients should become fatigued and less cooperative with prolonged testing.

The speed/accuracy question might be examined by comparing test diagnostic accuracy against duration of test administration. Since the latter parameter is not routinely measured in the clinical setting, more easily accessible surrogate measures might be used, such as the overall test score or by counting the total number of items/questions in the test (in each case higher = longer). Overall test scores may be the same despite different numbers of test items (eg, MMSE, Montreal Cognitive Assessment (MoCA)) depending on how each item is scored.

**Study aim**
The aim of the study presented here was to examine the relationship between overall test accuracy for the diagnosis of dementia versus no dementia of several commonly used CSIs, and the aforementioned surrogate measures of test administration time to see if there is a trade-off between speed and accuracy. The CSIs examined were: the MMSE,7 the Mini-Mental Parkinson (MMP);8 the Addenbrooke’s Cognitive Examination (ACE);9 the Addenbrooke’s Cognitive Examination-Revised (ACE-R);10 the 6CIT;11 DemTect;12 the MoCA,13 and the TYM test.6 None of these CSIs is a timed test, although some (MMP, ACE, ACE-R, DemTect and MoCA) include a timed verbal fluency item.

**Materials and methods**
Data from seven previous pragmatic diagnostic accuracy studies which examined the MMSE,14 MMP (used in this instance as a general cognitive screening instrument, not for cognitive impairments in Parkinson’s disease patients),14 ACE,15 ACE-R,16 6CIT,17 DemTect,18 MoCA,19 and TYM20 were re-analysed. All studies were undertaken in dedicated memory clinics based in secondary care settings, and all used a standardised methodology.21 Study details are shown in Table 1. In each of these studies, the criterion diagnosis was established by the judgement of an experienced clinician based on widely-accepted clinical diagnostic criteria. Because these were clinic-based pragmatic studies, there was no normal control group; the non-demented cases consisted of patients with at minimum subjective memory impairment, as well as patients with mild cognitive impairment insufficient to mandate a dementia diagnosis.

Overall test accuracy, defined as the sum of true positives and true negatives (MMSE) may be considered too long, and briefer instruments have been recommended for this setting such as the Six-item Cognitive Impairment Test (6CIT), the Memory Impairment Screen, and the Mini-Cog.5 Patient self-administered tests such as the Test Your Memory (TYM) test may be of particular value in situations where clinician time is limited, precluding performance of clinician-administered tests.9

A question arising from these considerations is whether there is a trade-off between CSI diagnostic accuracy and administration time, or, in other words, whether shorter CSIs are less accurate than longer ones which may sample more cognitive domains (eg attention, executive function, visuo-spatial function) and/or in greater depth. Conversely, another possibility is that longer tests might be less accurate than shorter ones, if patients should become fatigued and less cooperative with prolonged testing.

The speed/accuracy question might be examined by comparing test diagnostic accuracy against duration of test administration. Since the latter parameter is not routinely measured in the clinical setting, more easily accessible surrogate measures might be used, such as the overall test score or by counting the total number of items/questions in the test (in each case higher = longer). Overall test scores may be the same despite different numbers of test items (eg, MMSE, Montreal Cognitive Assessment (MoCA)) depending on how each item is scored.

**Study aim**
The aim of the study presented here was to examine the relationship between overall test accuracy for the diagnosis of dementia versus no dementia of several commonly used CSIs, and the aforementioned surrogate measures of test administration time to see if there is a trade-off between speed and accuracy. The CSIs examined were: the MMSE,7 the Mini-Mental Parkinson (MMP);8 the Addenbrooke’s Cognitive Examination (ACE);9 the Addenbrooke’s Cognitive Examination-Revised (ACE-R);10 the 6CIT;11 DemTect;12 the MoCA,13 and the TYM test.6 None of these CSIs is a timed test, although some (MMP, ACE, ACE-R, DemTect and MoCA) include a timed verbal fluency item.

**Materials and methods**
Data from seven previous pragmatic diagnostic accuracy studies which examined the MMSE,14 MMP (used in this instance as a general cognitive screening instrument, not for cognitive impairments in Parkinson’s disease patients),14 ACE,15 ACE-R,16 6CIT,17 DemTect,18 MoCA,19 and TYM20 were re-analysed. All studies were undertaken in dedicated memory clinics based in secondary care settings, and all used a standardised methodology.21 Study details are shown in Table 1. In each of these studies, the criterion diagnosis was established by the judgement of an experienced clinician based on widely-accepted clinical diagnostic criteria. Because these were clinic-based pragmatic studies, there was no normal control group; the non-demented cases consisted of patients with at minimum subjective memory impairment, as well as patients with mild cognitive impairment insufficient to mandate a dementia diagnosis.

Overall test accuracy, defined as the sum of true positives and true
negatives divided by the total number of patients tested (= output or effect, hence the dependent variable, y axis), was plotted against overall test score and against the number of items/questions in the test (inputs or causes, hence independent variables, x axis). Correlations between these parameters were calculated.

Results
Table 2 shows the approximate administration time of the examined CSIs, along with the surrogate measures of administration time (total test score, total number of test items/questions), along with overall diagnostic accuracy of each test obtained from the pragmatic diagnostic accuracy studies. These data are plotted in Figures 1 and 2 respectively.

There was a positive correlation between accuracy and total test score ($r=0.61$), a correlation which may be classified as moderate but which did not reach statistical significance ($t=1.86$, df=6, $p>0.1$).

Likewise, there was a positive correlation between accuracy and total number of items/questions ($r=0.72$) which may be classified as a high correlation and which did reach statistical significance ($t=2.52$, df=6, $p<0.05$).

Discussion
In this study, data from seven pragmatic diagnostic accuracy studies of CSIs$^{14-20}$ were analysed to examine the relationship between overall test diagnostic accuracy and surrogate measures of test administration time. Although overall test accuracy for the diagnosis of dementia versus no dementia was similar for all tests (range 0.78–0.89), nonetheless there were positive correlations between accuracy and the measures of administration time, though this was only statistically significant for the total number of items/questions and not for total test score. These observations suggest that tests with more items (ie longer tests) are more accurate, or in other words that there is a trade-off between speed and accuracy, the speedier (ie shorter) tests being less accurate. Incorrect diagnoses, the sum of false positives and false negatives (which if divided by the total number of patients tested = error rate or inaccuracy or $[1 – \text{accuracy}]$), may be twice as frequent in the least accurate test (DemTect, 22/100) versus the most accurate (ACE-R, 11/100).

Inter-study comparisons are problematic, notwithstanding the consistency of study protocols$^{21}$ and authorship of the studies examined here. Potential shortcomings of the analysis were the different sample characteristics for each of the seven studies (Table 1), particularly in terms of sample size.
(range 111–285) and dementia prevalence (21–52%). Sample size calculations were not performed, but a pragmatic approach to sample size estimates has suggested that normative ranges for sample sizes may be calculated for common research designs, with anything in the range of 25–400 being acceptable. A second limitation relates to the progressive fall in dementia prevalence that has been seen in patients referred to the clinic (and hence in the reported studies) over the past decade, perhaps reflecting national dementia directives which may have had the (unintended?) effect of prompting more individuals with subjective memory impairment to attend for assessment.

In conclusion, the findings of this study suggest that there is indeed a trade-off for CSIs between duration of administration and diagnostic accuracy, and therefore that the traditional policy of long outpatient clinic appointments (45–60 minutes) for patients with cognitive complaints is appropriate to permit adequate time for the administration of longer tests (if need be) and hence accurate patient diagnosis. Accordingly, any attempt to shorten outpatient clinic appointment times in order to increase patient throughput (is total number of patients seen) would not seem to be justified, unless and until such time as shorter CSIs can be shown to be as diagnostically accurate as longer tests.

Dr Larner is Consultant Neurologist at the Cognitive Function Clinic, Walton Centre for Neurology and Neurosurgery, Liverpool

Declaration of interests
There are no conflicts of interest declared.

References
22. Norman G, Monteiro S, Salama S. Sample size calculations: should the emperor’s clothes be off the peg or made to measure? BMJ 2012;345:e5728.