Chapter 2

The accuracy of family physicians’ dementia diagnoses at different stages of dementia: a systematic review

Pim van den Dungen
Harm W.M. van Marwijk
Henriëtte E. van der Horst
Eric P. Moll van Charante
Janet MacNeil Vroomen
Peter M. van de Ven
Hein P.J. van Hout

Published in: International Journal of Geriatric Psychiatry, 2012; 4: 342-354
Chapter 2

Summary

Objective: Optimising care for dementia patients and their informal carers is imperative in light of the impending dementia epidemic. An important aspect of care is accurate recognition and diagnosis of dementia. The aim of this review is to estimate family physicians’ diagnostic accuracy at the different stages of dementia.

Methods: Pubmed, Embase, CINAHL, PsycINFO and the Cochrane Library were searched for articles comparing family physicians’ ‘dementia’ and ‘cognitive impairment’ diagnoses in the primary care setting to reference standard dementia diagnoses.

Results: Data from six cross-sectional studies of moderate to reasonable methodological quality were extracted for the analysis. One study considered the accuracy of family physicians’ recollected diagnoses and three studies focussed on documented diagnoses. In these four studies, the sensitivity of family physicians’ combined diagnostic categories of ‘cognitive impairment’ together with ‘dementia’ was 0.48 - 0.67 for mild dementia and 0.76 - 0.85 for moderate to severe dementia. The sensitivity of their diagnostic category ‘dementia’ alone, was 0.14 - 0.33 for mild and 0.28 - 0.61 for moderate to severe dementia. Specificity was excellent for all severity stages in both comparisons.

Three studies examined the accuracy of family physicians’ judgement of cognition during consultation. Compared to the studies on recollection and documentation, these studies reported higher sensitivity and lower specificity.

Conclusion: Many individuals with dementia are not recognised or not diagnosed as such; particularly mild dementia is under-diagnosed. Collaboration within primary care and education focussing both on knowledge and attitude are recommended to improve the accuracy of family physicians’ dementia diagnosis.
The accuracy of family physicians’ dementia diagnoses: a systematic review

Introduction

A recent estimate of dementia prevalence was reported to be 35.6 million worldwide. This number is expected to quadruple to 115 million by 2050. In their effort to prepare for this dementia epidemic, national and international governments have recognised the need to gain more insight into the epidemiology of dementia. In gatekeeping primary health care systems such as in the Netherlands, knowing the accuracy of family physicians’ (FPs) dementia diagnoses is vital, as FPs play a pivotal role in the recognition and diagnosis of dementia.

For patients and families, diagnosing dementia has the potential to end of a period of uncertainty. Surveys among healthy elderly and individuals with (mild) dementia indicate that the vast majority (would) want to know this diagnosis. Communicating the diagnosis at early stages empowers the individual with dementia to comprehend the diagnosis and to participate in the planning of their own care. The latter may help the informal caregiver to prepare for the decisions he or she may face later on in the process when the person with dementia is no longer capable of deciding for him- or herself.

Informal caregivers of individuals with dementia are often heavily burdened and have an increased risk of developing a depressive disorder. For informal caregivers, the dementia diagnosis may allow psychoeducation, support and access to services and if needed care. Medical reasons for straightforward diagnosis of cognitive impairment and dementia are 1) the identification of potentially reversible causes and 2) the important prognostic and pharmacotherapeutic consequences a (nosological) dementia diagnosis may have.

Current evidence on the accuracy of FPs’ dementia diagnoses is conflicting. An earlier review on this subject also included studies targeting patients referred for specialist diagnosis and therefore does not strictly reflect FPs’ accuracy in the primary care setting. The methodological differences within the included studies were not considered in detail. In order to gain more insight into the accuracy of FPs’ dementia diagnoses in the primary care setting we performed a systematic review and critical appraisal of the literature on this subject. Considering the importance of early diagnosis, we compared the accuracy of diagnosis in individuals with mild dementia to the accuracy in individuals with moderate to severe dementia.
Chapter 2

Methods

We performed a systematic literature search and critical appraisal following the guidelines of the Cochrane Diagnostic Reviewers’ Handbook and the PRISMA statement for systematic reviews\textsuperscript{26,27}. The full review protocol can be obtained by corresponding with the first author.

Literature search

We performed our search on the 20\textsuperscript{th} of October 2009. We search Pubmed starting from 1951; Embase.com from 1974; CINAHL from 1982; PsycINFO from 1887 and the entire Cochrane Library. The search terms used were “dementia, diagnosis and general practitioner”. Synonyms and related terms were also used. We applied both controlled thesaurus terms and words in titles and abstracts to optimize the sensitivity of the search. Limits were not used. Additionally, the references of relevant articles were searched. A detailed description of the initial search is located in the appendix.

Eligibility criteria

The study inclusion and exclusion criteria are listed below.

Study inclusion criteria

- The target population consisted of individuals aged ≥ 55 living at home or in a home for the elderly.

- Study designs could comprise of cross-sectional studies and of prospective or longitudinal studies. Cross-sectional studies could measure FPs’ judgement on their complete elderly patient population or they could have FPs’ decide the cognitive functioning of elderly patients presenting for consultation.

- The method of case ascertainment, assuring or objectifying that dementia was (or was not) diagnosed by the FP, had to be standardised and clearly described. Case ascertainment could be based on medical record review or FPs could be presented with a form (for individual patients) or list (for all their elderly patients) on which they could document their judgement of the cognitive function of their patients.
• The index test (i.e. the FPs’ diagnosis) had to be compared with a reference standard diagnosis on the presence or absence of dementia. This could be done either for the complete population under study or for a random sample of it.

The index test was the FPs’ judgement on the presence or absence of ‘dementia’ in his/her patients in routine care and the FPs’ judgement on the presence or absence of ‘cognitive impairment’. We included this second diagnostic category because of the hypothesis that FPs do notice general mental dysfunction but are reluctant to specifically label such problems.

As a reference test, we accepted diagnosis by a multidisciplinary team or a specially trained physician, such as a geriatrician, or diagnosis based on a structured assessment for example, the CAMDEX and explicitly adhering to DSM III to IVR or CAMDEX criteria. For dementia severity staging, criteria of the DSM, Clinical Dementia Rating Scale (CDRS) and CAMDEX were considered valid.

**Study exclusion criteria**

• Studies in which a selective patient sample was analysed were excluded if information required for extrapolation of data to the general primary care practice population could not be extracted from the primary article or obtained otherwise.

• Studies in which a specific diagnostic intervention was performed, e.g. studies assessing the entire elderly population with a cognitive screening instrument to improve diagnostic accuracy of the FP.

• Studies on Mild Cognitive Impairment (MCI) were excluded because of the heterogeneity of the group with respect to clinical outcomes and poor predictive value of MCI for dementia in the general population.

**Critical appraisal and data extraction**

All titles and abstracts retrieved in the search were independently assessed for relevance by two of the authors (PD and HH or HM). Full text articles were obtained for all studies identified as possibly relevant by either of the assessors. Disagreement on which papers were to be included in our review was resolved through discussion.
Risk of bias was assessed at the study level. We chose to use the QUADAS instead of the STARD criteria for the assessment of methodological quality after consultation with the editors of the Cochrane Dementia and Cognitive Improvement Group. All studies were independently assessed by two of the authors (PD and HH or HM).

Data extraction was also performed independently by two authors (PD and HH or HM). Disagreement was resolved through discussion and if needed, a third author was consulted (HH or HM). We extracted data enabling us to compose two by two tables of FP diagnoses against the reference standard diagnoses for the different studies. We contacted authors through e-mail to obtain additional information on methodology and other data when required.

**Statistical analysis**

Confidence intervals for sensitivity and specificity were estimated using the normal approximation method. For one of the included studies we used the Wilson score interval method to obtain confidence intervals between 0 and 1. In studies in which a stratified random sample of patients was used for analysis, bootstrapping was performed to extrapolate the measures for diagnostic accuracy and their confidence intervals to the general population.

**Results**

After the removal of duplicates, the search yielded 1733 articles. Independent screening of title and abstract of these 1733 articles by two of the authors (PD and HH or HM) resulted in 30 potentially relevant articles. Independent assessment of relevance of the full text articles and references by two of the authors resulted in two new potentially relevant and six definitely relevant articles. No longitudinal or prospective studies met the inclusion criteria. In five of six relevant studies, dementia severity staging was part of the reference standard diagnostic procedure. The index test did not distinguish dementia severity stages in most studies; i.e. only one of the studies asked FPs to stage dementia. We did not take FPs’ staging into consideration in the presentation of the results.
Figure 1 provides a flow chart of the study selection process and table 1 provides a methodological summary of the six included studies. Meta-analysis of the included studies was considered inappropriate because of the substantial methodological heterogeneity of the included studies (Table 1). Table 2 and figure 2 provide an overview of the results.

Figure 1. Flow chart of study selection
### Table 1: Main characteristics of the included studies

<table>
<thead>
<tr>
<th>Country</th>
<th>Design</th>
<th>Number of GPs / PCP</th>
<th>Inclusion</th>
<th>Age in years (SD)</th>
<th>Method of case ascertainment</th>
<th>Reference standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooper</td>
<td>Germany</td>
<td>CS 21 PCPs</td>
<td>Random sample of all patient contacts, including home visits, during a four week period</td>
<td>≥ 65 mean 75.7</td>
<td><strong>FPs indicated on a form, directly after each patient contact:</strong> 1) <em>no cognitive impairment</em>, 2) <em>mild cognitive impairment</em>, 3) <em>mild dementia</em> or 4) <em>moderate-severe dementia</em></td>
<td>CAMDEX in stratified sample based on level of cognitive functioning as indicated by the FPs *</td>
</tr>
<tr>
<td>Eefsting</td>
<td>The Netherlands</td>
<td>CS 8 PCPs</td>
<td>All patients were asked to cooperate in the screening</td>
<td>≥ 65 median 73</td>
<td><strong>FPs were presented with a list of all their elderly patients and indicated:</strong> 1) <em>no cognitive impairment</em>, 2) <em>cognitive impairment</em> or 3) <em>dementia</em> based on recollection and medical records if needed</td>
<td>CAMDEX in stratified sample based on MMSE scores †</td>
</tr>
<tr>
<td>Lopponen</td>
<td>Finland</td>
<td>CS ---</td>
<td>All inhabitants of Lieto, Finland asked to cooperate in screening</td>
<td>≥ 64</td>
<td><strong>Medical record review</strong>, patients were categorised as 1) <em>documented dementia</em> when the word <em>dementia</em> in diagnosis list or physician’s notes 2) <em>memory impairment</em> if notes on suspicion of cognitive decline like “<em>forgetfulness</em>”</td>
<td>Semi-structured interview covering the items of the Hachinski ischemic index and the Clinical Dementia Rating scale (CDRS) in patients with an MMSE &lt; 24 or with any clinical suspicion or previous history of a dementing disorder †</td>
</tr>
<tr>
<td>Ollafsdottir</td>
<td>Sweden</td>
<td>CS 1 PCP</td>
<td>Random sample of all ≥ 70 mean 78.6 (SD 6.1) patients presenting for consultation</td>
<td></td>
<td><strong>Medical record review</strong> for entries on cognitive decline or dementia</td>
<td>Neuropsychiatric examination and proxy interview †</td>
</tr>
<tr>
<td>Pond</td>
<td>Australia</td>
<td>CS 13 FPs</td>
<td>All patients were asked to cooperate in the screening</td>
<td>≥ 70 mean 82.5 (SD 5.9)</td>
<td><strong>FPs indicated on a form</strong> for each patient attending whether they thought there was: <em>no dementia</em>, <em>dementia</em> or whether they were unsure</td>
<td>Canberra Interview for the Elderly ∆</td>
</tr>
<tr>
<td>Valcour</td>
<td>Hawaii, US</td>
<td>CS 1PCP 6 FPs</td>
<td>Sample of all patients presenting for consultation</td>
<td>≥ 65 mean 74.6</td>
<td><strong>FPs indicated on a form</strong> for part of patients attending whether they thought there was: <em>no dementia</em>, <em>dementia</em> or whether they were unsure</td>
<td>Cognitive Abilities Screening Instrument and the CDRS, proxy informant interview, review of medical records for documentation of cognitive problems □</td>
</tr>
</tbody>
</table>

CS = cross sectional study, PCP = primary care practice

* = Dementia diagnosis and staging according to CAMDEX criteria, † = Dementia diagnosis and staging according to DSM IIIR criteria, ‡ = Dementia diagnosis according to DSM IV criteria and staging
Table 1: Main characteristics of the included studies

<table>
<thead>
<tr>
<th>Country</th>
<th>Design</th>
<th>Number of patients</th>
<th>Inclusion age (mean, SD)</th>
<th>Method of case ascertainment</th>
<th>Reference standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canberra</td>
<td>Interview for the Elderly</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1)</td>
<td></td>
<td>no cognitive impairment, 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>mild cognitive impairment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>severe cognitive impairment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>moderate-severe cognitive impairment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>very severe cognitive impairment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopponen</td>
<td></td>
<td>All patients were asked to cooperate in the screening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>for all patient contacts, including home visits, during a four week period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooper</td>
<td></td>
<td>All patients were asked to cooperate in the screening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>for part of their elderly patients and indicated: 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>mild dementia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>moderate-severe dementia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ollafsdottir</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pond*</td>
<td></td>
<td>All patients were asked to cooperate in the screening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>for part of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>their elderly patients and indicated:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>mild dementia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>moderate-severe dementia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lieto, Finland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eefsting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valcour</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6 FPs

Table 2: Statistical summary of the included studies

Accuracy of FPs’ combined ‘dementia’ and ‘cognitive impairment’ diagnoses compared to reference standard diagnoses

<table>
<thead>
<tr>
<th>Study</th>
<th>Analysed sample</th>
<th>Mild dementia</th>
<th>Mod-sev dementia</th>
<th>Accuracy all stages sens</th>
<th>spec</th>
<th>Accuracy in persons with mod-sev dementia according to the reference standard sens</th>
<th>spec</th>
<th>PPV</th>
<th>NPV</th>
<th>NNS</th>
<th>DOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eefsting</td>
<td>375</td>
<td>18</td>
<td>53</td>
<td>0.69</td>
<td>0.94</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopponen</td>
<td>1,252</td>
<td>51</td>
<td>61</td>
<td>0.77</td>
<td>0.99</td>
<td>0.67</td>
<td>0.99</td>
<td>0.79</td>
<td>0.99</td>
<td>35.2</td>
<td>252.4</td>
</tr>
<tr>
<td>Valcour</td>
<td>297</td>
<td>14</td>
<td>12</td>
<td>0.35</td>
<td>1.00</td>
<td>0.21</td>
<td>1.00</td>
<td>1.00</td>
<td>0.96</td>
<td>95.2</td>
<td>147.8²</td>
</tr>
</tbody>
</table>

Accuracy of FPs’ ‘dementia’ diagnoses alone compared to reference standard diagnoses

<table>
<thead>
<tr>
<th>Study</th>
<th>Analysed sample</th>
<th>Mild dementia</th>
<th>Mod-sev dementia</th>
<th>Accuracy all stages sens</th>
<th>spec</th>
<th>Accuracy in persons with mild dementia according to the reference standard sens</th>
<th>spec</th>
<th>PPV</th>
<th>NPV</th>
<th>NNS</th>
<th>DOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eefsting</td>
<td>375</td>
<td>18</td>
<td>53</td>
<td>0.38</td>
<td>1.00</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopponen</td>
<td>1,252</td>
<td>51</td>
<td>61</td>
<td>0.48</td>
<td>1.00</td>
<td>0.33</td>
<td>1.00</td>
<td>0.81</td>
<td>0.97</td>
<td>70.1</td>
<td>142.0</td>
</tr>
<tr>
<td>Ollafsdottir</td>
<td>350</td>
<td>21</td>
<td>36</td>
<td>0.26</td>
<td>1.00</td>
<td>0.24</td>
<td>1.00</td>
<td>0.83</td>
<td>0.95</td>
<td>62.8</td>
<td>91.3</td>
</tr>
<tr>
<td>Valcour*</td>
<td>218</td>
<td>11</td>
<td>7</td>
<td>0.33</td>
<td>1.00</td>
<td>0.09</td>
<td>1.00</td>
<td>1.00</td>
<td>0.95</td>
<td>211.5</td>
<td>40.0²</td>
</tr>
<tr>
<td>Pond*</td>
<td>81</td>
<td>31</td>
<td>76</td>
<td>0.42</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooper*</td>
<td>407</td>
<td>42</td>
<td>76</td>
<td>0.78</td>
<td>0.87</td>
<td>0.60</td>
<td>0.85</td>
<td>0.30</td>
<td>0.95</td>
<td>19.5</td>
<td>8.25</td>
</tr>
</tbody>
</table>

¹ = Diagnoses and dementia stages based on the reference standards and staging instruments as described in Table 1.
² = In this study, there were no false positive cases. We assumed ‘0.5 false positive case’ in order to calculate the diagnostic odds ratio
* Study of recognition in consulting patients (Valcour et al. used two methods in their study)
sens = sensitivity, spec = specificity, PPV = positive predictive value, NPV = negative predictive value, NNS = number needed to screen, DOR = diagnostic odds ratio
Chapter 2

Figure 2: Above: Sensitivity (A) and specificity (B) of FPs’ ‘dementia’ diagnoses. Below: Sensitivity (C) and specificity (D): ‘dementia’ and ‘cognitive impairment’ merged.

Light grey marks = mild dementia. Dark grey marks = moderate to severe dementia
* = study of recognition in consulting patients (possible Hawthorne effect)
Studies of documented and recollected diagnoses

The data of the studies of documented and recollected diagnoses are presented in two ways. First, we present data on the diagnostic accuracy of FPs when persons diagnosed with ‘dementia’ are merged with persons recognised to have ‘cognitive impairment’ in the comparison with the reference standard. Second, we present data on the accuracy of the FPs more specific diagnostic label ‘dementia’ alone compared to reference standard diagnoses (see Figure 2).

‘Combined diagnostic category ‘cognitive impairment’ and ‘dementia’': Comparing all cases FPs labelled ‘cognitive impairment’ together with cases FPs labelled ‘dementia’ to the reference standard yielded a sensitivity that ranged from 0.21 to 0.67 for mild dementia and from 0.50 to 0.85 for moderate to severe dementia. The positive predictive value varied between 0.79 and 1.00 in mild dementia. In moderate to severe dementia, it ranged from 0.85 to 0.92. For all dementia stages together, specificity of this label ranged from 0.94 to 1.00.

Only half of mild dementia patients that FPs had recognised as ‘cognitively impaired’ were actually diagnosed with ‘dementia’ by them. Of the moderate to severe dementia patients that FPs recognised as ‘cognitively impaired’, the vast majority were diagnosed with ‘dementia’ by them.

‘Dementia’ alone: Comparing cases FPs diagnosed with ‘dementia’ with reference standard dementia diagnoses yielded a sensitivity of 0.14 to 0.30 for mild dementia and 0.28 to 0.61 for moderate to severe dementia. The positive predictive value of the FPs’ ‘dementia’ diagnosis was lower in mild dementia compared to moderate to severe dementia. The specificity of the FPs’ ‘dementia’ diagnosis was excellent (1.00) for both mild and moderate to severe dementia.

For both diagnostic labels, the diagnostic odds ratio was higher in moderate to severe dementia than in mild dementia.

Studies of recognition during consultation

In these studies in which FPs judged the cognitive function of consecutive contacting patients, there was no such diagnostic label as ‘cognitive impairment’. There was aside from
the category ‘dementia’, the category ‘unsure’ which presumably encompassed patients in whom FPs suspected cognitive impairment. If these cases would be regarded as positive diagnoses, this would substantially increase sensitivity. For example, in the study of Valcour et al., regarding cases labelled ‘unsure’ as positive, would increase the sensitivity from 0.33 to 0.67 (all dementia stages together)\textsuperscript{25}.

‘Dementia’ alone: In general, the studies exploring FPs’ judgment of cognitive functioning in consecutive patients reported a somewhat higher sensitivity compared to the studies that used the FPs’ documentation and/or recollection as method of case ascertainment. For all severity stages together, sensitivity ranged from 0.33 to 0.78. For mild dementia, sensitivity ranged from 0.09 to 0.60 and for moderate to severe dementia, sensitivity ranged from 0.71 to 0.95. The positive predictive value ranged 0.30 to 1.00 for mild dementia and from 0.41 to 1.00 for moderate to severe dementia. Specificity tended to be lower in these studies compared to the specificity found in the studies on documented and recollected diagnoses.

Sources of heterogeneity and potential sources of bias

Study design: The studies of recognition of cognitive impairment in consecutive consulting patients assessed the FPs’ accuracy in a selected group, namely patients who actively contacted them. The studies of documented and recollected diagnoses considered the entire primary care practice population and therefore, also patients that may not have had recent contact with their FP. In the latter case, patients that rarely contact their FP may negatively influence the FPs diagnostic accuracy. Moreover, the consecutive patient design would allow FPs to inquire into their patients’ cognitive function. This possibly introduces a Hawthorne effect\textsuperscript{43}. Therefore, the studies of recognition in consecutive contacting patients better reflect the level of diagnostic accuracy FPs might be able to achieve during their routine consultation hour, than their current diagnostic accuracy.

Case definition: In some studies, a case was defined as a ‘dementia’ diagnosis while in other studies cases were positive if for example, key-words indicative of cognitive impairment were found in the medical record review. To address this, we distinguished between the accuracy of FPs’ ‘dementia’ diagnoses and the broader diagnostic category of ‘dementia’ or ‘cognitive impairment’ (Graph A and B versus C and D in Figure 2).
Method of case ascertainment: The method of ‘ascertainment’ of a case also differed between studies and may have further contributed to heterogeneity. One study used the FPs’ recollection\textsuperscript{38}, whereas others reviewed medical records\textsuperscript{39,41} to ascertain that a diagnosis was made. In the studies of recognition in consulting patients, FPs filled in their judgement of their patients cognitive functioning on a form\textsuperscript{24,25,42}.

Reference standard and staging instrument: The use of different reference standards was another possible source of heterogeneity and bias. Previous research showed that a diagnosis according to DSM (both IIIR and IV) criteria may result in a higher prevalence of dementia than diagnosis according to the CAMDEX in the same patient group\textsuperscript{44–47}. Furthermore, staging of dementia can differ between instruments. A striking example is the prevalence of moderate to severe dementia which can be twice as high within the same group of patients when using DSM IIIR criteria for severity staging compared to the CAMDEX criteria\textsuperscript{45}.

Methodological quality of the studies and potential sources of bias
Table 3 provides an overview of the studies’ quality assessment.

Valid reference standard: Two studies did not apply the same reference standard to the entire study population but used a two-stage reference standard\textsuperscript{38,39}. In such a design, it is important that the first stage is sufficiently sensitive. We considered this to be likely the case in both studies. Inevitably, the chance of missing some cases of (mostly mild) dementia is somewhat higher for a two-stage reference standard compared to the reference standards used in the other studies. This potentially resulted in an underestimation of diagnostic accuracy.

Time between index test and reference standard: In two studies the time between the index test and reference standard diagnosis was unclear, possibly allowing time for disease progression which again may result in an underestimation of diagnostic accuracy\textsuperscript{24,41}. Considering the relatively slow progression of most dementia types, we considered a period of up to 6 months between index test and reference standard acceptable.
Table 3: overview of quality assessment

<table>
<thead>
<tr>
<th></th>
<th>Cooper</th>
<th>Eefsting</th>
<th>Löppönen</th>
<th>Ölafsdöttir</th>
<th>Pond</th>
<th>Valcour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representative patient sample</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
</tr>
<tr>
<td>Clear selection criteria</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Adequate reference standard</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adequate time between index test and reference standard</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Reference standard applied to whole sample or random selection</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Same reference standard regardless of index test result</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Reference standard independent of index test</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Index test clearly described</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Reference standard clearly described</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Index test interpreted without knowledge of reference standard results</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Reference standard interpreted without knowledge of index test results</td>
<td>Unclear</td>
<td>Yes</td>
<td>No</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Yes</td>
</tr>
<tr>
<td>Same clinical data available for reference standard and index test</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Uninterpretable / intermediate results reported</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Withdrawals explained</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Total numbers of met and unmet quality criteria</td>
<td>Yes</td>
<td>11</td>
<td>12</td>
<td>9</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Unclear</td>
<td>No</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

A complete description of QUADAS quality items can be obtained through the first author.
The accuracy of family physicians’ dementia diagnoses: a systematic review

Interpretation of reference standard without knowledge of the index test: Only three studies provided explicit information (after contacting the authors) on whether reference standard results were interpreted blind without knowledge of the FPs’ diagnosis which was methodologically preferable.25,38,39

Missing data: Pond et al. did not report on 17 patients sampled to undergo the reference standard criteria that were missing in later analysis. Unfortunately, we received no additional data from the authors.

Representative patient sample: Only one study included a patient sample that appeared representative of the elderly PCP population. Cooper et al. did investigate whether they included a representative patient sample and concluded that there was only a slight under representation of the youngest age group (65-70 years). Considering the high participation rates in three other studies, their patient samples presumably also represent the elderly PCP population fairly well.38,39,42 The participation rate in the study of Valcour et al. was rather low, increasing the probability of a selected sample that does not optimally represent the targeted patient population.

Discussion

Summary of evidence
All studies reported few false positive diagnoses for both dementia severity categories, indicating a high positive predictive value and specificity of FPs’ dementia diagnoses. However, in all of the included studies, a substantial number of individuals with dementia were not diagnosed in primary care. A dementia diagnosis was documented in only up to one third of mild dementia patients and in less than two thirds of moderate to severe dementia patients. The proportion of individuals in whom cognitive impairment was recognised yet without a diagnosis of dementia was higher in individuals with mild compared to moderate to severe dementia.

FPs identified a higher proportion of individuals with dementia in studies focusing on recognition during consultation compared to the proportion in studies of documented and recollected dementia but the sacrifice was increased false positive diagnoses.
Chapter 2

**Strengths and limitations**

The strengths of this review are that it exclusively reflects the accuracy of FPs’ judgement on cognitive functioning in the primary care population and that it considers methodological aspects of the included studies in detail. A reason to interpret our results with caution is the sizable methodological heterogeneity within the included studies. For example, the use of different reference standards and case ascertainment methods threatens the internal validity of various comparisons between studies. Second, the quality of some of the included studies was limited. For example, one study did not explain why part of the data was missing, in other studies reference standard assessors were not blinded to the FP diagnosis. Third, as the majority of studies were performed more than a decade ago, the results may not reflect the current level of diagnostic accuracy. However, recent studies support our findings.

**Interpretation**

Important to point out in the interpretation of the results is that the diagnostic accuracy is only partly a function of the FP. Bradford et al. distinguish four levels at which barriers to dementia diagnosis can occur; the patient level, the caregiver level, the level of the health care system and the physician level. Important barriers at the patient and caregiver level include misinterpretation or denial of symptoms, fear of stigmatisation, therapeutic nihilism and fear of institutionalisation. Important barriers at the level of the health care system are the limited reimbursement and time available per patient. Older patients may present with many co-morbid conditions, limiting FPs’ time to establish a well-founded dementia diagnosis even further. Important barriers at the physician level are: failure to recognise symptoms, diagnostic uncertainty, fear to stigmatise or harm and therapeutic nihilism. In conclusion, our results reflect FPs’ attitude, knowledge and skills but also barriers to dementia diagnosis imposed by the patient, caregiver and the health care system.

**Implications for practice**

Sensitivity of FPs’ dementia diagnoses will always be limited based on the multidimensional barriers mentioned above and considering the diagnostic difficulties created by the great variety in presentation and natural course of the syndrome. However, early diagnosis of dementia may benefit patients and their families. FPs play a pivotal role in the initial diagnosis of dementia and therefore also in initial disclosure, psychoeducation and support. Several interventions have proven to successfully increase the frequency of dementia
diagnosis in primary care: increasing awareness during consultation (this review), education of FPs, implementation of evidence based guidelines and structured assessment of older patients\(^{53,54}\). Furthermore, research suggests that collaboration of FPs with trained practice nurses increases the frequency of dementia diagnosis and improves the quality of dementia care\(^{55-57}\). FPs indicated that they welcome such collaboration\(^{11}\).

To provide fertile ground for such interventions, it is important that FPs explore their attitude towards diagnosing dementia and challenge the validity of their arguments not to diagnose. In light of proven benefits of therapy and support for both dementia patients and their informal caregivers, FPs’ diagnostic restraint seems no longer justifiable\(^{16,58,59}\).

Further potential to help FPs to accurately distinguish patients with mild cognitive impairment or dementia from individuals with cognitive decline normal for age, may lie in newly developed cognitive screening instruments that were shown to have superior psychometric properties, also in the primary care population, compared to the commonly used MMSE\(^{60-62}\).

Given the central role of FPs in diagnosing dementia, it is our opinion that not all patients need to be referred for diagnostic evaluation\(^{51}\). If FPs feel sufficiently confident about the diagnosis, and when the FP, patient and family agree that referral is of little value, they should have the liberty to forego it. Given FPs’ caution to mislabel dementia (this review), we expect them to refer in case of diagnostic uncertainty, for instance if an uncommon cause of cognitive decline is suggested by symptoms such as hallucinations or focal neurologic deficits.

**Future research**

Research on early detection may include the development and evaluation of ICT tools to identify individuals at risk of dementia (e.g. with higher age, diabetes mellitus) and individuals at risk of a missed dementia diagnosis (e.g. living alone), based on risk factors readily available in primary care\(^{48,63,64}\). Another area deserving further attention is the development of educational interventions with more emphasis on FPs’ perceptions of their own suitability and capability to diagnose dementia, their skills to communicate the diagnosis and the importance of early diagnosis\(^{65,66}\). Finally, additional research is recommended to develop and test criteria that support FPs in the decision to refer or not refer patients for specialist diagnostic evaluation.
References


The accuracy of family physicians’ dementia diagnoses: a systematic review


Appendix: Search strategy

Date of search: October, 20th 2009

Overview of databases searched and results:

<table>
<thead>
<tr>
<th>Database</th>
<th>Number of retrieved articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>PubMed</td>
<td>996</td>
</tr>
<tr>
<td>Embase</td>
<td>1183</td>
</tr>
<tr>
<td>Cinahl</td>
<td>311</td>
</tr>
<tr>
<td>PsycINFO</td>
<td>344</td>
</tr>
<tr>
<td>Cochrane Library</td>
<td>31</td>
</tr>
<tr>
<td><strong>Total before removal of duplicates</strong></td>
<td><strong>2865</strong></td>
</tr>
</tbody>
</table>

Overview of search terms per database:

**PubMed:**


**EMBASE:**

(‘dementia’/exp/dm_di OR (‘dementia’/exp OR dementia:ab,ti) AND (‘diagnosis’/exp OR ‘sensitivity and specificity’/exp OR sensitiv*:ab,ti OR diagnos*:ab,ti)) AND (‘general practice’/exp OR ‘general practitioner’/exp OR ‘primary health care’/exp OR ‘family practice’:ab,ti OR ‘general practice’:ab,ti OR ‘family practices’:ab,ti OR ‘general practices’:ab,ti OR ‘family practices’:ab,ti OR ‘family practices’:ab,ti)
The accuracy of family physicians’ dementia diagnoses: a systematic review

practitioner':ab,ti OR ‘general practitioner':ab,ti OR ‘family practitioners':ab,ti OR ‘general practitioners':ab,ti OR ‘primary care':ab,ti OR ‘nurse practitioner'/exp OR ‘physician assistant'/exp OR ‘nurse practitioner':ab,ti OR ‘nurse practitioners':ab,ti OR ‘physician assistant':ab,ti OR ‘physician assistants':ab,ti)

CINAHL:

(MH "Dementia+/DI" or ((MH “Dementia+” or TI dementia or AB dementia) AND (MH “Diagnosis, Differential” or MH “Sensitivity and Specificity” or MH “Diagnosis” or TI sensitiv* or AB sensitiv* or TI diagnos* or AB diagnos*))) AND (MH “Family Practice” or MH “Physicians, Family” or MH “Primary Health Care” or TI (“family practice” OR “general practice” OR “family practices” OR “general practices” OR “family practitioner” OR “general practitioner” OR “family practitioners” OR “general practitioners” OR “primary care”) OR AB (“family practice” OR “general practice” OR “family practices” OR “general practices” OR “family practitioner” OR “general practitioner” OR “family practitioners” OR “general practitioners” OR “primary care”) OR MH “Nurse Practitioners+” or MH “Physician Assistants” or TI (“Nurse Practitioner” or “Nurse Practitioners” or “Physician Assistant” or “Physician Assistants”) or AB (“Nurse Practitioner” or “Nurse Practitioners” or “Physician Assistant” or “Physician Assistants”))

PsycINFO:

(DE=(“dementia” or “aids dementia complex” or “dementia with lewy bodies” or “presenile dementia” or “alzheimers disease” or “creutzfeldt jakob syndrome” or “picks disease” or “senile dementia” or “senile psychosis” or “vascular dementia”) OR TI=dementia OR AB=dementia) AND (DE=(“diagnosis” or “computer assisted diagnosis” or “differential diagnosis” or “educational diagnosis” or “galvanic skin response” or “medical diagnosis” or “biopsy” or “cardiography” or “electrocardiography” or “dexamethasone suppression test” or “echoencephalography” or “electro oculography” or “electroencephalography” or “alpha rhythm” or “delta rhythm” or “theta rhythm” or “electromyography” or “electronystagmography” or “electroplethysmography” or “electroretinography” or “encephalography” or “pneumoencephalography” or “rheoencephalography” or “hiv testing” or “ophthalmologic examination” or “plethysmography” or “prenatal diagnosis” or “roentgenography” or “angiography” or “mammography” or “tomography” or “magnetic
resonance imaging” or “functional magnetic resonance imaging” or “positron emission
tomography” or “single photon emission computed tomography” or “urinalysis” or
“psychodiagnosis” or “psychodiagnostic interview” or “diagnostic interview schedule”
or “structured clinical interview”) OR TI=sensitiv* or TI=diagnos* or AB=sensitiv* or
AB=diagnos*) AND (DE=(“primary health care” or “family medicine” or “family physicians”
or “general practitioners”) OR TI=(“family practice” OR “general practice” OR “family
practices” OR “general practices” OR “family practitioner “ OR “general practitioner”
OR “family practitioners “ OR “general practitioners” OR “primary care” OR “Physician
Assistant” OR “Physician Assistants” OR “Nurse Practitioner” OR “Nurse Practitioners”) OR
AB=(“family practice” OR “general practice” OR “family practices” OR “general practices”
OR “family practitioner “ OR “general practitioner” OR “family practitioners “ OR “general
practitioners” OR “primary care” OR “Physician Assistant” OR “Physician Assistants” OR
“Nurse Practitioner” OR “Nurse Practitioners”))

Cochrane Library:

((dementia):ti or (dementia):ab) AND ((diagnos*):ti or (diagnos*):ab or (sensitiv*):ti or
(sensitiv*):ab) AND ((“family practice” OR “general practice” OR “family practices” OR “general
practices” OR “family practitioner “ OR “general practitioner” OR “family practitioners
“ OR “general practitioners” OR “primary care” OR “Physician Assistant” OR “Physician
Assistants” OR “Nurse Practitioner” OR “Nurse Practitioners”):ti or (“family practice” OR
“general practice” OR “family practices” OR “general practices” OR “family practitioner “ OR
“general practitioner” OR “family practitioners “ OR “general practitioners” OR “primary
care” OR “Physician Assistant” OR “Physician Assistants” OR “Nurse Practitioner” OR “Nurse
Practitioners”):ab)
The accuracy of family physicians’ dementia diagnoses: a systematic review