PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Reliability of Coded Data to Identify Earliest Indications of Cognitive Decline, Cognitive Evaluation, and Alzheimer's Disease Diagnosis: A Pilot Study in England
AUTHORS	Dell'Agnello, Grazia; Desai, Urvi; Kirson, Noam; Wen, Jody; Meiselbach, Mark; Reed, Catherine; Belger, Mark; Lenox-Smith, Alan; Martinez, Carlos; Rasmussen, Jill

VERSION 1 – REVIEW

REVIEWER	Dr Terry Quinn Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK
REVIEW RETURNED	14-Nov-2017

GENERAL COMMENTS	Introduction Well written and sets the scene well. However, I would quibble with a couple of points. Line 21-22 diagnosis requires cognitive impairment and functional limitation (not, as the authors suggest, behavioural symptoms). Line 42-44 Tests such as Addenbrooke's and Mini-Mental State are not used for diagnosis (or at least should not be used in this way) There are also some statements of fact in the introduction that perhaps require supporting citation eg 'this information may often not be captured in existing, structured, real-world data sources' and 'These supplemental data elements are generally not available to researchers' Aims: There is inconsistency in the terminology used to describe the primary aims; is this a study of feasibility, reliability, validity? Various terms are used and they are not synonymous. Methods: Certain aspects of the methods need more detail and justification. From reading the methods, the process seems to have been iterative and dynamic. This is acceptable in feasibility work but it would be reassuring to see the original protocol and for the authors to be explicit around which aspects of the study were pre-defined and which were post-hoc. The 'substrate' for the research was information from CPRD including free text notes and linked secondary care materials. As the authors acknowledge, the primary care free text notes are no longer

	available and processes/governance for accessing secondary care data have also evolved. Thus, this study is looking at a method of data driven research that is no longer available to researchers. The authors correctly describe the issues with generalisability of the CPRD data. In fact for this study the potential poor external validity is even greater. The sampling frame is a highly selected group who followed a particular dementia 'journey; and had high quality data with adequate linkage. The dementia pathway studies is atypical as it does not include memory clinics yet these are the major dementia diagnostic service providers in the UK. With all this in mind, the authors need to be very cautious in any extrapolation to a broader, clinical population. The process for 'random sampling' needs elaborated upon. The syntax used to search the free text needs to be better described – both development and validation. I found the process difficult to follow in places and wonder if a figure or flow diagram may help the reader. If I understand correctly, only those with dementia were included. If this is the case then nothing can be said about false positives and related metrics (ie those with free text phrases suggesting cognitive decline who did not develop dementia). I have misunderstood then this speaks to my point above about the methods being difficult to follow. Using 'AD-related medication' (not defined in text) as a marker of Alzheimer's diagnosis is usually valid, but there are frequent examples of off licence use of these medications for other neurological issues. The cognitive assessments that were searched for do not include informant based tests such as IQCODE and AD8, yet these are suited to primary care and used in certain parts of the UK. Results For a novel approach to data analysis, I would have liked some metrics around the work involved. If a similar approach is going to be considered for future studies, researchers will want to know how many hours of text mining, data analysis etc will be required. A lot of the results text is available in the tabulated materials; the results section could be shortened with signposting to the relevant tables I disagree with the description '[time] between first cognitive symptoms and …..'. The dataset is reliant on symptoms volunteered in primary care consultations and many with subjective cognitive impairment will have symptoms for a long time before they seek help.

REVIEWER	Jacques Hugon Center of Cognitive Neurology Lariboisiere FW Hospital APHP University of Paris Diderot 75010 Paris France
REVIEW RETURNED	06-Dec-2017

GENERAL COMMENTS	This is an interesting study focused more on public health than on research about clinical diagnosis of AD. The main question is : is it possible to determine with an appropriate accuracy the date of the first symptoms of AD using an elaborate algorithm of coded information's or a more usual approach based on medical records extracted from GP's practice mainly. There are several remarks that can be addressed 1- The precise procedures used in this article are more difficult to assess for a non-British European reviewer than for a British reviewer. Each country has its own medical assessment and a comparison with other European countries would be an advantage 2- The diagnosis of AD is complex and I did not find out on which criteria the diagnosis was made in the 50 AD patients. 3- Why the medical records extracted from Memory Clinics were not available? As mentioned in the text this would be a real asset for the differential diagnosis between AD and other dementias 4- As MMSE seems to be still one of the best test depicted in this study, How reliable is the test performed in general population by GP's or other health professionals? Are there comparison studies with MMSE performed in Memory Clinics? 5- Finally in the discussion the authors should include a more accurate view of the diagnostic procedures performed in patients with cognitive disorders. Currently it is not so the date of the dementia symptoms which is crucial but how to detect patients with early mild cognitive impairment. May be including new tools in GP's management of patients with memory complaints could bring about a new way to detect individuals who could slowly evolved to AD symptoms.

VERSION 1 – AUTHOR RESPONSE

Dear Editor,

On behalf of all the authors, thank you for the very thoughtful review of our manuscript entitled "Reliability of Coded Data to Identify Earliest Indications of Cognitive Decline, Cognitive Evaluation, and Alzheimer's Disease Diagnosis: A Pilot Study in England". Below we have provided point-by-point responses to the comments raised. We have also attached all requested files with this submission.

Please let me know if I can provide any additional clarifications.
We look forward to your decision.

Sincerely,
Urvi Desai

++

Response to Reviewer Comments:

Reviewer: 1
Reviewer Name: Dr Terry Quinn

Institution and Country: Institute of Cardiovascular and Medical Sciences, University of Glasgow, UK
Please state any competing interests or state 'None declared': I am co-author on a systematic review looking at data driven approaches to dementia diagnosis. I am co-ordinating editor of the Cochrane dementia group with a particular interest in diagnostic test accuracy.

1. Introduction Well written and sets the scene well. However, I would quibble with a couple of points.
a. Line 21-22 diagnosis requires cognitive impairment and functional limitation (not, as the authors suggest, behavioural symptoms).

RESPONSE: Thank you for the suggestion. We agree that the diagnosis is primarily based on evaluation of cognitive and functional abilities. However, we would like to note that certain diagnosis guidelines, such as the US-based NIH/NIA guidelines, suggest that patients' behavioral characteristics should also be evaluated to ensure that the observed symptoms are indeed attributable to a dementia/AD etiology and not to other conditions such as depression. We have updated the manuscript accordingly to further clarify this point (see p. 5).

b. Line 42-44 Tests such as Addenbrooke's and Mini-Mental State are not used for diagnosis (or at least should not be used in this way)

RESPONSE: Thank you - we have updated the description to clarify that findings from these tests provide information that can in turn help with the diagnosis.

c. There are also some statements of fact in the introduction that perhaps require supporting citation eg 'this information may often not be captured in existing, structured, real-world data sources' and 'These supplemental data elements are generally not available to researchers'

RESPONSE: Thank you for the suggestion. We have added a citation to support these statements on pp. 6-7 of the manuscript.

2. Aims: There is inconsistency in the terminology used to describe the primary aims; is this a study of feasibility, reliability, validity? Various terms are used and they are not synonymous.

RESPONSE: As noted on p. 6 of the manuscript, this was a study to determine the reliability of using a code-based algorithm to identify the timing of symptomatic onset, cognitive assessment, and formal diagnosis of AD, using the comprehensive information from both codes and free text data fields as a reference. We have reviewed and updated the manuscript to ensure consistent use of this terminology.

3. Methods: Certain aspects of the methods need more detail and justification.
a. From reading the methods, the process seems to have been iterative and dynamic. This is acceptable in feasibility work but it would be reassuring to see the original protocol and for the authors to be explicit around which aspects of the study were pre-defined and which were post-hoc.

RESPONSE: As noted on page 17 of the manuscript, the study protocol was approved by ISAC (Protocol # 16_043R). We have attached this protocol with the revised manuscript.

b. The 'substrate' for the research was information from CPRD including free text notes and linked secondary care materials. As the authors acknowledge, the primary care free text notes are no longer available and processes/governance for accessing secondary care data have also evolved. Thus, this study is looking at a method of data driven research that is no longer available to researchers.

RESPONSE: While we agree that the replication or extrapolation of such an approach in the future is limited by availability of text notes from primary/secondary care, we believe that the study provides important information about the quality of the coded data for identifying the earliest indications of AD and related events. In particular, our study findings suggest that the coded data are reliable for identifying earliest record of AD. However, the reliability of these data for identifying the onset of symptoms and work-up leading up to the AD diagnosis is questionable.

c. The authors correctly describe the issues with generalisability of the CPRD data. In fact for this study the potential poor external validity is even greater. The sampling frame is a highly selected group who followed a particular dementia 'journey; and had high quality data with adequate linkage. The dementia pathway studies is atypical as it does not include memory clinics yet these are the major dementia diagnostic service providers in the UK. With all this in mind, the authors need to be very cautious in any extrapolation to a broader, clinical population.

RESPONSE: Thank you. As you note here, we have acknowledged this limitation in the paper and have noted that additional research using larger patient populations is necessary to further test the reliability and generalizability of the algorithm (see p. 14).

d. The process for 'random sampling' needs elaborated upon.

RESPONSE: Of the patients meeting the selection criteria, a random sample of 50 patients were selected based on a computer-generated randomization algorithm. In particular, using the SAS software (SAS Institute, Cary, NC), all patients were assigned a random number. Following this, the first 50 patients with the smallest values for the randomly assigned numbers were selected from the dataset. No other selection criteria were imposed. We have clarified this on p. 8 of the manuscript.

e. The syntax used to search the free text needs to be better described – both development and validation.

RESPONSE: As described on p. 8 of the manuscript, the free text was manually reviewed to identify the key phrases suggestive of the earliest markers of symptoms related to cognitive impairment, cognitive assessment, and AD diagnosis. Please refer to the newly added Appendix table 2 that describes all the key phrases identified from this process.

f. I found the process difficult to follow in places and wonder if a figure or flow diagram may help the reader.
RESPONSE: Thank you for the comment. We have added a flow diagram illustrating the study approach in the Appendix (see Appendix Figure 1).

g. If I understand correctly, only those with dementia were included. If this is the case then nothing can be said about false positives and related metrics (ie those with free text phrases suggesting cognitive decline who did not develop dementia). I have misunderstood then this speaks to my point above about the methods being difficult to follow.

RESPONSE: The reviewer is correct that the study focused on patients diagnosed with AD only, and does not shed light on the lack of an AD or dementia diagnosis among some patients with evidence of cognitive symptoms in the free text fields. However, the aim of the study was to assess the reliability of identifying the timing of the AD diagnosis and the events leading up to the diagnosis among those diagnosed. We have clarified the terminology in the manuscript.

h. Using 'AD-related medication' (not defined in text) as a marker of Alzheimer's diagnosis is usually valid, but there are frequent examples of off licence use of these medications for other neurological issues.

RESPONSE: The point is well-taken. However, all patients in our study had a diagnosis code for AD and no indication for other dementia etiologies (see p. 7 of the manuscript). As such, we have no reason to believe that the prescribed AD-related medication were for other conditions. To that effect, on p. 8 of the manuscript, we have clarified that AD-related medications included cholinesterase inhibitors and memantine.

i. The cognitive assessments that were searched for do not include informant based tests such as IQCODE and AD8, yet these are suited to primary care and used in certain parts of the UK.

RESPONSE: The search for types of cognitive assessments was not limited to specific tests. As noted on p. 13 of the manuscript, nearly two-thirds of the patients had information available on the type of cognitive assessments in the text-based data and/or Read codes; however, the tests noted above were not captured within these records.

4. Results
a. For a novel approach to data analysis, I would have liked some metrics around the work involved. If a similar approach is going to be considered for future studies, researchers will want to know how many hours of text mining, data analysis etc will be required.

RESPONSE: Thank you for the question. Given the iterative nature of the study approach, it is difficult to quantify the time for text mining and data analysis. Additionally, the precise effort required for a study of similar design would be highly dependent on the scope of the project as well as the resources available.

b. A lot of the results text is available in the tabulated materials; the results section could be shortened with signposting to the relevant tables

RESPONSE: Thank you for the suggestion. We believe that the narrative of the results provides additional context for readers not familiar with this type of research, and helps with the interpretations. For the time being, no changes were made to this section, but we will defer to the Editor on whether the results section should be trimmed.

c. I disagree with the description '[time] between first cognitive symptoms and …..'. The dataset is reliant on symptoms volunteered in primary care consultations and many with subjective cognitive impairment will have symptoms for a long time before they seek help.

RESPONSE: The point is well taken. We have added this to the study limitations on p. 15 of the revised manuscript.

Reviewer: 2
Reviewer Name: Jacques Hugon
Institution and Country: Center of Cognitive Neurology, Lariboisiere FW Hospital APHP, University of Paris Diderot, 75010 Paris France Please state any competing interests or state 'None declared': None declared

This is an interesting study focused more on public health than on research about clinical diagnosis of AD. The main question is: is it possible to determine with an appropriate accuracy the date of the first

symptoms of AD using an elaborate algorithm of coded information's or a more usual approach based on medical records extracted from GP's practice mainly. There are several remarks that can be addressed

1. The precise procedures used in this article are more difficult to assess for a non-British European reviewer than for a British reviewer. Each country has its own medical assessment and a comparison with other European countries would be an advantage

RESPONSE: Thank you for the comment. We agree that future studies should evaluate whether similar approaches can be implemented in other countries, and even within the UK, for larger populations.

2. The diagnosis of AD is complex and I did not find out on which criteria the diagnosis was made in the 50 AD patients.

RESPONSE: The diagnosis of AD is indeed complex. For this study, as noted on p.6 of the manuscript, the diagnosis was identified using coded data (ICD-10 codes and Read codes specific to UK CPRD; see Appendix Table 1), and the clinical information about the diagnostic process was not available.

3. Why the medical records extracted from Memory Clinics were not available? As mentioned in the text this would be a real asset for the differential diagnosis between AD and other dementias

RESPONSE: The dataset captures information about care provided in primary care and hospital settings as well as whether patients were referred to specialists. However, the data do not contain any information regarding visits to other secondary care settings such as Memory Clinics. Because we had no visibility into whether the patients were in fact evaluated in Memory Clinics (or other secondary care settings), it was not feasible to request the detailed patient records from these settings.

4. As MMSE seems to be still one of the best test depicted in this study, How reliable is the test performed in general population by GP's or other health professionals? Are there comparison studies with MMSE performed in Memory Clinics?

RESPONSE: The properties of MMSE in different settings have been studied extensively. Findings from a recent meta-analysis that evaluated the sensitivity and specificity of MMSE in different settings including the memory clinics can be found here: Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. J Psychiatr Res. 2009 Jan;43(4):411-31. doi: 10.1016/j.jpsychires.2008.04.014. PMID: 18579155.

5. Finally in the discussion the authors should include a more accurate view of the diagnostic procedures performed in patients with cognitive disorders. Currently it is not so the date of the dementia symptoms which is crucial but how to detect patients with early mild cognitive impairment. May be including new tools in GP's management of patients with memory complaints could bring about a new way to detect individuals who could slowly evolved to AD symptoms.

RESPONSE: We agree that the process to facilitate an early identification and diagnosis of cognitive impairment is important, and that future research should evaluate the implications of changes in procedures and policies surrounding the diagnosis and management of patients with cognitive impairment. However, the purpose of this study was to determine whether coded data can be used to understand the sequencing of events leading up to an AD diagnosis. Having this information could in turn facilitate real-world research into the implications of earlier identification of cognitive impairment on subsequent outcomes (e.g., healthcare resource use and costs).

VERSION 2 – REVIEW

REVIEWER	Dr Terence J Quinn Institute of Cardiovascular and Medical Sciences, University of Glasgow
REVIEW RETURNED	24-Jan-2018

GENERAL COMMENTS	The authors have responded to all the comments and the paper is improved. I have no further suggestions or feedback.