

Research

Michael Pentzek, Michael Wagner, Heinz-Harald Abholz, Horst Bickel, Hanna Kaduszkiewicz, Birgitt Wiese, Siegfried Weyerer, Hans-Helmut König, Martin Scherer, Steffi G Riedel-Heller, Wolfgang Maier and Alexander Koppara, for the AgeCoDe Study Group

The value of the GP's clinical judgement in predicting dementia:

a multicentre prospective cohort study among patients in general practice

Abstract

Background

Clinical judgement is intrinsic to diagnostic strategies in general practice; however, empirical evidence for its validity is sparse.

Aim

To ascertain whether a GP's global clinical judgement of future cognitive status has an added value for predicting a patient's likelihood of experiencing dementia.

Design and setting

Multicentre prospective cohort study among patients in German general practice that took place from January 2003 to October 2016.

Method

Patients without baseline dementia were assessed with neuropsychological interviews over 12 years; 138 GPs rated the future cognitive decline of their participating patients. Associations of baseline predictors with follow-up incident dementia were analysed with mixed-effects logistic and Cox regression.

Results

A total of 3201 patients were analysed over the study period (mean age = 79.6 years, 65.3% females, 6.7% incident dementia in 3 years, 22.1% incident dementia in 12 years). Descriptive analyses and comparison with other cohorts identified the participants as having frequent and long-lasting doctor–patient relationships and being well known to their GPs. The GP baseline rating of future cognitive decline had significant value for 3-year dementia prediction, independent of cognitive test scores and patient's memory complaints (GP ratings of very mild [odds ratio (OR) 1.97, 95% confidence intervals [95% CI] = 1.28 to 3.04]; mild [OR 3.00, 95% CI = 1.90 to 4.76]; and moderate/severe decline [OR 5.66, 95% CI = 3.29 to 9.73]). GPs' baseline judgements were significantly associated with patients' 12-year dementia-free survival rates (Mantel–Cox log rank test $P < 0.001$).

Conclusion

In this sample of patients in familiar doctor–patient relationships, the GP's clinical judgement holds additional value for predicting dementia, complementing test performance and patients' self-reports. Existing and emerging primary care-based dementia risk models should consider the GP's judgement as one predictor. Results underline the importance of the GP–patient relationship.

Keywords

clinical decision making; cognitive decline; dementia; general practice; intuition; validity.

INTRODUCTION

Health policy and research have provided increased awareness about identifying individuals at risk of dementia in primary care.¹ This is justified by strategies for risk reduction,² and the opportunity for patient-centred management³ in case of cognitive deterioration, for example, home visits, adjusted communication, family involvement, checking medication, and alertness for risks such as social isolation. For older people with memory concerns, the GP is a significant contact.⁴ A patient's predicted cognitive status can add to the GP's holistic view of the patient and provide relevant information for ongoing patient management. Parameters discussed for identifying dementia risk include biomarkers,⁵ neuropsychological constructs such as mild cognitive impairment,⁶ short tests like the Mini-Mental State Examination (MMSE),⁷ and patient self-reports of subjective cognitive decline (SCD).⁸ Several dementia risk scores for use in general practice have been published recently,^{9–12} but none of these considers the GP's judgement as an important source of prognostic power, therefore neglecting the

expertise and relevance of GPs as experts on their patients. Ford *et al* suggested that easily available predictive information from primary care should also be explored in future studies because it '... *might increase the predictive validity of a model*'.¹³

Particularly in the absence of general screening activity,¹⁴ clinical judgement is an essential part of early diagnostic strategies used in primary care, arising from a hermeneutic approach to individual symptoms, changes, complaints, and patterns.¹⁵ Continuous and familiar doctor–patient relationships in primary care foster the integration of longitudinal clinical, biographical, and psychosocial information for clinical judgement.¹⁶ As Thornton emphasised, '... *at the heart of evidence based medicine is good judgement*'.¹⁷ In other medical settings, expert judgement (alone or combined with objective measures) has shown several advantages and good validity, for example, in the emergency department¹⁸ and in medical education.¹⁹ With regard to early dementia recognition and prediction, the GP's judgement and the patient's subjective cognitive decline

M Pentzek, PhD, senior researcher; **H-H Abholz**, MD, professor, Institute of General Practice (ifam), Medical Faculty, University of Düsseldorf, Düsseldorf. **M Wagner**, PhD, professor, Department of Neurodegenerative Diseases and Geriatric Psychiatry, University of Bonn, Bonn. **H Bickel**, PhD, senior researcher, Department of Psychiatry, Technical University of Munich, Munich. **H Kaduszkiewicz**, MD, professor, Institute of General Practice, Medical Faculty, Kiel University, Kiel. **B Wiese**, MSc, biostatistician, Institute for General Practice, Hannover Medical School, Hannover. **S Weyerer**, PhD, professor, Central Institute of Mental Health, Medical Faculty, Mannheim/Heidelberg University, Mannheim. **H-H König**, MD, professor, Department of Health Economics and Health Services Research, Hamburg Center for Health Economics; **M Scherer**, MD, professor, Department of Primary Medical Care, University Medical Center Hamburg-Eppendorf, Hamburg. **SG Riedel-Heller**, MD, professor, Institute of Social Medicine, Occupational Health

and Public Health, University of Leipzig, Leipzig. **W Maier**, MD, professor; **A Koppara**, PhD, research associate, Department of Neurodegenerative Diseases and Geriatric Psychiatry, University of Bonn, Bonn; German Center for Neurodegenerative Diseases (DZNE), Bonn.

Address for correspondence

Michael Pentzek, Institute of General Practice (ifam), Medical Faculty, Heinrich-Heine University Düsseldorf, Moorenstr. 5, 40225 Düsseldorf, Germany.

Email: pentzek@med.uni-duesseldorf.de

Submitted: 2 February 2019; **Editor's response:** 11 March 2019; **final acceptance:** 8 May 2019

©British Journal of General Practice

This is the full-length article (published online 8 Oct 2019) of an abridged version published in print. Cite this version as: **Br J Gen Pract** 2019; DOI: <https://doi.org/10.3399/bjgp19X706037>

How this fits in

Several dementia risk scores have been constructed recently that do not include the GP's expertise. This study found that a GP's global clinical judgement of future cognitive status has an independent value for predicting a patient's dementia, yielding additional information beyond test scores and patient reports. This finding supports the potential of intuitive and heuristic clinical judgement as a diagnostic strategy.

are the most important parameters in the initial diagnostic trigger stage, whereas cognitive tests and laboratory parameters become relevant in the subsequent disease-oriented diagnostic stage.²⁰ According to the stimulating and substantiating factors of intuitive clinical judgement,²¹ dementia prediction is a prime example of intuition as an appropriate initial diagnostic strategy in primary care.

The aim of this study was to address the research question: does the GP's clinical judgement contribute additional information for the prediction of dementia within 3 years and time of dementia onset over 12 years, beyond the patient's self-report and test performance?

METHOD

The German Study on Ageing, Cognition, and Dementia (AgeCoDe) is a multicentre prospective cohort study of German patients in primary care.²² Demographic, clinical, and neuropsychological data were collected from patients during home visits by interviewers at intervals of 18 months (at baseline and in eight follow-up assessments I to VIII). Interviewers were researchers, trained and supervised in the use of the assessment procedures. Independent from these patient interviews, GPs gave information about their participating patients. The analyses presented in this article consist of data from baseline and two follow-ups (with regard to dementia prediction over 3 years) and eight follow-ups (with regard to prediction of time to dementia onset over 12 years), respectively.

GPs were recruited by local researchers at six metropolitan study centres. Patients were recruited from the 138 participating GP practices between 1 January 2003 and 30 November 2004. Inclusion criteria for patients were age ≥ 75 years, absence of dementia in the GP's view, and at least one contact with the GP within the previous 12 months. Exclusion criteria

were consultations only by home visits, residence in a nursing home, severe illness the GP would deem fatal within 3 months, insufficient German language skills, deafness, blindness, lacking the ability to consent, and not being a regular patient of the participating GP practice. Researchers and GPs together compiled the lists of eligible patients from all practice patients; researchers then randomly chose up to 60 eligible patients per practice using random number tables.

Outcome variable: incident dementia

Interviewers conducted a SIDAM interview (structured interview for the diagnosis of dementia of the Alzheimer type, multi-infarct dementia, and dementias of other aetiology),²³ including a neuropsychological test battery (the SIDAM-Score or SISCO), at baseline and all follow-ups. DSM-IV dementia criteria²⁴ were included in a diagnostic algorithm in the SIDAM by assessing cognitive performance and activities of daily living, including an informant interview in cases with suspected dementia. Dementia status in patients without personal interview at follow-up (for example, because of refusal or death) was based on additional informant interviews with relatives, professional carers, and GPs, using the Global Deterioration Scale (GDS)²⁵ and Blessed Dementia Scale.²⁶ All diagnoses were discussed in consensus conferences with the interviewers and experienced geriatric psychiatrists or geriatricians. Dementia diagnoses could be verified by follow-up results.

Predictor variables

Education. The level of education received by the patients was assessed in at-home interviews and categorised as low (elementary education), intermediate (intermediate vocational/general qualification), or high (tertiary education) using an international educational classification.²⁷

GPs' clinical judgement. At baseline and without information about the results of the research interviews, GPs rated their patients' future cognitive status using the seven-point ordinal GDS (1 = no subjective or objective impairment to 7 = severe dementia)²⁵ and answering the question 'Which GDS stage do you expect the patient will have in 3 years?' This GP rating was analysed as a global clinical judgement of future cognitive status.

The MMSE (score 0–30)⁷ was used as a short test of the patient's global cognition. A patient's SCD⁸ was assessed with the questions: 'Do you feel like your memory is becoming worse?' (yes/no); in case of

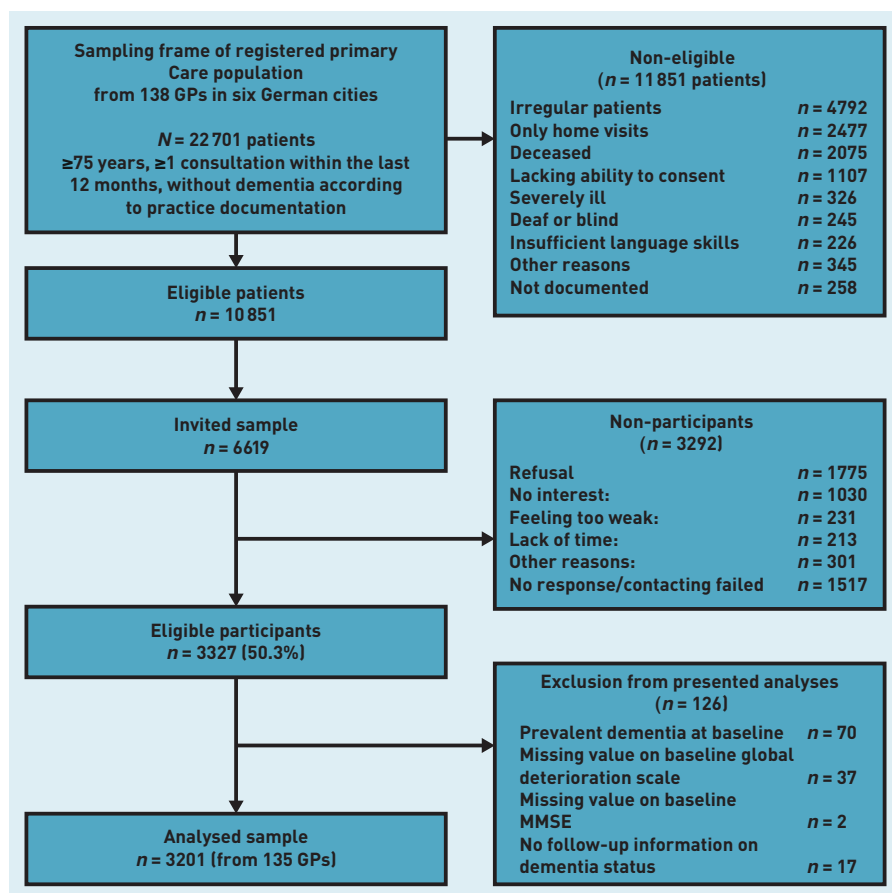


Figure 1. Patient recruitment flow chart.
MMSE = Mini-Mental State Examination.

response 'yes': 'Does this worry you?' (yes/no).²⁸

Characteristics of patients, doctor-patient relationship, and GPs' decision making. Beside the SISCO, subtests of the CERAD-NP (neuropsychological assessment of the Consortium to Establish a Registry for Alzheimer's Disease)²⁹ were used to measure memory performance. All drugs a patient was currently taking were documented in the interviews. GPs gave information about patients' comorbidities and were asked to describe their doctor-patient relationship in terms of familiarity with the patient (from 1 = very poor to 4 = very high) and years of doctor-patient contact.

To evaluate the relevance of cognitive tests in GPs' diagnostic repertoire, all GPs in the follow-up phases III-VI reported on whether they conducted or requested any cognitive testing for each patient.

Statistical analysis

Predictor variables were checked for multicollinearity in terms of bivariate correlations and all coefficients were found to be negligible and well below 0.50 (see Supplementary Table 1 for details).

To evaluate the independent effects of predictors on the development of incident dementia within 3 years, multivariable mixed-effects logistic regression was applied, considering the cluster effect of GP practices. Odds ratios (OR) were calculated with 95% confidence intervals (CI).

Multivariable mixed-effects Cox regression (again including the cluster effect of GP practices) was used to analyse the independent effects of predictors on time of dementia onset over 12 years. Hazard ratios (HR) were calculated with 95% CI.

For all models, the Bayesian information criterion was chosen as a robust and conservative indicator of model fit, with lower values indicating better fit.

Differences in dementia-free survival times between participants with different GP-predicted GDS stages were illustrated by Kaplan-Meier survival curves and tested for significance using the log rank test.

Analyses were performed with SPSS (version 24) and SAS (version 9.3).

RESULTS

At baseline, 138 GPs (19–29 per study centre) recruited a total of 3327 participants (mean number = 24.1, SD = 11.7). Most of the recruiting GPs were male (74.5%), their mean age was 50.9 years (SD = 6.5), and the mean of their years in practice was 14.9 (SD = 6.7). Figure 1 gives details of the patient recruitment process. Consenting patients ($n = 3327$) were slightly younger (mean = 80.1 years [SD = 3.8] versus 80.7 years [SD = 3.8]) and more often male (34.5% versus 31.1%) than refusing or non-responding patients ($n = 3292$). The analysed sample for all regression models consists of $n = 3201$ participants with complete baseline information (face-to-face patient interview and GP questionnaire) and at least one valid item of follow-up information on dementia status (based on face-to-face or informant interviews), from 135 GPs (for three GPs no patient was eligible for the analysis).

The proportion of participants with personal versus informant follow-up interview is more than 80% versus <20% for each follow-up assessment, and personally interviewed patients slightly differ from non-interviewed patients (see Supplementary Tables 2 and 3 for details). The main reasons for missing personal follow-up interviews were death and refusal.

Characteristics of patients and doctor-patient relationships

Table 1 gives baseline characteristics for the participants. A high generalisability is supported by the low prevalence setting

Table 1. Baseline characteristics of the patients in the study ($n = 3201$, all without dementia at baseline according to both reference standard and GP judgement)

Variable		Results
Age [years]: mean (SD), median (IQR)		79.6 (3.6), 79.0 (5.0)
Sex, n (%)	Female	2090 (65.3)
	Male	1111 (34.7)
Education, n (%)	Low	1982 (61.9)
	Intermediate	875 (27.3)
	High	344 (10.7)
GP's prediction of patient's cognitive status in 3 years from baseline [GDS stage, stated on a questionnaire at baseline], n (%)	GDS 1	1312 (41.0)
	GDS 2	1080 (33.7)
	GDS 3	586 (18.3)
	GDS ≥ 4	223 (7.0)
SCD at baseline (patient report in interview), n (%)	No SCD	1339 (41.8)
	SCD without worry	1323 (41.3)
	SCD with worry	539 (16.8)
MMSE (score 0–30), mean (SD), median (IQR)		27.5 (1.9), 28.0 (3.0)
CERAD-NP (memory scores)	Verbal fluency (named animals), mean (SD), median (IQR)	19.5 (5.5), 19.0 (7.0)
	Immediate recall (three 10-word list learning trials [0–10]), mean (SD), median (IQR)	6.2 (1.6), 6.3 (2.0)
	Delayed recall (10-word list [0–10]), mean (SD), median (IQR)	5.4 (2.2), 6.0 (3.0)
	Recognition (sum of hits and correct rejections [0–20]), mean (SD), median (IQR)	18.8 (1.7), 19.0 (2.0)
Familiarity with the patient (GP rating), n (%)	Very high	1433 (44.8)
	High	1599 (50.0)
	Poor	156 (4.9)
	Very poor	13 (0.4)
Duration of doctor–patient relationship (GP rating in years), mean (SD), median (IQR)		11.6 (7.3), 10.0 (12.0)
GP consultations per participant in the last 12 months (GP information available from $n = 560$ patients from 29 GPs of the Düsseldorf subsample), mean (SD), median (IQR)		13.9 (11.4), 12.0 (13.5)
Comorbidity (GP information)		
Diabetes, n (%)		719 (22.5)
Hypertension, n (%)		2274 (71.0)
Coronary heart disease, n (%)		1098 (34.3)
Depression, n (%)		374 (11.7)
Number of comorbidities, mean (SD), median (IQR)		3.6 (2.1), 3.0 (3.0)
Number of drugs taken: prescribed and OTC (information from interviews at patients' home), mean (SD), median (IQR)		4.6 (2.8), 4.0 (3.0)
Participants taking ≥ 5 drugs: prescribed and OTC (information from interviews at patients' home), n (%)		1506 (47.0)
Follow-up information		
Incident dementia, n (%)	Until FU2	216 (6.7)
	Until FU8	707 (22.1)
Patients for which GPs indicated 'cognitive testing' as diagnostic procedure, n (%)	FU3 ($n = 2263$ with GP questionnaire)	179 (7.9)
	FU4 ($n = 1851$ with GP questionnaire)	123 (6.6)
	FU5 ($n = 1122$ with GP questionnaire)	54 (4.8)
	FU6 ($n = 763$ with GP questionnaire)	75 (9.8)

CERAD-NP = neuropsychological assessment of the Consortium to Establish a Registry for Alzheimer's Disease. FU = follow-up assessment. GDS = Global Deterioration Scale (7-point scale from 1 to 7 indicating increasing cognitive impairment). IQR = interquartile range. MMSE = Mini-Mental State Examination. OTC = over the counter. SCD = subjective cognitive decline. SD = standard deviation.

with realistic incidence rates of dementia. However, compared with other, younger cohorts,³⁰ CERAD memory performance on verbal fluency and immediate word list recall in the AgeCoDe baseline sample was considerably better. GPs report excellent patient knowledge, and the mean consultation rate is higher than in representative German cohorts.³¹ The number of drugs taken and

the proportion of patients taking five or more drugs is comparable, but slightly smaller than in an unselected German primary care sample of community-dwelling older adults.³² Except for the high rate of coronary heart disease, prevalence rates for chronic conditions in the AgeCoDe cohort very closely correspond to the rates described in representative German cohorts.^{33,34}

Table 2. Independent effects of baseline predictors for the 3-year prediction of dementia (multivariable mixed-effects logistic regression, $n = 3201$)

Baseline predictors	Model with GP+SCD+MMSE		Model with GP+MMSE		Model with SCD+MMSE		Model with GP+SCD	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
Age (per 1 year older)	1.07 (1.03 to 1.11)	<0.001	1.07 (1.03 to 1.11)	0.001	1.09 (1.05 to 1.13)	<0.001	1.09 (1.05 to 1.13)	<0.001
Female (ref. male)	1.00 (0.72 to 1.38)	0.98	1.05 (0.76 to 1.46)	0.75	0.98 (0.71 to 1.36)	0.92	1.05 (0.76 to 1.45)	0.76
Education								
Intermediate	1.21 (0.84 to 1.74)	0.31	1.21 (0.84 to 1.73)	0.31	1.14 (0.80 to 1.62)	0.48	0.86 (0.61 to 1.22)	0.41
(ref. low) High	1.55 (0.92 to 2.62)	0.098	1.67 (0.99 to 2.79)	0.052	1.57 (0.95 to 2.62)	0.080	0.98 (0.59 to 1.62)	0.94
GP judgement								
GDS 2	1.97 (1.28 to 3.04)	0.002	2.14 (1.39 to 3.31)	0.001	— ^a	—	2.32 (1.49 to 3.59)	<0.001
(ref. GDS 1) GDS 3	3.00 (1.90 to 4.76)	<0.001	3.59 (2.27 to 5.69)	<0.001	—	—	4.11 (2.58 to 6.56)	<0.001
GDS 4+	5.66 (3.29 to 9.73)	<0.001	6.40 (3.72 to 10.99)	<0.001	—	—	8.76 (5.06 to 15.17)	<0.001
SCD (ref. no SCD)								
– Worry	1.47 (1.01 to 2.13)	0.042	—	—	1.59 (1.10 to 2.29)	0.013	1.52 (1.06 to 2.19)	0.025
+ Worry	3.41 (2.30 to 5.05)	<0.001	—	—	3.93 (2.68 to 5.77)	<0.001	3.53 (2.41 to 5.19)	<0.001
MMSE (per 1 point less)	1.46 (1.36 to 1.57)	<0.001	1.47 (1.37 to 1.57)	<0.001	1.51 (1.41 to 1.62)	<0.001	—	—
Model fit (BIC)^b	1394.55		1417.06		1418.09		1495.72	

^aThe dashes indicate that some predictors are not part of the model, for example, the model termed 'Model with GP+MMSE' shows dashes in the two rows for 'SCD', as SCD is not included in this model. ^bLower values indicating better model fit. BIC = Bayesian information criterion. CI = confidence interval. GDS = Global Deterioration Scale (7-point scale from 1 to 7 indicating increasing cognitive impairment). MMSE = Mini-Mental State Examination. OR = odds ratio. SCD = subjective cognitive decline.

Prediction of incident dementia within 3 years

In all models, the GP's judgement yielded additional and significant information for the 3-year prediction of dementia, independent of the patient's sex, age, education, SCD, and MMSE scores (Table 2).

The regression model with GP/SCD/MMSE shows a better model fit than all other models based on the Bayesian information criterion (lowest BIC).

Prediction of time of dementia onset over 12 years

In all models, the GP's judgement added specific information for the prediction of time to dementia onset in a 12-year period,

independent of the patient's sex, age, education, SCD, and MMSE scores (Table 3). The regression model with GP/SCD/MMSE outperformed all other models in terms of model fit based on the Bayesian information criterion (lowest BIC).

Multivariable mixed-effects Cox regression found that dementia-free survival significantly varied as a function of the GP's rating of a patient's future cognitive status (Figure 2).

DISCUSSION

Summary

In the study sample of patients with a close doctor–patient relationship, the GP's clinical judgement holds an added value

Table 3. Independent effects of baseline predictors for the prediction of time of dementia onset over 12 years (multivariable mixed-effects Cox regression, $n = 3201$)

Baseline predictors	Model with GP+SCD+MMSE		Model with GP+MMSE		Model with SCD+MMSE		Model with GP+SCD	
	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value	HR (95% CI)	P-value
Age (per 1 year older)	1.10 (1.08 to 1.12)	<0.001	1.10 (1.08 to 1.12)	<0.001	1.11 (1.09 to 1.14)	<0.001	1.12 (1.09 to 1.14)	<0.001
Female (ref. male)	1.07 (0.90 to 1.27)	0.45	1.10 (0.93 to 1.30)	0.27	1.06 (0.89 to 1.25)	0.51	1.09 (0.92 to 1.29)	0.32
Education								
Intermediate	1.01 (0.85 to 1.21)	0.91	1.02 (0.86 to 1.22)	0.80	0.98 (0.82 to 1.17)	0.80	0.84 (0.71 to 1.01)	0.056
(ref. low) High	1.15 (0.89 to 1.51)	0.29	1.18 (0.90 to 1.54)	0.22	1.12 (0.86 to 1.46)	0.39	0.85 (0.65 to 1.10)	0.21
GP judgement								
GDS 2	1.58 (1.29 to 1.92)	<0.001	1.65 (1.36 to 2.01)	<0.001	—	—	1.63 (1.34 to 1.98)	<0.001
(ref. GDS 1) GDS 3	1.98 (1.58 to 2.49)	<0.001	2.15 (1.71 to 2.69)	<0.001	—	—	2.32 (1.85 to 2.90)	<0.001
GDS ≥4	3.45 (2.59 to 4.59)	<0.001	3.73 (2.81 to 4.95)	<0.001	—	—	4.19 (3.16 to 5.55)	<0.001
SCD (ref. no SCD)								
– Worry	1.24 (1.03 to 1.48)	0.020	—	—	1.27 (1.07 to 1.52)	0.007	1.26 (1.05 to 1.50)	0.012
+ Worry	2.11 (1.72 to 2.57)	<0.001	—	—	2.31 (1.89 to 2.81)	<0.001	2.15 (1.76 to 2.62)	<0.001
MMSE (per 1 point less)	1.31 (1.26 to 1.36)	<0.001	1.31 (1.26 to 1.37)	<0.001	1.33 (1.28 to 1.39)	<0.001	—	—
Model fit (BIC)^a	5840.42		5873.17		5872.06		5960.15	

^aLower values indicating better fit. BIC = Bayesian information criterion. GDS = Global Deterioration Scale (7-point scale from 1 to 7 indicating increasing cognitive impairment). HR = hazard ratio. MMSE = Mini-Mental State Examination. SCD = subjective cognitive decline.

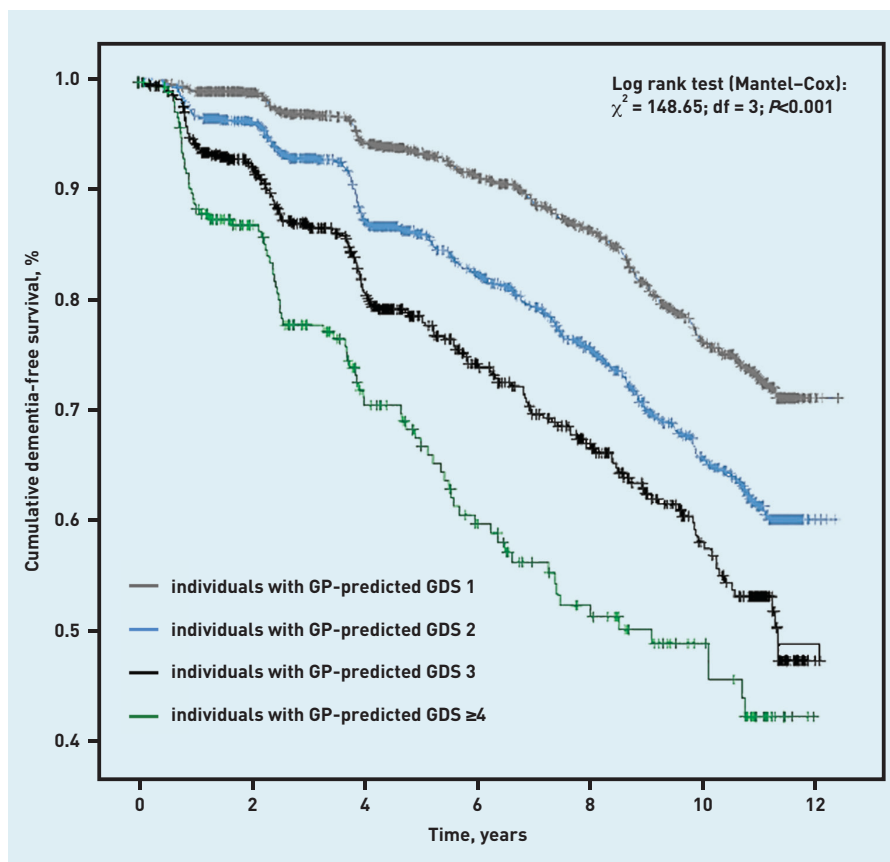


Figure 2. Kaplan-Meier curves of dementia-free survival in patients with different GP ratings on future cognitive status (GDS stage). GDS = Global Deterioration Scale.

Funding

This study is part of the German Research Network on Dementia (KND), the German Research Network on Degenerative Dementia (KNDD), and the Health Service Research Initiative, and was funded by the German Federal Ministry of Education and Research [grant numbers: grants KND: 01GI0102, 01GI0420, 01GI0422, 01GI0423, 01GI0429, 01GI0431, 01GI0433, 01GI0434; grants KNDD: 01GI0710, 01GI0711, 01GI0712, 01GI0713, 01GI0714, 01GI0715, 01GI0716; grants Health Service Research Initiative: 01GY1322A, 01GY1322B, 01GY1322C, 01GY1322D, 01GY1322E, 01GY1322F, 01GY1322G]. The funder had no influence on study design, collection, analysis, and interpretation of data, or the writing of the article and the decision to submit it for publication. Researchers are independent from funders.

for the prediction of dementia, apparently by providing information beyond cognitive test performance and patient's self-report. This applies to both short-term prediction of incident dementia within 3 years and long-term prediction of dementia onset over 12 years. The results underline the importance of the GP-patient relationship.

Strengths and limitations

The presented analyses and conclusions are restricted to the applied parameters; there are many alternative measures of SCD, test performance, and clinical dementia rating.

The strong effect of the MMSE may be partly explained by a bias in terms of methodological proximity of predictor MMSE and gold standard. The MMSE, as the only one of the examined parameters, is an explicit part of the diagnostic SIDAM algorithm for the gold standard of clinical dementia diagnosis.

Compared with other cohorts, participants show better cognitive and verbal abilities (CERAD) and a closer doctor-patient relationship in terms of GP familiarity, duration of doctor-patient contact, and number of consultations. These biases may be the result of the study's inclusion and exclusion criteria, as well as GPs' involvement

in the initial patient recruitment process. The effect of the GP's judgement may have benefited from this biased selection of participants with a more involved kind of doctor-patient contact. In other aspects, such as somatic comorbidities, drug intake, and dementia incidence, the analysed sample seems not to be biased.

The independent association of the GP's judgement was analysed and model fit indices of several prediction models were compared. These analyses can be interpreted as an indication of content validity, that is, the fundamental insight that a GP's judgement on patient cognition is independently associated with dementia risk. For analyses of predictive validity, studies following STARDdem criteria³⁵ are needed, incorporating the GP's judgement and other predictors for generating practically relevant parameters such as positive and negative predictive values.

The main strength of the presented study is its prospective design. Especially in cross-sectional diagnostic studies on cognitive impairment, the risk of incorporation bias is high.³⁵ A prospective design implies decoupling the reference standard from the index tests, thus reducing the risk of incorporation bias as well as diagnostic review bias. With regard to epidemiological studies, an elaborate diagnostic reference standard was used, mainly (the interview) being applied to most participants and thus minimising verification bias. The validity of the reference standard was further facilitated by re-evaluating a dementia diagnosis in follow-up visits (in cases with personal interviews), thereby reducing the risk of misclassification. However, dementia status in non-participants was derived from informant interviews, also including information from the GP. Apart from the validity aspects of informant reports, this could have introduced a bias to the benefit of the GP's judgement.

Another advantage of the study is a valid GP judgement in terms of explicit questions about cognitive status; the study did not rely on GPs' documentation of cognitive status in medical records.

Comparison with existing literature

To the authors' knowledge, this is the first analysis of the validity of GPs' clinical judgement regarding a patient's likelihood of future dementia. Two previous studies report relevant validity of an expert clinical rating for future dementia,^{36,37} but ratings were based on singular lengthy interviews (90–120 minutes) by 'dementia experts' and not GPs. Primary care decision making with its

Ethical approval

Approval was obtained from each study centre: Ethics Commission of the Medical Association Hamburg (OB/08 /022817/2007); Ethics Commission of the University of Bonn (050/02258/07); Medical Ethics Commission II, University of Heidelberg at the University Medical Center of Mannheim (0226.4/2002 2007-253E-MA); Ethics Commission at the Medical Center of the University of Leipzig (143/2002 309/2007); Ethics Commission of the Medical Faculty of the Heinrich Heine University Düsseldorf (2079/2002 2999/2008); and Ethics Committee of the TUM School of Medicine, Munich (713/02 713/02 E).

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Contributors

Further co-authors involved in the study were Jürgen in der Schmitt, Dagmar Weeg, Silke Mamone, Tobias Luck, Jochen Werle, Dagmar Lühmann, Christian Brettschneider, and Luca Kleineidam. Members of the AgeCoDe Study Group: Wolfgang Maier and Martin Scherer (principal investigators), Hendrik van den Bussche (principal investigator 2002–2011), Heinz-Harald Abholz, Christian Brettschneider, Cadja Bachmann, Horst Bickel, Wolfgang Blank, Hendrik van den Bussche, Sandra Eifflaender-Gorfer, Marion Eisele, Annette Ernst, Angela Fuchs, Kathrin Heser, Frank Jessen, Hanna Kaduszkiewicz, Teresa Kaufeler, Mirjam Köhler, Hans-Helmut König, Alexander Koppa, Carolin Lange, Diana Lubisch, Tobias Luck, Dagmar Lühmann, Melanie Lupp, Manfred Mayer, Edelgard Mösch, Michael Pentzek, Tina Posselt, Jana Prokein, Steffi Riedel-Heller, Susanne Röhr, Anna Schumacher, Janine Stein, Susanne Steinmann, Franziska Tebarth, Michael Wagner, Klaus Weckbecker, Dagmar Weeg, Jochen Werle, Siegfried Weyerer, Birgitt Wiese, Steffen Wolfgruber, and Thomas Zimmermann.

Acknowledgements

The authors thank all participants and GPs for their collaboration.

Discuss this article

Contribute and read comments about this article: bjgp.org/letters

specific diagnostic concepts, strategies, and holistic perspective has been described in theory as carrying diagnostic and prognostic potential.^{15,16} However, empirical evidence is sparse. The results of the current study underline the value of clinical judgement as a central parameter in the trigger stage: asking and testing patients is in contrast to what GPs normally do in the initial diagnostic phase.³⁸ As a first step, they usually evaluate complex constructs, such as the risk of dementia, in a more intuitive way, based on a close, trusting, and continuous doctor–patient relationship.

Throughout the literature, there exists a strong belief in the superiority of questionnaires and performance-based tests over global, intuitive clinical judgement. As generalist concepts may be distinct from specialist classifications (reported not only for dementia,³⁹ but also for depression⁴⁰ and headache),⁴¹ there is some questionable justification for assuming superiority of one over the other. Innovative study designs⁴² and

elaborate methods are needed to compare different approaches in a valid manner, to reveal the diagnostic yield of GPs' clinical judgement, and to disentangle the cognitive mechanisms behind it.

Implications for research

The presented analyses provide a first indication for the value of the GP's judgement and may stimulate rigorous studies on GPs' dementia recognition and prediction. From the results of this study and the subsequent discussion, it seems justified to consider the GP's judgement in prediction models as one parameter for the identification of individuals at risk of dementia, at least in familiar doctor–patient relationships.

As one of the central parameters in the diagnostic trigger stage, the GP's judgement should be further investigated beyond dementia. Addressing attitudinal and cognitive biases of the GP's judgement⁴³ may help to improve its validity.

REFERENCES

- Naylor MD, Karlawish JH, Arnold SE, *et al*. Advancing Alzheimer's disease diagnosis, treatment, and care: recommendations from the Ware Invitational Summit. *Alzheimers Dement* 2012; **8**(5): 445–452.
- Rosenberg A, Ngandu T, Rusanen M, *et al*. Multidomain lifestyle intervention benefits a large elderly population at risk for cognitive decline and dementia regardless of baseline characteristics: the FINGER trial. *Alzheimers Dement* 2018; **14**(3): 263–270.
- Mead N, Bower P. Patient-centred consultations and outcomes in primary care: a review of the literature. *Patient Educ Couns* 2002; **48**(1): 51–61.
- Livingston G, Baio G, Sommerlad A, *et al*. Effectiveness of an intervention to facilitate prompt referral to memory clinics in the United Kingdom: cluster randomised controlled trial. *PLoS Med* 2017; **14**(3): e1002252.
- Weinstein G, Seshadri S. Circulating biomarkers that predict incident dementia. *Alzheimers Res Ther* 2014; **6**(1): 6.
- Petersen RC, Lopez O, Armstrong MJ, *et al*. Practice guideline update summary: mild cognitive impairment. *Neurology* 2018; **90**(3): 126–135.
- Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state': A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; **12**(3): 189–198.
- Jessen F, Amariglio RE, van Boxtel M, *et al*. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement* 2014; **10**(6): 844–852.
- Licher S, Leening MJG, Yilmaz P, *et al*. Development and validation of a dementia risk prediction model in the general population: an analysis of three longitudinal studies. *Am J Psychiatry* 2019; **176**(6): 543–551.
- Walters K, Hardoon S, Petersen I, *et al*. Predicting dementia risk in primary care: development and validation of the Dementia Risk Score using routinely collected data. *BMC Med* 2016; **14**: 6.
- Barnes DE, Beiser AS, Lee A, *et al*. Development and validation of a brief dementia screening indicator for primary care. *Alzheimers Dement* 2014; **10**(6): 656–665.e1.
- Li J, Ogronik M, Devine S, *et al*. Practical risk score for 5-, 10-, and 20-year prediction of dementia in elderly persons: Framingham Heart Study. *Alzheimers Dement* 2018; **14**(1): 35–42.
- Ford E, Greenslade N, Paudyal P, *et al*. Predicting dementia from primary care records: a systematic review and meta-analysis. *PLoS One* 2018; **13**(3): e0194735.
- Moyer VA. Screening for cognitive impairment in older adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2014; **160**(11): 791–797.
- Heneghan C, Glasziou P, Thompson M, *et al*. Diagnostic strategies used in primary care. *BMJ* 2009; **338**: b946.
- Gulbrandsen P, Fugelli P, Hjortdahl P. General practitioners' knowledge of their patients' socioeconomic data and their ability to identify vulnerable groups. *Scand J Prim Health Care* 1998; **16**(4): 204–210.
- Thornton T. Tacit knowledge as the unifying factor in evidence based medicine and clinical judgement. *Philos Ethics Humanit Med* 2006; **1**(1): E2.
- Body R, Cook G, Burrows G, *et al*. Can emergency physicians 'rule in' and 'rule out' acute myocardial infarction with clinical judgement? *Emerg Med J* 2014; **31**(11): 872–876.
- Ilgen JS, Ma IW, Hatala R, Cook DA. A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Med Educ* 2015; **49**(2): 161–173.
- Pentzek M, Vollmar HC, Wilm S, Leve V. Putting dementia awareness into general practice: the CADIF approach. *Z Gerontol Geriatr* 2017; **50**(suppl 2): 44–47.
- Hamm RM. Clinical intuition and clinical analysis: expertise and the cognitive continuum. In: Dowie J, Elstein AS, eds. *Professional judgment: a reader in clinical decision making*. New York, NY: Cambridge University Press, 1988: 78–105.
- Koppara A, Wagner M, Lange C, *et al*. Cognitive performance before and after the onset of subjective cognitive decline in old age. *Alzheimers Dement (Amst)* 2015; **1**(2): 194–205.
- Zaudig M, Hiller W. SIDAM-Handbuch. *Strukturiertes Interview für die Diagnose einer Demenz vom Alzheimer Typ, der Multi-Infarkt- (oder vaskulären) Demenz und Demenzen anderer Ätiologie nach DSM-III-R, DSM-IV und ICD-10*. [SIDAM handbook: a structured interview for the diagnosis of dementia of the Alzheimer type, multi-infarct dementia and dementias of other aetiology according to DSM-III-R, DSM-IV and ICD-10]. Bern: Hans Huber, 1996.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th edn. Washington, DC: APA, 2000.
- Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *Am J Psychiatry* 1982; **139**(9): 1136–1139.
- Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry* 1968; **114**(512): 797–811.
- Brauns H, Steinmann S. Educational reform in France, West Germany and the United Kingdom: updating the CASMIN educational classification. *ZUMA Nachrichten* 1999; **23**(44): 7–44.
- Geertlings MI, Jonker C, Bouter LM, *et al*. Association between memory complaints and incident Alzheimer's disease in elderly people with normal baseline cognition. *Am J Psychiatry* 1999; **156**(4): 531–537.
- Morris JC, Heyman A, Mohs RC, *et al*. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989; **39**(9): 1159–1165.
- Fillenbaum GG, Heyman A, Huber MS, *et al*. Performance of elderly African American and White community residents on the CERAD Neuropsychological Battery. *J Int Neuropsychol Soc* 2001; **7**(4): 502–509.
- Grobe TG, Steinmann S, Szecseny J, Barmer GEK. Arztreport 2016. [Physician report 2016]. Siegburg: Asgard-Verlagsservice GmbH, 2016.
- Junius-Walker U, Theile G, Hummers-Pradier E. Prevalence and predictors of polypharmacy among older primary care patients in Germany. *Fam Pract* 2007; **24**(1): 14–19.
- Göfßwald A, Schienkiewitz A, Nowossadeck E, Busch MA. [Prevalence of myocardial infarction and coronary heart disease in adults aged 40–79 years in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1)]. [Article in German]. *Bundesgesundheitsbl* 2013; **56**(5–6): 650–655.
- Heidemann C, Du Y, Schubert I, *et al*. [Prevalence and temporal trend of known diabetes mellitus: results of the German Health Interview and Examination Survey for Adults (DEGS1)]. [Article in German]. *Bundesgesundheitsbl* 2013; **56**(5–6): 668–677.
- Noel-Storr AH, McCleery JM, Richard E, *et al*. Reporting standards for studies of diagnostic test accuracy in dementia: the STARDdem Initiative. *Neurology* 2014; **83**(4): 364–373.
- Okereke OI, Pantoja-Galicia N, Copeland M, *et al*. The SIST-M: predictive validity of a brief structured clinical dementia rating interview. *Alzheimer Dis Assoc Disord* 2012; **26**(3): 225–231.
- Daly E, Zaitchik D, Copeland M, *et al*. Predicting conversion to Alzheimer disease using standardized clinical information. *Arch Neurol* 2000; **57**(5): 675–680.
- Donner-Banzhoff N, Seidel J, Sikeler AM, *et al*. The phenomenology of the diagnostic process: a primary care-based survey. *Med Decis Making* 2017; **37**(1): 27–34.
- Dhedhi SA, Swinglehurst D, Russell J. 'Timely' diagnosis of dementia: what does it mean? A narrative analysis of GPs' accounts. *BMJ Open* 2014; **4**(3): e004439.
- Sielk M, Abholz HH. [Why do general practitioners characterize other patients as more depressive than psychiatrists do?]. [Article in German]. *Z Allg Med* 2005; **81**(11): 486–490.
- Bösner S, Hartel S, Diederich J, Baum E. Diagnosing headache in primary care: a qualitative study of GPs' approaches. *Br J Gen Pract* 2014; DOI: <https://doi.org/10.3399/bjgp14X681325>.
- Donner-Banzhoff N, Haasenritter J, Hüllermeier E, *et al*. The comprehensive diagnostic study is suggested as a design to model the diagnostic process. *J Clin Epidemiol* 2014; **67**(2): 124–132.
- Sapostnik G, Redelmeier D, Ruff CC, Tobler PN. Cognitive biases associated with medical decisions: a systematic review. *BMC Med Inform Decis Mak* 2016; **16**: 138.