American Journal of Epidemiology

Vol. 143, No. 3
Printed in U.S.A.

Predicting the Onset of Alzheimer's Disease Using Bayes' Theorem

Martin J. Prince

Bayes' theorem describes the effect of new information (e.g., a test result) on the probability of an outcome (e.g., a disease). Likelihood ratios for separate tests can be combined to assess the joint effect of their results on disease probability. This approach has been used to develop a test package for Alzheimer's disease that consists of some simple cognitive tests (Paired Associate Learning Test, Trailmaking Test, and Raven's Progressive Matrices) combined with age and family history of dementia. A total of 1,454 subjects who had been recruited into the Medical Research Council Elderly Hypertension Trial between 1983 and 1985 completed cognitive tests at entry to the trial (when they were without signs of dementia) and 1 month later. Their dementia status was ascertained in 1990–1991. The test package identified 52% of Alzheimer's disease cases with a 9% false-positive rate or 90% of Alzheimer's disease cases with a 29% false-positive rate. The author proposes the use of a similar test package in conjunction with a test for apolipoprotein E e4 status, which is a powerful risk factor for late-onset Alzheimer's disease, as a likelihood ratio approach to the prospective identification of Alzheimer's disease cases. This approach could be followed by ethically sound trials of new therapeutic agents for subjects who have a high probability of developing Alzheimer's disease. *Am J Epidemiol* 1996;143:301–8.

Alzheimer's disease; Bayes theorem; dementia; epidemiologic methods

SCREENING FOR ALZHEIMER'S DISEASE, APOLIPOPROTEIN E E4, AND BAYES' THEOREM

Recent reports (1–4) of a powerful and robust association between the apolipoprotein E e4 (apoE e4) allele and Alzheimer's disease have generated interest in the possibility of screening (5) before the disease is clinically evident. Enthusiasm has been tempered by the recognition that there is little point in early detection of a condition unless early intervention conveys some therapeutic advantage. Nonetheless, there is recent evidence that tetrahydroaminoacridine (tacrine) can improve cognition in some patients with established Alzheimer's disease (6–8); and trials of other compounds with the potential to modify the course of dementia are in progress. A useful therapeutic lead time may be achieved by implementing potential treatment trials on asymptomatic individuals who have a high probability of developing Alzheimer's disease in the near future.

Unfortunately, apoE e4 is an insufficiently specific test to be used on its own to screen for Alzheimer's disease. A recent paper (1) has suggested that 3 percent of controls (vs. 13 percent of cases) may be homozygous and 19 percent (vs. 50 percent of cases), heterozygous for apoE e4. Use of these prevalence rates and the presence of any apoE e4 allele as the test criterion suggest a test with 78 percent specificity and 63 percent sensitivity for a diagnosis of Alzheimer's disease. However, sensitivity and specificity are meaningful only if assessed in relation to the prevalence of the condition to be detected by screening. Using Bayes' theorem (9), we can calculate the posttest probability of disease given knowledge of the pretest probability (in this case, disease prevalence) and the likelihood ratio associated with different test results. The likelihood ratio is defined as the probability of test result in diseased persons/probability of test result in nondiseased persons. The likelihood ratios derived from the apoE e4 frequencies given above are 4.3 for apoE e4 homozygosity, 2.5 for heterozygosity, and 0.48 for absence of an apoE e4 allele. The likelihood ratio for a given test result is related to the pretest and posttest probability of disease, i.e., pretest odds of disease × likelihood ratio = posttest odds of disease. If the pretest probability of disease is 0.10 (a generous

Received for publication September 15, 1994, and in final form October 9, 1995.

Abbreviations: apoE e4, apolipoprotein E e4; CI, confidence interval; LR, likelihood ratio; MRC, Medical Research Council; NART, New Adult Reading Test; PALT, Paired Associate Learning Test; RM, Raven's Matrices; ROC, receiver operating characteristic; TMT, Trail Making Test.

From the Institute of Psychiatry, London, England.

Reprint requests to Dr. Prince, Section of Epidemiology and General Practice, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England.

estimate for the eventual lifetime prevalence for those who have already survived to ≥60 years of age), then the pretest odds are 0.1/(1 − 0.10) = 0.11. If a subject is then found to be homozygous for apoE e4, his or her posttest odds become 0.11 × 4.3 = 0.47. This translates into a posttest probability of disease (positive predictive value for the test) of 0.47/(1 + 0.47) = 0.32. The posttest probabilities for heterozygosity and for no apoE e4 allele are 0.22 and 0.05, respectively. The positive predictive values (0.32 and 0.22) encompass too much uncertainty to be of use to screened subjects and their clinicians. One reason for this shortcoming is the low prevalence rate of Alzheimer's disease. For a test with given predictive power, the posttest probability of disease is crucially dependent on the pretest probability. Rarely can a single test be used as an early indicator of a disease with as low a population prevalence as Alzheimer's disease. It can be shown that a test with a given likelihood ratio will provide a maximum "gain" of posttest diagnostic probability when the pretest probability is in the range of 0.4–0.6 (10). One solution then might be to apply the test to a target population with a known significant lifetime prevalence of the disease; in the case of apoE e4 and Alzheimer's disease, for example, the test might work satisfactorily in subjects with a strong family history of Alzheimer's disease. Alternatively, Bayes' theorem can be used to combine a number of moderately predictive tests into a more effective package. Given the assumption of conditional independence (i.e., that the results of the second test do not depend on the results of the first), then pretest odds × likelihood ratio (test 1) × likelihood ratio (test 2) = posttest odds (tests 1 and 2). Hence, apoE e4 status might be developed into a clinically useful diagnostic test for dementia if used in combination with tests for other characteristics associated with preclinical Alzheimer's disease. These might be gene markers but could, as in this analysis, consist of cognitive tests, age, and the presence or absence of a family history of dementia.

MATERIALS AND METHODS

Sources of data

Three studies provided data for this analysis. Details of the study designs are described in the cited references, and brief details follow.

1) The Medical Research Council (MRC) Hypertension Treatment Trial in older subjects (aged 65–74) (11) was conducted between 1982 and 1986. A total of 4,396 individuals were recruited in a randomized, single blind, placebo-controlled trial to discover whether reduction of moderate hypertension reduced mortality from stroke. Subjects were followed up for an average of 5.8 years. The trial was conducted in 226 general practices (primary care units) throughout the United Kingdom.

2) The MRC cognitive substudy (12) was included in the main trial to assess whether reduction in blood pressure altered rates of cognitive decline during the follow-up period. The 2,651 participants were recruited sequentially as all consenting MRC trial entrants between February 1983 and October 1985. The psychometric test package consisted of the Paired Associate Learning Test (PALT) (13), Trail Making Test (TMT) (14), Raven's Matrices (RM) (15), and New Adult Reading Test (NART) (16) and was administered at entry and at months 1, 9, 21, and 54. The cognitive tests were administered to subjects in the individual practices by specially trained MRC research nurses.

3) The Alzheimer's disease and dementia case-control study (based on the MRC hypertension treatment trial) (17) was conducted in 1990–1991. A total of 1,545 subjects were screened from 71 of the practices participating in the MRC trial to identify all cases of dementia incident since the beginning of the trial according to DSM III-R criteria (18) (including Alzheimer's disease, vascular dementia, mixed vascular and Alzheimer's dementia, and amnestic syndrome) and possible or probable Alzheimer's disease according to the criteria of the National Institute of Neurological Disorders and Stroke-Alzheimer's Disease and Related Disorders Research Association (NINCDS-ADRDA) (19). The 71 practices were those whose trial participants included at least one whose PALT score had fallen into the impaired range over the course of the cognitive substudy. Subjects who had shown no evidence of decline on the PALT over the 5 years of the MRC trial were thought unlikely to have developed dementia in the 2 years since the end of that trial. Subjects selected for screening by a research psychiatrist were therefore those who had shown evidence of decline on the PALT, those whose general practitioner or research nurse was in any way concerned regarding their cognitive status, and one control per practice who showed no evidence of PALT decline to check the assumption that such subjects were likely to be free of dementia. On this basis, 293 subjects were selected for further screening. A total of 216 (74 percent) were seen by the research psychiatrist who administered the Mini-Mental State Examination (20). Of the original 293 subjects, 23 had died, 13 had moved away, and the remainder declined to participate. Thirty-nine of these subjects scored below the Mini-Mental State Examination

dementia screening threshold of 24/25, indicating the possible presence of dementia. These potential cases were definitively diagnosed by means of CAMCOG (21), GMS-AGECAT (22), and a general clinical examination. Appropriate investigations were used to exclude secondary dementia. Thirty-five individuals were found to meet diagnostic criteria for dementia. With the use of medical records and informants, the inquiry was extended as far as possible to include subjects who had died or moved from their practice between the beginning of the MRC trial and the 1991 survey. An additional 15 dementia cases were ascertained in this manner. Of the 50 dementia cases thus identified, 31 were probable or possible Alzheimer's disease cases according to the criteria of the National Institute of Neurological Disorders and Stroke. Data on family history of dementia were ascertained from 315 subjects during the 1991 dementia case-control study, comprising all dementia cases, all those whose PALT scores had declined significantly during the MRC trial, and a random subsample of 223 noncases (17). These data were obtained from an informant, or from the subject if no informant was available, and were coded as a dichotomous variable: convincing history of dementia in at least one first degree relative—yes/no.

The presence of dementia was an exclusion criterion for entry into the MRC trial allowing the assumption that subjects were free of clinically apparent dementia at its outset. The coupling of the baseline psychometric test data with subsequent knowledge of dementia case status allows assessment of the potential of some simply administered tests (NART, RM, PALT, and TMT) as screening tests for the onset of frank dementia.

The sample used for the analysis was finally reduced to 1,454 by the elimination of 91 subjects (four dementia cases and 87 noncases) who had not provided at least one cognitive test score at entry to the MRC trial.

Analysis

Selection of potential screening tests. Given the interest in the predictive power of the cognitive tests, the analysis was restricted to tests carried out at entry or in the first month of the MRC trial, before clinical signs were apparent in those who subsequently developed dementia. These tests were RM and NART (administered at entry) and PALT and TMT (administered at entry and at 1 month) (12). Two methods were used to determine the predictive potential of the psychometric test data and other variables for the later onset of dementia and Alzheimer's disease in the study sample. Dichotomous variables (e.g., family history of dementia) were cross-tabulated with dementia case status. To assess the effect of the time between initial cognitive testing and dementia ascertainment in 1990–1991 on the likelihood of being identified as a case of dementia, dementia case status was stratified by year of entry into the MRC trial. To examine the effect of dropout from the MRC trial, dementia case status was also stratified by the number of MRC trial follow-up data points collected. The frequency distributions of subjects' ages at entry to the MRC trial, their PALT and TMT scores at entry and at 1 month (PALT0, PALT1, TMT0, TMT1), the RM and the NART, together with their means and standard deviations were compared for the dementia case and noncase groups. To assess the extent of independent association between the test variables and dementia status, the same variables were next entered individually and then simultaneously (forced entry) into a logistic regression model with dementia and Alzheimer's disease case status as the dependent variables. Some variables might have shown a nonlinear association with dementia or Alzheimer's disease outcome. Therefore the test package indicated by the logistic regression exercise was extended by each of the remaining variables in turn to determine whether their inclusion improved the overall predictive power of the package.

Calculation of likelihood ratios. Continuous variables associated with the diagnoses of dementia and Alzheimer's disease were divided into quintiles or quartiles as appropriate. For dichotomous variables, likelihood ratios were calculated for both exposure levels. Subjects with missing values for a test level were allotted a neutral likelihood ratio of 1. As the results of the PALT at entry and at 1 month were significantly correlated, separate likelihood ratios were calculated for every possible combination of these two test scores to fulfill the assumption of conditional independence. As some combinations of PALT0 and PALT1 scores occurred in the noncase but not in the case group, these noncases were allotted a likelihood ratio calculated on the basis of their PALT1 score only. Although other tests were less associated with each other, it was assumed that their results were effectively independent of each other.

Combination of tests into a package. For each disease group (dementia and Alzheimer's disease), summary likelihood ratios were obtained for each subject by calculating the product of the individual likelihood ratios for PALT0/PALT1, RM, TMT1, age, and family history of dementia.

Assessment of the predictive potential of the test package. The distribution of summary likelihood ra-

tios was used to divide the sample into deciles, allowing estimation of the Alzheimer's disease and dementia case detection rates for each 10 percent increment in the false-positive rate of the test. An overall measure of the predictive power of the test is given by the area under the receiver operator characteristic (ROC) curve. The ROC curve was obtained by plotting sensitivity against the false-positive rate for each likelihood ratio decile. The area under the curve was calculated according to the algorithm proposed by Hanley and McNeil (23). In an ideal test, the area under the ROC curve would equal 1 and in a useless test, 0.5.

Split-half reliability. To assess the reliability of these results, the sample was next divided at random into two halves. Likelihood ratios for PALT0/1, TMT1, RM, age, and family history of dementia were derived from one half of the sample (development data set) and then were applied to the other half (test data set). Summary likelihood ratios were calculated as before for each subject, and their distribution was divided into deciles. The predictive power of the test package derived from the development data set was then compared when applied first with the development data set and then with the test data set. The indices of predictive power were again the case detection rates at 10 percent increments of the false-positive rate and the areas under the ROC curves.

RESULTS

Characteristics of the study sample

Of the subjects, 58.9 percent were female. The mean age of subjects at entry to the MRC trial was 70.5 years (range, 65–75; standard deviation 2.7). At the beginning of the trial, they were hypertensive by definition (mean systolic blood pressure of 185 mmHg; range, 160–214; standard deviation 12.9). The mean NART score at entry was 30.0 (standard deviation 10.9). This approximates an intelligence quotient score of 111, which is 11 points more than the norm for the general population according to Wechsler's Adult Intelligence Scale.

Individual screening test variables

Response rates for the cognitive test scores ranged from 94 percent (TMT1) to 99 percent (TMT0), and age was recorded in all except nine subjects. Family history of dementia was recorded only in 315 subjects (22 percent) (see Methods). The significant differences for the means of age and of all cognitive test scores (except the NART) between noncases and the two case groups are presented in table 1. The future dementia cases were older and performed less well on

TABLE 1. Means and standard deviations (SDs) for potential test variables in noncase, dementia, and Alzheimer's disease groups, with *t* tests for difference between the case and noncase group means, Medical Research Council Elderly Hypertension Treatment Trial, 1982–1986, and Dementia Case-Control Study, 1990–1991, United Kingdom

	Noncases	Cases			Dementia vs. noncases			Alzheimer's disease vs. noncases		
		Dementia	Alzheimer's disease							
	Mean (SD) (n = 1,325–1,402)	Mean (SD) (n = 44–46)	Mean (SD) (n = 27–28)		*df*	*t* value	*p* value	*df*	*t* value	*p* value
Age (years)	70.4 (2.8)	71.7 (2.3)	71.6 (2.5)		1,446	3.2	0.002	1,429	2.3	0.02
Test score										
Entry PALT*	17.0 (1.7)	16.2 (2.0)	16.3 (2.0)		1,431	3.1	0.002	1,414	2.1	0.03
PALT (at 1 month)	16.6 (1.9)	15.4 (2.7)	15.1 (2.6)		1,364	3.8	<0.001	1,347	3.8	<0.001
Entry TMT*	57.2 (26.3)	69.4 (35.9)	70.9 (36.6)		1,437	3.1	0.002	1,420	2.7	0.006
TMT (at 1 month)	53.4 (24.1)	66.0 (36.7)	71.3 (44.6)		1,367	3.3	0.001	1,350	3.7	<0.001
RM*	15.3 (4.0)	13.3 (3.8)	13.4 (4.2)		1,417	3.2	0.001	1,401	2.4	0.015
NART*	30.1 (10.8)	28.0 (11.3)	29.4 (12.0)		1,426	1.3	0.21	1,409	0.3	0.74

* PALT, Paired Associate Learning Test; TMT, Trail Making Test; RM, Raven's Matrices; NART, New Adult Reading Test.

the tests carried out at entry and at 1 month into the MRC trial.

Neither year of entry into the MRC trial (chi-square test for trend = 0.001 (p = 0.98)) nor duration of follow-up during the MRC trial (chi-square test for trend = 1.14 (p = 0.29)) was associated with dementia status. Those who had a family history of dementia included 29.4 percent of dementia cases and 28 percent of Alzheimer's disease cases compared with 12.5 percent of noncases (chi-square = 5.8 (p = 0.016) for dementia, and chi-square = 3.46 (p = 0.063) for Alzheimer's disease). The coefficients of linear association between test variables and the log odds of Alzheimer's disease status derived from logistic regression are given in table 2. Associations with dementia followed a similar pattern and are not reported. PALT1, TMT1, and age are clearly independently associated with Alzheimer's disease. PALT0, TMT0, and RM dropped out of the equation on forced simultaneous entry. RM and PALT0 were subsequently reinstated into the test package on the empirical grounds that doing so significantly improved the predictive value of the test, indicating nonlinear association with dementia outcome. The final test package consisted of PALT0/PALT1, TMT1, RM, age, and family history of dementia. To illustrate the calculation of likelihood ratios for individual tests, likelihood ratios for the diagnoses of dementia and Alzheimer's disease for the quintiles of the Raven's matrices (RM) test score distribution are given in table 3, and the 16 combinations of the entry and 1-month PALT scores are presented in table 4.

TABLE 2. Logistic regression linear coefficients for association between test variables and the log odds of Alzheimer's disease outcome, Medical Research Council Elderly Hypertension Treatment Trial, 1982–1986, and Dementia Case-Control Study, 1990–1991, United Kingdom

Variable	Coefficient			
	Unadjusted	p value	Adjusted*	p value
Age (years)	0.187	0.02	0.160	0.04
Test score				
Entry PALT†	−0.134	0.12	0.010	0.92
PALT (at 1 month)	−0.247	<0.01	−0.199	0.01
Entry TMT†	0.009	0.06	−0.007	0.34
TMT (at 1 month)	0.016	<0.01	0.015	0.03
RM†	−0.090	0.06	−0.020	0.72

* Adjusted for age and scores from the tests listed in the left column.

† PALT, Paired Associate Learning Test; TMT, Trail Making Test; RM, Raven's Matrices.

TABLE 3. Frequencies of noncases, Alzheimer's disease cases, and dementia cases by Raven's Progressive Matrices test level with associated likelihood ratios, Medical Research Council Elderly Hypertension Treatment Trial, 1982–1986, and Dementia Case-Control Study, 1990–1991, United Kingdom

Raven's Matrices level (test score range)	Noncases	Cases			
		Dementia		Alzheimer's disease	
		No.	Likelihood ratio	No.	Likelihood ratio
1 (<13)	338	19	1.72	13	1.82
2 (13–14)	225	11	1.49	7	1.47
3 (15–17)	366	8	0.667	2	0.260
4 (18–19)	228	6	0.804	6	1.25
5 (>19)	217	1	0.141	1	0.218
Total	1,374	45		29	

Performance of the combined screening test package

Dementia and Alzheimer's disease case detection rates at various false-positive rate levels are given in table 5. The false-positive rates have been set at the same level for the prediction of dementia and Alzheimer's disease, allowing direct comparison of the case detection rates for the two diagnoses. Cases that are detected with a 49 percent false-positive rate include 87 percent of dementia cases and 93 percent of Alzheimer's disease cases. At lower, more realistic false-positive rates, the test package appears to detect cases of Alzheimer's disease better than those of dementia. The inclusion of family history of dementia (albeit ascertained in only one fifth of the sample) contributed usefully to the sensitivity of the test package, increasing the detection rate for Alzheimer's disease from 45 to 52 percent at the 9 percent false-positive level. Omitting family history of dementia from the test package, the area under the ROC curve for the prediction of Alzheimer's disease was 0.82 (95 percent CI 0.75–0.88). Including family history of dementia increased the area under the curve for the prediction of Alzheimer's disease to 0.84 (95 percent CI 0.78–0.90). The area under the ROC curve for the prediction of dementia was 0.77 (95 percent CI 0.71–0.84). The results of the split-half reliability exercise for Alzheimer's disease are given in table 6. False-positive rates were set at the same level for the two data sets so that the dementia detection rates could be compared directly. It can be seen that although there is some loss of sensitivity at certain levels of specificity, by and large the test performs as well in the test data

TABLE 4. Frequency of dementia cases by entry PALT* score and PALT score at 1 month with associated likelihood ratios (LRs) for dementia, Medical Research Council Elderly Hypertension Treatment Trial, 1982–1986, and Dementia Case-Control Study, 1990–1991, United Kingdom

Entry PALT level (actual score)	PALT score level at 1 month (actual score							
	1 (<16)		2 (16)		3 (17)		4 (18)	
	No.	LR	No.	LR	No.	LR	No.	LR
1 (<16)	6/95	2.03	2/38	1.65	3/34	2.88	1/47	0.65
2 (16)	6/46	4.46	1/18	1.75	0/26	1.13†	0/44	0.60†
3 (17)	3/64	1.46	1/35	0.88	2/53	1.17	3/86	1.08
4 (18)	3/101	0.91	1/87	0.35	1/150	0.20	11/432	0.78

* PALT, Paired Associate Learning Test.
† Calculated on the basis of entry PALT score only.

set as in the development data set from which it was derived. This is reflected in the similarity of the areas under the two ROC curves: 0.82 (95 percent CI 0.72–0.92) for the development data set and 0.78 (95 percent CI 0.70–0.86) for the test data set.

DISCUSSION

It is hoped that this paper will draw attention to the potential application of the Bayes' theorem approach to the early identification of Alzheimer's disease, although it is not purported to present a thoroughly examined or practical test package. Two important reservations must be expressed. First, circularity is implicit in this analysis. The characteristics of dementia and Alzheimer's disease cases and noncases have been examined, quantified, and then applied to the same population to identify the same cases. Chance would suggest that if the likelihood ratios derived in this population were applied to another, then less impressive screening results would be obtained. However, this issue has been addressed in a split-half reliability exercise in which likelihood ratios were derived from one half of the sample and tested on the other. Fortunately, the characteristics of the first sample appear to apply adequately to the second. Second, this study sample may not be representative of the population that is older than 65 years. All subjects were moderately hypertensive and, as volunteers for a medical trial, are not typical of their peers. Their volunteerism is perhaps reflected in their mean premorbid intelligence quotient, which is 11 points more than the norm. For this reason, likelihood ratios de-

TABLE 5. Case detection rates for Alzheimer's disease and dementia at different false positive rates*, Medical Research Council Elderly Hypertension Treatment Trial, 1982–1986, and Dementia Case-Control Study, 1990–1991, United Kingdom

False-positive rate† (%)	Detection rate of cases	
	Dementia(%) (n = 44)	Alzheimer's disease (%) (n = 29)
49	87	93
39	78	90
29	70	90
19	52	66
9	44	52

* The test package, derived from and applied to the whole data set, was based on a summary likelihood ratio calculated from age, family history of dementia, and scores from the following tests: Raven's Matrices, Trail Making Test (at 1 month), and Paired Associate Learning Tests (at entry and at 1 month combined).
† n = 1,408 noncases.

TABLE 6. Split half reliability of Alzheimer's disease case detection rates in two data sets at different false positive rates*, Medical Research Council Elderly Hypertension Treatment Trial, 1982–1986, and Dementia Case-Control Study, 1990–1991, United Kingdom

False-positive rate (%)	Detection rate	
	Development data set (%) (n = 17 cases, 700 noncases)	Test data set (%) (n = 12 cases, 708 noncases)
49	94	100
39	88	92
29	88	58
19	53	58
9	53	17

* The test package was derived from the development data set and applied to the development and test data sets.

rived from this population should not be generalized indiscriminately.

However, two factors may have led to an underestimation of the screening potential of the test package. First, even though every effort was made to ascertain all cases of dementia and Alzheimer's disease arising among the 1,545 subjects recruited in the 71 MRC practices for the purposes of the case-control study, it is likely that some cases were missed. Follow-up was not complete, particularly when subjects had died or moved away between the beginning of the MRC trial and 1990–1991. Misclassification of disease status, whether caused by loss to follow-up or by failure of the dementia process to manifest itself during the follow-up period, is likely to have been randomly distributed with respect to test results. Such misclassification would tend to shift likelihood ratios toward unity, reducing the discriminating power of the test package. In practice, misclassification bias was not particularly apparent. There was no relation between the number of MRC trial follow-up screening sessions attended and the likelihood of being ascertained as a case of dementia. Again, although one might expect a relation between date of entry into the MRC trial and dementia case status in 1990–1991, this did not appear to be a factor for subjects who entered the trial in 1983 with 2 more years to develop dementia compared with those who entered the study in 1985. Second, exposure to family history of dementia emerged as an important part of the test package. Because this exposure was recorded in only 22 percent of the sample, one might expect that more complete ascertainment would improve the performance of the test package.

The Bayesian approach to disease diagnosis is well established in cardiology practice (24). It has a potential application in any situation in which two or more tests may be independently predictive of the presence or onset of the disease but no single test can provide adequate predictive power. Bayes' theorem allows us to estimate the odds of disease given a particular test result. To that extent, it resembles a logistic model in which the log odds of disease are estimated given a particular exposure. In this analysis, the Bayes' theorem approach has been preferred because of its practicality and transparency. A likelihood ratio may be calculated from a cross-tabulation with a pocket calculator. Always assuming conditional independence of the tests, a likelihood ratio pertaining to one test can be multiplied with another to produce a likelihood ratio summarizing the combined effect of the two test results. In the foreseeable future, a clinician who has a valid table of likelihood ratios may be able to answer with some confidence a query from a 64-year-old man regarding his lifetime probability of developing Alzheimer's disease when he is heterozygous for apoE e4 and has a father already affected by the disease.

In addition to advising subjects at known high risk of developing Alzheimer's disease, the Bayes' theorem approach could be used to screen a wider population regularly after the age of 65 using a combination of a single apoE e4 test to determine innate risk followed by annual or biennial cognitive tests. It is probable that the cognitive tests that form a large part of this test package are impaired as a preclinical manifestation of the dementing process. This places the test package in a different category from that of apoE e4 or other genetic markers, which could conceivably be used as tests at any time from birth on. This procedure might identify the pathological process of Alzheimer's disease in its early preclinical phase, giving a useful lead time when plausible therapies might be tried ethically but before the clinical onset of dementia and the major structural damage associated with it.

This is an important area for future research. The tests described in this paper, together with other plausible screening tests including apoE e4, need to be refined in longitudinal studies of unselected populations of older subjects so that a practical and reliable test package with a positive predictive value well in excess of 50 percent might be developed. Examples from the literature of other risk factors with potential for inclusion in a likelihood ratio (LR) test package include head injury with loss of consciousness (25) (LR^+ 5.3, LR^- 0.96) and the presence of extrapyramidal signs (26) (LR^+ 3.2, LR^- 0.71).

ACKNOWLEDGMENTS

The author expresses appreciation to the following individuals: Professor Anthony Mann (under whose supervision this work has been completed); from the Medical Research Council (MRC) Working Party and the MRC Epidemiology and Medical Care Unit, Director Professor T. Meade (for permission to investigate trial subjects); from the MRC Framework General Practitioners and Research Nurses, Coordinator W. Browne (for data collection), Dr. Rowan Harwood (for suggesting the likelihood ratio approach), Dr. Simon Lovestone (for helpful comments), Dr. Anne Bird (for coordinating the cognitive substudy), Patrick Brennan and Robert Blizard (for computing support), and Dr. Paul Williams and Glaxo Pharmaceuticals (for funding of the dementia case-finding study).

Dr. Prince has held an Epidemiology Training Fellowship, supported by the Wellcome Trust, while carrying out this work.

REFERENCES

1. Poirier J, Davignon J, Bouthillier D, et al. Apolipoprotein E polymorphism and Alzheimer's disease. Lancet 1993;342: 697–9.
2. Saunders AM, Strittmatter WJ, Schmechel D, et al. Association of apolipoprotein E allele e4 with late-onset familial and sporadic Alzheimer's disease. Neurology 1993;43:1467–72.
3. Mayeux R, Stern Y, Ottman R, et al. The apolipoprotein e4 allele in patients with Alzheimer's disease. Ann Neurol 1993; 34:752–4.
4. Saunders AM, Schmader K, Breitner JCS, et al. Apolipoprotein E e4 allele distributions in late-onset Alzheimer's disease and in other amyloid forming diseases. Lancet 1993;342: 710–11.
5. Scott J. Apolipoprotein E and Alzheimer's disease. Lancet 1993;342:696.
6. Farlow M, Gracon S, Hershey LA, et al. A controlled trial of tacrine in Alzheimer's disease. JAMA 1992;268:2523–9.
7. Knapp MJ, Knopman DS, Solomon PR, et al. A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease. JAMA 1994;271:985–91.
8. Byrne EJ, Arie T. Tetrahydroaminoacridine and Alzheimer's disease. BMJ 1994;308:868–9.
9. Gorry GA, Barnett GO. Sequential diagnosis by computer. JAMA 1968;205:849–54.
10. Sox HC. Probability theory in the use of diagnostic tests. Ann Intern Med 1986;104:60–5.
11. MRC Working Party. Medical Research Council Trial of treatment of hypertension in older adults: principal results. BMJ 1992;304:405–12.
12. Bird AS, Blizard AR, Mann AH. Treating hypertension in the older person: an evaluation of the association of blood pressure level and its reduction with cognitive performance. J Hypertens 1990;8:147–52.
13. Inglis J. A paired associate learning test for use with elderly psychiatric patients. J Ment Sci 1959;105:440–3.
14. Reltan RM. Validity of the Trail Making Test as an indication of organic brain damage. Percept Mot Skills 1959;8:271–6.
15. Raven JC. Matrix tests. Ment Health 1940;1:10–18.
16. Nelson H, O'Connell A. Dementia: the estimation of premorbid intelligence levels using the New Adult Reading Test. Cortex 1978;14:234–44.
17. Prince MJ, Cullen MC, Mann AH. Risk factors for Alzheimer's disease and dementia: a case-control study based on the MRC Elderly Hypertension Trial. Neurology 1994;44: 97–104.
18. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 3rd ed revised. APA: Washington, DC, 1987.
19. McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–44.
20. 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;12:189–98.
21. Roth M, Tym E, Mountjoy CQ, et al. CAMDEX: a standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. Br J Psychiatry 1986;149:698–709.
22. Copeland JRM, Dewey ME, Griffith-Jones HM. A computerised psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGECAT. Psychol Med 1986;16: 89–99.
23. Hanley JA, McNeil BJ. The meaning and use of the area under the receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
24. Rifkin RO, Hood WB. Bayesian analysis of electrocardiographic stress testing. N Engl J Med 1977;297:681–6.
25. Mayeux R, Ottman R, Tang M, et al. Genetic susceptibility and head injury as risk factors for Alzheimer's disease among community-dwelling elderly persons and their first-degree relatives. Ann Neurol 1993;33:494–501.
26. Richards M, Stern Y, Mayeux R. Subtle extrapyramidal signs can predict the development of dementia in elderly individuals. Neurology 1993;43:2184–8.

APPENDIX

In an attempt to assess the potential for combining apoE e4 status with the test package presented in this paper, apoE e4 status was assigned to dementia cases and noncases on a random basis with probability proportional to the prevalence data previously cited (1, p. 1). The Alzheimer's disease detection rates for the test package alone and for apoE e4 combined with the test package are presented in Appendix table 1. It should be stressed that this projection, which was obtained by randomly assigning allele prevalences from another study to cases and noncases from this study, does *not* relate to apoE e4 data ascertained directly from this sample. This random assignation assumes conditional independence between apoE e4 and the other tests in the package, an assumption unlikely to be fulfilled in practice. Nonetheless, the results are extremely promising. The Alzheimer's disease detection rate is boosted from 52 to 66 percent, at the 9 percent false-positive level, and the area under the ROC curve is increased from 0.84 (95 percent CI 0.78–0.90) to 0.88 (0.84–0.93).

APPENDIX TABLE 1. Case detection rates for Alzheimer's disease at different false-positive rates for the basic test package* showing the projected effect of the inclusion of apolipoprotein E e4, Medical Research Council Elderly Hypertension Treatment Trial, 1982–1986, and Dementia Case-Control Study, 1990–1991, United Kingdom

False-positive rate† (%)	Detection rate of cases	
	Basic test package (%) (n = 29)	Basic test package + apolipoprotein E e4 (%) (n = 29)
49	93	100
39	90	97
29	90	86
19	66	76
9	52	66

* The test package was based on a summary likelihood ratio calculated from age, family history of dementia, and scores from the following tests: Raven's Matrices, Trail Making Test (at 1 month), and Paired Associate Learning Test (at entry and at 1 month combined).
† n = 1,408 noncases.