International Psychogeriatrics, Vol. 12, No. 2, 2000, pp. 231-247

Tracking Cognitive Decline in Alzheimer's Disease Using the Mini-Mental State Examination: A Meta-Analysis

LING HAN, MARTIN COLE, FRANCOIS BELLAVANCE, JANE MCCUSKER, AND FRANCOIS PRIMEAU

ABSTRACT. *Objectives*: To estimate the annual rate of change scores (ARC) on the Mini-Mental State Examination (MMSE) in Alzheimer's disease (AD) and to identify study or population characteristics that may affect the ARC estimation. *Methods*: MEDLINE was searched for articles published from January 1981 to November 1997 using the following keywords: AD and longitudinal study or prognosis or cognitive decline. The bibliographies of review articles and relevant papers were searched for additional references. All retrieved articles were screened to meet the following inclusion criteria: (a) original study; (b) addressed cognitive decline or prognosis or course of AD; (c) published in English; (d) study population included AD patients with ascertainable sample size; (e) used either clinical or pathological diagnostic criteria; (f) longitudinal study design; and (g) used the MMSE as one of the outcome measures. Data were systematically abstracted from the included studies, and a random effects regression model was employed to synthesize relevant data across studies and to evaluate the effects of study methodology on ARC estimation and its effect size. *Results*: Of the 439 studies screened, 43 met all the inclusion criteria. After 6 studies with inadequate or overlapping data were excluded, 37 studies involving 3,492 AD patients followed over an average of 2 years were included in the meta-analysis. The pooled estimate of ARC was 3.3 (95% confidence interval [CI]: 2.9-3.7). The observed variability in ARC across studies could not be explained with the covariates we studied, whereas part of the variability in the effect size of ARC could be explained by the minimum MMSE score at entry and number of assessments. *Conclusions*: A pooled average estimate of ARC in AD patients was 3.3 points (95% CI: 2.9-3.7) on the MMSE. Significant heterogeneity of ARC estimates existed across the studies and cannot be explained by the study or population characteristics investigated. Effect size of ARC was related to the initial MMSE score of the study population and the number of assessments.

From the Departments of Psychiatry (L. Han, MD; M. Cole, MD, FRCPC; and F. Primeau, MD, FRCPC) and Epidemiology (J. McCusker, MD, DrPH), St. Mary's Hospital and McGill University, Montreal, Canada, and the Department of Quantitative Methods, Ecole des Hautes Etudes Commerciales (F. Bellavance, PhD), Montreal, Canada.

Offprints. Requests for offprints should be directed to Martin Cole, MD, FRCPC, Department of Psychiatry, St. Mary's Hospital Center, 3830 Lacombe Ave., Montreal, PQ H3T 1M5 Canada.
e-mail: mih6@musica.mcgill.ca

Received June 15, 1999
Accepted September 7, 1999

Progressive cognitive decline is a cardinal feature of Alzheimer's disease (AD) and an essential criterion for establishing the clinical diagnosis of the disorder (American Psychiatric Association, 1987; McKhann et al., 1984). Knowledge of the rate of cognitive decline in AD is fundamental to understanding the natural history of the disorder, planning patient care, allotting medical and social resources, and evaluating the effectiveness of clinical interventions (Galasko et al., 1991; Yesavage & Brooks, 1991).

During the past decade, longitudinal studies measuring cognitive decline in AD patients have proliferated (Galasko et al., 1991; Yesavage & Brooks, 1991). These studies have provided information on the rate of cognitive decline in AD patients as measured by annual rate of change scores (ARC) on mental status examinations or global cognitive tests. Unfortunately, there has been great variability in the ARC estimates across studies. Because these studies differed in the selection and diagnosis of AD patients, choice of instruments, follow-up length, number of assessments, and statistical methods, it is difficult to determine whether the observed variability reflects true heterogeneity of AD course, methodological differences, or both. Thus, we decided to conduct a meta-analysis of ARC by systematically reviewing studies addressing cognitive decline in AD using the Mini-Mental State Examination (MMSE; Folstein et al., 1975), and we identify methodological characteristics of studies or differences in patient populations that might have contributed to the observed variability in the ARC.

METHODS

The review process, modified from the one described by Oxman and colleagues (1994), involved systematic selection of articles, abstraction of data, descriptive presentation of the characteristics of studies, and quantitative synthesis of the results using a random effect model (Cooper & Hedges, 1994).

Selection of Articles

To locate relevant studies, we first searched MEDLINE for studies published from January 1981 to November 1997, using the keywords "Alzheimer's disease" and "longitudinal study" or "prognosis" or "cognitive decline." Next, we searched the bibliographies of relevant papers and review articles for additional papers. Finally, the abstracts of all retrieved articles were screened by one of the authors (L. H.) to meet the following seven inclusion criteria: (a) original study; (b) addressed cognitive decline or prognosis or course of AD; (c) published in English; (d) study population included AD patients with ascertainable sample size; (e) used either clinical or pathological diagnostic criteria; (f) longitudinal study design; and (g) used the MMSE as one of the outcome measures. If the study met all the inclusion criteria or a decision could not be made based on its abstract, the original paper was then retrieved. To avoid excluding relevant studies, a sample of 64 studies was independently evaluated by another author (M. C.). Interreviewer agreement on application of exclusion/inclusion criteria was satisfactory (kappa = .79 for 58 abstracts, and 1.0 for 6 papers). Subsequently, a

list of all the included papers was sent to two experts for comment and suggestion for additional papers.

Abstraction of Data

One author (L. H.) used a standard form to abstract the following information from included studies: (a) Study features—research facilities and settings, sources and referral pattern of the patients, follow-up methods, diagnostic and inclusion criteria, length of follow-up, and number of assessments during follow-up; (b) Study population—number of subjects and subgroups, age, gender and education, age at onset, duration and severity of AD at entry; and (c) Study results—MMSE scores at baseline and during the follow-up period and corresponding standard deviations (*SD*s), estimated ARC, and its *SD*; in addition, test statistics, such as *t* or *F*, and *p* values were also retrieved for computing unavailable ARC estimate or its variance. When relevant data were presented by dividing the same patient group in different ways, we used only the data that were the most complete and in which subgroup sample sizes were the most equal. Abstracted data were checked for accuracy by two authors: M.C. for study and population characteristics and F. B. for study results.

Statistical Analysis

The major objective and one of the advantages of a quantitative meta-analysis is its ability to summarize results from many different studies. To synthesize the estimates of ARC across studies, we first tried to use the original values of ARC and its *SD* provided in the paper. When they were not provided, we calculated ARC by dividing the difference of baseline and endpoint mean scores of MMSE of the follow-up group with mean interval in year between the two times of assessments. To calculate *SD* of ARC, we used either the *t* or *p* value of the paired *t* test in the follow-up groups if they were available, or pooled *SD*s of baseline and endpoint MMSE scores of the follow-up groups as a conservative approximation.

Because methodology and population characteristics differed from study to study, we used a multiple random effects regression model to synthesize data and to evaluate the impact of study characteristics on ARC estimates. Such a random effects model would, according to Cooper and Hedges (1994), allow for the true ARC to vary from study to study and for residual heterogeneity of ARC to be explained by a random error after taking into account known or suspected study characteristics. In addition, because the accuracy of the ARC estimates may also vary across studies, a weighted regression analysis was employed in which studies with high accuracy (i.e., low *SD* of ARC) were given more weight (Cooper & Hedges, 1994).

The following variables were included as covariates in the model: mean age in years, years of education, age at onset, length of follow-up in months, number of assessments, and number of study centers involved. Because most studies reported the mean and/or minimum MMSE scores of subjects at entry, we used only the MMSE scores as indices of dementia severity at entry, though other severity measures might also have been used in some studies. Within each study population, we calculated the percentages

of probable AD of female patients and of subjects who did not complete follow-up. We also created two categorical variables to denote the source of the ARC/*SD* estimate (presented in the original paper, or calculated by us) and study design (retrospective, involved both retrospective and prospective components, or prospective). To avoid underestimating within-study variation of ARC, we treated subgroups of patients as an independent sample.

In addition to ARC, we ran the same analysis with the effect size (ES) of ARC, which is a ratio of ARC to its *SD* (Cooper & Hedges, 1994). ES is frequently utilized for sample size calculation in clinical trials or for comparing outcome measures derived from different instruments. Based on the same rationale, we judged that ES might serve as an index of reliability or precision of ARC measure for each study. Indeed, a study with a large ARC may have a small ES if there is large variability associated with the ARC estimate as measured by the *SD*. Thus, modeling ES would provide additional information to our understandings of ARC variation in terms of the reliability or precision of the measurement.

We first evaluated each covariate individually in simple random effect regression models. Then we fitted a multiple random effect model by including all significant covariates ($p < .05$) in the one-covariate model. Following a backwards selection procedure, we reduced this multivariate model by deleting the least significant (i.e., highest p value greater than .05) covariate at a time, until all the covariates left in the model were statistically significant. Each covariate was evaluated in both continuous and categorical format. Categorization of continuous or proportional variables was made by using tertile or median values of the study population or clinically relevant criteria as cutoff points.

All the statistical analyses were conducted using SAS IML software, version 6.12 (SAS, 1997).

RESULTS

A total of 439 potentially relevant studies was identified by the aforementioned search strategy; based on the title and abstract, 142 were retrieved for more detailed evaluation. Of these, 99 were further excluded due to one or more of the following reasons: (a) not original study ($n = 3$); (b) not addressing cognitive decline or prognosis or course of AD ($n = 38$); (c) not published in English ($n = 2$); (d) not including ascertainable AD patients ($n = 6$); (e) not using established diagnostic criteria ($n = 6$); (f) not longitudinal design ($n = 6$); or (g) not using the MMSE as a longitudinal outcome measure ($n = 78$).

Forty-three studies that had met all the seven inclusion criteria were retrieved for this meta-analysis. Of these, 34 (79%) studies were conducted in the USA, especially in university-affiliated AD research centers. Thirty-eight (88%) were published between 1990 and 1997. Study designs included prospective cohort or clinical follow-up studies ($n = 37$) or retrospective chart review ($n = 6$). The most frequently used diagnostic criteria were NINCDS/ADRDA ($n = 37$), followed by ICD-10 ($n = 2$) (World Health Organization, 1992) and DSM-III ($n = 1$). Characteristics of study population and follow-up period varied greatly across the studies, but can be summarized as follows: mean age at entry ranged from

56 to 82 years (median: 72.5); percentage of female subjects ranged from 0 to 87.5% (median: 62%); mean education ranged from 8.7 to 16.5 years (median: 12.7); mean MMSE at entry ranged from 7.2 to 26 (median: 18.4); mean follow-up length ranged from 10 to 60 months (median: 21); and number of assessments ranged from 2 to 8 (median: 3).

Of the 43 included studies, 6 were not used for quantitative meta-analysis due to either lack of adequate data to compute *SD* of ARC ($n = 2$, data not shown) or potential overlapping of the study population with other included studies ($n = 4$, data not shown). For the remaining 37 studies, which consisted of 65 subgroups of AD patients with a total sample size of 3,492, the main methodological and population characteristics, estimates of ARC, and estimates of ES are presented in Table 1. Of the 37 studies, the ARC estimates ranged from 0.9 to 5.7 and the ES estimates from 0.3 to 6.0.

When fitting a simple random effect regression model on the ARCs with inclusion of a single covariate, none of the covariates was statistically significant (all p values were greater than .05, data not shown). Thus, final random effect model included an intercept and a random effect only. The estimate of the intercept was 3.3 (95% confidence interval [CI], 2.9 to 3.7), corresponding to the pooled estimate of ARC across studies. The random effect was statistically greater than zero ($p < .0001$), suggesting significant unexplained variability of ARCs across studies.

The modeling results of ES are presented in Table 2. In the one-covariate models, the minimum MMSE at entry, proportion of female subjects, number of assessments, source of ARC/*SD*, length of follow-up, age at entry, and age at onset were significantly related to ES. When running multiple regression analysis, started with simultaneous inclusion of all these significant covariates except for age at onset due to its small sample size, we ended up with a final model that included two significant covariates, minimum MMSE at entry (categorized into ≤5, 6-14, and ≥15) and number of assessments (categorized into 2, 3, and 4 or more) (Table 2). The random effect was significantly greater than zero ($p < .0001$), suggesting significant unexplained heterogeneity remaining.

DISCUSSION

Based on our review of the 37 longitudinal studies of AD published during the last 10 years, we estimated the average ARC to be 3.3 MMSE points (95% CI, 2.9-3.7). Because this meta-analysis was conducted in a large sample of published studies involving 3,492 AD patients followed over an average period of 2 years, our pooled ARC estimate provides a better approximation of population ARC in AD patients than that from a single study. Given that use of the MMSE is almost universal in dementia clinics (Galasko et al., 1991; Tombaugh & McIntyre, 1992), such an estimate provides clinicians with a guide to assess the deterioration of patients and counsel their families. In addition, this combined ARC estimate may be useful in assessing effects of interventions hypothesized to halt AD progression, and in evaluating the representativeness of the change over time of the placebo groups in clinical trials. Of course, we acknowledge the great variability in ARC measures across studies, as evidenced by the significant random effect term in the regression model.

TABLE 1. Characteristics of the 37 Studies Included in the Meta-Analysis

Study and Publication Year	Facility/Country	AD Diagnostic Criteria/ Groups of Patients[a]	Follow-Up Months/# Assessments	# Patients: Enrolled/ Followed Up[b] (% Female)	Mean Age ± *SD* at Entry, Years	Mean MMSE ± *SD* at Entry (Minimum)	ARC ± *SD*	ES
Becker et al., 1988	University AD research program/USA	Prob	13/2	86/44 (52.0)	67.0 ± 9.4	23.7 ± 4.2	1.8 ± 1.7*	1.1
Becker et al., 1992	University AD research program/USA	Prob						
		1. Nonfocal	12/2	140/55 (64.9)[c]	70.5 ± 8.5	17.4 ± 5.4	4.2 ± 9.4**	0.4
		2. Amnesic	12/2	43/5 (64.9)[c]	72.3 ± 8.5	20.0 ± 4.2	4.8 ± 5.5**	0.9
		3. Dysexecutive	12/2	8/6 (64.9)[c]	67.9 ± 10.1	23.9 ± 2.6	4.3 ± 4.4**	1.0
Burns et al., 1991	2 psychiatric hospitals in a catchment area/UK	Prob (total $N = 110$)						
		1. MMSE ≤ 9	12/2	na/42 (79.0)[c]	80.7[c]	(0)	2.4 ± 8.3**	0.3
		2. MMSE 10–15	12/2	na/26 (79.0)[c]	80.7[c]	(10)	5.2 ± 8.3**	0.6
		3. MMSE ≥ 16	12/2	na/17 (79.0)[c]	80.7[c]	(16)	3.3 ± 8.3**	0.4
Butters et al., 1996	University AD research program/USA	Prob						
		1. Mild AD	24/2	107/59 (64.5)	70.7 ± 8.1	18.2 ± 4.4	3.7 ± 3.9**	1.0
		2. Temporal lobe AD	24/2	32/18 (62.5)	74.3 ± 8.6	22.5 ± 6.7	1.0 ± 3.3**	0.3
Coen et al., 1996	Hospital memory clinic/Ireland	Prob						
		1. Poorer category fluency	11/2	25/15 (78.8)[c]	72.5 ± 5.5	21.9 ± 2.7	2.2 ± 2.3	1.0
		2. Poorer letter fluency	11/2	15/11 (78.8)[c]	70.5 ± 6.2	23.1 ± 2.4	4.1 ± 4.5	0.9
Corey-Bloom et al., 1993	9 AD research centers/USA	1. Poss	12/2	291/58 (63.0)	74.6 ± 8.6	17.9 ± 7.4	3.0 ± 4.3	0.7
		2. Prob	12/2	967/244 (70.6)	74.4 ± 7.8	13.9 ± 7.5	3.1 ± 3.9	0.8
Dukoff & Sunderland, 1997	Clinical research center/USA	Prob	37/2	59/35 (62.3)	69.1 ± 9.5	21.7 ± 5.3	2.8 ± 3.3**	0.8
Ferris et al., 1997	27 centers in an AD cooperative study/USA	Prob (total $N = 242$)						
		1. MMSE >20	12/3	50/44 (54.0)	73.2 ± 8.5	23.7 ± 2.4	3.9 ± 4.1	1.0
		2. MMSE 16–20	12/3	47/41 (51.1)	73.3 ± 9.0	18.1 ± 1.3	4.8 ± 5.3	0.9
		3. MMSE 10–15	12/3	46/34 (65.2)	72.4 ± 7.8	13.2 ± 1.8	5.0 ± 4.3	1.2
		4. MMSE 5–9	12/3	49/33 (65.3)	71.8 ± 10.8	7.1 ± 1.5	3.8 ± 3.0	1.3
		5. MMSE 0–4	12/3	50/31 (68.0)	70.9 ± 8.7	2.2 ± 1.4	1.1 ± 1.9	0.6
Forstl et al., 1996	Memory disorder clinic/UK	Prob or Poss	24/2	55/55 (62.0)	70 ± 9.0	na	3.0 ± 4.0***	0.8

(continued)

TABLE 1. Continued

Study and Publication Year	Facility/Country	AD Diagnostic Criteria/ Groups of Patients[a]	Follow-Up Months/# Assessments	# Patients: Enrolled/ Followed Up[b] (% Female)	Mean Age ± *SD* at Entry, Years	Mean MMSE ± *SD* at Entry (Minimum)	ARC ± *SD*	ES
Fritz et al., 1995	Clinical research center/USA	Prob or Poss (total $N = 64$)						
		1. Pet-exposed	12/2	na/12 (59.4)[c]	73.3 ± 8.1[c]	18.4 ± 7.3	3.8 ± 3.8	1.0
		2. Non-Pet exposed	12/2	na/10 (59.4)[c]	73.3 ± 8.1[c]	15.0 ± 8.7	4.2 ± 3.7	1.1
Galasko et al., 1995	21 CERAD centers/USA	Prob (2-year group, $N = 343$)						
		1. IADL ≥ 5	24/3	na/139 (53.4)[c]	70.9 ± 7.9[c]	18.3 ± 4.1	4.4 ± 2.4	1.8
		2. IADL <5	24/3	na/175 (53.4)[c]	70.9 ± 7.9[c]	19.9 ± 3.9	2.6 ± 2.3	1.1
Gallagher-Thompson et al., 1992	Hospital clinic/USA	Prob	42/7	na/35 (31.4)	63.2	na	3.3 ± 2.1***	1.6
Goldblum et al., 1994	Not specified/France	Prob (total $N = 21$)						
		1. Slower decliner	12/2	na/8 (87.5)	78.5	19.1 ± 4.2	1.0 ± 0.4**	2.8
		2. Faster decliner	12/2	na/8 (75.0)	78.6	22.9 ± 4.2	3.9 ± 0.6*	6.0
Haxby et al., 1992	Neuroscience laboratory/USA	Prob or Poss						
		1. Isolated memory impairment	60/6	na/6 (0.0)	58	26.0 ± 2.0	4.4 ± 3.6	1.2
		2. Language and visuospatial impairment	52/6	na/10 (40.0)	56	24.0 ± 4.0	4.4 ± 2.5	1.8
Holmes et al., 1996	Specialist agencies for the elderly/UK	Prob or Poss	16/3	164/107 (na)	81.9 ± 6.6	13.0 ± 5.1	2.6 ± 3.5	0.7
Kurz et al., 1996	Memory clinic/ Germany	ICD-10						
		1. Apo E4+	12/4	na/44 (na)	73	16.8 ± 4.7	4.3 ± 2.8**	1.5
		2. Apo E4–	12/4	na/20 (na)	73	18.0 ± 4.9	4.5 ± 3.2**	1.4
Kuskowski et al., 1991	Not specified/USA	DSM-III-R	12/3	33/8 (50.0)	64 ± 6.2	22.4 ± 3.7	3.0 ± 1.2*	2.5
Lopez et al., 1990	University AD research program/USA	Prob (total $N = 62$)						
		1. With depression	13/2	na/10 (80.0)	67.4 ± 6.6	21.3 ± 2.3	4.1 ± 2.6*	1.6
		2. Without depression	12/2	na/10 (80.0)	67.5 ± 11.6	21.4 ± 2.2	3.8 ± 4.4**	0.9

(continued)

TABLE 1. Continued

Study and Publication Year	Facility/Country	AD Diagnostic Criteria/ Groups of Patients[a]	Follow-Up Months/# Assessments	# Patients: Enrolled/ Followed Up[b] (% Female)	Mean Age ± *SD* at Entry, Years	Mean MMSE ± *SD* at Entry (Minimum)	ARC ± *SD*	ES
McShane et al., 1995	Not specified/UK	Prob						
		1. With hallucination and LB	32/8	na/5 (51.2)[c]	81[c]	13[c]	3.6 ± 2.9**	1.2
		2. With hallucination	32/8	na/8 (51.2)[c]	81[c]	13[c]	3.3 ± 2.7**	1.2
		3. With LB	32/8	na/3 (51.2)[c]	81[c]	13[c]	3.6 ± 3.0**	1.2
		4. Without hallucination or LB	32/8	na/25 (51.2)[c]	81[c]	13[c]	2.1 ± 4.2**	0.5
Miller et al., 1991	University clinical research center/USA	Prob						
		1. With EPS at entry	21/5	na/24 (38.3)[c]	na	18	4.5 ± 1.8	2.5
		2. Without EPS at entry	21/5	na/57 (38.3)[c]	na	20	2.7 ± 2.1	1.3
Mortimer et al., 1992	VA medical center/USA	Prob	33/6	76/65 (21.5)	63.8 ± 8.5	17.2 ± 4.8	4.5 ± 3.1***	1.4
Murphy et al., 1996	VA medical center/USA	Prob	43/6	na/86 (38.3)[c]	73.5 ± 7.0	(15)	4.1 ± 2.4***	1.7
Obara et al., 1994	VA medical center/USA	Prob	27/2	na/18 (61.1)	75.4 ± 5.2	14.6 ± 8.1	2.8 ± 2.6*	1.1
Pearlson et al., 1989	Psychiatric inpatient service/USA	Prob	25/2	13/7 (76.9)	70.6 ± 1.7	15.4 ± 2.4	3.5 ± 3.1**	1.1
Peavy et al., 1996	University-affiliated research center/USA	Prob or Poss	12/2	41/12 (50.0)	75.2 ± 6.3	7.2 ± 3.8	1.3 ± 2.0	0.6
Peters et al., 1988	University geriatric medical clinic/USA	Clinical criteria (total $N = 38$)						
		1. With hearing impairment	9/2	na/10 (87.0)[c]	81.4 ± 5.2	17.0 ± 5.3	2.5 ± 4.6	0.5
		2. Without hearing impairment	9/2	na/7 (87.0)[c]	74.4 ± 6.8	19.2 ± 5.4	3.2 ± 5.7	0.6
Pollman et al., 1995	Dementia clinic/ Germany	ICD-10	12/2	90/80 (72.2)	73.2	17.0 ± 4.7	4.3 ± 8.5**	0.5
Rasmusson, 1996	University-affiliated research center/USA	Prob	30/7	210/132 (63.6)	70.1 ± 4.0	16.7 ± 4.0	4.4 ± 2.3***	2.0
Rebok et al., 1990	University-affiliated research center/USA	Prob or Poss	24/5	209/51 (68.6)	67.3 ± 7.9	15.4 ± 5.5	3.7 ± 5.7**	0.7

(continued)

TABLE 1. Continued

Study and Publication Year	Facility/Country	AD Diagnostic Criteria/ Groups of Patients [a]	Follow-Up Months/# Assessments	# Patients: Enrolled/ Followed Up [b] (% Female)	Mean Age ± *SD* at Entry, Years	Mean MMSE ± *SD* at Entry (Minimum)	ARC ± *SD*	ES
Salmon et al., 1990	University-affiliated research center/USA	Prob (2-year group)	12/3	92/55 (62.0)	72.6 ± 7.4	18.5 ± 4.6	2.8 ± 4.3	0.7
Shear et al., 1995	VA medical center/USA	Prob	25/2	117/41 (36.6)	70.7 ± 7.6	17.5 ± 5.5	3.2 ± 2.7	1.2
Taylor et al., 1996	University-affiliated research center/USA	Prob or Poss	24/3	na/40 (40.0)	73.0 ± 6.7	21.1 ± 4.5	3.2 ± 4.2**	0.8
Teri et al., 1990	University outpatient clinics/USA	Prob	24/3	200/106 (71.7)	77.0 ± 6.7	18.2 ± 6.7	2.8 ± 4.6	0.6
Uhlmann et al., 1986	University internal medicine clinic/USA	Clinical criteria						
		1. With hearing impairment	12/2	na/36 (64.3)	78.9 ± 5.2	20.0 ± 6.5	3.9 ± 4.8	0.8
		2. Without hearing impairment	12/2	na/120 (64.8)	75.8 ± 7.7	17.2 ± 7.4	2.2 ± 5.0	0.4
Wolfe et al., 1995	University dementia clinic/USA	Prob or Poss	24/2	na/29 (48.3)	70.7 ± 7.9	22.0 ± 4.5	3.4 ± 4.3	0.8
Yesavage et al., 1988	VA medical center/USA	Prob	12/2	na/30 (na)	na	16.5 ± 7.6	4.2 ± 4.8	0.9
Yesavage et al., 1993	9 clinical centers/USA	Prob						
	State cohort	1. Without aphasia	40/3	na/27 (76.0)	na	22.5 ± 4.0	2.4 ± 2.1	1.1
	(*N* = 1,251)	2. With early aphasia	43/4	na/11 (80.0)	na	21.3 ± 3.1	5.7 ± 2.4	2.4
		3. With late aphasia	46/4	na/32 (72.0)	na	20.6 ± 3.3	4.5 ± 2.4	1.9
	Palo Alto cohort	4. Without aphasia	41/7	na/6 (67.0)	na	22.2 ± 4.7	0.9 ± 0.3	3.0
	(*N* = 140)	5. With early aphasia	36/6	na/32 (62.0)	na	19.6 ± 3.6	3.6 ± 2.4	1.5
		6. With late aphasia	49/8	na/19 (79.0)	na	23.8 ± 2.2	3.0 ± 2.1	1.4

Note. AD = Alzheimer's disease; ARC = annual rate of change scores; ES = effect size; CERAD = Consortium to Establish a Registry for Alzheimer's Disease; VA = Veterans Administration; na = not available.

[a]Prob = probable and/or definite AD; Poss = possible AD, by the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA); MMSE = Mini-Mental State Examination; Pet = positron emission tomography; IADL = instrumental activities of daily living; DSM-III-R = the 3rd edition, revised, of the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 1987); ICD-10 = the 10th revision of the International Classification of Diseases (draft) (World Health Organization, 1992); Clinical criteria = any clinical criteria other than NINCDS/ADRDA, ICD, or DSM; LB = (cortical) Lewy bodies; EPS = extrapyramidal symptoms. [b]When baseline sample sizes for subgroups were not available, a total *N* of study population is indicated in parentheses in the column AD Diagnostic Criteria/Groups of Patients. [c]These values were derived from study population data at baseline.

*These *SD*s were estimated from paired *t* test or paired *p* value. **These *SD*s were estimated from unpaired *t* test or *p* value. ***These *SD*s were either intercept or β for covariate of multiple regression models provided in original report.

TABLE 2. Random Effects Regression Models on Effect Size of ARC

Covariate		One Covariate Model β (*SE*)	*p*	Full Model[a] β (*SE*)	*p*	Final Model (*n* = 61) β (*SE*)	*p*
1. Minimum MMSE at entry	≤5	Reference category					
(*n* = 61)	6–14	.27 (.16)	.09	.37 (.22)	.09	.30 (.14)	.04
	≥15	.49 (.17)	.005	.42 (.28)	.15	.46 (.15)	.003
2. Proportion (%) of female	≤50	Reference category					
patients (*n* = 61)	50–65	−.41 (.19)	.03	−.03 (.23)	.91		
	>65	−.36 (.20)	.07	.04 (.27)	.88		
3. Proportion (%) of probable AD	<100	Reference category					
(*n* = 65)	100	.24 (.16)	.13				
4. Number of assessments	2	Reference category					
(*n* = 65)	3	.26 (.14)	.059	.12 (.19)	.52	.22 (.14)	.12
	4 or more	.64 (.14)	<.001	.36 (.29)	.22	.59 (.15)	<.001
5. Study design (*n* = 65)	Retrospective	Reference category					
	Mixed	.04 (.25)	.88				
	Prospective	.11 (.26)	.68				

(continued)

TABLE 2. Continued

Covariate		One Covariate Model β (*SE*)	*p*	Full Model[a] β (*SE*)	*p*	Final Model (*n* = 61) β (*SE*)	*p*
6. Proportion (%) of patients lost to follow-up (*n* = 48)	>60	Reference category					
	60–30	−.01 (.20)	.97				
	<30	−.18 (.19)	.37				
7. Numbers of study centers (*n* = 65)	1	Reference category					
	≥2	−.03 (.16)	.85				
8. Source of ARC/*SD* (*n* = 65)	Given in article	Reference category					
	Calculated by us	.38 (.14)	.007	.04 (.17)	.82		
9. Follow-up length in months (*n* = 65)		.02 (.01)	<.001	.004 (.01)	.75		
10. Age at entry in years (*n* = 55)		−.04 (.01)	.004	−.002 (.02)	.94		
11. Years of education (*n* = 44)		.002 (.06)	.97				
12. Mean MMSE at entry (*n* = 59)		.03 (.02)	.09	−.002 (.02)	.93		
13. Age at onset in years (*n* = 24)		−.06 (.02)	.02				

Note. ARC = annual rate of change scores; MMSE = Mini-Mental State Examination; AD = Alzheimer's disease.
[a]This model (*n* = 38, *p* < .0001) included all the significant covariates of the one covariate model except for age at onset, which was not included because of small sample size.

One of the objectives of this meta-analysis was to determine the underlying source(s) of observed heterogeneity of ARC estimates. Although none of the study characteristics we evaluated appeared to be significant in explaining the variability of ARC, we observed a positive correlation between ES of ARC with number of assessments and minimum MMSE score at entry. These results seem consistent with findings of previous studies that an increase in number of assessments or length of follow-up would improve the reliability of ARC estimates (Morris et al., 1993; Stern et al., 1992; van Belle et al., 1990), or suggest that more number of assessments would increase the likelihood of observing a cognitive decline of AD patients.

On the other hand, the observed association between ES and baseline MMSE score may suggest that initial cognitive function would also affect the reliability of ARC measurements, in addition to predicting cognitive decline of AD patients, as reported by previous studies (Burns et al., 1991; Drachman et al., 1990; Haxby et al., 1992; Jacobs et al., 1994; Morris et al., 1993; Rich et al., 1995; Salmon et al., 1990; Teri et al., 1990). Thus, to improve the reliability of ARC measurement and to facilitate comparisons of ARC estimates across studies, future studies should probably use stratum-specific ARCs by baseline cognitive function of patients, instead of an overall ARC, as indices of cognitive decline. Based on this meta-analysis and the popularity of the MMSE, we propose that the MMSE be used as a standard instrument for estimating stratum-specific ARC, as have several authors (Burns et al., 1991; Drachman et al., 1990; Ferris et al., 1997) and that uniform cutoff points be used to standardize such stratification procedure. However, what cutoff points should be used needs to be determined in light of both clinical significance and statistical justification.

Reasons that may underlie our failure to detect significant predictors of ARC variability include:

1. The covariates we studied may have no consistent effect on ARC and our results correctly reflect the heterogeneity of prognostic findings in AD patients. For example, a most intensively studied prognostic factor, initial severity of dementia, has been reported to predict faster decline (Burns et al., 1991; Drachman et al., 1990; Morris et al., 1993; Teri et al., 1990) or slower decline (Rich et al., 1995) or to have no effect (Goldblum et al., 1994; Jacobs et al., 1994; Salmon et al., 1990).
2. Covariates other than those included in our study may be more important in explaining ARC variation. We did not evaluate some potentially important predictors of cognitive decline, e.g., Apo E gene (Holmes et al., 1996; Kurz et al., 1996), aphasia (Becker et al., 1988; Goldblum et al., 1994; Kurz et al., 1996; Mortimer et al., 1992; Yesavage et al., 1993), and extrapyramidal signs (Corey-Bloom et al., 1993; Miller et al., 1991), because they were not included in most of the studies.
3. Measurement error in coding the covariates or potential overestimating of the *SD* of ARC may have hidden the effects. However, we evaluated each continuous covariate using both its actual and categorical value in the regression models. The results did not differ significantly. Similarly, the source of ARC/*SD* estimates for individual studies was not related to the ARC.

4. Observed ARC estimates across studies may be too variable; consequently the covariate's effect was undetectable. This notion is partially supported by the different results between our two separate analyses. When the ARC was used as a dependent variable, no covariate was statistically significant. However, when ES was used as a more reliable and standardized outcome measure, two significant predictors, minimum MMSE at baseline and number of assessments, were found. These differences may have such a methodological implication that future studies should make more effort to improve the accuracy or reliability of ARC measurement before the true effect of any predictor of ARC can be determined. Although meta-analysis is a good method to synthesize research findings across studies, it cannot eliminate the methodological flaws of the original measurements.

This review and meta-analysis has limitations. First, we may have missed some relevant studies that were unpublished, published in languages other than in English, or excluded based on abstracts only. However, our rigorous search strategy and cross-checking procedure make it unlikely that we missed important papers. Second, our evaluation of study methodology and population characteristics was selective: We focused on some predictors while neglecting others. However, the frequency of the predictors being evaluated across the studies probably reflected their recognized importance in planning a natural history study in an AD population, regardless of the particular interests of the researchers. Finally, our study focused on the MMSE, which has been criticized for its inability to detect change in severely demented patients or to depict potential nonlinearity of the AD course (Morris et al., 1993; Tombaugh & McIntyre, 1992). Development of measures that can describe the full course of AD deterioration has been suggested (Cole & Dastoor, 1996; Galasko et al., 1991). Nevertheless, the MMSE performs well in patients with mild to moderate dementia, who are the target of most predictor and intervention studies. The MMSE is also the most widely used instrument in both clinical and research settings. A good understanding of the advantages and disadvantages of the MMSE in measuring the ARC of AD patients would serve as a starting point in understanding the usefulness of other instruments.

To conclude, great variation of ARC estimates existing across studies cannot be explained with the study or population characteristics we evaluated. Such unexplained ARC variability warrants further effort to improve the reliability and precision of ARC measurement. Stratum-specific ARC by baseline MMSE scores may be useful to serve this purpose. Given the potential limitations of this meta-analysis and of the MMSE, our combined ARC estimate may be most applicable to the course of mild or moderate AD during the first 1 or 2 years following the initial examination. However, we acknowledge that even though two patients may have the same ARC on MMSE scores, the content of their decline may differ dramatically. Finally, our combined estimate of ARC is based solely on MMSE data. Its relevance for other cognitive or function tests needs to be examined in both clinical and research settings.

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.-rev.). Washington, DC: Author.

Becker, J. T., Bajulaiye, O., & Smith, C. (1992). Longitudinal analysis of a two-component model of the memory deficit in Alzheimer's disease. *Psychological Medicine, 22,* 437-445.

Becker, J. T., Huff, F. J., Nebes, R. D., Holland, A. L., & Boller, F. (1988). Neuropsychological function in Alzheimer's disease. Pattern of impairment and rates of progression. *Archives of Neurology, 45,* 263-268.

Burns, A., Jacoby, R., & Levy, R. (1991). Progression of cognitive impairment in Alzheimer's disease. *Journal of the American Geriatrics Society, 39,* 39-45.

Butters, M. A., Lopez, O. L., & Becker, J. T. (1996). Focal temporal lobe dysfunction in probable Alzheimer's disease predicts a slow rate of cognitive decline. *Neurology, 46,* 687-692.

Coen, R. F., Maguire, C., Swanwick, G. R., Kirby, M., Burke, T., et al. (1996). Letter and category fluency in Alzheimer's disease, a prognostic indicator of progression? *Dementia, 7,* 246-250.

Cole, M., & Dastoor, D. (1996). The hierarchic dementia scale, conceptualization. *International Psychogeriatrics, 8,* 202-212.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis.* New York: Russell Sage Foundation.

Corey-Bloom, J., Galasko, D., Hofstetter, C. R., Jackson, J. E., & Thal, L. J. (1993). Clinical features distinguishing large cohorts with possible AD, probable AD, and mixed dementia. *Journal of the American Geriatrics Society, 4,* 31-37.

Drachman, D. A., O'Donnell, B. F., Lew, R. A., & Swearer, J. M. (1990). The prognosis in Alzheimer's disease: 'How far' rather than 'how fast' best predicts the course. *Archives of Neurology, 47,* 851-856.

Dukoff, R., & Sunderland, T. (1997). Durable power of attorney and informed consent with Alzheimer's disease patients, a clinical study. *American Journal of Psychiatry, 154,* 1070-1075.

Ferris, S. H., Mackell, J. A., Mohs, R., Schneider, L. S., Galasko, D., et al. (1997). A multicenter evaluation of new treatment efficacy instruments for Alzheimer's disease clinical trials, overview and general results. *Alzheimer Disease and Associated Disorders, 11*(Suppl. 2), S1-S12.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189-198.

Forstl, H., Sattel, H., Besthorn, C., Daniel, S., Geoger-Kabisch, C., et al. (1996). Longitudinal cognitive, electroencephalographic and morphological brain changes in ageing and AD. *British Journal of Psychiatry, 168,* 280-286.

Fritz, C. L., Farver, T. B., Kass, P. H., & Hart, L. A. (1995). Association with companion animals and the expression of noncognitive symptoms in Alzheimer's patients. *Journal of Nervous and Mental Disease, 183,* 459-463.

Galasko, D., Corey-Bloom, J., & Thal, L. J. (1991). Monitoring progression in Alzheimer's disease. *Journal of the American Geriatrics Society, 39,* 932-941.

Galasko, D., Edland, S. D., Morris, J. C., Clark, C., Mohs, R., et al. (1995). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part XI. Clinical milestones in patients with Alzheimer's disease followed over 3 years. *Neurology, 45,* 1451-1455.

Gallagher-Thompson, D., Brooks, J. O. III, Bliwise, D., Leader, J., & Yesavage, J. A. (1992). The relations among caregiver stress, "sundowning" symptoms, and cognitive decline in Alzheimer's disease. *Journal of the American Geriatrics Society, 40,* 807-810.

Goldblum, M. C., Tzortzis, C., Michot, J. L., Panisset, M., & Boller, F. (1994). Language impairment and rate of cognitive decline in Alzheimer's disease. *Dementia, 5,* 334-338.

Haxby, J. V., Raffael, K., Gillette, J., Schapiro, M. B., & Rapoport, S. I. (1992). Individual trajectories of cognitive decline in patients with dementia of the Alzheimer type. *Journal of Clinical and Experimental Neuropsychology, 14,* 575-592.

Holmes, C., Levy, R., McLoughlin, D. M., Powell, J. F., & Lovestone, S. (1996). Apolipoprotein E, non-cognitive symptoms and cognitive decline in late onset Alzheimer's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 61,* 580-583.

Jacobs, D., Sano, M., Marder, K., Bell, K., Bylsma, F., et al. (1994). Age at onset of Alzheimer's disease, relation to pattern of cognitive dysfunction and rate of decline. *Neurology, 44,* 1215-1220.

Kurz, A., Egensperger, R., Haupt, M., Lautenschlager, N., Romero, B., et al. (1996). Apolipoprotein E epsilon 4 allele, cognitive decline, and deterioration of everyday performance in Alzheimer's disease. *Neurology, 47,* 440-443.

Kuskowski, M. A., Morley, G. K., Malone, S. M., & Okaya, A. J. (1991). Longitudinal measurements of brainstem auditory evoked potentials in patients with dementia of the Alzheimer type. *International Journal of Neuroscience, 60,* 79-84.

Lopez, O. L., Boller, F., Becker, J. T., Miller, M., & Reynolds, C. F. III. (1990). Alzheimer's disease and depression, neuropsychological impairment and progression of the illness. *American Journal of Psychiatry, 147,* 855-860.

McKhann, G., Drachman, D., Katzman, R., Price, D., & Stadian, E. M. (1984). 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, 34,* 939-944.

McShane, R., Gedling, K., Reading, M., McDonald, B., Esiri, M. M., et al. (1995). Prospective study of relations between cortical Lewy bodies, poor eyesight, and hallucinations in Alzheimer's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 59,* 185-188.

Miller, T. P., Tinklenberg, J. R., Brooks, J. O. 3d, & Yesavage, J. A. (1991). Cognitive decline in patients with Alzheimer disease, differences in patients with and without extrapyramidal signs. *Alzheimer Disease and Associated Disorders, 5,* 251-256.

Morris, J. C., Edland, S., Clark, C., Galasko, D., Koss, E., et al. (1993). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part IV. Rates of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology, 43,* 2457-2465.

Mortimer, J. A., Ebbitt, B., Jun, S. P., & Finch, M. D. (1992). Predictors of cognitive and functional progression in patients with probable Alzheimer's disease. *Neurology, 42,* 1689-1696.

Murphy, G. M. Jr., Yang, L., Yesavage, J. A., & Tinklenberg, J. R. (1996). Rate of cognitive decline in Alzheimer's disease is not affected by the alpha-1-antichymotrypsin A allele or the CYP2D6 B mutant. *Neuroscience Letters, 217,* 200-202.

Obara, K., Meyer, J. S., Mortel, K. F., & Muramatsu, K. (1994). Cognitive declines correlate with decreased cortical volume and perfusion in dementia of Alzheimer type. *Journal of the Neurological Sciences, 127,* 96-102.

Oxman, A. D., Cook, D. J., & Guyatt, G. H. (1994). Users' guides to the medical literature, V. How to use an article about prognosis. *Journal of the American Medical Association, 272,* 234-237.

Pearlson, G. D., Rabins, P. V., Kim, W. S., Speedie, L. J., Moberg, P. J., et al. (1989). Structural brain CT changes and cognitive deficits in elderly depressives with and without reversible dementia ('pseudodementia'). *Psychological Medicine, 19,* 573-584.

Peavy, G. M., Salmon, D. P., Rice, V. A., Galasko, D., Samuel, W., et al. (1996). Neuropsychological assessment of severely demented elderly, the severe cognitive impairment profile. *Archives of Neurology, 53,* 367-372.

Peters, C. A., Potter, J. F., & Scholer, S. G. (1988). Hearing impairment as a predictor of cognitive decline in dementia. *Journal of the American Geriatrics Society, 36,* 981-986.

Pollman, S., Haupt, M., & Kurz, A. (1995). Changes of the relative severity of naming, fluency and recall impairment in the course of dementia of the Alzheimer type. *Dementia, 6,* 252-257.

Rasmusson, D. X. (1996). Predicting rate of cognitive decline in probable Alzheimer's disease. *Brain and Cognition, 3,* 133-147.

Rebok, G., Brandt, J., & Folstein, M. (1990). Longitudinal cognitive decline in patients with Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology, 3,* 91-97.

Rich, J. B., Rasmusson, D. X., Folstein, M. F., Carson, K. A., Kawas, C., et al. (1995). Nonsteroidal anti-inflammatory drugs in Alzheimer's disease. *Neurology, 45,* 51-55.

Salmon, D. P., Thal, L. J., Butter, N., & Heindel, W. C. (1990). Longitudinal evaluation of dementia of the Alzheimer type, a comparison of 3 standardized mental status examinations. *Neurology, 40,* 1225-1230.

SAS. (1997). *SAS version 6.12.* Cary, NC: Author.

Shear, P. K., Sullivan, E. V., Mathalon, D. H., Lim, K. O., Davis, L. F., et al. (1995). A longitudinal volumetric computed tomographic analysis of regional brain changes in normal aging and Alzheimer's disease. *Archives of Neurology, 52,* 392-402.

Stern, R. G., Mohs, R. C., Bierer, L. M., Silverman, J. M., Schmeidler, J., et al. (1992). Deterioration on the Blessed test in Alzheimer's disease, longitudinal data and their implications for clinical trials and identification of subtypes. *Psychiatric Research, 42,* 101-110.

Taylor, K. I., Salmon, D. P., Rice, V. A., Bondi, M. W., Hill, L. R., et al. (1996). Longitudinal examination of American National Adult Reading Test (AMNART) performance in dementia of the Alzheimer type (DAT), validation and correction based on degree of cognitive decline. *Journal of Clinical and Experimental Neuropsychology, 18,* 883-891.

Teri, L., Hughes, J. P., & Larson, E. B. (1990). Cognitive deterioration in Alzheimer's disease: Behavioral and health factors. *Journal of Gerontology. Psychological Sciences, 45,* 58-63.

Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini-Mental State Examination, a comprehensive review. *Journal of the American Geriatrics Society, 40,* 922-935.

Uhlmann, R. F., Larson, E. B., & Koepsell, T. D. (1986). Hearing impairment and cognitive decline in senile dementia of the Alzheimer's type. *Journal of the American Geriatrics Society, 34,* 207-210.

van Belle, G., Uhlmann, R. F., Hughes, J. P., & Larson, E. B. (1990). Reliability of estimates of changes in mental status test performance in senile dementia of the Alzheimer's type. *Journal of Clinical Epidemiology, 43,* 589-595.

Wolfe, N., Reed, B. R., Eberling, J. L., & Jagust, W. J. (1995). Temporal lobe perfusion on single photon emission computed tomography predicts the rate of cognitive decline in Alzheimer's disease. *Archives of Neurology, 52,* 257-262.

World Health Organization. (1992). *The international classification of diseases, classification of mental and behavioral disorders* (10th ed.). Geneva, Switzerland: Author.

Yesavage, J. A., & Brooks, J. O. (1991). On the importance of longitudinal research in Alzheimer's disease. *Journal of the American Geriatrics Society, 39,* 942-944.

Yesavage, J. A., Brooks, J. O. III, Taylor, J., & Tinklenberg, J. (1993). Development of aphasia, apraxia, and agnosia and decline in Alzheimer's disease. *American Journal of Psychiatry, 150,* 742-747.

Yesavage, J. A., Poulsen, S. L., Sheikh, J., & Tanke, E. (1988). Rate of change of common measures of impairment in senile dementia of the Alzheimer's type. *Psychopharmacology Bulletin, 24,* 531-534.

Acknowledgments. This work was made possible by a Research Fellowship in Geriatric Psychiatry provided by

Novartis Pharmaceuticals Canada Inc. to the first author. We thank Drs. S. Gauthier, S. Vida, and M. Elie for their review and helpful comments on our reference list and/or an earlier version of this manuscript. This work was presented in part at the 151st annual meeting of the American Psychiatric Association, Toronto, Canada (May 31-June 4, 1998), and at the International Psychogeriatric Association-Beijing Joint Meeting, Beijing, China (April 12-14, 1999).