Published in final edited form as:

J Am Geriatr Soc. 2008 October ; 56(10): 1893–1897. doi:10.1111/j.1532-5415.2008.01922.x.

Variable Judgments of Decisional Capacity in Cognitively Impaired Research Subjects

Carol B. Stocking, PhD^{1,2}, Gavin W. Hougham, PhD¹, Deborah D. Danner, PhD³, Marion B. Patterson, PhD⁴, Peter J. Whitehouse, MD⁴, and Greg A. Sachs, MD⁵

¹Section of Geriatrics, Department of Medicine, University of Chicago, Chicago, Illinois

²MacLean Center for Clinical Medical Ethics, Department of Medicine, University of Chicago, Chicago, Illinois

³University of Kentucky, Lexington, Kentucky

⁴Case Western Reserve University, Cleveland, Ohio

⁵Indiana University Center for Aging Research, Indianapolis, Indiana

Abstract

OBJECTIVES—To compare assessments of the decisional capacity of cognitively impaired patients by research assistants (RAs) and by family caregiver/proxies; and to determine whether either or both groups judge capacity differently depending on the specific (hypothetical) research enrollment decision being made.

DESIGN—Cross sectional

SETTING—Three clinics, one each in Ohio, Kentucky and Illinois

PARTICIPANTS—149 patients with established dementia diagnoses and their caregiver/proxies

MEASUREMENTS—As part of a longer interview, patients were asked about future enrollment in five hypothetical research projects with varying risks and benefits. After patients reported each decision and reasons for that decision, RAs (using Applebaum & Grisso's four standards) indicated whether the patient was competent to make that decision on their own. In separate interviews, caregiver/proxies were asked for a similar appraisal based on life experience with the patient. RA and proxy judgments were compared.

RESULTS—Capacity judgments by RAs and by caregivers differed by specific project for most patients. Agreement between RA and proxy judgments varied by project but agreement was only "fair" when tested by kappa (range in data .21 - .39). There were 50 patients who were appraised as competent for all decisions by proxies and 47 who were so assessed by RAs. Of these only 24 were the same patients.

CONCLUSION—If capacity were assessed anew for each enrollment decision, more potential participants could maintain authority in making those decisions.

Keywords

decisional capacity; dementia research; ethics

Corresponding author: Carol B. Stocking, Ph.D., MacLean Center for Clinical Medical Ethics & Section of Geriatrics, Department of Medicine MC 6098, University of Chicago Medical Center, 5841 S. Maryland Avenue, Chicago IL 60637, Telephone: 773-702-1648, Fax: 773-702-0090; **Alternate Corresponding Author:** Greg A. Sachs, M.D., Sachs, Greg [sachsg@iupui.edu].

Author Contributions: All authors contributed to study concept and design, acquisition of data and review of the final manuscript

Sponsor's Role: NIA had no role in the design, methods, subject recruitment, data collection, analysis or preparation of the paper.

INTRODUCTION

Once a potential subject is found to meet the basic eligibility criteria for a research trial, his or her consent for enrollment is sought. But an early step is deciding whether the subject has the capacity to make the enrollment decision at hand or whether the assistance of a proxy must be sought. Although there are more formal methods established for making this determination,¹⁻⁶ many judgments about decisional capacity for treatment and research are made ad hoc and very quickly depending more or less on the gravity of the decision in question.

In this paper we will consider a set of capacity judgments, comparing those made by trained research assistants and those made by family caregiver/proxies. We suggest that the trained research assistants may be analogous to persons who seek informed consent for enrollment in many clinical trials. There are two primary research questions: 1) how closely will the decisional capacity assessments of research assistants and appraisals of family caregivers match, and 2) do either or both groups making the judgments assess capacity differently depending on the specific decision being made?

METHODS

A convenience sample of 149 persons with established diagnoses of dementia (hereafter referred to as patients) and their family caregiver/proxies were interviewed separately about patient wishes for future enrollment in five fully described (hypothetical) research studies with varying levels of risk and benefit. The sample was drawn for the DRIPAC project from two Alzheimer Research Centers and a Memory Center at different geographic locations.⁷⁻⁹ The research described by the vignettes ranged from simple blood draws (for a diagnostic test, or for a genetic marker), to a new medicine (with possible benefits and possible negative side effects), to a lumbar puncture for development of a diagnostic test (which involves some risk and no benefit), and the implantation of special cells in the brain (with more risk and some chance of benefit). After each vignette describing a research study, patients were asked a series of questions including whether they would enroll in such a study in the future and their reasons for that decision. The research assistant (RA) asked the structured questions and wrote down the patient's responses.

Immediately after the patient made each (of the five) enrollment decision and gave reasons for it, the RA wrote down his or her own responses to a series of questions about the patient's competence (based on Applebaum & Grisso's four standards) to make a decision to enroll in that specific project.¹⁰ In separate interviews, caregiver/proxies also were asked to appraise the patient's competence to make each of the five enrollment decisions. In this paper we use the term caregiver/proxy (or proxy, for brevity) because although these persons are not necessarily legally designated proxies, they have been designated as the persons who would, if necessary, assist in research decision making.

Research Assistant Training

Research assistants (RAs) at three sites were specifically trained by one Project Coordinator in the application of Applebaum & Grisso's four standards: communication of a choice, understanding relevant information, appreciation of the situation and consequences, and rational manipulation of information. Immediately after the patient made each hypothetical research enrollment decision and gave his/her reasons the research assistant was asked to indicate whether the patient had, in that decision, met each of these four standards. After making an assessment of each of those factors, RAs were asked to "rate the competence of the patient regarding consent for research: Would you be willing to allow this patient to

decide for him/herself if he or she were being approached for consent to this research?" The possible responses provided were definitely competent, probably competent, probably incompetent, and definitely incompetent. We used the first (definitely competent) when we compared RA assessments and proxy appraisals since that would be the appropriate standard if such a decision were to be made by the patient alone.

Family caregiver/proxies separately responded to a similar interview about the five types of research and were asked whether in the future they would give permission for the patient to enroll in each project. Caregiver/proxies were then asked the series again but were asked to indicate whether they thought the patient would choose (give his/her own permission) to enroll. Finally, proxies were asked "Do you think that PATIENT currently would understand enough about the research involving RESEARCH TYPE that you could leave the decision up to him/her?" Possible answers provided were yes, no, and unsure.

Since we will be comparing the appraisals made by proxies to the assessments made by RAs it is important to bear in mind that proxies are basing their appraisals that 'the patient would currently understand enough about (each research project) so that the decision could be left to him/her 'on their life experience with the patient; RAs are basing their assessments as 'definitely competent' on Applebaum and Grisso's four standards.

Differences in responses to categorical variables were assessed by chi square tests; differences in means were assessed by analysis of variance. Kappa was used to assess the concordance between RA and proxy judgments. Binary logistic regression was used to assess multivariate effects.

The DRIPAC study was approved by Institutional Review Boards at the University of Chicago, Case Western Reserve University and the University of Kentucky. Patient and proxy informed consent (or patient assent with proxy consent) was obtained for participation in the DRIPAC study.

RESULTS

The mean age of the 149 patients enrolled in this study was 78.6 years (range 52-94); the mean MMSE score was 19.8 (range 2 - 29). About two thirds (67.1%) were white and a similar proportion (62.4%) were female. Analyses of variance confirmed that the patients at Center 1 were significantly older than those at the other centers, and patients in Center 3 had significantly lower MMSE scores. There were also more African Americans at Center 1 and more males in Center 2 (when assessed by chi-square). (See Table 1)

When the Centers are considered together, the proxy appraisals appear somewhat more consistent across trials than do those of the RAs. For the three low risk projects, proxies consider a higher proportion of patients to be competent than do RAs. The opposite is true for the lumbar puncture and cell implantation trials. When assessed by Kappa, agreement between RA and proxy judgments is 'fair' (.276 for tube of blood; .251 for genetic test; .367 for new medicine; .218 for lumbar puncture; and .394 for implanted cells).¹¹ The highest proportion of RA assessments as 'definitely competent' is associated with cell implants, the trial that is clearly the most dangerous. (See Table 2)

Capacity judgments by both RAs and caregivers differ by specific project, providing an indication that decisional capacity (or the judgment of capacity) does depend on the specific decision to be made in the eyes of the two groups of assessors.

There were differences between the Centers in mean MMSE, mean age and in ethnic and gender composition. (See Table 1) Any of these could affect judgments of competence by

the RAs and by the proxies. We thus examined the Centers separately. (See Table 2) There is a reasonably consistent set of differences between judgments between Centers, seen most clearly in the “both” and “neither” columns of Table 2. For each of the vignettes, Center 1 had the lowest, Center 2 the middle, and Center 3 the highest proportion of patients judged to be competent. Binary logistic regression showed that only MMSE scores relate to judgments by both RAs and proxies in all five trials.

Although patients at Center 1 were older, age does not appear to be a major factor in the assessments made by RAs. The mean age of all Center 1 patients is 82.1. The mean age of those judged definitely competent for each enrollment decision is 83.0 for the blood draw; 80.4 for the genetic test; 82.0 for the new medicine; 82.1 for the lumbar puncture; and 81.1 for the implantation of special cells.

Neither the ratings of competence by the RAs nor those of the proxies differed by patient gender in either the overall sample or the sample at Center 2, where there were significantly more males.

Since for each trial almost half of the patients are considered by both RAs and proxies to have decisional capacity it seems possible that the same set of patients was found to be competent for all five decisions. There were 50 patients who were judged competent for all decisions by proxies and 47 who were judged definitely competent by RAs. Half of those patients (24) were consistently judged across the five trials as definitely competent by both groups. At the other end of the spectrum, 46 patients were not judged competent to make any decisions by RAs and 41 patients were so judged by proxies. Interestingly, only twenty-four patients (coincidentally the same number as those judged competent) were assessed by both RAs and proxies as not being able to make any of the five decisions on their own. (See Table 3)

DISCUSSION

Judgments of decisional capacity by research assistants and by family caregivers do not match as closely as we might have expected. Research assistants (who may be analogous to persons recruiting patients to trials) and family caregiver/proxies have different knowledge bases from which to judge a patient's competence to make a specific decision. In this study, for example, the RAs had just heard the patient make a specific participation decision and explain his/her reasons for that decision. Then the RA, using four standards, assessed the competence of the patient to make that specific enrollment decision. This may explain why RAs found a higher proportion of patients competent to decide about the projects with greater risks. The RAs had heard that (usually) the patient would not wish to enroll in risky projects and gave context-appropriate reasons for that choice. RAs could then correctly assess the patient as competent. In contrast, proxies didn't know how the patient would answer those enrollment questions and may have assumed a prophylactic protective stance.

Although the RAs in this study may have had no particular stake in finding patients competent to decide to enroll in hypothetical projects, in other circumstances persons seeking consent may have an interest in finding as many patients as possible competent for recruitment. In real research recruitment situations, timely trial accrual may be critically important to the success of the study, and may even involve substantial financial rewards to an investigator or to an institution. In such circumstances there may be substantial motivation to be more generous in assessments of competence. And recruitment may often involve fairly brief encounters in which to make an assessment of capacity.

In other settings the persons doing the enrollment may already know the potential research subject from participation in previous research trials or from non-research clinic or

laboratory visits and thus have a broader frame of reference within which to make capacity assessments.

Family caregiver/proxies have a wealth of day to day information about the patients on which to base their capacity appraisals. Proxies too may have incentives to foster (and perhaps even overestimate) whatever capacities their relatives may have. For example, they may well be able to interpret correctly gestures and sounds that are ambiguous or even incomprehensible to others and truly sustain the patient's abilities. But caregivers may also over-interpret, hoping to sustain abilities which have faded.

Could the strengths of these two knowledge bases be combined to sustain the autonomy of the patient as long as is ethically possible? It is hard to imagine a circumstance in which this could actually occur. However, our data suggest that more people would be able to make decisions about enrollment in a particular research project without direct proxy involvement if either an RA or a proxy judgment of adequate decisional capacity were to be considered sufficient for a particular enrollment decision. (There would be about 100 eligible patients for each study, as can be seen by summing the figures for each study in Table 2.) In contrast, if both RA and proxy judgments as competent were necessary, the number of patient eligible to decide without direct proxy involvement would be reduced to about 50 for each project. The differences in our data between RA and proxy judgments raise troubling ethical questions.

We did indeed find evidence that both research assistants and family caregiver/proxies judge capacity differently depending on the specific decision being made. This provides further empirical evidence for those who have argued that decisional capacity is not a generalized ability but one that relates to specific decisions and can be considered to be best measured on a sliding scale.¹²

How are we to understand the unanticipated differences in assessments of competence by RAs among the three Centers? Why were such a high proportion of patients at Center 3 judged competent when by MMSE score they seem less likely to be competent? (There were no differences among the Centers in patient education.) Or why were so few at Center 1 judged to be competent? We can only speculate. First, Centers 2 and 3 were both federally-funded Alzheimer Disease Research Centers and the RAs there may have had more experience talking with patients with mild to moderate dementia than did RAs at Center 1. They may have already encountered these patients in other research studies, or during clinic visits, and thus have had a wider array of information available than did RAs at Center 1. They may have had previous experience enrolling mildly impaired people in research projects and be comfortable applying a more flexible standard.

The patients at Center 1 were older and it is possible that RAs at that Center had unconscious ageist notions about competence, shading their assessments. However, there is no evidence by statistical tests that age was a major factor in the assessments they made. There were more African-American patients at Center 1 but there is no way to know if RAs there unconsciously assessed them stringently because they had lower prior expectations of their abilities or for some other reason. Proxies at Center 1 also rated competence lower than did proxies at other Centers, although the differences are smaller than those between RA assessments.

Although these data shed light on our two main questions, there are limitations in the data that prevent adequate exploration of the possible effects of ethnicity. We cannot extricate the effects of race from those of the consistently low ratings by both RAs and proxies of patient decisional capacity in Center 1. Logistic regression indicates that patient MMSE is related to judgments of decisional competence in all five trials. But mean MMSE scores are

significantly lower at Center 3, not Center 1. We are only able to speculate about why both the RAs and proxies at Center 1 assessed such low percentages of patients to be decisionally competent.

It would be possible to assume that some difference in RA training as it was implemented in the three sites accounts for the systematic differences in RA competence assessments at those sites. But then how would we account for the differences in proxy appraisals? Center 1 is always lowest and Center 2 highest (although only very slightly higher than Center 3). Although it is possible to speculate, our data do not provide sufficient clarification of this issue.

Our data do provide empirical evidence to support the commonsensical, but not widely tested, notion that standards for judgment of decisional competence change as the risks associated with the decision become greater and the balance between risk and benefit shifts. Accepting such a sliding scale would place an additional burden on those assessing capacity to consent but would give prospective research participants the benefit of maintaining decisional authority in as many situations as are ethically possible.

Acknowledgments

Conflict of Interest:

The editor in chief has reviewed the conflict of interest checklist provided by the authors and has determined that the authors have no financial or any other kind of personal conflicts with this paper.

Dr. Sachs is an unpaid consultant to Caremark. He has received grants from Hartford, NIA, Reynolds and HRSA

The research reported in this paper was supported by grant RO1AG15317 from the National Institute on Aging.

REFERENCES

1. Grisso, T.; Applebaum, PS. *Assessing Competence to Consent to Treatment: A Guide for Physicians and Other Health Professionals*. Oxford University Press; New York: 1998.
2. Marson DC, Ingram KK, Cody HA, et al. Assessing the Competency of Patients with Alzheimer's Disease under Different Legal Standards: A prototype instrument. *Arch Neurol*. 1995; 52:949–954. [PubMed: 7575221]
3. Applebaum PS, Grisso T. The MacArthur Treatment Competence Study, I: mental illness and competence to consent to treatment. *Law Human Behav*. 1995; 19:105–126.
4. Janofsky JS, McCarthy RJ, Folstein MG. The Hopkins Competence Assessment Test: a brief method of evaluating patients' capacity to give informed consent. *Hosp Community Psychol*. 1992; 43:132–136.
5. Kim SYH, Karlawish JHT, Caine ED. Current State of Research on Decision-Making Competence of Cognitively Impaired Elderly Persons. *Am J Geriatric Psych*. 2002; 10:151–165.
6. Dunn LB, Nowrangi MA, Palmer BW, et al. Assessing Decisional Capacity for Clinical Research or Treatment: a Review of Instruments. *Am J Psychiat*. 2006; 163:1323–1334. [PubMed: 16877642]
7. Brauner DJ, Merel SE. How a Model Based on Linguistic Theory Can Improve the Assessment of Decision-Making Capacity for Persons with Dementia. *J Clin Ethic*. 2006; 17:139–148.
8. Stocking CB, Hougham GW, Danner DD, et al. Speaking of Research Advance Directives. *Neurology*. 2006; 66:1361–1366. [PubMed: 16682668]
9. Stocking CB, Hougham GW, Danner DD, et al. Empirical assessment of a Research Advance Directive for Persons with Dementia and Their Proxies. *J Am Geriatr Soc*. 2007; 55:1609–1612. [PubMed: 17714461]
10. Applebaum PS, Grisso T. Assessing Patients' Capacities to Consent to Treatment. *New Engl J Med*. 1988; 319:1635–1638. [PubMed: 3200278]

11. Landis RJ, Koch GG. The Measurement of Observed Agreement for Categorical Data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
12. Drane JF. Competency to Give Informed Consent. A model for making clinical assessments. *J Amer Med Assoc*. 1984; 252:925–927.

Table 1
Patient Characteristics

	Center 1 N = 59	Center 2 N = 29	Center 3 N = 61	Data set N = 149
Mean age (range 52-94)	82.18*	74.6	77.11	78.6
Mean MMSE score (range 2-29)	20.64	22.34	17.74*	19.8
Percent high school graduate or less	52.5%	41.4	44.2	46.9
Percent White	25.4%*	82.7	98.3	67.1
Percent Female	69.4%	34.3*	68.8	62.4

Mean age and Percent White at Center 1 differed from age and percent white at the other two centers (mean age, ANOVA, Duncan, $P < .05$) (percent white, chi sq, $P < .001$).

MMSE scores at Center 3 differed from those at Centers 1 and 2 (ANOVA, Duncan, $P < .05$).

Center 2 differed from the other Centers in Percent Female (chi sq, $P < .005$).

Table 2
Number and Percent of Respondents at each Center Judged Competent by Research Assistants by Proxies, by Both and by Neither

	RA only	Proxy only	Both	Neither
Tube of blood				
Center 1 N = 59	3 (5.08%)	21 (35.59%)	12 (20.33%)	23 (39.98%)
Center 2 N = 29	3 (10.34%)	8 (27.58%)	11 (37.93%)	7 (24.13%)
Center 3 N = 59	11 (18.64%)	8 (13.55%)	28 (47.45%)	12 (20.33%)
Total N = 147	17 (11.56%)	37 (25.17%)	51 (34.69%)	42 (28.57%)
Genetic test				
Center 1 N = 59	4 (6.77%)	21 (35.59%)	10 (16.94%)	24 (40.67%)
Center 2 N = 29	3 (10.34%)	7 (24.13%)	12 (41.37%)	7 (24.13%)
Center 3 N = 59	14 (23.72%)	6 (10.16%)	31 (52.54%)	8 (13.55%)
Total N = 147	21 (14.28%)	34 (23.12%)	53 (36.05%)	39 (26.53%)
New medicine				
Center 1 N = 59	5 (8.47%)	15 (25.42%)	11 (18.64%)	28 (47.45%)
Center 2 N = 28	2 (7.14%)	5 (17.85%)	14 (50.00%)	7 (25.00%)
Center 3 N = 59	12 (20.33%)	7 (11.86%)	31 (52.54%)	9 (15.25%)
Total N = 146	19 (13.01%)	27 (18.49%)	56 (38.35%)	44 (30.13%)
Lumbar puncture				
Center 1 N = 59	7 (11.86%)	14 (23.72%)	7 (11.86%)	31 (52.54%)
Center 2 N = 29	5 (17.24%)	5 (17.24%)	10 (34.48%)	9 (31.03%)
Center 3 N = 59	18 (30.50%)	8 (13.55%)	21 (35.59%)	12 (20.33%)
Total N = 147	30 (20.40%)	27 (18.36%)	38 (25.85%)	52 (35.37%)
Implanted cells				
Center 1 N = 59	12 (20.33%)	7 (11.86%)	12 (20.33%)	28 (47.45%)
Center 2 N = 28	2 (7.14%)	4 (14.28%)	13 (46.42%)	9 (32.14%)
Center 3 N = 59	18 (30.50%)	2 (3.38%)	26 (44.06%)	13 (22.03%)
Total N = 147	32 (21.91%)	13 (8.90%)	51 (34.93%)	50 (34.24%)

Two proxies at Center 3 did not provide any assessments; thus assessments at Center 3 are based on 59 cases.

One proxy at Center 2 did not respond to the new medicine assessment; one proxy at Center 2 did not respond to the new medicine appraisal; one proxy at Center 2 did not respond to the implanted cells appraisal; thus the percentages for those trials at Center 2 are based on 28 cases.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript