

# Using an Automated Recruitment Process to Generate an Unbiased Study Sample of Multiple Sclerosis Patients

Deborah M. Miller, Ph.D.,<sup>1</sup> R. Fox, M.D.,<sup>1</sup> A. Atreja, M.D., M.P.H.,<sup>2</sup> S. Moore, Ph.D.,<sup>3</sup> J.-C. Lee, M.S.,<sup>4</sup> A.Z. Fu, Ph.D.,<sup>4</sup> A. Jain, M.D.,<sup>2</sup> W. Saupe, B.S.,<sup>5</sup> S. Chakraborty, M.Eg.,<sup>4</sup> M. Stadler,<sup>1</sup> and R.A. Rudick, M.D.<sup>1</sup>

<sup>1</sup>Mellen Center for Treatment and Research in Multiple Sclerosis, Neurological Institute, Cleveland Clinic, Cleveland, Ohio.

<sup>2</sup>Digestive Diseases Institute, Cleveland Clinic, Cleveland, Ohio.

<sup>3</sup>Frances Payne Bolton School of Nursing, Case Western Reserve University, Cleveland, Ohio.

<sup>4</sup>Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio.

<sup>5</sup>eCleveland Clinic, Cleveland Clinic, Cleveland, Ohio.

## Abstract

The objective of this study was to test the efficiency of an automated recruitment methodology developed as a component of a practical controlled trial to assess the benefits of a Web-based personal health site to guide self-management of multiple sclerosis symptoms called Mellen Center Care On-line. We describe the study's automated recruitment methodology using clinical and administrative databases and assess the comparability between subjects who completed informed consent (IC) forms, and individuals who were invited to participate but did not reply, designated as patient nonresponders (PNR). The IC and PNR groups were compared on demographics, number of physician or advanced practice nurse/physician assistant visits during the 12 months prior to the initial invitation, and level of disability as measured by the Charlson Comorbidity Index (CCI). Out of a total dynamic potential pool of 2,421 patients, 2,041 had been invited to participate, 309 had become ineligible to participate during the study, and 71 individuals remained in the pool at the end

of recruitment. The IC group had a slightly greater proportion of females. Both groups were predominantly white with comparable marital status. The groups had comparable mean household income, education level, and commercial insurance. The computed mean CCI was similar between the groups. The only significant difference was that the PNR group had fewer clinic visits in the preceding 12 months. The subjects were highly representative of the target population, indicating that there was little bias in our selection process despite a constantly changing pool of eligible individuals.

**Key words:** e-health, telehealth, multiple sclerosis, self-management, Mellen Center Care On-line

## Introduction

An important goal of subject recruitment for clinical trials is that the study subjects are representative of the patients likely to receive the intervention, thus assuring generalizability to the population of interest.<sup>1-3</sup> This is particularly important when the investigation evaluates routine processes of care. These studies are sometimes referred to as practical controlled trials.<sup>4,5</sup> In this article, we describe Mellen Center Care On-line (MCCO), a randomized electronic recruitment method that was integrated into a telemedicine intervention. MCCO was designed to generate a representative study sample of the target clinical population. It also allowed us to determine the extent to which that representative recruitment occurred.

In 1993, the National Institutes of Health (NIH) Revitalization Act was signed by President Clinton. Section 492B(d)(2)(B) of this act<sup>6</sup> directed the NIH to establish guidelines for the inclusion of women and minorities in clinical research so that "In the case of any clinical trial in which women or members of minority groups will be included as subjects, the Director of NIH shall ensure that the trial is designed

and carried out in a manner sufficient to provide for valid analysis of whether the variables being studied in the trial affect women or members of minority groups, as the case may be, differently than other subjects in the trial." Nonetheless, many trials continue to recruit nonrepresentative samples.<sup>7-9</sup> Recent research evaluating the demographics of trial participants demonstrates that there is under-recruitment of racial and ethnic minorities, women, and the elderly,<sup>7,9-11</sup> which raises methodological and ethical issues.<sup>12</sup> In response to these concerns, there has been a call to fully report the recruitment process in study manuscripts,<sup>3</sup> and for this reporting to include the socio-economic status and race of trial participants.<sup>2</sup>

As to the use of Web-based recruitment, several electronic methods have been developed to assist in the clinical trial recruitment process. Some examples include the development of volunteer patient registries,<sup>13</sup> recruiting from population-based registries for cancer trials,<sup>14</sup> or recruiting through the use of clinical trial alert systems built into a commercial electronic health record (EHR) system.<sup>15</sup> Smith<sup>16</sup> describes a trial that recruited subjects from EHR searches and a study Web site. Both recruitment methods generated comparable subjects. At one private group practice, EHRs are used to successfully screen for sponsored trials.<sup>15</sup> While these methods have been useful in accelerating trial recruitment, none reports how closely the recruited subjects reflect the patient population and thus do not address the issues of trial generalizability. Additionally, there are no sociodemographic comparisons reported between potential patients and recruited subjects.

## Methods

We designed a randomized, controlled clinical trial to assess possible benefits of a Web-based personal health site designed to guide self-management of multiple sclerosis (MS) symptoms (NIH Grant 5R01LM008154), and approached this trial using practical controlled trial methodology proposed by Tunis.<sup>5</sup> The Web-based personal health site is called MCCO. The Cleveland Clinic Mellen Center is one of the leading patient care and research centers for MS in the United States, with approximately 7,000 patient visits per year. Our study's automated recruitment methodology that was designed for this application allowed for comparison between subjects randomized to the trial and patients who met the trial inclusion criteria but did not respond to the study invitation.

## Subjects

Our inclusion criteria required only that the subjects have clinically definite MS, reside in the county where the Mellen Center is located or any of the five surrounding counties, and that they had completed at least two appointments with a physician or advanced practice

nurse/physician assistant (APN/PA) at our Center in the previous 12 months. The geographic restriction was necessary, as we installed computers and Internet access for participants who did not have them. The restriction to patients with at least 2 physician or APN/PA appointments in the previous 12 months was an attempt to include patients who received their primary neurological care at our Center. Subjects were recruited over a period of 36 months, and individuals could move in and out of eligibility depending on a change of residence or the number of Mellen Center visits during the preceding 12 months. This report describes the degree to which the automated recruitment system resulted in a sample that is representative of patients who were eligible for the study but did not respond.

## Recruitment

As approved by the Cleveland Clinic Institutional Review Board, our EHR system (EpicCare; Epic Systems Corporation, Verona, WI) was regularly queried using an automated report to identify Mellen Center patients who met the inclusion criteria over the course of the study period. The medical record numbers (MRNs) of those patients were placed in an electronic "dynamic potential patient pool" (DPPP). It is important to emphasize that this potential patient pool was "dynamic," as patients could move in and out of eligibility depending on the number of physician or APN/PA visits they completed in the preceding 12 months, and whether the patients moved in or out of the designated geographical area. The recruitment process is documented in *Figure 1*. Using a random number generator program, approximately every 2 weeks, 20 potential subjects were selected by their MRN to be invited to participate in the study. Standardized letters of invitation for those MRNs were generated and personalized by the study coordinator, who then mailed them. A study number was generated for the potential subject, and that number and the MRN were linked in a protected electronic file to guard the confidentiality of the patient information. The study number was the identifier for subjects throughout the trial. Patients were asked to contact the Center if they were either interested in study participation or if they actively wished to decline participation. If patients did not respond to the first letter within 2 weeks, a follow-up letter was mailed. If there was no response to the second letter, no further contact was attempted.

## Assessment of Target Population and Subject Recruitment Concordance

With enrollment completed, we compared two groups to assess the generalizability of our study sample: (1) patients who enrolled in the study, designated as those who completed informed consent (IC) forms; and (2) patient nonresponders (PNRs), those who did not

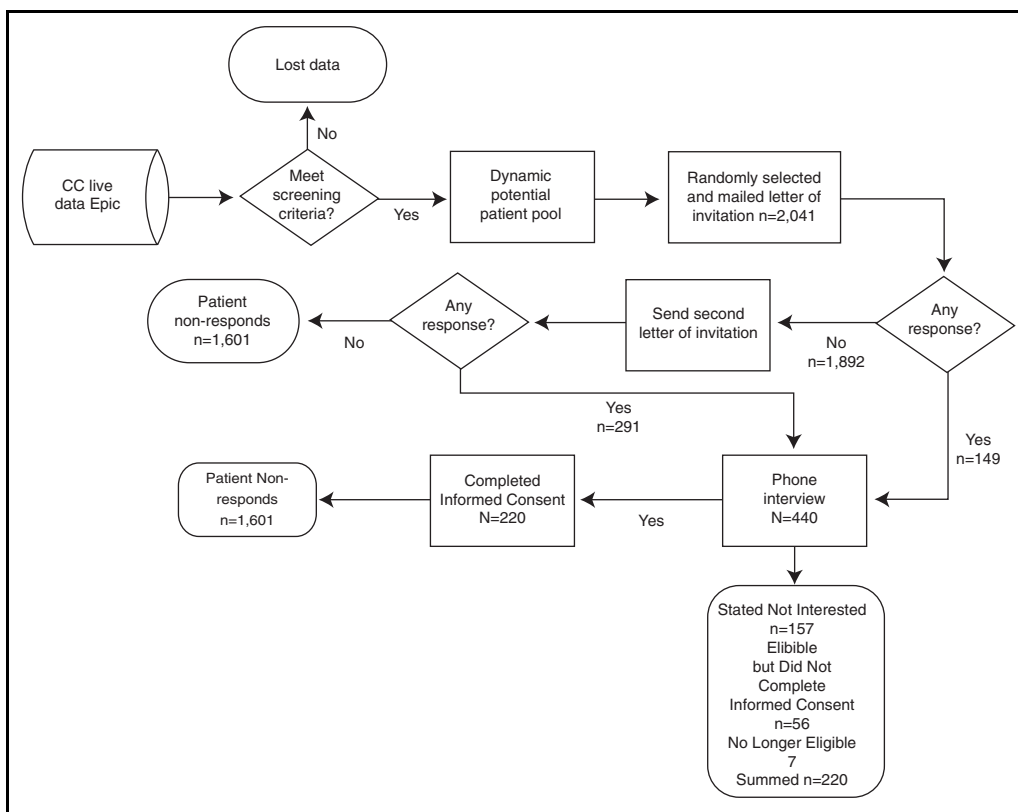


Fig. 1. Enrollment flow diagram.

respond to the letters of invitation. Because the number of subjects in the PNR group was more than seven times greater than those in the IC group, we considered the PNR group as representative of our study population. We consider the PNR group as representative of the study population. A third group of patients declined participation or failed the phone screening (not eligible/not interested). The Cleveland Clinic Institutional Review Board (IRB) allowed database review for clinical characteristics of the PNRs, but withheld this permission for individuals who actively declined participation or who contacted us but were not eligible. The IRB based its reasoning for this restriction on the fact that those individuals had directly advised us that they did not wish to participate. Thus, no further analysis can be conducted for this third group of individuals.

Comparability between IC and PNR groups was assessed based on demographic factors, healthcare use as measured by the median number of visits to the Cleveland Clinic Mellen Center for a physician or APN/PA during the 12 months prior to the initial invitation mailing,

and by the Charlson Comorbidity Index (CCI).<sup>18</sup> The CCI, a weighted list of summed diagnoses, was included as a surrogate for MS disease severity because we did not have specific information about MS captured in Epic, our EHR. These data used for this calculation were obtained from the clinical database by reviewing all diagnoses during the 12 months before letters of invitation for the study were mailed. Differences in household income, education level, and percentage with commercial insurance between the two groups were computed using zip code level data obtained from U.S. Census 2000 data.

Chi-square tests were used for categorical variable comparisons between the groups, and *t*-tests or the Wilcoxon rank-sum test were calculated for continuous variables. A sample size of 112 subjects in each of the two groups was calculated. Based on *Table 1* of the Sickness Impact Profile

(SIP) User's Manual<sup>19</sup> and Interpretation Guide, we estimated that the mean  $\pm$  SD for the SIP in our MS patients was  $21.0 \pm 12.0$  and that a realistic and clinically meaningful improvement would be to reduce this mean by 3 points, to 18. Under these rather conservative assumptions and using a one-tailed, 0.05-level analysis of covariance *t*-test,  $N = 112 + 112$  gives powers from 0.85 to 0.93.

## Results

Of 2,041 patients in the DPPP, 213 patients declined participation and 7 were no longer eligible at the phone screening, and of these, 220 (10.7%) potential subjects were excluded from further analysis due to the IRB restriction. Of the remaining 1,821 patients, 220 completed the IC; the other 1,601 individuals were classified as PNRs. The first subject enrolled on November 4, 2004, and enrollment was completed on October 9, 2007. Demographic data are compared between the IC and PNR groups are presented in *Table 1*. The groups were similar in regard to sex distribution, age, race, and marital status. Using the U.S. Census

**Table 1. Comparison of Characteristics Between Patients Who Did Not Respond to an Invitation to Participate and Those Who Completed an Informed Consent Form<sup>a</sup>**

COMPARATOR	IC (N = 220)	PNR = 1,601	P VALUE
% Female	78.2%	74.1%	0.19
Mean (SD) age	48.6 (9.7)	47.2 (11.3)	0.06
% White	75.9	78.0	0.21
% Married	61.4	57.7	0.30
Census-based mean (SD) household income <sup>b</sup>	\$68,896.00 (\$23,349.00)	\$69,714.31 (\$23,544.30)	0.63
Census-based % with at least a college degree <sup>b</sup>	27.2	26.8	0.68
Census-based % with commercial health insurance <sup>c</sup>	65.1	61.1	0.52
Mean (SD) Charlson Comorbidity Index	0.90 (1.20)	0.95 (1.30)	0.55
% Charlson = 0	59.1	58.4	0.85
Median (interquartile range) number of Mellen Center physician or advanced practice nurse/physician's assistant visits in preceding 12 months	9.0 (5.0, 18.0)	6.0 (3.0, 12.0)	0.001 <sup>d</sup>

<sup>a</sup>Data based on date "Letter of Invitation" was mailed.

<sup>b</sup>Based on zip code data from the 2000 US Census.

<sup>c</sup>Data missing for 46 patients (5 from IC, 41 from PNR).

<sup>d</sup>Significant.

IC, informed consent; PNR, patient nonresponder; SD, standard deviation.

2000 zip code data, the groups were similar in regard to household income, level of education, and commercial insurance. The computed mean CCI was similar between groups (0.90 [SD = 1.2] versus 0.95 [SD = 1.30];  $p = 0.55$ ). The only significant difference between the two groups was that the IC group had a greater median number of physician/APN visits in the preceding 12 months (median [interquartile range]: 9.0 [5.0, 18.0] versus 6.0 [3.0, 12.0];  $p = 0.001$ ).

An important aspect of this recruitment process involved maintaining the DPPP in which patients moved in and out of eligibility. *Figure 2* shows at 6-month intervals the number of individuals who met the inclusion criteria, the number of individuals who received invitation letters during that time period, and the number of patients who remained active in the potential subject pool at each of these time points. The number of patients who met the inclusion criteria for the DPPP remained fairly constant over time, the number of patients invited varied from quarter to quarter, and the number of uninvited members of the DPPP

diminished over time. At the end of enrollment, a total of 2,421 patients had been members of the DPPP; 2,041 patients had been invited, and 71 patients remained in the pool, indicating that 309 patients had become ineligible for participation over the study period.

## Discussion

Our findings indicate that the automated recruitment process developed and implemented for the MCCO study, when applied to existing clinical care and administrative databases, resulted in the enrollment of 220 subjects from a pool of 2,041 at a single site within 36 months for a 10.7% enrollment rate. Furthermore, the enrolled subjects were highly representative of the target population, indicating that there was little bias in our selection process. This included a representative proportion of females and nonwhites in comparison to our clinic population and the general MS population. Thus, we avoided the reported concerns about the adequate representation of

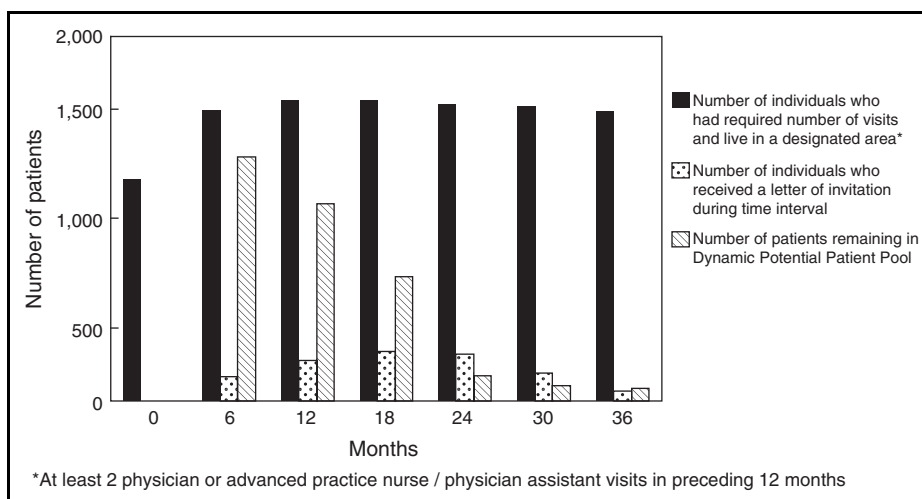


Fig. 2. Characteristics of dynamic potential patient pool.

women and minorities in clinical trials.<sup>7,9,12</sup> Individuals in the IC and PNR groups were comparable regarding demographic factors and comorbidities. Those subjects who chose to participate in the study had approximately 30% more visits with Mellen Center physicians or APNs/PAs in the 12 months prior to receiving their letter of invitation than those who did not respond ( $p = 0.001$ ).

While our automated recruitment methodology typically recruited 20 patients every 2 weeks, we had the flexibility to adjust that recruitment rate with seasonal response (e.g., fewer responders in winter months) and the number of available staff to work on the study, making its management more efficient. As is illustrated in *Figure 2*, after month 0, the number of Mellen Center patients who met the enrollment criteria remained relatively steady. The fewer number of eligible patients at Month 0 was a programming artifact. Over the course of the study, the number of individuals who were mailed letters of invitation was varied based on study personnel availability and expected seasonal availability of subjects at different times of the year. With only 71 patients remaining in the DPPPP at the study's end, it is clear that we would have exhausted the number of potential subjects had the estimated sample size been any larger. Information about the number of potential subjects and the likely response rate can be very useful when planning the methodology for an investigation of this sort. As shown in *Figure 1*, more than half ( $n = 291$ ) of the 440 patients who completed phone interviews responded after receiving their second study invitation letter. It is not known how quickly enrollment would have been completed if the study protocol allowed for a third follow-up letter to be sent.

There is limited and conflicting evidence in the literature to suggest why those in the IC group had a greater median number of physician/APN visits in the 12 months preceding enrollment than the PNR group (median of nine visits versus six). While at least one study has shown that enrollment in similar personal health Internet sites is greater among patients who are healthier,<sup>20</sup>—which is counter to what these data indicate because more visits would suggest the patients were sicker—other studies demonstrate that once enrolled, disabled or elderly subjects use such systems more frequently than younger, healthier enrollees.<sup>21,22</sup> Further research will be required to explore the motivations and characteristics of individuals who enroll and utilize

similar personal health sites.

The automated recruitment process did not require involvement of clinicians, thus precluding concerns about patient–clinician/patient–researcher relationships.<sup>23</sup> Our recruitment methodology, in fact, made that aspect of recruitment the least labor intensive. Because clinicians were not directly involved in the recruitment process, we avoided, as reported by Subramanian et al.,<sup>10</sup> the possible exclusion of otherwise eligible patients who are considered poor study candidates by their care providers.

There are three important limitations to this study. First, we did not have access to MS-specific data during the recruitment phase. While the CCI has been used as a surrogate measure of disease burden in MS,<sup>24</sup> it is not an ideal measure in this population, since more than 50% of both groups had CCI scores of 0. Development of other automated measures of MS disease burden should be considered. Second, we do not have comparison information for those individuals who actively declined participation. These are individuals who contacted us and actively expressed their intention to not participate. It is possible that these individuals may have differed from other nonresponders in their ethnicity, language spoken, gender, age, socio-economic status, understanding of research, and other factors. It is not possible to predict the potential impact that excluding these individuals might have on the conclusions of this study. The third important limitation of this study is the use of 2000 census data for household income, insurance type, and educational attainment. While these data were assessed at the zip-code level, which allows a high degree of granularity, it is possible that there have been some

demographic shifts at the zip-code level from 2000 when the data were collected and 2004 to 2007, when the study was conducted.

There is clear evidence that many trials do not recruit representative samples,<sup>7,9-11</sup> which gives rise to both methodological and ethical concerns.<sup>12</sup> In response to these concerns, there has been a call to fully report the recruitment process<sup>3</sup> and include the socio-economic status and race of trial participants.<sup>2</sup> Development and implementation of automated algorithms applied to clinical and administrative databases to recruit trial subjects have been successful,<sup>16,17</sup> but sample representativeness has not been reported. The automated recruitment method used as a part of the MCCO investigation was an approach that was designed to demonstrate a transparent comparison of study sample and patient population on socio-economic status and race representativeness.

## Acknowledgments

Data for this article were supported by NIH Grant 5R01LM008154.

## Disclosure Statement

No competing financial interests exist.

## REFERENCES

1. Gotay CC. Increasing trial generalizability. *J Clin Oncol* 2006;24:846-847.
2. Lee SJ, Kavanaugh A. A need for greater reporting of socioeconomic status and race in clinical trials. *Ann Rheum Dis* 2004;63:1700-1701.
3. Wright JR, Bouma S, Dayes I, et al. The importance of reporting patient recruitment details in phase III trials. *J Clin Oncol* 2006;24:843-845.
4. Glasgow RE, Wagner EH, Schaefer J, Mahoney LD, Reid RJ, Greene SM. Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). *Med Care* 2005;43:436-444.
5. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: Increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 2003;290:1624-1632.
6. NIH Revitalization Act, Subtitle B, Part 1, Section 131-133. June 10, 1993.
7. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: Race-, sex-, and age-based disparities. *JAMA* 2004;291:2720-2726.
8. Friedman LS, Simon R, Foulkes MA, et al. Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993—the perspective of NIH clinical trialists. *Cont Clin Trials* 1995;16:277-285; discussion 279-286.
9. Baquet CR, Commiskey P, Daniel Mullins C, Mishra SI. Recruitment and participation in clinical trials: Socio-demographic, rural/urban, and health care access predictors. *Cancer Detect Prev* 2006;30:24-33.
10. Subramanian U, Hopp F, Lowery J, Woodbridge P, Smith D. Research in home-care telemedicine: Challenges in patient recruitment. *Telemed J E Health* 2004;10:155-161.
11. Ashburn A, Pickering RM, Fazakarley L, Ballinger C, McLellan DL, Fitton C. Recruitment to a clinical trial from the databases of specialists in Parkinson's disease. *Parkinsonism Rel Disord* 2007;13:35-39.
12. Wasserman J, Flannery MA, Clair JM. Raising the ivory tower: The production of knowledge and distrust of medicine among African Americans. *J Med Ethics* 2007;33:177-180.
13. Harris PA, Lane L, Biaggioni I. Clinical research subject recruitment: The Volunteer for Vanderbilt Research Program www.volunteer.mc.vanderbilt.edu *J Am Med Informatics Assoc* 2005;12:608-613.
14. Andersen MR, Schroeder T, Gaul M, Moinpour C, Urban N. Using a population-based cancer registry for recruitment of newly diagnosed patients with ovarian cancer. *Am J Clin Oncol* 2005;28:17-20.
15. Embi PJ, Jain A, Harris CM. Physicians' perceptions of an electronic health record-based clinical trial alert approach to subject recruitment: a survey. *BMC Med Inform Decision Making* 2008;8:13.
16. Smith KS, Eubanks D, Petrik A, Stevens VJ. Using web-based screening to enhance efficiency of HMO clinical trial recruitment in women aged forty and older. *Clin Trials* 2007;4:102-105.
17. Miller JL. The EHR solution to clinical trial recruitment in physician groups. *Health Manage Technol* 2006;27:22-25.
18. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chron Dis* 1987;40:373-383.
19. Bergner M, Bobbitt RA, Pollard WE, Martin DP, Gilson BS. The sickness impact profile: Validation of a health status measure. *Med Care* 1976;14:57-67.
20. Weingart SN, Rind D, Tofias Z, Sands DZ. Who uses the patient Internet portal? The PatientSite experience [see comment]. *J Am Med Informatics Assoc* 2006;13:91-95.
21. Miller H, Vandenbosch B, Ivanov D, Black P. Determinants of personal health record use: A large population study at Cleveland Clinic. *J Healthcare Inf Manage* 2007;21:44-48.
22. Kim E-H, Stolyar A, Lober WB, et al. Usage patterns of a personal health record by elderly and disabled users. *AMIA Annual Symposium Proceedings/AMIA Symposium*, 2007:409-413.
23. Ohmann C, Kuchinke W. Meeting the challenges of patient recruitment a role for electronic health records. *Int J Pharm Med* 2007;21:8.
24. Vickrey BG, Edmonds ZV, Shatin D, et al. General neurologist and subspecialist care for multiple sclerosis: Patients' perceptions. *Neurology* 1999;53:1190-1197.

Address correspondence to:

Deborah M. Miller, Ph.D.

Mellen Center for Treatment and Research  
in Multiple Sclerosis  
Neurological Institute  
Cleveland Clinic  
9500 Euclid Avenue, U-10  
Cleveland, OH 44195

E-mail: millerd@ccf.org

Received: June 8, 2009

Revised: August 24, 2009

Accepted: August 26, 2009