Published in final edited form as:
Clin Ther. 2008 August ; 30(8): 1553–1557. doi:10.1016/j.clinthera.2008.08.001.

Optimal Frequency for Measuring Health Care Resource Utilization in Parkinson's Disease Using Participant Recall: The FS-TOO Resource Utilization Substudy

Patrick D. Mauldin, PhD[1], **Paulo Guimaraes, PhD**[2], **Roger L. Albin, MD**[3], **E. Ray Dorsey, MD, MBA**[4], **Jacquelyn L. Bainbridge, PharmD**[5], and **Andrew Siderowf, MD, MSCE**[6] **for the NINDS NET-PD Investigators***

[1]Department of Clinical Pharmacy and Outcomes Sciences, Medical University of South Carolina, and the Ralph H. Johnson VA Medical Center, Charleston, South Carolina

[2]Department of Biostatistics, Bioinformatics, and Epidemiology, Medical University of South Carolina, Charleston, South Carolina

[3]Department of Neurology, University of Michigan Health System, Ann Arbor, Michigan

[4]Department of Neurology, University of Rochester Medical Center, Rochester, New York

[5]Department of Clinical Pharmacy, University of Colorado at Denver and Health Sciences Center, Denver, Colorado

[6]Department of Neurology, University of Pennsylvania Health System, Philadelphia, Pennsylvania

Abstract

Objective—The aim of this substudy was to determine the agreement between 2 approaches for measuring health care resource utilization (eg, number of hospital visits, number of primary care physician visits) in trial participants with Parkinson's disease (PD).

Methods—A substudy of the 1-year multicenter futility trial of GPI-1485 and coenzyme Q_{10} (FS-TOO) was performed to assess health care resource utilization agreement by measuring participant utilization recall after 12 months versus measuring participant utilization recall at regular 3-month intervals. Trial participants were selected from patients in the National Institutes of Health-sponsored FS-TOO multicenter study. Persons aged ≥30 years with confirmed PD diagnosis within the previous 5 years were eligible for inclusion in the substudy. Participants were also required to have at least 2 of 3 cardinal manifestations of PD (tremor, rigidity, and bradykinesia). Participants were excluded from the study if they had presence of atypical Parkinson's syndromes due to drugs, metabolic identified neurogenetic disorders, encephalitis, or other degenerative diseases. Agreement was determined using Lin's concordance and Cohen's kappa statistics.

Results—Between March and July of 2004, a total of 424 potential subjects were identified and evaluated for trial eligibility. Of these, 213 subjects (139 men, 74 women; mean [SD] age, 61.5 [10.3] years) met entry criteria and were included in the study. Trial participants were randomized equally to 1 of 3 groups. The 3 groups had similar baseline characteristics in terms of demographic data (age, race, sex, employment status, and annual income), total Unified Parkinson Disease Rating Scale (UPDRS) score, and UPDRS subscores. In this substudy, 141 participants had a *true*

Address correspondence to: Patrick D. Mauldin, PhD, Department of Clinical Pharmacy and Outcomes Sciences, Medical University of South Carolina, College of Pharmacy, PO Box 250144, QF213B, Charleston, SC 29425. mauldinp@musc.edu.

*National Institute of Neurological Disorders and Stroke, National Institutes of Health Exploratory Trials in Parkinson's Disease.

baseline visit, indicating a clinical baseline date, and 182 participants completed the Baseline Resource Utilization Form within 3 months of the true baseline visit. The comparison of concordance between the summed information over 3-month recalls and the 12-month recall from baseline was derived from these 182 participants. The level of agreement between the 2 approaches was high, ranging from 64.4% to 95.1%. Where disagreement was identified, the more frequent measurement approach (every 3 months) led to higher estimates, ranging from 20.4% to 77.4%.

Conclusion—The results of this trial indicate internal consistency with the self-reported measures of health care resource utilization, suggesting that these simple measures might provide reliable information about units of health care resource utilization in the context of clinical trials for PD.

Keywords

health care resources; Parkinson's disease; clinical trials

INTRODUCTION

The evaluations of cost-effectiveness for therapeutic interventions are increasingly important. These evaluations are especially important even if the new drug being tested is not the first in its class. Health care resource utilization by the patient may or may not be equivalent across all drugs that are chemically similar, despite their being a part of the same pharmaceutical class. The US Food and Drug Administration does not mandate such evaluations, but in some instances, the Centers for Medicare & Medicaid Services and other payers may use cost-effectiveness analyses as part of their decision-making process for reimbursing new diagnostic and therapeutic interventions. To evaluate the costs and benefits of new therapies, researchers and policymakers need to collect patient health care resource utilization data from clinical trials. Such data are critically important with regard to therapies for neurodegenerative conditions, such as Parkinson's disease (PD), where the health care costs of the illness are high.[1]

There has been no clear evidence supporting either archival records or self-report for the retrieval of health care resource utilization data, and in other complex conditions such as substance abuse and AIDS, patient recall of health care resource utilization (eg, office visits, diagnostic examinations) every 3 months appears to produce reliable estimates of health care utilization.[2-4] However, patient recall of health care services beyond 3 months in patients with PD is less clear. Therefore, the National Institute of Neurological Disorders and Stroke sponsored a 1-year, multicenter futility trial[5] of GPI-1485 and coenzyme Q10 (CoQ_{10}) in early PD (FS-TOO).[6]

The aim of this substudy was to determine whether an assessment of health care resource utilization at 12 months (ie, a single assessment) is comparable to more frequent assessments at regular 3-month intervals in determining the agreement between these 2 approaches for measuring health care resource utilization.

STUDY PARTICIPANTS AND METHODS

Participants

Subjects aged ≤30 years, with a confirmed diagnosis of PD made within the last 5 years and who were not requiring any medication for the treatment of their symptoms at the time of study entry, were eligible to participate in the study. Participants were also required to have at least 2 of 3 cardinal manifestations of PD (tremor, rigidity, and bradykinesia). Participants were excluded from the study if they had presence of atypical Parkinson's syndromes due to

drugs (eg, metoclopramide, flunarizine), metabolic identified neurogenetic disorders (eg, Wilson's disease), encephalitis, or other degenerative diseases (eg, progressive supranuclear palsy). In addition, patients with any clinically significant medical condition (eg, active gastrointestinal illnesses, angina, active neoplasm) or laboratory abnormality, which would, in the judgment of the investigator, interfere with the subject's ability to participate in the study or to be followed, were also excluded.

Collection of Health Care Resource Utilization Data

The primary objective of the FS-TOO study was to compare the impact of GPI-1485 and CoQ10 on the progression of PD to assess whether it was nonfutile to proceed with further study of these agents. The futility design allowed for the comparison of a single treatment arm against a predetermined threshold value reflective of a clinically meaningful change over a short period of time.[5] The trial was conducted in 43 sites across the United States and in 2 sites in Canada. We collected trial participant recall data on health care resource utilization. For the FS-TOO Health Care Resource Utilization substudy, follow-up health care resource utilization data were obtained from responses to survey questions on separate case-report forms (CRFs) administered to trial participants by the study coordinator or nurse. The trial participants were asked to recall their utilization of major cost drivers for PD, including hospitalizations, procedures, length of stay, time to first hospitalization/procedure, diagnostic procedures, rehabilitation stays, outpatient visits, and personal patient costs. We did not collect data on whether the trial participant or a surrogate responded. However, for a separate questionnaire assessing total functional capacity used in the trial, 60 participants required assistance from a caregiver or companion, suggesting that up to 60 surveys may have been completed by surrogates. The baseline visit CRF, standard 3-month follow-up CRF (3-month recall), and standard 12-month follow-up CRF (12-month recall) were each 1 page, and capture follow-up information on hospitalizations, physician/professional visits, employment information, and personal subject costs (can be provided by corresponding author on request). The surveys were designed by selecting questions from forms used in the Agency for Healthcare Research and Quality, and Research and Development co-sponsored HIV Cost and Services Utilization Study,[7] and a National Institutes of Health-sponsored study comparing nelfinavir and ritonavir in HIV patients.[8] Minor modifications in wording were made due to the clinical nature of PD compared with HIV (eg, the recording of Other Health Professional Visits, including physical, occupational, and speech therapists).

This trial received Institutional Review Board approval at all participating sites and all trial participants signed informed-consent forms.

Statistical Analysis

Reliability was assessed by examining the agreement between the cumulative data reported in the 3-month and the 12-month follow-up CRFs. For most questions, agreement between the cumulative results reported in the 3- and 12-month follow-up CRFs was assessed using Lin's concordance correlation coefficient, an adequate method for continuous variables and preferable when evaluating reproducibility.[9] For questions with discrete variables, as in question 1, only 3 different values (0, 1, and 2) could be selected. Therefore, the degree of agreement was measured using Cohen's kappa statistics, which is more appropriate for categoric data.[10]

To assess systematic difference between the information reported in the 2 questionnaires, an exact binomial test was used and the distribution of points that fell above and below the 45-degree line of perfect agreement was tested. For those variables that revealed systematic biases, an additional analysis that examined factors that might explain the observed discrepancies between the values reported in the 3- and 12-month recall CRFs was

performed. A variable for the difference between the value reported in the 12-month recall form and the sum of the values reported in the 3-month recall follow-up forms was computed. We then estimated linear regressions using the absolute difference as a dependent variable and different specifications that accounted for individual baseline characteristics (age, race, and sex), the total Unified Parkinson Disease Rating Scale (UPDRS) and measures of cognition such as the Frontal Assessment Battery (FAB), and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). To account for the anomalous nature of the dependent variable, the SEs were bootstrapped using 1000 replications. All calculations were performed with the statistical package Stata version 9.2 Stata-Corp LP, College Station, Texas).

RESULTS

Between March and July of 2004, 424 potential subjects were identified and evaluated for trial eligibility. A total of 213 subjects met study entry criteria and were randomized equally to 1 of 3 groups. The groups had similar baseline characteristics, in terms of demographic data (age, race, sex, employment status, and annual income) (Table I), total UPDRS, and UPDRS subscores. Approximately 2 months after study initiation, the utilization substudy began. In the substudy, 141 participants had a *true* baseline visit, indicating a clinical baseline date, and 182 participants completed the Baseline Resource Utilization Form within 3 months of the true baseline visit. The comparison of concordance between the summed information over 3-month recalls and the 12-month recall from baseline was based on these 182 participants.

As shown in Table II, the concordance between measuring health care resource utilization at 12 months (ie, once a year) versus every 3 months ranged from 64.4% to 95.1% for all 7 questions on health care resource utilization. Also, 5 of the 7 questions asked had <50% of responses that were not zero.

Restricting our sample to the nonzero responses, the measures of agreement were slightly reduced, with the largest decrease occurring in the measure of agreement for question 5, from 77.5% to 68.4%. The results of the test for systematic bias provided no evidence of bias in questions 1, 2, and 6 with a 95% CI. However, for the remaining questions (questions 3, 4, 5, and 7), we found that when there was a disagreement between the 2 measures, individuals tended to underestimate their health care resource utilization on their 12-month recall questionnaire.

The regression analysis was performed only on those variables that had a majority of nonzero responses (questions 3 and 4). The outcomes of the regression analysis were similar for the 2 questions, so we discuss them jointly. We first performed unadjusted regressions in each of the covariates and then specified models with all covariates. For all the different specifications that were tested, no covariate was found to be statistically significant. Age ($P = 0.47$, $P = 0.09$), race ($P = 0.27$, $P = 0.60$), and sex ($P = 0.25$, $P = 0.99$) did not show relation to the dependent variables in the unadjusted regressions. The total UPDRS ($P = 0.21$, $P = 0.76$) as well as the 2 measures of cognition, FAB ($P = 0.25$, $P = 0.23$) and RBANS ($P = 0.63$, $P = 0.58$), were not statistically significant in any of the specifications. Likewise, in the multiple linear regression, age ($P = 0.52$, $P = 0.29$), race ($P = 0.20$, $P = 0.75$), sex ($P = 0.17$, $P = 0.96$), total UPDRS ($P = 0.34$, $P = 0.95$), FAB ($P = 0.10$, $P = 0.28$), and RBANS ($P = 0.31$, $P = 0.86$) were not statistically significant.

DISCUSSION

Our analysis of the FS-TOO Resource Utilization substudy data shows that there is moderate to good agreement between assessing health care resource utilization at 12 months and at 3-month intervals. The results indicate internal consistency with these self-reported measures of health care resource utilization, suggesting that these simple measures will provide reliable information (as defined by the concordance measures) about units of health care resource utilization in this population. Such data may be useful in assessing the economic impact of novel therapies for PD and other neurodegenerative disorders in future Phase III trials where minimal collection of data is preferred. This self-reported data might be combined with other data sets to produce estimates of the costs of health care resource utilization in this study population. For example, combining these measures of health care resource utilization with regional average measures of the costs of physician visits and hospitalizations would give crude but useful estimates of the direct costs incurred by PD patients or their insurers.

Limitations

This substudy did have some unavoidable limitations. Health care resource utilization is self-reported and the magnitude of an episode of health care resource utilization is not measured. In addition, although we combined items from 2 validated instruments, we did not revalidate the new combination. Our trial participants tended to be somewhat younger and better educated than the typical PD population, reflecting the tendency of trials in specialty centers to recruit more educated and motivated patients. Our trial participants were also early in their disease course, as they were required to be untreated with symptomatic therapy at entry to the trial. This subpopulation of PD trial participants is more likely to self-report health care resource utilization accurately. A high percentage of our trial participants were working, and as individuals with early, mild PD, may be less likely to have high levels of health care resource utilization, which might make their self-reporting more accurate.

CONCLUSIONS

The results of our study indicate that simple methods may capture important information about health care resource utilization in the context of clinical trials for PD and other neurodegenerative disorders. Such data are likely to become increasingly important in evaluating the impact of novel therapies.

Acknowledgments

This study was sponsored by the National Institute of Neurological Disorders and Stroke (U01NS043127, U01NS043128, and U10NS44415 through 4455).

References

1. Noyes K, Liu K, Li Y, et al. Economic burden associated with Parkinson’s disease on elderly Medicare beneficiaries. Mov Disord. 2006; 21:362–372. [PubMed: 16211621]
2. Epstein AM, Seage G III, Weissman JS, et al. Costs of medical care and out-of-pocket expenditures for persons with AIDS in the Boston Health Study. Inquiry. 1995; 32:211–221. [PubMed: 7601519]
3. Killeen TK, Brady KT, Gold PB, et al. Comparison of self-report versus agency records of service utilization in a community sample of individuals with alcohol use disorders. Drug, Alcohol Depend. 2004; 73:141–147. [PubMed: 14725953]
4. Tourangeau, R.; Rasinski, KA. Evaluation of Data Collection Frequency and the Use of a Summary in the National Medical Utilization and Expenditure Survey, Series A. Washington, DC: US Government Printing Office; 1987.

5. Tilley BC, Palesch YY, Kieburtz K, et al. for the NET-PD Investigators. Optimizing the ongoing search for new treatments for Parkinson disease: Using futility designs. Neurology. 2006; 66:628–633. [PubMed: 16534099]

6. The NINDS NET-PD Investigators. A randomized clinical trial of coenzyme Q10 and GP1-1485 in early Parkinson disease. Neurology. 2007; 68:20–28. [PubMed: 17200487]

7. Shapiro MF, Morton SC, McCaffrey DF, et al. Variations in the care of HIV-infected adults in the United States: Results from the HIV Cost and Services Utilization Study. JAMA. 1999; 281:2305–2315. [PubMed: 10386555]

8. Perez G, MacArthur RD, Walmsley S, et al. for the Terry Beirn Community Programs for Clinical Research on AIDS; Canadian Trials Network. A randomized clinical trial comparing nelfinavir and ritonavir in patients with advanced HIV disease (CPCRA 042/CTN 102). HIV Clin Trials. 2004; 5:7–18. [PubMed: 15002082]

9. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 1989; 45:255–268. [PubMed: 2720055]

10. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960; 20:37–46.

Table I

Baseline characteristics of the FS-TOO[6] health care resource utilization trial participants (N = 182). Unless otherwise specified, data are %.

Variable	Value
Age, y	
Mean (SD)	61.5 (10.0)
Range (median)	32-87 (62)
Race, white	92.9
Sex, male	68.1
Employment status	
Full-time	43.0
Part-time	11.0
Retired	35.5
Other	10.5
Annual income	
<$35,000	33.1
$35,000-$70,000	32.6
>$70,000	34.3

Table II

Concordance between measuring health care resource utilization at 12 months versus 3-month intervals. Data are %.

Question	Percent of Responses That Were Not Zero (N = 182)	Measure of Agreement (95% CI)
1. Have you been in the hospital overnight or longer since your baseline visit? (number of admissions)	10	82.9 * (60.2-93.3)
2. Total number of hospital nights since your baseline visit?	10	95.1 † (93.6-96.5)
3. How many visits have you had to your primary care physician since your baseline visit?	88	74.3 † (67.7-80.9)
4. How many visits have you had to a specialist since your baseline visit?	74	74.0 † (67.5-80.5)
5. How many visits have you had to a therapist since your baseline visit?	27	77.5 † (72.5-82.6)
6. How many “other” professional visits have you had since your baseline visit?	38	90.1 † (87.3-92.9)
7. Number of days forced to take away from occupation due to illness?	27	64.4 † (58.6-70.3)

* Data determined by Cohen’s kappa.

† Data determined by Lin’s concordance coefficient.