

Is Personality a Key Predictor of Missing Study Data? An Analysis From a Randomized Controlled Trial

Anthony Jerant, MD¹

Benjamin P. Chapman, PhD²

Paul Duberstein, PhD²

Peter Franks, MD^{1,3}

¹Department of Family and Community Medicine, University of California Davis School of Medicine, Sacramento, California

²Laboratory of Personality and Development, Department of Psychiatry, University of Rochester Medical Center, Rochester, New York

³Center for Healthcare Policy and Research, Department of Family and Community Medicine, University of California Davis School of Medicine, Sacramento, California

ABSTRACT

PURPOSE Little is known regarding the effects of psychological factors on data collection in research studies. We examined whether Five Factor Model (FFM) personality factors—Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness—predicted missing data in a randomized controlled trial (RCT).

METHODS Individuals (N = 415) aged 40 years and older with various chronic conditions, plus basic activity impairment, depressive symptoms, or both, were recruited from a primary care network and enrolled in a 6-week RCT of an illness self-management intervention, delivered by means of home visits or telephone calls or usual care. Random effects logistic regression modeling was used to examine whether FFM factors predicted missing illness management self-efficacy data at any scheduled follow-up (2, 4, and 6 weeks, and 6 and 12 months), controlling for disease burden, study arm, and sociodemographic characteristics.

RESULTS Across all follow-up points, the missing data rate was 4.5%. Higher levels of Openness (adjusted odds ratio [AOR] for 1-SD increase = 0.24; 95% CI, 0.12-0.46; $P < .001$), Agreeableness (AOR = 0.29; CI 0.14-0.60; $P = .001$), and Conscientiousness (AOR = 0.24; CI 0.15-0.50; $P < .001$) were independently associated with fewer missing data. Accuracy of the missing data prediction model increased when personality variables were added (change in area under the receiver operating characteristic curve from 0.71 to 0.77; $\chi^2_1 = 6.6$; $P = .01$).

CONCLUSIONS Personality was a powerful predictor of missing study data in this RCT. Assessing personality could inform efforts to enhance data completion and adjust analyses for bias caused by missing data.

Ann Fam Med 2009;7:148-156. DOI: 10.1370/afm.920.

Conflicts of interest: none reported

CORRESPONDING AUTHOR

Anthony Jerant, MD
Department of Family and Community
Medicine
UC Davis School of Medicine
4860 Y St, Suite 2300
Sacramento, CA 95817
afjerant@ucdavis.edu

INTRODUCTION

Despite advantages compared with observational studies, randomized controlled trials (RCTs) have limitations.¹ One threat to their validity is bias resulting from nonrandom missing data. Commonly measured sociodemographic²⁻⁶ and health status⁷⁻⁹ variables have been associated with missing data caused by study attrition. Because such variables may also affect study outcomes, bias in the assessment of treatment effects may occur, though statistical adjustment can mitigate bias.

Psychological factors also likely affect study participants' decisions to drop out from, keep data collection appointments for, or complete all questionnaire items in clinical research studies, yet they are rarely measured. This omission may be critical in assessing the validity and applicability of studies, since adherence to study protocols, including placebos, can have profound effects on outcomes, comparable to active treatment effects.¹⁰⁻¹²

Although earlier studies have explored attrition resulting from psychological factors in patients with psychiatric disorders¹³⁻¹⁵ and arthritis,¹⁶ finding associations with anxiety,^{13,14} depression,¹⁶ disordered personality,¹⁴

and disease attributions and treatment beliefs,¹⁵ the role of psychological factors in predicting attrition in other RCT settings remains unclear. The personality factors in the Five Factor Model (FFM)—Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness—are particularly promising targets for study.¹⁷⁻¹⁹ The result of more than 7 decades of research,¹⁷ FFM factors are empirically derived, broad clusters of behavioral and dispositional tendencies (Table 1). They capture the major axes of psychological and behavioral variation in humans and are associated with an array of important health behaviors and outcomes.²⁰⁻²⁴ Higher levels of Conscientiousness have been consistently associated with better adherence to chronic illness therapies.²⁵⁻³¹ Similar associations might exist between Conscientiousness and incomplete trial data, but no studies have addressed this question, and associations between other FFM factors and missing data are also unexplored.

We examined the relationship of FFM personality factors to missing data in a RCT of Homing in on Health (HIOH), which was developed to make

the Chronic Disease Self-Management Program (CDSMP)^{32,33} available to persons less able or willing to participate in group training. The goals were to determine whether in-home and telephone-delivered HIOH would enhance self-efficacy, or confidence for executing self-management tasks, and improve outcomes in persons with chronic conditions.

Based on prior FFM research,³⁴ we examined the following hypotheses: Missing data will be more likely for participants higher in (1) Neuroticism because they will experience difficulties bearing the emotional distress of RCT participation; but missing data will be less likely for participants higher in (2) Extraversion, which is associated with gregariousness and positive mood,¹⁷⁻¹⁹ potentially enhancing enjoyment of trial participation; (3) Openness, which may stimulate interest in novel experiences,¹⁷⁻¹⁹ such as RCT participation; (4) Agreeableness, which is associated with adherence and altruism¹⁷⁻¹⁹; and (5) Conscientiousness, which is associated with reliability, diligence, and health-consciousness.¹⁷⁻¹⁹ Further analyses examined whether the FFM factors provide incremental power in predicting

missing data beyond sociodemographic and other commonly measured variables²⁻⁹ or confound apparent associations of those variables with attrition.

Table 1. Dispositional Tendencies Within 5 Factor Model (FFM) Personality Factors

FFM Factor	Dispositional Tendencies	Example Personality Questionnaire Items
Agreeableness	Cooperative, compassionate	I make people feel at ease. I sympathize with others' feelings. I take time out for others. I am not interested in others people's problems. ^a I am not really interested in others. ^a
Conscientiousness	Painstaking, careful, planning, achievement-driven	I am always prepared. I am exacting in my work. I follow a schedule. I get chores done right away. I like order.
Extraversion	Stimulation seeking, tendency to experience positive emotions, sociable	I feel comfortable around people. I start conversations. I talk to a lot of different people at parties. I am quiet around strangers. ^a I laugh a lot.
Neuroticism	Tendency to experience negative emotions	I am easily disturbed. I change my mood a lot. I get irritated easily. I get stressed out easily. I get upset easily.
Openness	Explore new ideas and experiences, intellectually curious	I have a vivid imagination. I have excellent ideas. I spend time reflecting on things. I use difficult words. I am not interested in abstract ideas. ^a

^a Reverse scored items.

METHODS

The study was conducted from July 2004 through December 2007 and approved by our institutional review board. We recruited participants from the 12 offices in the University of California Davis Primary Care Network. In CDSMP studies, the intervention effect size for self-efficacy was around 0.3.^{32,33} Our power calculations estimated that 120 participants per group (360 total) would provide a power of 0.80 to detect an effect size of 0.2.

Using billing information, we identified patients aged 40 or older with arthritis, asthma, chronic obstructive pulmonary disease, congestive heart failure, depression, or diabetes mellitus and contacted them by mail and telephone. Interested patients were further screened for ability to speak and read English; resi-

dence in a private home with an active telephone; adequate eyesight and hearing to participate; and at least 1 basic activity impairment, as assessed by the Health Assessment Questionnaire³⁵ or a score of 4 points or greater on the 10-item Center for Epidemiologic Studies Depression Scale.³⁶ A study nurse visited eligible individuals in their homes to obtain informed consent, administer the baseline questionnaire, and implement randomized allocation in blocks of 12 participants using sealed opaque envelopes containing printed group assignments.

Procedures

Study Intervention

HIOH was delivered over 6 weekly sessions, and the content provided was essentially identical to that for the original CDSMP. HIOH differed from the CDSMP terms of delivery process, however, as only a single trained layperson provided all 6 intervention sessions to each participant. It also differed in delivery setting, being provided in participants' homes or by telephone, using the same script for both intervention groups, rather than in centralized groups. Four non-health care professionals underwent intensive week-long training to deliver the intervention. Further details about HIOH are available from the authors upon request.

Usual-Care Control Group

Usual-care participants completed the same baseline and follow-up telephone questionnaires as intervention participants. Otherwise, they received the care delivered by their usual clinicians.

Study outcomes, including efficacy for illness self-management, were measured through telephone calls at 2, 4, 6, 26, and 52 weeks. Data collection personnel were blinded to participant group assignment. Participants were paid \$25 after each scheduled data collection (\$125 if all 5 were completed).

Measures

Five Factor Model Personality Factors

At baseline, participants completed the 60-item NEO Five-Factor Inventory (NEO-FFI).³⁴ Its 5, 12-item scales tap the FFM personality factors of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Scores were standardized (mean = 0, SD = 1) to facilitate interpretation. Cronbach's α for the 5 scales ranged from 0.70 to 0.87 in this sample.

Covariates

Covariates measured by participant self-report at baseline were age in years, sex, race/ethnicity (white, black, Hispanic, Asian, other), health insurance (private, public, uninsured), education level (non-high school

graduate, graduated high school, some college, graduated college, greater than college), and disease burden (1, 2, 3, 4, or more chronic diseases).

Missing Study Data

Missing study data were defined as missing for illness management self-efficacy data—a primary outcome of the RCT—at any study follow-up point for any reason. Study data could have been missing because of study dropout, missed data collection appointments without permanent dropout, and failure to complete the illness management self-efficacy questionnaire at 1 or more data collection appointments. Missing data were assessed at each of the 5 scheduled follow-up points for each participant.

Analyses

Analyses were conducted using Stata, version 10.0 (StataCorp LP, College Station, Texas). We used mixed-effects logistic regression to model the status of illness management self-efficacy data collection at each of the 5 scheduled study follow-up points, including random effects to account for repeated measures. Independent variables of interest were the FFM personality factors. Adjustment for health status involved covariates for disease burden, with indicators for 2, 3, or 4 or more chronic diseases (reference = 1). Demographic covariates included age, an indicator variable for female sex (reference = male), and an indicator variable for racial/ethnic minority category (reference category = white). Socioeconomic covariates included lack of insurance and education, with indicators for high school, some college, college, and greater than college (reference category = less than high school). We adjusted for length of follow-up through inclusion of indicators for the 2-, 4-, 6-, and 26-week follow-up (reference = 52 weeks). Study group was also included, with home group as the reference. To assess the degree to which personality factors might enhance power to predict missing study data over other commonly measured covariates alone, we compared the area under the receiver operating characteristic curve (sensitivity plotted against 1-specificity) for a model with only commonly measured covariates with that for a model with commonly measured covariates plus personality factors.

RESULTS

Figure 1 shows the flow of participants through the RCT. In all, 415 participants were randomized (home intervention = 138, telephone intervention = 139, usual care = 138). Of these, 321 (77%) were female, with a mean age of 60 years (range 41 to 95 years). Most participants (59%) reported 2 or more chronic conditions.

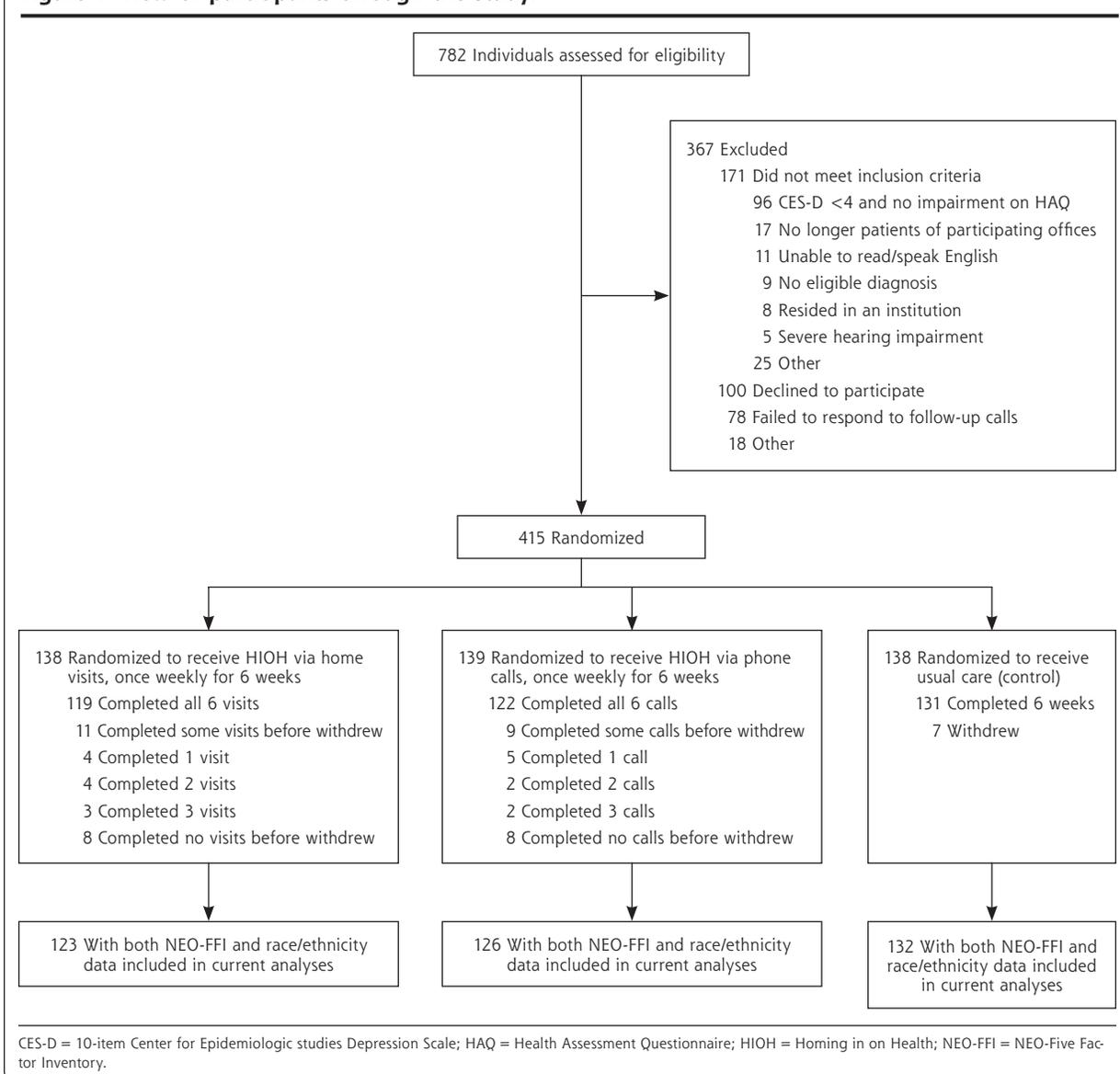
Figure 1. Flow of participants through the study.

Table 2 summarizes participants' baseline characteristics and rates of missing data at each follow-up point. Nearly all (94%, or 384) completed the NEO-FFI at baseline, but 3 did not provide education or race/ethnicity data, resulting in 381 participants included in analyses. Across all follow-up points, there were 86 instances of missing data (4.5%). Three instances of missing data (3%) were due to death (home intervention = 1, usual care = 2), all occurring after the 6-month assessment, whereas 2 were due to severe illness, both in the home group and both precluding data collection at 2 weeks.

Table 3 shows multivariate models predicting missing data from (1) health status, demographic, socioeco-

nom, and follow-up duration predictors only, and (2) the same covariates plus personality factors. Lower odds of missing data was associated with higher Openness (adjusted odds ratio [AOR]=0.24; 95% confidence interval [CI] 0.12-0.46; $P < .001$), higher Agreeableness (AOR = 0.29; 95% CI 0.14-0.60; $P = .001$), and higher Conscientiousness (AOR=0.28; 95% CI 0.15-0.50; $P < .001$). Adding personality factors increased the area under the receiver operating characteristic curve from 0.711 in model 1 to 0.774 in model 2, a significant improvement in predictive accuracy ($\chi^2_1 = 6.6$, $P = .01$). Figure 2 depicts the improvement in area under the curve resulting from the addition of FFM personality factors.

Adjusting for personality in model 2 also attenuated

Table 2. Characteristics of Study Participants

Characteristic	Home (n = 138)	Telephone (n = 139)	Usual-Care Control (n = 138)
Age, mean (SD), years	59.8 (11.2)	61.2 (11.6)	60.1 (11.7)
Female, No. (%)	108 (78)	109 (78)	104 (75)
Race/ethnicity, No. (%)			
Non-Hispanic white	103 (75)	110 (79)	115 (83)
Black	20 (15)	11 (8)	15 (11)
Other	14 (9)	14 (10)	7 (5)
Declined to answer	1 (1)	4 (3)	1 (1)
Education level, No. (%)			
High school or less	19 (14)	20 (14)	22 (16)
Some college	53 (38)	50 (36)	58 (42)
College graduate or greater	66 (47)	65 (47)	57 (41)
Declined to answer	0 (0)	4 (3)	1 (1)
Income level, No. (%)			
<40,000	41 (30)	42 (31)	44 (32)
40,000-79,999	42 (30)	37 (27)	43 (31)
≥80,000	22 (16)	27 (20)	22 (16)
Declined to answer	33 (24)	33 (24)	29 (21)
Married, No. (%)	79 (57)	79 (57)	76 (55)
Uninsured, No. (%)	3 (2)	5 (4)	2 (2)
Chronic conditions, No. (%)			
1	55 (40)	72 (51)	43 (31)
2	51 (37)	40 (29)	65 (47)
3	18 (13)	21 (15)	21 (15)
>4	14 (10)	6 (4)	9 (7)
Study diagnoses, No. (%) ^a			
Arthritis	83 (60)	73 (52)	77 (55)
Depression	59 (43)	64 (46)	70 (51)
Diabetes	64 (46)	50 (36)	58 (42)
Asthma	34 (25)	25 (18)	39 (28)
Chronic obstructive lung disease	15 (11)	11 (8)	17 (12)
Congestive heart failure	17 (12)	17 (12)	14 (10)
Personality factors, mean (SD) ^b			
Neuroticism	20.6 (9.9)	22.1 (8.9)	21.9 (9.6)
Extraversion	25.9 (8.0)	26.1 (6.9)	25.9 (7.6)
Openness	28.2 (6.6)	28.6 (6.3)	28.9 (6.3)
Agreeableness	34.4 (4.8)	33.3 (5.6)	33.3 (5.7)
Conscientiousness	31.3 (7.5)	32.0 (6.4)	31.9 (6.7)
Participants missing self-efficacy data, No. (%)			
2-week follow-up	4 (3)	13 (10)	2 (2)
4-week follow-up	10 (8)	10 (8)	5 (4)
6-week follow-up	5 (4)	6 (5)	1 (1)
6-month follow-up	6 (5)	6 (5)	2 (2)
1-year follow-up	8 (6)	9 (7)	4 (3)

^a Percentages exceed 100 because many participants had more than 1 condition.
^b Scored on a range from 0 to 48.

Table 3. Multivariate Models of Predictors of Missing Study Data Status

Predictor	Model 1 AOR (95% CI)	Model 2 AOR (95% CI)
Number of chronic diseases ^a		
2	5.34 (1.12-25.43) ^b	1.31 (0.24-7.10)
3	0.67 (0.14-3.19)	0.68 (0.61-1.53)
>4	0.19 (0.01-2.45)	0.01 (0.00-0.09) ^c
Demographic factors		
Age	0.99 (0.94-1.03)	1.02 (0.95-1.08)
Ethnic/racial minority ^d	2.18 (0.67-7.12)	13.96 (1.99-98.08) ^e
Female ^f	0.05 (0.01-0.23) ^e	0.58 (0.15-2.32)
No health insurance ^g	0.00 (0.00-0.00)	0.00 (0.00-0.00)
Education level ^h		
High school graduate	0.01 (0.00-0.01) ^e	0.01 (0.00-0.06) ^e
Some college	0.01 (0.00-0.01) ^e	0.02 (0.01-0.18) ^c
College graduate	0.01 (0.00-0.01) ^e	0.01 (0.00-0.12) ^c
Post-college	0.01 (0.00-0.01) ^e	0.01 (0.00-0.04) ^e
Intervention arm ⁱ		
Telephone intervention	0.40 (0.11-1.36)	0.95 (0.94-2.46)
Usual care	0.01 (0.00-0.07) ^e	0.32 (0.06-1.89)
Assessment point ^j		
2 Weeks	0.74 (0.25-2.18)	0.73 (0.58-2.42)
4 Weeks	1.68 (0.61-4.60)	1.68 (0.61-4.60)
6 Weeks	0.15 (0.04-0.63) ^c	0.11 (0.02-0.57) ^c
6 Months	0.27 (0.08-0.96) ^b	0.23 (0.06-0.92) ^b
Five Factor Model personality factors ^k		
Neuroticism	–	0.96 (0.93-4.00)
Extraversion	–	0.97 (0.95-4.01)
Openness	–	0.24 (0.12-0.46) ^c
Agreeableness	–	0.29 (0.14-0.60) ^c
Conscientiousness	–	0.28 (0.15-0.50) ^c
Model predictive precision (area under curve)	0.711	0.774 ^b

AOR = adjusted odds ratio; CI = confidence interval.

Model 1: adjusted for disease burden, age, sex, race/ethnicity, insurance, education, study group, and length of follow-up.

Model 2: adjusted for covariates in model 1, plus Five Factor Model personality factors (Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness).

^a Reference category, 1 chronic disease.

^b $P < .05$.

^c $P < .01$.

^d Reference category, non-Hispanic white.

^e $P < .001$.

^f Reference category, male.

^g Reference category, have health insurance.

^h Reference category, less than high school graduate.

ⁱ Reference category, home intervention.

^j Reference category, 1 year.

^k Personality factors are scaled in standard deviation units.

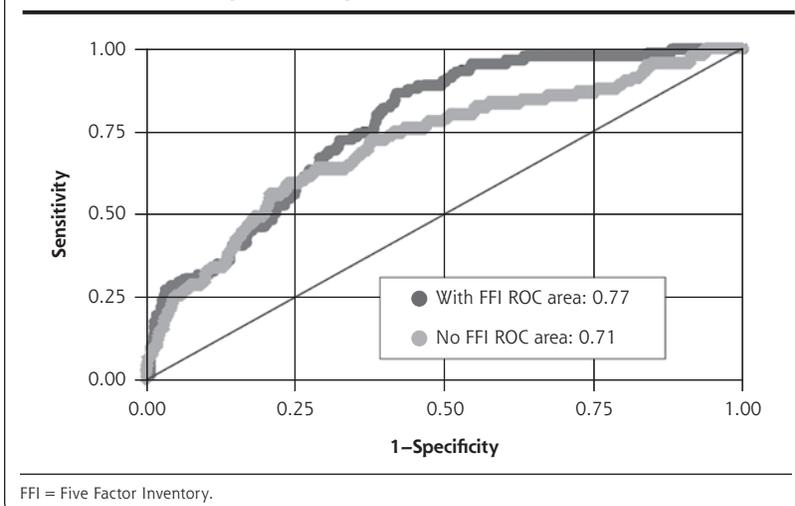
and having 4 or more chronic diseases, both of which decreased the risk of missing data.

DISCUSSION

We found missing data were significantly less likely in participants with higher levels of Openness, Conscientiousness,

the effect associated with sex of the participant. Other significant covariates in model 2 included minority status, which significantly increased the risk of missing data, and high school or greater level of education

Figure 2. Receiver operating characteristic (ROC) curves for models with and without personality factors included.



tiousness, and Agreeableness. In previous studies, FFM personality factors have been associated with other important health behaviors and outcomes.²⁰⁻³¹ Thus, personality effects could have important consequences for the validity and applicability of RCT findings.

Although this study is the first to find significant associations between FFM factors and missing trial data, the nature and direction of the associations were generally consistent with prior FFM research. For example, higher levels of Openness have been linked to receptivity to novel activities,³⁷ which could increase interest in RCT completion. Similarly, individuals with higher levels of Agreeableness tend to be more compliant and eager to please,^{17,34} characteristics rendering RCT completion more likely. Finally, higher levels of Conscientiousness have been associated with greater dutifulness³⁴ and treatment adherence,²⁵⁻³¹ characteristics likely to reduce RCT attrition. These findings suggest a discrete psychological profile predicting missing trial data may exist.

Accounting for FFM personality factors also enhanced prediction of missing data beyond that achieved by commonly measured sociodemographic,²⁻⁶ disease burden-related,^{7,8} and study-related correlates of missing data in prediction models. The area under the receiver operating characteristic curve for random guesses about who will fail to complete any given data collection is 0.5. In the current study, the combined effects of disease, sociodemographic, and intervention covariates improved the area under the curve to 0.711, an increase of 0.211. Accounting for FFM personality factors resulted in further improvement of 0.063, increasing the total area under the curve to 0.774. Thus, the additional increment in predictive precision

provided by FFM personality factors over other covariates was about 30% ($[0.063/0.211] \times 100$) as much as the initial increment over chance alone of health status, sociodemographic factors, intervention arm, and follow-up duration combined.

Inclusion of personality variables also attenuated the association between sex of the participant and missing data. Higher levels of Agreeableness and Neuroticism in women are 2 of the most universally observed associations in personality research, replicated in more than 27 countries and across age-groups.^{38,39} The current findings suggest differences by sex in missing data in some previous studies^{3,4,6} may be partially explained by personality.

Our study findings raise concerns about whether results of RCTs that rest on the psychological characteristics of enrollees, yet fail to measure and adjust for bias resulting from such characteristics, are applicable to the general population. Policy decisions stemming from such RCTs may lead to the expenditure of valuable resources supporting health interventions targeted to the general population, yet our findings suggest the interventions may only be effective in or adhered to by a subset of patients.⁴⁰

Our findings also have important methodological ramifications for both RCTs and observational studies. Regarding RCTs, currently available methods for handling missing data, such as intention-to-treat analysis and last observation carried forward, assume dropout occurs at random⁴¹ and ignore the possibility that unmeasured variables driving missing data might interact with interventions to worsen outcomes. Thus, currently available methods for handling missing data do not address and may actually compound bias because of nonrandom missing data. By contrast, measuring and including the status of FFM personality factors in analyses may aid more precise modeling of missing data mechanisms,⁴² thereby improving their applicability to general practice. Integral to this aim will be the use of reliable, valid personality measures. A number of brief FFM measures are available,⁴³⁻⁴⁵ but they require broader use and testing to better establish their reliability and validity.

Our study findings further suggest that greater efforts to retain participants identified as being at-risk for missing data collection based on their psychological characteristics may be an important step toward improving the conduct of research studies. For exam-

ple, we found participants with lower levels of Conscientiousness were more likely than those with higher levels to have missing data. Thus, additional structure in the data collection process (eg, reminder telephone calls, printed schedules, and other prompts) might help increase data collection individuals with lower levels of Conscientiousness, though research studies are required to examine this hypothesis formally. In light of the well-established association between Conscientiousness and treatment adherence,²⁵⁻³¹ such approaches might also be considered for improving treatment adherence in clinical practice.

Finally, our findings raise the intriguing possibility that FFM personality factors may moderate participants' responses to interventions.⁴⁶ Substantial variation in response to RCT interventions may exist, with some participants improving while others do not respond or are even harmed as a result of the moderating effects of one or more variables.⁴⁷ We recently conducted analyses, to be summarized in an additional report, showing that several FFM factors moderated the effects of HIOH on self-efficacy. These findings suggest the need to explore tailoring interventions to participant personality. They also underscore the need for broader use of brief FFM personality measures in both primary care research and practice⁴³⁻⁴⁵ to better assess their reliability, validity, and potential clinical utility.

More speculatively, it is also likely that personality is associated with decisions to enroll in clinical research studies⁴⁸ and with adherence to study protocols once enrolled. Personality variables have been associated with important health outcomes,²⁰⁻²⁴ and ongoing faithful study participation can itself have potent effects on outcomes.^{10-12,49} The interrelationships among personality, study enrollment and retention, protocol adherence, and health outcomes suggest personality factors may be key variables confounding interpretation of the relationship between study interventions and outcomes. Further work exploring these issues will be important.

Also, although observational studies may offer greater external validity than RCTs, they are more susceptible to bias caused by participant characteristics. Group assignment is nonrandom, and those in the intervention arm are typically already adhering to the intervention when enrolled. This powerful healthy adherer effect may bias observational studies of interventions toward positive findings and may help explain why their findings are often contradicted by RCTs.⁵⁰⁻⁵² The current findings suggest that accounting for personality could mitigate the healthy adherer effect and enhance the internal validity of observational studies. At minimum, routine inclusion of personality data in analytic models, similar to the routine inclusion of demographics, could characterize potential bias.

Our study had some limitations. The sample was composed of participants with common chronic conditions, plus functional impairment or depressive symptoms. Women were also slightly overrepresented compared with the general primary care population, in part because they had a higher prevalence of depression (1 of our 6 study diagnoses) relative to men.⁵³ Patients' FFM scores differed from national norms, probably as a result of these inclusion criteria. For instance, Neuroticism scores ranged from 0.20 to 0.36 SDs above national norms, and people with depression tend to measure high in neuroticism.⁵⁴ Likewise, the lack of significant effects for Neuroticism and Extraversion could have been due to the range restriction in the levels of these factors related to study entry criteria, which would tend to underestimate the effects of personality on attrition. These observations suggest the FFM personality profile that we found predicted missing data could be specific to this RCT. Confirmatory analyses from other RCTs are needed, ideally involving a wide array of participant selection characteristics, interventions, and outcomes.

Another limitation was the relatively low rate of missing data from our RCT, which means the analytic models were likely overfitted. Additionally, predictors of missing data might vary as a function of the rate of missing data. For example, higher rates of missing data might be attributable primarily to problems with study implementation and follow-up, rather than patient characteristics. Thus, similar analyses should be repeated using data from RCTs with larger samples and more missing data. A final limitation was that a small number of study participants did not complete the personality inventory at baseline.

Our findings suggest personality-related missing data may affect the results of some RCTs. Clinical researchers must strive to reduce sources of bias to enhance the meaning and relevance of research.¹ The current study provides a starting point for investigations to determine the degree to which personality bias may compromise the validity and applicability of RCTs and observational studies beyond the constraints imposed by their more widely acknowledged design limitations.

To read or post commentaries in response to this article, see it online at <http://www.annfammed.org/cgi/content/full/7/2/148>.

Submitted April 4, 2008; submitted, revised, June 25, 2008; accepted July 29, 2008.

Key words: Bias (epidemiology); data interpretation, statistical; patient compliance; personality; randomized controlled trials

Funding support: Funded in part by Agency for Healthcare Research and Quality grant number R01HS013603, and National Institute of Health grant numbers T32 MH073452 and K24MH072712.

Trial Registration: Homing in on Health: Study of a Home Delivered Chronic Disease Self Management Program. www.ClinicalTrials.gov identifier NCT00263939.

References

- Ioannidis JP. Why most published research findings are false. *PLoS Med.* 2005;2(8):e124.
- Edlund MJ, Wang PS, Berglund PA, Katz SJ, Lin E, Kessler RC. Dropping out of mental health treatment: patterns and predictors among epidemiological survey respondents in the United States and Ontario. *Am J Psychiatry.* 2002;159(5):845-851.
- Honas JJ, Early JL, Frederickson DD, O'Brien MS. Predictors of attrition in a large clinic-based weight-loss program. *Obes Res.* 2003;11(7):888-894.
- Snow WM, Connert JE, Sharma S, Murray RP. Predictors of attendance and dropout at the Lung Health Study 11-year follow-up. *Contemp Clin Trials.* 2007;28(1):25-32.
- Van Beijsterveldt CE, van Boxtel MP, Bosma H, Houx PJ, Buntinx F, Jolles J. Predictors of attrition in a longitudinal cognitive aging study: the Maastricht Aging Study (MAAS). *J Clin Epidemiol.* 2002;55(3):216-223.
- Warden D, Trivedi MH, Wisniewski SR, et al. Predictors of attrition during initial (citalopram) treatment for depression: a STAR*D report. *Am J Psychiatry.* 2007;164(8):1189-1197.
- Benoit SR, Ji M, Fleming R, Philis-Tsimikas A. Predictors of dropouts from a San Diego diabetes program: a case control study. *Prev Chronic Dis.* 2004;1(4):A10.
- Inelmen EM, Toffanello ED, Enzi G, et al. Predictors of drop-out in overweight and obese outpatients. *Int J Obes (Lond).* 2005;29(1):122-128.
- Garcia M, Fernandez E, Schiaffino A, Borrell C, Marti M, Borrás JM. Attrition in a population-based cohort eight years after baseline interview: the Cornella Health Interview Survey Follow-up (CHIS.FU) Study. *Ann Epidemiol.* 2005;15(2):98-104.
- Gallagher EJ, Viscoli CM, Horwitz RI. The relationship of treatment adherence to the risk of death after myocardial infarction in women. *JAMA.* 1993;270(6):742-744.
- Horwitz RI, Horwitz SM. Adherence to treatment and health outcomes. *Arch Intern Med.* 1993;153(16):1863-1868.
- Horwitz RI, Viscoli CM, Berkman L, et al. Treatment adherence and risk of death after a myocardial infarction. *Lancet.* 1990;336(8714):542-545.
- Ross L, Thomsen BL, Boesen EH, Johansen C. In a randomized controlled trial, missing data led to biased results regarding anxiety. *J Clin Epidemiol.* 2004;57(11):1131-1137.
- Tedlow JR, Fava M, Uebelacker LA, Alpert JE, Nierenberg AA, Rosenbaum JF. Are study dropouts different from completers? *Biol Psychiatry.* 1996;40(7):668-670.
- Grilo CM, Money R, Barlow DH, et al. Pretreatment patient factors predicting attrition from a multicenter randomized controlled treatment study for panic disorder. *Compr Psychiatry.* 1998;39(6):323-332.
- Shaw WS, Cronan TA, Christie MD. Predictors of attrition in health intervention research among older subjects with osteoarthritis. *Health Psychol.* 1994;13(5):421-431.
- Goldberg LR. The structure of phenotypic personality traits. *Am Psychol.* 1993;48(1):26-34.
- Costa PT Jr, McCrae RR. Stability and change in personality assessment: the revised NEO Personality Inventory in the year 2000. *J Pers Assess.* 1997;68(1):86-94.
- Marshall GN, Wortman CB, Vickers RR Jr, Kusulas JW, Hervig LK. The Five-Factor Model of personality as a framework for personality-health research. *J Pers Soc Psychol.* 1994;67(2):278-286.
- Bogg T, Roberts BW. Conscientiousness and health-related behaviors: a meta-analysis of the leading behavioral contributors to mortality. *Psychol Bull.* 2004;130(6):887-919.
- Chapman B, Duberstein P, Lyness JM. Personality traits, education, and health-related quality of life among older adult primary care patients. *J Gerontol B Psychol Sci Soc Sci.* 2007;62:P343-P352.
- Chapman BP, Lyness JL, Duberstein PD. Personality and medical illness burden among older adults in primary care. *Psychosom Med.* 2007;69(3):277-282.
- Friedman HS. Long-term relations of personality and health: dynamics, mechanisms, tropisms. *J Pers.* 2000;68(6):1089-1107.
- Roberts BW, Kuncel NR, Shiner R, Caspi A, Goldberg LR. The power of personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspect Psychol Sci.* 2007;2(4):313-345.
- O'Cleirigh C, Ironson G, Weiss A, Costa PT Jr. Conscientiousness predicts disease progression (CD4 number and viral load) in people living with HIV. *Health Psychol.* 2007;26(4):473-480.
- Christensen AJ, Smith TW. Personality and patient adherence: correlates of the five-factor model in renal dialysis. *J Behav Med.* 1995;18(3):305-313.
- Cohen NL, Ross EC, Bagby RM, Farvolden D, Kennedy SH. The 5-factor model of personality and antidepressant medication compliance. *Can J Psychiatry.* 2004;49(2):106-113.
- Stillely CS, Sereika S, Muldoon MF, Ryan CM, Dunbar-Jacob J. Psychological and cognitive function: predictors of adherence with cholesterol lowering treatment. *Ann Behav Med.* 2004;27(2):117-124.
- Hyphantis T, Kaltsouda A, Triantafyllidis J, et al. Personality correlates of adherence to type 2 diabetes regimens. *Int J Psychiatry Med.* 2005;35(1):103-107.
- Galluccio-Richardson RM. *Predicting Adherence to a Weight Loss Regimen Using the NEO Personality Inventory—Revised* [dissertation]. Rutherford, NJ: Fairleigh Dickinson University; 2003.
- Wilson ST. *Personality, Insight, and Compliance: A Longitudinal Study of Treatment Adherence in Outpatients With Schizophrenia and Schizoaffective Disorder* [dissertation]. New York, NY: Columbia University; 2003.
- Lorig KR, Ritter P, Stewart AL, et al. Chronic disease self-management program: 2-year health status and health care utilization outcomes. *Med Care.* 2001;39(11):1217-1223.
- Lorig KR, Sobel DS, Stewart AL, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Med Care.* 1999;37(1):5-14.
- Costa PT, McCrae RR. *Revised NEO Personality Inventory and NEO Five Factor Inventory: Professional Manual.* Odessa, FL: Psychological Assessment; 1992.
- Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum.* 1980;23(2):137-145.
- Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiologic Studies Depression Scale (CES-D). *Arch Intern Med.* 1999;159(15):1701-1704.
- McCrae RR. Openness to experience: expanding the boundaries of factor V. *Eur J Pers.* 1994;8(4):251-272.
- Chapman BP, Duberstein PD, Sorensen S, Lyness JL. Gender differences in five factor model personality traits in an elderly cohort. *Pers Individ Dif.* 2007;43(6):1594-1603.
- Costa PT Jr, Terracciano A, McCrae RR. Gender differences in personality traits across cultures: robust and surprising findings. *J Pers Soc Psychol.* 2001;81(2):322-331.

40. Greenfield S, Kravitz R, Duan N, Kaplan SH. Heterogeneity of treatment effects: implications for guidelines, payment, and quality assessment. *Am J Med.* 2007;120(4):S3-S9.
41. Krause MS, Howard KI. What random assignment does and does not do. *J Clin Psychol.* 2003;59(7):751-766.
42. Little RJA. Modeling the drop-out mechanism in repeated measures studies. *JASA.* 1995;90:1112-1121.
43. de Raad B, Perugini M, eds. *Big Five Assessment.* Ashland, OR: Hogrefe & Huber; 2002.
44. Gallacher JE. Methods of assessing personality for epidemiological study. *J Epidemiol Community Health.* 1992;46(5):465-469.
45. Denissen J, Geenen R, Selfhout M, Van Aken M. Single item Big Five ratings in a social network design. *Eur J Pers.* 2008;22:37-54.
46. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry.* 2002;59(10):877-883.
47. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA.* 2007;298(10):1209-1212.
48. Lonqvist J-E, Paunonen S, Verkasalo M, et al. Personality characteristics of research volunteers. *Eur J Pers.* 2007;21(8):1017-1030.
49. Simpson SH, Eurich DT, Majumdar SR, et al. A meta-analysis of the association between adherence to drug therapy and mortality. *BMJ.* 2006;333(7557):15.
50. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA.* 2001;286(7):821-830.
51. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA.* 2005;294(2):218-228.
52. MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AM. A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. *Health Technol Assess.* 2000;4(34):1-154.
53. Kuehner C. Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. *Acta Psychiatr Scand.* 2003;108(3):163-174.
54. Enns MW, Cox BJ. Personality dimensions and depression: review and commentary. *Can J Psychiatry.* 1997;42(3):274-284.