1. Supplementary Material

This supplementary material provides additional information to accompany the paper entitled 'An imputation strategy for sequential multiple assignment randomized trials" by Shortreed, Laber, et al. Here we provide more details about the data collected in the Clinical Antipsychotic Trials of Intervention and Effectiveness (CATIE) study, as well as more algorithmic details about the imputation process we implemented in the CATIE study. We provide code that can be used to impute missing data from an *artificial* data set designed to mimic the CATIE trial (this artificial data set has fewer variables and observations); this *artificial* data is provided online at {`include link Supplementary Material example data set`}. Additionally, we provide diagnostic plots used to assess the quality of the imputations produced when the imputation strategy outlined in the main body of the paper is implemented in the CATIE data set.

1.1. Description of individuals who completed and did not complete the CATIE study

We describe the CATIE study as well as the missing data patterns that resulted in the data collected from CATIE in Section 2.2 of the main manuscript. In Tables 1 and 2 of this appendix we compare those individuals who completed the CATIE study and those who dropped out of the study early on baseline characteristics, both demographic and disease-status covariates. As noted in the main body of the paper those individuals who complete the CATIE study are very similar on baseline characteristics to those who dropped out of the study early.

1.2. Example of time-ordered data structure

Here we provide an example of the time-order data structure outlined in Section 3.1 of the main body of the manuscript using a small example data set collected from a hypothetical SMART study. Consider a SMART study, with 3 study visits (baseline, one follow-up visit, and an end-of-study visit), two treatment stages, and two time-varying covariates. Gender (G_0), Body mass index (BMI. W_0, and Positive and negative symptom scale (PANSS, P_0) are collected at baseline, $t = 0$ prior to determining assigned treatment (A_1); all of this baseline information is known for all participants. BMI and PANSS were scheduled to be collected on all participants at all follow-up visits. By the first follow-up, $t = 1$, some trial participants had dropped out of the study, and two individuals did not provide adequate information to calculate a PANSS score. At this first follow-up visit, if current treatment was inadequate participants could choose to transition into the second treatment stage; we use C_1 to denote the indicator for if an individual transitioned into the second treatment stage. We use A_2 to denote the assigned second stage treatment for those individuals who transitioned to stage 2, note A_2 is structurally missing for those individuals who chose to continue on their first assigned treatment (i.e. $C_1 = 0$). By the end-of-stage visit ($t = 2$), additional individuals had dropped out of the study and one participant refused to provide weight information to study investigators. Prior to imputing the data, we restructure participant information in the data set so the covariates are in the following order: $G_0, W_0, P_0, A_1, W_1, P_1, C_1, A_2, P_2, W_2$. This example data set restructured in the time-ordered nested format described in Section 3.1 of the main body of the manuscript is given in Table 3.

1.3. Algorithmic details of CATIE imputation procedure

After the CATIE data set was assembled in the time-ordered data structure, we created twenty-five complete data sets with missing information imputed using the procedure described in Section 4 of the main body of the manuscript. We used the `R` software packages `mice` [1] and `pan` [2] for implementing the imputation strategy in the CATIE study. The package `mice` was used for estimating the traditional univariate fully specified conditional models, while the `pan` was used for estimating a longitudinal Bayesian mixed effects model (BMEM) for PANSS. We used the flat (noninformative) priors, independent both across models and across parameters, that are the default prior distributions available in the software packages.

Copyright © 0000 John Wiley & Sons, Ltd.

Table 1. Comparison of individuals who completed the CATIE study versus those who did not complete the CATIE study on baseline demographic and disease-status covariates. Means (standard deviation) reported for continuous variables; percentages for categorical variables.

	Completed the CATIE study.	
Baseline covariates	Yes (n=755)	No (n=705)
Age	39.6 (11.3)	41.6 (10.8)
Male	73.6	74.3
Race		
Black	40.0	29.9
White	54.4	65.7
Marital status: married	11.9	10.9
Patient education		
College graduate	7.5	9.4
Community college or technical school degree	6.4	6.5
Some college, did not graduate	23.7	25.1
GED/High school diploma	35.2	35.0
Did not complete high school	27.2	24.0
Employment status		
Did not work	85.4	84.2
Full time	7.0	6.6
Part time	7.6	9.3
Site type		
Private practice	12.6	13.2
State mental health	17.7	19.9
University clinic	24.9	21.0
VA	14.2	12.8
Combination	30.6	33.2
Tardive Dyskenisia at Study entry	15.8	15.9
Hospitalized in 3 mos. prior to enrollment in CATIE	32.6	22.1
Treatment upon CATIE entry		
Newly treated	31.3	25.2
Same medication as that taken prior to enrollment	12.6	16.5
Switched medication from that taken prior to enrollment	56.2	58.3
Years on prescription medication prior to CATIE	14.3 (10.3)	14.5 (11.1)
Stage 1 assigned treatment		
Olanzapine	19.7	26.5
Perphenazine	17.1	18.7
Quetiapine	24.6	21.4
Risperidone	24.4	22.3
Ziprasidone	14.2	11.1

Included in the supplementary material is an artificial CATIE data set, `CATIE_artificial.Rdata`, with missing information. We provide example code that implements the proposed imputation strategy on this artificial data; for more details on the exact function calls used, please see this `R` code: `Impute_artificial_CATIE.r`. We provide here, the general overview on implementing the imputation process described in the main body of the manuscript. This code is a modification of the code that we used to perform the CATIE imputations, updated to run on `R` version 3.0 using version 2.18 of `mice` and version 0.9 of `pan`. Additionally we made changes to reflect the artificial CATIE data set as it is smaller (fewer variables and time points).

We used the "norm" argument in `mice` for all continuous variables; this specification assumes that the conditional distribution of the variable is assumed Gaussian. No transformations were needed for continuous-valued variables in the artificial CATIE data set. Although a continuous variable, reflecting movement disorder symptoms is only applicable

 Copyright © 0000 John Wiley & Sons, Ltd.

Table 2. Cont'd. Comparison of individuals who completed the CATIE study versus those who did not complete the CATIE study on baseline demographic and disease-status covariates. Means (standard deviation) reported for continuous variables; percentages for categorical variables.

	Completed the CATIE study.	
Baseline covariates	Yes (n=755)	No (n=705)
PANSS (total score)	74.3 (18.12)	76.0 (17.2)
Mental health short form score	40.8 (11.6)	41.1 (11.7)
Physical health short form score	48.1 (10.3)	48.3 (10.0)
BMI	29.6 (7.1)	30.0 (7.0)
Quality of life (total score)	2.7 (1.0)	2.8 (1.1)
Calgary depression score	4.7 (4.4)	4.4 (4.4)
Clinical Global Impression Score		
Not ill or minimally ill	4.2	7.8
Mildly ill	23.3	19.7
Moderately ill	48.9	46.2
Markedly ill	19.4	21.3
Severely or very severely ill	4.2	5.0
Illicit drug use (hair test)		
No Drugs	55.3	67.3
At least 1 illicit drug found	44.7	32.7
Illegal drug use (clinician-reported) (CS14)		
Abstinent	69.6	82.3
Use without impairment	17.5	11.5
Abuse	9.3	5.2
Dependence	3.6	1.0
Alcohol use (clinician-reported)		
Abstinent	62.2	67.7
Use without impairment	27.7	27.4
Abuse	6.9	3.3
Dependence	3.2	1.7
Simpson-Angus EPS Scale - Presence of symptoms	44.6	54.2
Simpson-Angus EPS Scale - Symptom severity score*	0.2 (0.3)	0.2 (0.3)
Barnes Akathisia Scale - Presence of symptoms	40.0	39.3
Barnes Akathisia Scale - Symptom severity score*	1.0 (1.6)	1.1 (1.6)
Abnormal Involuntary Movement scale - Presence of symptoms	37.3	39.3
Abnormal Involuntary Movement scale - Symptom severity score*	1.7 (3.2)	1.6 (2.9)

to those individual who are exhibiting such symptoms, was semi continuous. Thus for the imputation process this is broken up into two variables; 1) a binary variable indicating the presence of symptoms and 2) the continuous-valued score on the movement disorder symptom scale for those exhibiting symptoms. Conditional models for binary and categorical variables were estimated using the "logreg" and "polyreg" arguments in `mice` respectively [1]. The imputation for PANSS was performed with `pan` using a random intercept. We used the default priors, which includes an invert Wishart prior on the variance compnents of the random intercept and the error terms and flat improper priors for the fixed effect terms.

1.4. Assessing the CATIE Imputations

Imputation methodology rests on the untestable assumption that missing data values are missing at random (MAR) and thus can be predicted using imputation models estimated from the observed data. While this assumption can never truly be validated, diagnostics are usually performed to compare the imputed values with the observed values [3, 4]. If the distributions are very different, this could be a sign that the missing at random assumption has been violated, or that the imputation models are not a good fit to the observed data. Here we briefly use visual diagnostics to assess the validity of the CATIE imputations. We first evaluate the singly imputed transition times, the multiply imputed stage 2

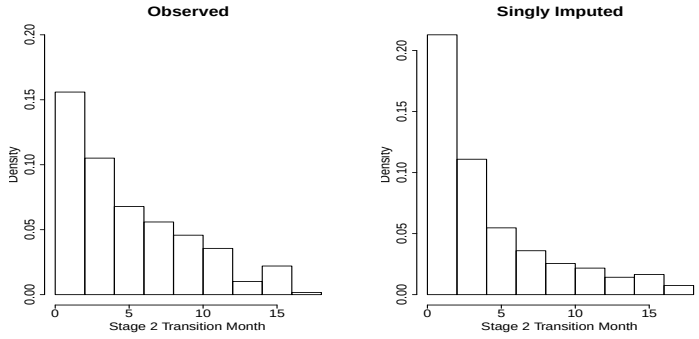
Copyright © 0000 John Wiley & Sons, Ltd.

Table 3. Example data set in the time-ordered data structure. NA refers to structural missingness, while blank cells represent missing information.

G_0	W_0	P_0	A_14	W_1	P_1	C_1	A_2	P_2	W_2
Female	31.8	103	Perphenazine	23.4	77	SWITCHED	Ziprasidone	86	26.9
Male	29.4	108	Risperidone	18.2	102	STAYED	NA	88	19
Male	32.6	63	Olanzapine	35.2		STAYED	NA	85	38.2
Female	30.4	102	Quetiapine	34.6	99	SWITCHED	Olanzapine	77	
Female	23.2	100	Risperidone	20.8	96	SWITCHED	Olanzapine	71	31.6
Male	38.1	86	Perphenazine	28.7	75	STAYED	NA		
Female	31.1	80	Risperidone	22.8	89	SWITCHED	Clozapine		
Female	31.6	71	Olanzapine	21.1	74	STAYED	NA		
Male	25.1	77	Perphenazine	19.7		STAYED	NA		
Male	37.9	64	Olanzapine			STAYED	NA		
Female	28.7	91	Risperidone						
Male	37.8	65	Perphenazine						

treatment assignments, and then use three variables (PANSS, BMI, and medication adherence) as examples in assessing the imputations of time-varying variables collected in the CATIE trial.

We assess the imputations for missing stage transition times in the CATIE data set by comparing the distribution of the observed transition times to the singly imputed transition times. Figure 1 presents two histograms to help evaluate the differences in these distribution. The month of entry into stage 2 was observed for 539 CATIE participants. Figure 1:(a) presents a histogram of the observed month of entry into stage 2 for these CATIE participants. Additionally, 509 individuals had their stage 2 transition month singly imputed with month of drop out. Figure 1:(b) shows the histogram of these singly imputed stage 2 transition times is similar to the distribution of the observed transition times. To assess the multiply imputed treatment assignments, Table 4 presents the number and percentage of individuals observed to follow each treatment sequence in the complete case sample used in the analysis presented in Section 5 of the main body of the manuscript and averaged over 25 imputed data sets. The overall proportion of those individuals who were observed to remain on their stage 1 treatment compared to those who opted to transition in to stage 2 are similar between the complete case and imputed data sets. Additionally, the proportion of individuals who were assigned each of the stage 2 treatments is also similar between the complete case and imputed data sets. These results are not surprising, because our imputed stage 2 treatment assignment used the randomization probabilities from the CATIE protocol.



(a) (b)

Figure 1. Histograms for (a) observed month of entry into stage 2 of CATIE, (b) singly imputed stage 2 transition time for those CATIE participants who dropped out of CATIE before transitioning into stage 2.

Copyright © 0000 John Wiley & Sons, Ltd.

Table 4. Number and percentage of individuals observed to follow each treatment sequence. The columns entitled "Complete Case" reports these numbers for the sample used in the analyses reported in Section 5 of the main body of the manuscript The columns entitled "Multiple Imputation" presents the mean number of people observed to follow each treatment sequence averaged over 25 imputations.

	Complete Case N=321		Multiple Imputation N=1014	
Treatment Regimes	n	%	n	%
Olanzapine until end-of-study	83	25.9	123.0	12.1
Olanzapine followed by quetiapine	11	3.4	63.3	6.2
Olanzapine followed by risperidone	14	4.4	63.8	6.3
Olanzapine followed by clozapine or ziprasidone	20	6.2	85.9	10.6
Quetiapine until end-of-study	35	10.9	62.0	6.1
Quetiapine followed by olanzapine	17	5.3	83.4	8.2
Quetiapine followed by risperidone	16	5.0	84.1	8.3
Quetiapine followed by clozapine or ziprasidone	20	6.2	107.5	10.6
Risperidone until end-of-study	60	18.7	91.0	9.0
Risperidone followed by olanzapine	17	5.3	76.5	7.5
Risperidone followed by quetiapine	14	4.4	77.8	7.7
Risperidone followed by clozapine or ziprasidone	14	4.4	95.7	9.4

In order to assess the imputation models for PANSS, BMI and treatment adherence, we present QQ-plots of the observed values and the imputed values over 25 imputation. Figure 2 presents QQ-plots of the observed PANSS scores at every month of scheduled observation, as well as those collected at the end-of-stage visit, versus the multiply imputed scores. The similarity between the distribution of the multiply imputed PANSS scores and the distribution of the observed PANSS scores supports the validity of our imputation models for PANSS. The QQ-plots for the observed BMI values versus the imputed BMI values are presented in Figure 3. While the distribution of the multiply imputed BMI values has a few more individuals in the lower end of the distribution than the distribution of the observed population, overall the distribution of BMI values are similar, suggesting that our imputation models are adequately modeling BMI values.

Figure 4 shows the QQ-plots of observed values versus imputed values for monthly medication adherence, as measured by the proportion of pills taken since the last visit. In this case, we notice non-trivial differences between the observed and imputed distributions. In particular, many more people have lower adherence in the imputed data than in the observed. While this certainly raises a red flag, it does not necessarily mean that the imputations are not valid [4]. Recall that CATIE participants were allowed to discontinue treatment, or drop out of the study, for any reason including adherence. In fact, this aspect of the CATIE protocol resulted in many non-adherent patients switching into the next treatment stage, or dropping out of the study, rather than remaining non-adherent to their current treatment This resulted in very high recorded adherence rates throughout the CATIE study, with the median recorded adherence at each month ranging from 75% to 100%. CATIE participants with adherence below 50% at a monthly visit compared to those who had adherence higher than 50% had a log odds ratio of dropping out of the study before the next monthly visit of 1.82, with a standard error 0.16. This high rate of drop out among non-adherent participants resulted in a semi-continuous distribution for treatment adherence, with many participants having recorded adherence of 100%, a few at 0% adherence, and some with varying levels of partial compliance. Research involving patients with schizophrenia has shown that current treatment adherence to antipsychotic medication is the strongest predictor of future treatment adherence [5]. Thus, it is reasonable that there are more individuals with low adherence in the imputed data, as the observed population has many non-adherent participants removed due to study drop out.

Copyright © 0000 John Wiley & Sons, Ltd.

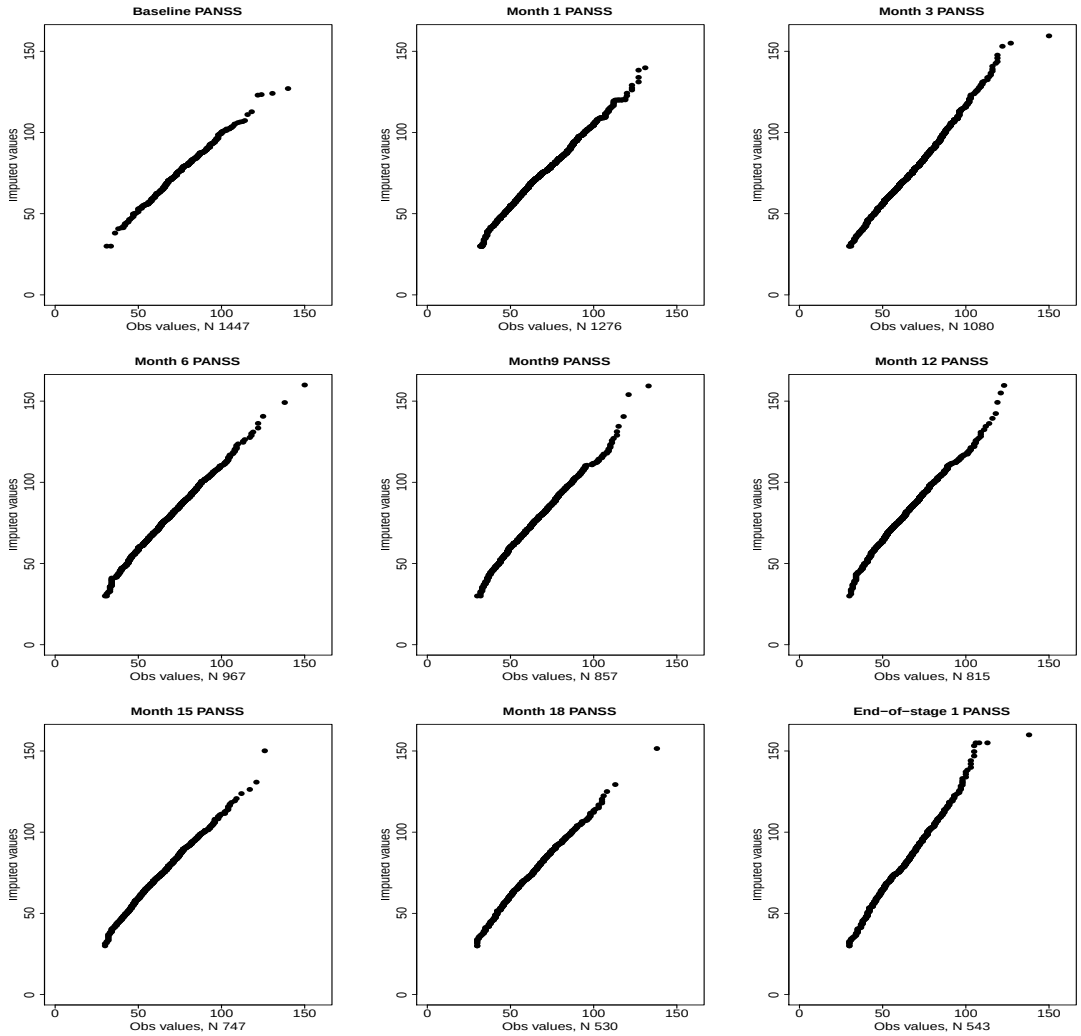


Figure 2. QQ-plots of imputed versus observed PANSS scores measured at all months in which PANSS was scheduled to be collected and all end-of-stage PANSS scores. The missing data distribution contains the imputed values from twenty-five imputations (and none of the observed values).

References

1. van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software* 2011; **forthcoming**.
2. Schafer JL. *Multiple imputation for multivariate panel or clustered data* 2009. URL `http://cran.r-project.org/web/packages/pan/pan.pdf`, r package version 0.2-6.
3. Abayomi K, Gelman A, Levy M. Diagnostics for multivariate imputations. *Journal of the Royal Statistical Society, Series C* 2008; **57**:273–291.
4. Stuart EA, Azur M, Frangakis C, Leaf P. Multiple imputation with large data sets: A case study of the children's mental health initiative. *American Journal of Epidemiology* 2009; **169**(9):1133–1139.
5. Novick D, Haro JM, Suarez D, Perez V, Dittmann R, PM H. Predictors and clinical consequences of non-adherence with antipsychotic medication in the outpatient treatment of schizophrenia. *Psychiatry Research* 2010; **173**(2):109–113.

Copyright © 0000 John Wiley & Sons, Ltd.

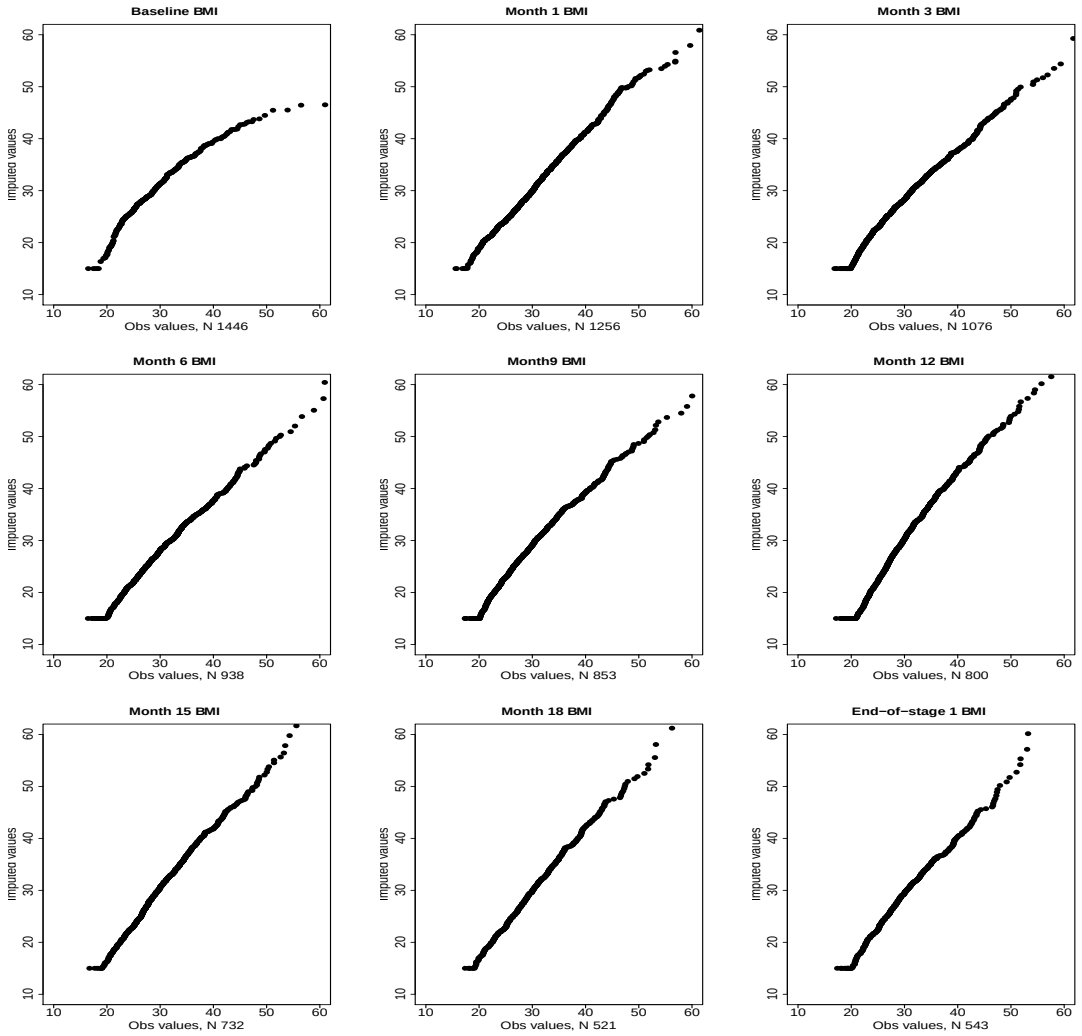


Figure 3. QQ-plots of imputed versus observed BMI values measured at all months in which BMI was scheduled to be collected as well as all end-of-stage BMI values. The missing data distribution contains the imputed values from twenty-five imputations (and none of the observed values).

Copyright © 0000 John Wiley & Sons, Ltd.

Prepared using simauth.cls

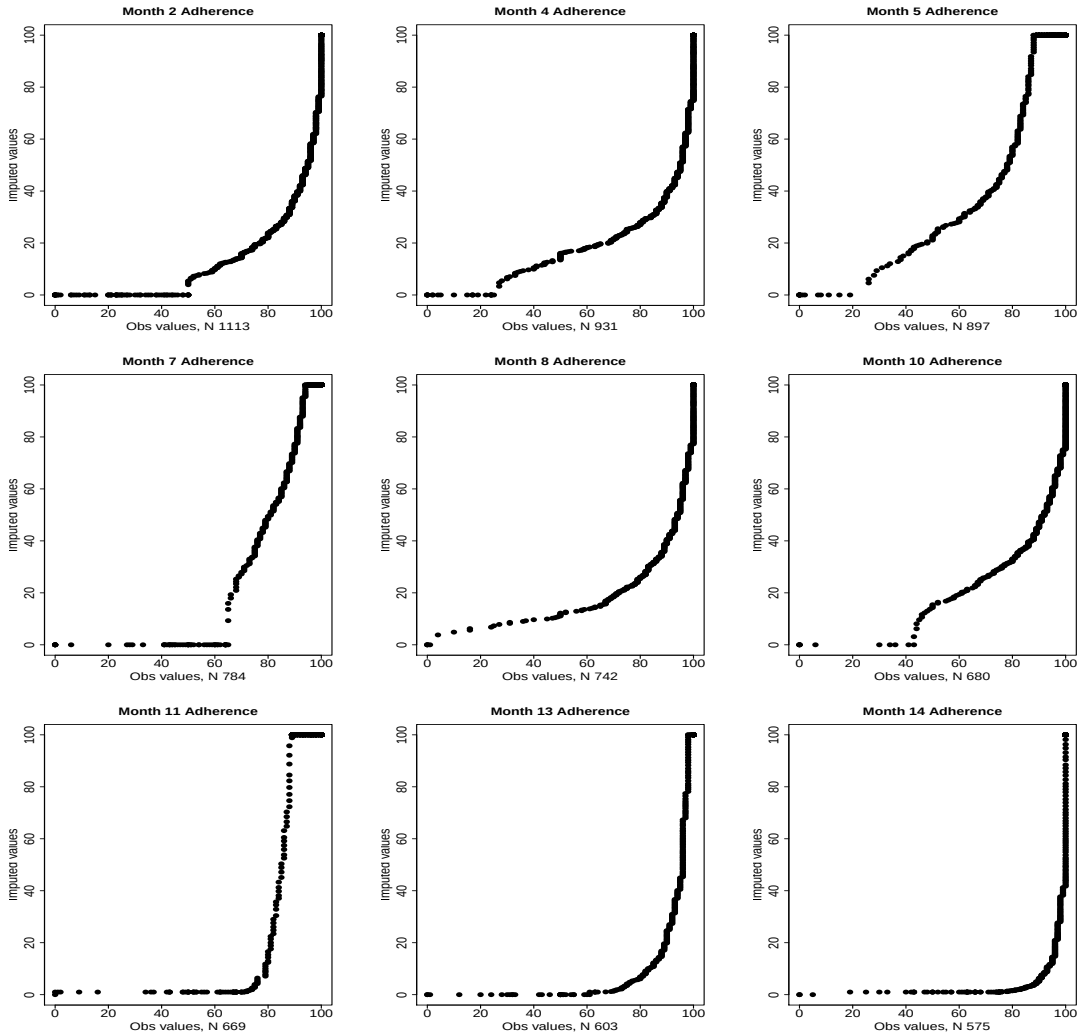


Figure 4. QQ-plots of imputed versus observed treatment adherence as measured by pill counts for selected months. We can see that the imputation models for adherence are not very accurate. Instead of using the continuous adherence measure as a predictor, we use a categorical variable indicating no adherence, partial adherence, or complete adherence as a predictor in imputation models of all other variables. The missing data distribution contains the imputed values from twenty-five imputations (and none of the observed values).

Copyright © 0000 John Wiley & Sons, Ltd.