Covariate Adjustment in Analyses of Time-to-Event End Points

Lian Mao,[1] Ibrahim Turkoz,[1] Larry Alphs[2]

[1]Janssen Research & Development, LLC, Titusville, NJ, USA; [2]Janssen Scientific Affairs, LLC, Titusville, NJ, USA

INTRODUCTION

- In long-term studies of subjects with severe mental illness, a significant percentage frequently discontinue study participation prematurely without reaching a study end point. The problem is compounded if dropout patterns differ for treatment versus control subjects
- Missing data make it challenging to estimate the true treatment differences that would exist at study termination in a counterfactual setting in which all subjects are followed to the predetermined end point
- Baseline covariates are routinely collected in randomized clinical trials
  - An unadjusted analysis provides valid treatment comparisons in randomized studies, but covariate-adjusted analyses are often implemented to increase statistical power or to offset the influence of random imbalances between treatment groups for the covariates that may have strong relationships to the primary outcome[1]
  - Therefore, adjustment for important prognostic covariates may better define differential treatment responses in patient subpopulations, thus improving the precision of estimation and statistical inferences or generalizability of the results
- In patients with serious mental illness such as schizophrenia, numerous prognostic factors have been reported in the literature as potential predictors of relapse or treatment discontinuation[2,3]
- Analyses ignoring heterogeneity in patient subsamples represent only a crude estimate of treatment effect
  - Unadjusted methodologies are particularly problematic for assessing the robustness of active-control trials because the relative efficacy is usually much smaller than the absolute efficacy from placebo-controlled studies
- This poster illustrates the importance of adjustment for important prognostic factors in the sensitivity analysis for the time-to-event end point, using data from a randomized, active-controlled study in patients with schizophrenia and a history of incarceration (NCT01157351)[4]

METHODS

- An important assumption in most time-to-event analyses is that the censoring mechanism is noninformative or ignorable. This means that the censoring of an observation does not provide any information regarding the prospect of event time of that particular subject beyond the censoring time. However, in many studies a noninformative censoring assumption may not hold
- The proposed sensitivity analyses address the implications of departures from noninformative censoring by imputing event times for dropouts under various plausible scenarios
- This approach enables assessment of the robustness of the results from primary analyses when censoring of follow-up times for patients with early discontinuation has occurred
- To objectively identify important covariates post hoc, we applied the principled approaches proposed by Tsiatis et al[5] to model the covariate-outcome relationship separately within each treatment group, using Cox regression with a stepwise model-selection procedure
  - The following prognostic baseline variables were analyzed: age, sex, race, duration of illness, baseline Personal and Social Performance (PSP) scale score, baseline Clinical Global Impression of Severity (CGI-S) scale score, multiple (≥2) prior incarcerations (yes/no), history of substance abuse (yes/no), prior health insurance coverage (yes/no), and being randomly assigned to the same antipsychotic (AP) medication received before randomization (yes/no)
  - Predictive covariates that met the model-selection criteria from these independent models were retained in the final Cox regression model
- To assess the impact of covariate adjustment, we compared the results from the covariate-adjusted model against the results without covariate adjustment in the context of sensitivity analysis to address the issue of informative dropouts. Event times for dropouts were imputed using 3 algorithms:
  1. Model 1 (naive imputation): It was assumed that a fraction of dropouts without an event would experience an event at the time of withdrawal. To implement this algorithm, x% (10%–90%) of the dropouts were randomly sampled, and their event times and censoring times were imputed. These imputed event times were combined with the existing event times, and the combined data were analyzed using Cox regression with and without covariate adjustment. This process was repeated 1000 times to obtain 1000 estimates of the hazard ratio (HR) and its corresponding 95% confidence interval (CI)
  2. Model 2: Event times for dropouts were simulated using an exponential distribution. To implement this imputation, the exponential distribution that would yield x% (10%–90%) failure rate for the dropouts was first derived. Using the derived exponential distribution and inversion method, a random event time was simulated for each dropout. The simulated event times with the existing event times were combined to form a new data set, and the analysis was rerun using Cox regression with and without covariate adjustment. The simulation process was repeated 1000 times for each exponential distribution
  3. Model 3: Event times for dropouts were simulated using a Weibull distribution separately for subjects lost to withdrawal of consent or to follow-up. The Weibull distribution is a more general parametric distribution with 2 parameters; thus, it is better able to fit a wide range of survival data. An empirical Weibull distribution was derived from the existing data for different subgroups of subjects. Using these empirical Weibull distributions and an inversion method, event times were simulated. The simulated event times were combined with the existing event times to form a new data set for analysis with and without covariate adjustment; 1000 simulations were performed. To assess the robustness of the primary study findings, the simulations were repeated by inflating the observed HR for the treatment arm

RESULTS

- For the reference study, the primary interest was to compare the efficacy of an experimental treatment with an active control on time to treatment failure (TTF). The treatment failure time was determined by a blinded event-monitoring board, independent of the study sponsor
- The disposition of subjects is summarized in **Table 1**
  - A total of 444 intent-to-treat (ITT) subjects were included in the primary analysis
  - Of the 444 ITT subjects, 181 (40.8%) completed the study and 124 (27.9%) discontinued the study early but had experienced the primary study event. A total of 139 (31.3%) subjects discontinued from the study without a primary event, of whom 81 (35.8%) received experimental treatment and 58 (26.6%) received an active control
- Subject demographics and baseline characteristics, symptom scores, and psychiatric history are summarized in **Table 2**
- A total of 90 subjects (39.8%) in the treatment group and 117 subjects (53.7%) in the control group had a treatment failure. Treatment was superior to control in delaying time to first treatment failure (HR, 1.43; 95% CI, 1.09–1.88; *P* = 0.011) (**Table 3**)
- The model-selection procedure yielded a final Cox regression model with the following covariates identified as important predictors of TTF, in addition to study treatment: gender, multiple prior incarcerations, history of substance abuse (yes/no), prior health insurance coverage (yes/no), and being randomized to the same AP medication received before randomization (**Table 4**)

Table 1. Subject Disposition[4]

	Treatment	Control	Total
Subjects in the ITT analysis set[a]	226	218	444
Subjects who completed study, n (%)	93 (41.2)	88 (40.4)	181 (40.8)
With event	38 (16.8)	45 (20.6)	83 (18.7)
Without event	55 (24.3)	43 (19.7)	98 (22.1)
Subjects who discontinued study early, n (%)	133 (58.8)	130 (59.6)	263 (59.2)
With event	52 (23.0)	72 (33.0)	124 (27.9)
Without event	81 (35.8)	58 (26.6)	139 (31.3)

[a]ITT subjects are those who were randomly assigned to treatment.
Note: All percentages are based on the ITT population.

Table 2. Demographic and Baseline Characteristics (ITT analysis set)[4]

	Treatment n = 226	Control n = 218	Total N = 444
Age, years, mean (SD)	37.7 (10.6)	38.6 (10.4)	38.1 (10.5)
Male, n (%)	193 (85.4)	190 (87.2)	383 (86.3)
Race, n (%)	n = 226	n = 217	n = 443
White	73 (32.3)	74 (34.1)	147 (33.2)
Black/African-American	145 (64.2)	130 (59.9)	275 (62.1)
Other	8 (3.5)	13 (6.0)	21 (4.7)
CGI-S total, mean (SD)	n = 226 3.8 (0.8)	n = 217 3.9 (0.7)	n = 443 3.8 (0.8)
PSP total, mean (SD)	n = 226 54.8 (12.8)	n = 215 55.0 (12.7)	n = 441 54.9 (12.8)
Age at first psychiatric diagnosis, years, mean (SD)	n = 226 22.0 (9.6)	n = 216 22.1 (9.8)	n = 442 22.1 (9.7)
Duration of illness ≤5 years, n (%)	n = 226 42 (18.6)	n = 216 35 (16.2)	n = 442 77 (17.4)
Prior psychiatric diagnosis, n (%)	n = 220	n = 216	n = 436
Anxiety	23 (10.5)	30 (13.9)	53 (12.2)
Schizophrenia	197 (89.5)	198 (91.7)	395 (90.6)
Bipolar disorder	42 (19.1)	48 (22.2)	90 (20.6)
Depression	70 (31.8)	72 (33.3)	142 (32.6)
Schizoaffective disorder	26 (11.8)	29 (13.4)	55 (12.6)
Other	64 (29.1)	62 (28.7)	126 (28.9)
Number of psychiatric hospitalizations within ≤12 months, n (%)	n = 176	n = 173	n = 349
0	107 (60.8)	88 (50.9)	195 (55.9)
≥1	69 (39.2)	85 (49.1)	154 (44.1)
History of substance abuse, yes, n (%)	208 (92.0)	202 (92.7)	410 (92.3)
Multiple prior incarcerations, yes, n (%)	83 (36.7)	71 (32.6)	154 (34.7)
Randomized to the prior AP medication treatment, yes, n (%)	0	27 (12.4)	27 (6.1)
Insurance coverage prior to study entry, yes, n (%)	126 (55.8)	111 (50.9)	237 (53.4)

Percentages are based on the number of subjects in the ITT population with a nonmissing value for the parameter.

Table 3. Primary Outcome of the Study: First Treatment Failure[4]

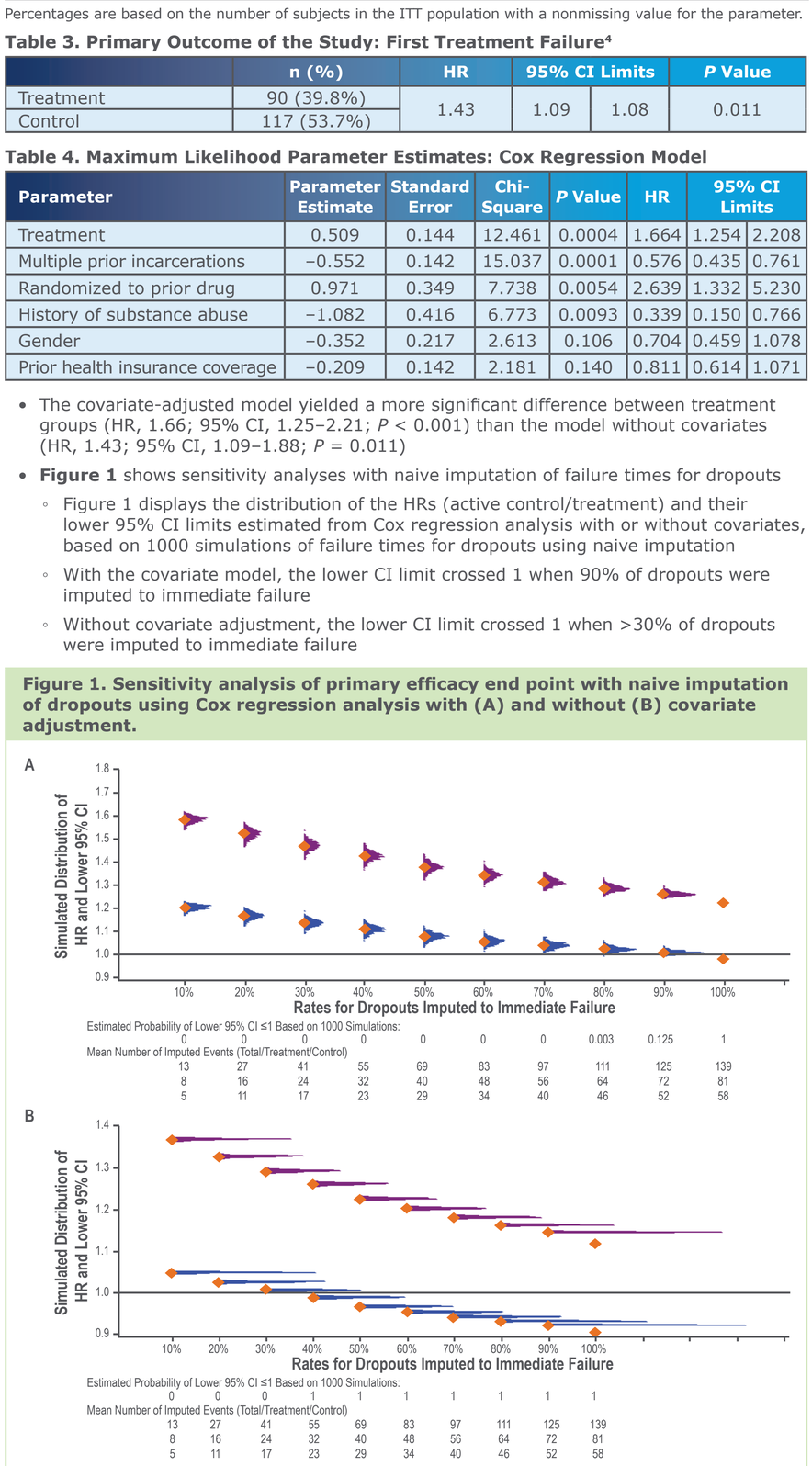
	n (%)	HR	95% CI Limits		*P* Value
Treatment	90 (39.8%)	1.43	1.09	1.08	0.011
Control	117 (53.7%)				

Table 4. Maximum Likelihood Parameter Estimates: Cox Regression Model

Parameter	Parameter Estimate	Standard Error	Chi-Square	*P* Value	HR	95% CI Limits	
Treatment	0.509	0.144	12.461	0.0004	1.664	1.254	2.208
Multiple prior incarcerations	−0.552	0.142	15.037	0.0001	0.576	0.435	0.761
Randomized to prior drug	0.971	0.349	7.738	0.0054	2.639	1.332	5.230
History of substance abuse	−1.082	0.416	6.773	0.0093	0.339	0.150	0.766
Gender	−0.352	0.217	2.613	0.106	0.704	0.459	1.078
Prior health insurance coverage	−0.209	0.142	2.181	0.140	0.811	0.614	1.071

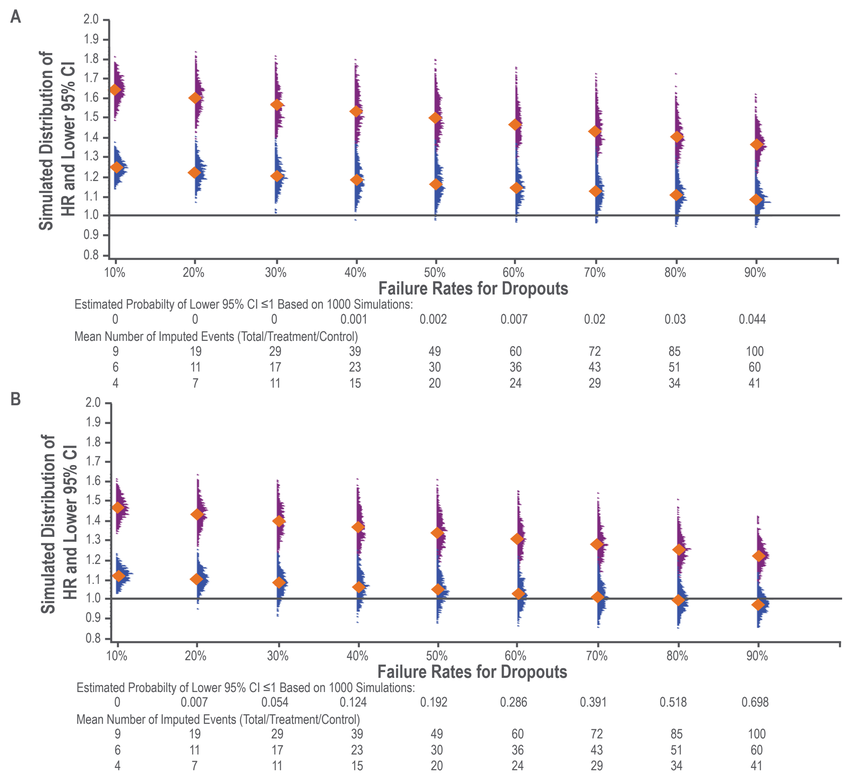
- The covariate-adjusted model yielded a more significant difference between treatment groups (HR, 1.66; 95% CI, 1.25–2.21; *P* < 0.001) than the model without covariates (HR, 1.43; 95% CI, 1.09–1.88; *P* = 0.011)
- **Figure 1** shows sensitivity analyses with naive imputation of failure times for dropouts
  - Figure 1 displays the distribution of the HRs (active control/treatment) and their lower 95% CI limits estimated from Cox regression analysis with or without covariates, based on 1000 simulations of failure times for dropouts using naive imputation
  - With the covariate model, the lower CI limit crossed 1 when 90% of dropouts were imputed to immediate failure
  - Without covariate adjustment, the lower CI limit crossed 1 when >30% of dropouts were imputed to immediate failure

Figure 1. Sensitivity analysis of primary efficacy end point with naive imputation of dropouts using Cox regression analysis with (A) and without (B) covariate adjustment.



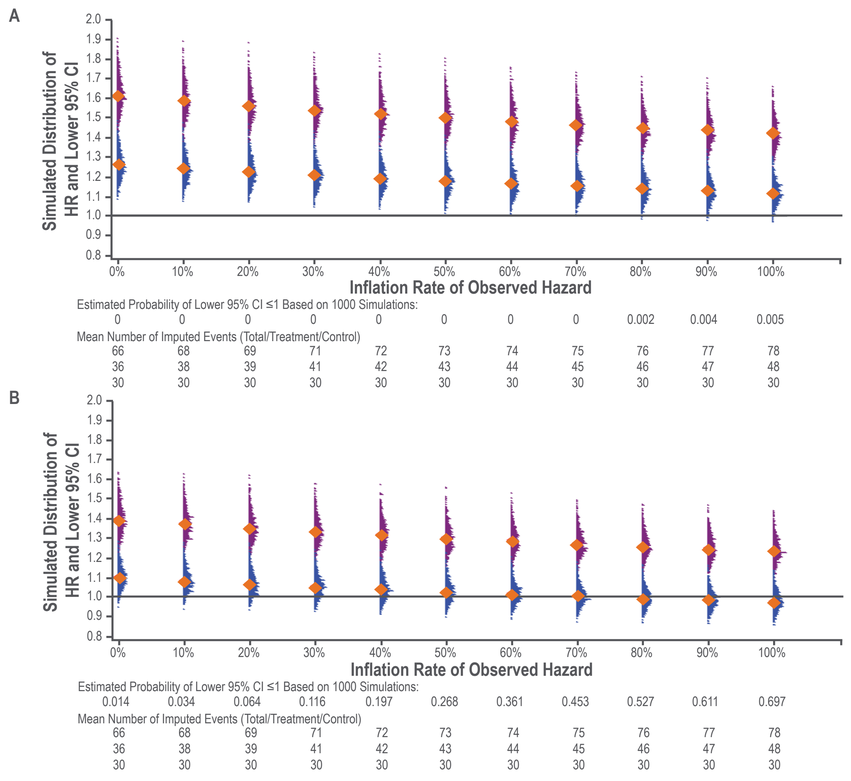
- **Figure 2** presents the sensitivity analysis with exponential simulation of failure time for dropouts
  - Figure 2 displays the distribution of the HRs (active control/treatment) and their lower 95% CI limits estimated from Cox regression analysis with covariates, based on 1000 simulations of failure time for dropouts using the exponential distribution
  - With the covariate-adjusted model, the lower CI limit crossed 1 when assuming >90% of dropouts would fail after the dropout time
  - Without covariate adjustment, the lower CI limit crossed 1 when assuming 30% of dropouts would fail after the dropout time

Figure 2. Sensitivity analysis of primary efficacy end point with exponential imputation of dropouts using Cox regression analysis with (A) and without (B) covariate adjustment.



- **Figure 3** presents the results using Weibull simulation of failure times for dropouts
  - Figure 3 displays the distribution of HRs (control/treatment) and their lower 95% CI limits estimated from Cox regression analysis with or without covariate adjustment, based on 1000 simulations of failure time for dropouts using Weibull distribution
  - With the covariate-adjusted model, the lower CI limit did not cross 1 even when the HR for the treatment arm was inflated by 100%
  - Without covariate adjustment, the lower CI limit crossed 1 when the HR for the treatment arm was inflated by ~20%

Figure 3. Sensitivity analysis of primary efficacy end point with Weibull imputation of dropouts using Cox regression with (A) and without (B) covariate adjustment.



DISCUSSION/LIMITATIONS

- The comparison of results using data from the referenced active-control trial supports the value of adjusting for covariates. For this data set, with adjustment for covariates, treatment estimates from the Cox model strengthened the primary result and gave greater confidence in the generalizability of the estimation
- Although the analysis with no adjustment for prognostic covariates was statistically significant, the significance could disappear when a moderate number of dropouts were imputed as treatment failure
- Although there are obvious benefits of including prognostic covariates in an analysis, it is important to ensure that covariates identified post hoc are based on systematic, unbiased methods and are clinically meaningful and clinically identifiable
- The appropriate application of covariate-adjusted Cox regression models depends on several assumptions, such as correct model specification and proportional hazards for each variable in the model
  - When the proportional hazards assumption is not satisfied and the Cox model is adjusted for covariates that are related to the outcome, the type I error is inflated[6]

CONCLUSIONS

- **Covariate adjustment is an important method to quantify treatment effect more precisely and to increase confidence in the generalizability of the results. It is especially relevant when conducting analyses of randomized, active-control clinical trials**
- **Post hoc identification of covariates should be based on objective criteria in order to minimize biases**

REFERENCES

1. Tangen CM, Koch GG. *Stat Med.* 2000;19(8):1039–1058. **2.** Hall DL et al. *J Am Acad Psychiatry Law.* 2012;40(2):221–231. **3.** Essock SM et al. *Am J Psychiatry.* 2006;163(12):2090–2095. **4.** Alphs L et al. *J Clin Psychiatry.* 2015;76(5):554–561. **5.** Tsiatis AA et al. *Stat Med.* 2008;27(23):4658–4677. **6.** Jiang H et al. *Stat Med.* 2008;27(28):5850–5860.

Disclosures

L. Mao and I. Turkoz are employees of Janssen Research & Development, LLC, and Johnson & Johnson stockholders. L. Alphs is an employee of Janssen Scientific Affairs, LLC, and a Johnson & Johnson stockholder.

Acknowledgments

The authors thank Matthew Grzywacz, PhD, for his writing and editorial assistance, which was supported by Janssen Scientific Affairs, LLC.

Supported by Janssen Scientific Affairs, LLC

Poster presented at the International Society for CNS Clinical Trials and Methodology 2015 Autumn Conference; August 27–29, 2015; Amsterdam, The Netherlands

INTRODUCTION

- In long-term studies of subjects with severe mental illness, a significant percentage frequently discontinue study participation prematurely without reaching a study end point. The problem is compounded if dropout patterns differ for treatment versus control subjects
- Missing data make it challenging to estimate the true treatment differences that would exist at study termination in a counterfactual setting in which all subjects are followed to the predetermined end point
- Baseline covariates are routinely collected in randomized clinical trials
  - An unadjusted analysis provides valid treatment comparisons in randomized studies, but covariate-adjusted analyses are often implemented to increase statistical power or to offset the influence of random imbalances between treatment groups for the covariates that may have strong relationships to the primary outcome[1]
  - Therefore, adjustment for important prognostic covariates may better define differential treatment responses in patient subpopulations, thus improving the precision of estimation and statistical inferences or generalizability of the results
- In patients with serious mental illness such as schizophrenia, numerous prognostic factors have been reported in the literature as potential predictors of relapse or treatment discontinuation[2,3]
- Analyses ignoring heterogeneity in patient subsamples represent only a crude estimate of treatment effect
  - Unadjusted methodologies are particularly problematic for assessing the robustness of active-control trials because the relative efficacy is usually much smaller than the absolute efficacy from placebo-controlled studies
- This poster illustrates the importance of adjustment for important prognostic factors in the sensitivity analysis for the time-to-event end point, using data from a randomized, active-controlled study in patients with schizophrenia and a history of incarceration (NCT01157351)[4]

METHODS

- An important assumption in most time-to-event analyses is that the censoring mechanism is noninformative or ignorable. This means that the censoring of an observation does not provide any information regarding the prospect of event time of that particular subject beyond the censoring time. However, in many studies a noninformative censoring assumption may not hold
- The proposed sensitivity analyses address the implications of departures from noninformative censoring by imputing event times for dropouts under various plausible scenarios
- This approach enables assessment of the robustness of the results from primary analyses when censoring of follow-up times for patients with early discontinuation has occurred
- To objectively identify important covariates post hoc, we applied the principled approaches proposed by Tsiatis et al[5] to model the covariate-outcome relationship separately within each treatment group, using Cox regression with a stepwise model-selection procedure
  - The following prognostic baseline variables were analyzed: age, sex, race, duration of illness, baseline Personal and Social Performance (PSP) scale score, baseline Clinical Global Impression of Severity (CGI-S) scale score, multiple (≥2) prior incarcerations (yes/no), history of substance abuse (yes/no), prior health insurance coverage (yes/no), and being randomly assigned to the same antipsychotic (AP) medication received before randomization (yes/no)
  - Predictive covariates that met the model-selection criteria from these independent models were retained in the final Cox regression model
- To assess the impact of covariate adjustment, we compared the results from the covariate-adjusted model against the results without covariate adjustment in the context of sensitivity analysis to address the issue of informative dropouts. Event times for dropouts were imputed using 3 algorithms:
  1. Model 1 (naive imputation): It was assumed that a fraction of dropouts without an event would experience an event at the time of withdrawal. To implement this algorithm, x% (10%–90%) of the dropouts were randomly sampled, and their event times and censoring times were imputed. These imputed event times were combined with the existing event times, and the combined data were analyzed using Cox regression with and without covariate adjustment. This process was repeated 1000 times to obtain 1000 estimates of the hazard ratio (HR) and its corresponding 95% confidence interval (CI)
  2. Model 2: Event times for dropouts were simulated using an exponential distribution. To implement this imputation, the exponential distribution that would yield x% (10%–90%) failure rate for the dropouts was first derived. Using the derived exponential distribution and inversion method, a random event time was simulated for each dropout. The simulated event times with the existing event times were combined to form a new data set, and the analysis was rerun using Cox regression with and without covariate adjustment. The simulation process was repeated 1000 times for each exponential distribution
  3. Model 3: Event times for dropouts were simulated using a Weibull distribution separately for subjects lost to withdrawal of consent or to follow-up. The Weibull distribution is a more general parametric distribution with 2 parameters; thus, it is better able to fit a wide range of survival data. An empirical Weibull distribution was derived from the existing data for different subgroups of subjects. Using these empirical Weibull distributions and an inversion method, event times were simulated. The simulated event times were combined with the existing event times to form a new data set for analysis with and without covariate adjustment; 1000 simulations were performed. To assess the robustness of the primary study findings, the simulations were repeated by inflating the observed HR for the treatment arm

RESULTS

- For the reference study, the primary interest was to compare the efficacy of an experimental treatment with an active control on time to treatment failure (TTF). The treatment failure time was determined by a blinded event-monitoring board, independent of the study sponsor
- The disposition of subjects is summarized in **Table 1**
  - A total of 444 intent-to-treat (ITT) subjects were included in the primary analysis
  - Of the 444 ITT subjects, 181 (40.8%) completed the study and 124 (27.9%) discontinued the study early but had experienced the primary study event. A total of 139 (31.3%) subjects discontinued from the study without a primary event, of whom 81 (35.8%) received experimental treatment and 58 (26.6%) received an active control
- Subject demographics and baseline characteristics, symptom scores, and psychiatric history are summarized in **Table 2**
- A total of 90 subjects (39.8%) in the treatment group and 117 subjects (53.7%) in the control group had a treatment failure. Treatment was superior to control in delaying time to first treatment failure (HR, 1.43; 95% CI, 1.09–1.88; *P* = 0.011) (**Table 3**)
- The model-selection procedure yielded a final Cox regression model with the following covariates identified as important predictors of TTF, in addition to study treatment: gender, multiple prior incarcerations, history of substance abuse (yes/no), prior health insurance coverage (yes/no), and being randomized to the same AP medication received before randomization (**Table 4**)

Table 1. Subject Disposition[4]

	Treatment	Control	Total
Subjects in the ITT analysis set[a]	226	218	444
Subjects who completed study, n (%)	93 (41.2)	88 (40.4)	181 (40.8)
With event	38 (16.8)	45 (20.6)	83 (18.7)
Without event	55 (24.3)	43 (19.7)	98 (22.1)
Subjects who discontinued study early, n (%)	133 (58.8)	130 (59.6)	263 (59.2)
With event	52 (23.0)	72 (33.0)	124 (27.9)
Without event	81 (35.8)	58 (26.6)	139 (31.3)

[a]ITT subjects are those who were randomly assigned to treatment.
Note: All percentages are based on the ITT population.

Table 2. Demographic and Baseline Characteristics (ITT analysis set)[4]

	Treatment n = 226	Control n = 218	Total N = 444
Age, years, mean (SD)	37.7 (10.6)	38.6 (10.4)	38.1 (10.5)
Male, n (%)	193 (85.4)	190 (87.2)	383 (86.3)
Race, n (%)	n = 226	n = 217	n = 443
White	73 (32.3)	74 (34.1)	147 (33.2)
Black/African-American	145 (64.2)	130 (59.9)	275 (62.1)
Other	8 (3.5)	13 (6.0)	21 (4.7)
CGI-S total, mean (SD)	n = 226 3.8 (0.8)	n = 217 3.9 (0.7)	n = 443 3.8 (0.8)
PSP total, mean (SD)	n = 226 54.8 (12.8)	n = 215 55.0 (12.7)	n = 441 54.9 (12.8)
Age at first psychiatric diagnosis, years, mean (SD)	n = 226 22.0 (9.6)	n = 216 22.1 (9.8)	n = 442 22.1 (9.7)
Duration of illness ≤5 years, n (%)	n = 226 42 (18.6)	n = 216 35 (16.2)	n = 442 77 (17.4)
Prior psychiatric diagnosis, n (%)	n = 220	n = 216	n = 436
Anxiety	23 (10.5)	30 (13.9)	53 (12.2)
Schizophrenia	197 (89.5)	198 (91.7)	395 (90.6)
Bipolar disorder	42 (19.1)	48 (22.2)	90 (20.6)
Depression	70 (31.8)	72 (33.3)	142 (32.6)
Schizoaffective disorder	26 (11.8)	29 (13.4)	55 (12.6)
Other	64 (29.1)	62 (28.7)	126 (28.9)
Number of psychiatric hospitalizations within ≤12 months, n (%)	n = 176	n = 173	n = 349
0	107 (60.8)	88 (50.9)	195 (55.9)
≥1	69 (39.2)	85 (49.1)	154 (44.1)
History of substance abuse, yes, n (%)	208 (92.0)	202 (92.7)	410 (92.3)
Multiple prior incarcerations, yes, n (%)	83 (36.7)	71 (32.6)	154 (34.7)
Randomized to the prior AP medication treatment, yes, n (%)	0	27 (12.4)	27 (6.1)
Insurance coverage prior to study entry, yes, n (%)	126 (55.8)	111 (50.9)	237 (53.4)

Percentages are based on the number of subjects in the ITT population with a nonmissing value for the parameter.

Table 3. Primary Outcome of the Study: First Treatment Failure[4]

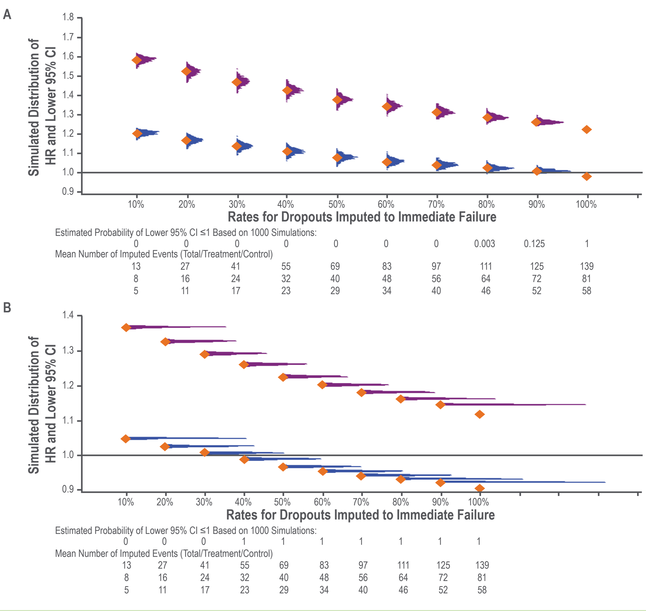
	n (%)	HR	95% CI Limits		*P* Value
Treatment	90 (39.8%)	1.43	1.09	1.08	0.011
Control	117 (53.7%)				

Table 4. Maximum Likelihood Parameter Estimates: Cox Regression Model

Parameter	Parameter Estimate	Standard Error	Chi-Square	*P* Value	HR	95% CI Limits	
Treatment	0.509	0.144	12.461	0.0004	1.664	1.254	2.208
Multiple prior incarcerations	–0.552	0.142	15.037	0.0001	0.576	0.435	0.761
Randomized to prior drug	0.971	0.349	7.738	0.0054	2.639	1.332	5.230
History of substance abuse	–1.082	0.416	6.773	0.0093	0.339	0.150	0.766
Gender	–0.352	0.217	2.613	0.106	0.704	0.459	1.078
Prior health insurance coverage	–0.209	0.142	2.181	0.140	0.811	0.614	1.071

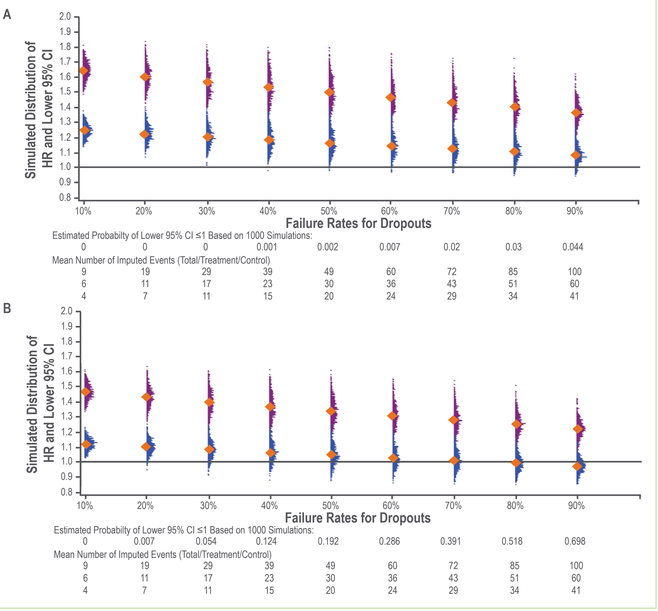
- The covariate-adjusted model yielded a more significant difference between treatment groups (HR, 1.66; 95% CI, 1.25–2.21; *P* < 0.001) than the model without covariates (HR, 1.43; 95% CI, 1.09–1.88; *P* = 0.011)
- **Figure 1** shows sensitivity analyses with naive imputation of failure times for dropouts
  - Figure 1 displays the distribution of the HRs (active control/treatment) and their lower 95% CI limits estimated from Cox regression analysis with or without covariates, based on 1000 simulations of failure times for dropouts using naive imputation
  - With the covariate model, the lower CI limit crossed 1 when 90% of dropouts were imputed to immediate failure
  - Without covariate adjustment, the lower CI limit crossed 1 when >30% of dropouts were imputed to immediate failure

Figure 1. Sensitivity analysis of primary efficacy end point with naive imputation of dropouts using Cox regression analysis with (A) and without (B) covariate adjustment.



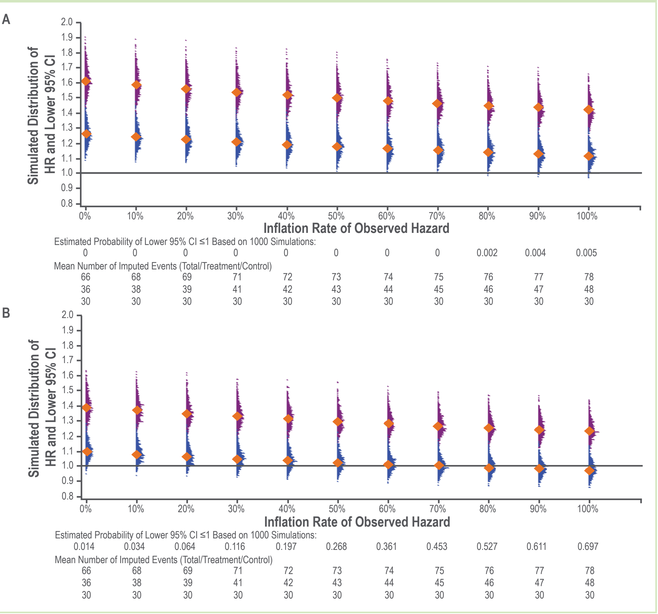
- **Figure 2** presents the sensitivity analysis with exponential simulation of failure time for dropouts
  - Figure 2 displays the distribution of the HRs (active control/treatment) and their lower 95% CI limits estimated from Cox regression analysis with covariates, based on 1000 simulations of failure time for dropouts using the exponential distribution
  - With the covariate-adjusted model, the lower CI limit crossed 1 when assuming >90% of dropouts would fail after the dropout time
  - Without covariate adjustment, the lower CI limit crossed 1 when assuming 30% of dropouts would fail after the dropout time

Figure 2. Sensitivity analysis of primary efficacy end point with exponential imputation of dropouts using Cox regression analysis with (A) and without (B) covariate adjustment.



- **Figure 3** presents the results using Weibull simulation of failure times for dropouts
  - Figure 3 displays the distribution of HRs (control/treatment) and their lower 95% CI limits estimated from Cox regression analysis with or without covariate adjustment, based on 1000 simulations of failure time for dropouts using Weibull distribution
  - With the covariate-adjusted model, the lower CI limit did not cross 1 even when the HR for the treatment arm was inflated by 100%
  - Without covariate adjustment, the lower CI limit crossed 1 when the HR for the treatment arm was inflated by ~20%

Figure 3. Sensitivity analysis of primary efficacy end point with Weibull imputation of dropouts using Cox regression with (A) and without (B) covariate adjustment.



DISCUSSION/LIMITATIONS

- The comparison of results using data from the referenced active-control trial supports the value of adjusting for covariates. For this data set, with adjustment for covariates, treatment estimates from the Cox model strengthened the primary result and gave greater confidence in the generalizability of the estimation
- Although the analysis with no adjustment for prognostic covariates was statistically significant, the significance could disappear when a moderate number of dropouts were imputed as treatment failure
- Although there are obvious benefits of including prognostic covariates in an analysis, it is important to ensure that covariates identified post hoc are based on systematic, unbiased methods and are clinically meaningful and clinically identifiable
- The appropriate application of covariate-adjusted Cox regression models depends on several assumptions, such as correct model specification and proportional hazards for each variable in the model
  - When the proportional hazards assumption is not satisfied and the Cox model is adjusted for covariates that are related to the outcome, the type I error is inflated[6]

CONCLUSIONS

- **Covariate adjustment is an important method to quantify treatment effect more precisely and to increase confidence in the generalizability of the results. It is especially relevant when conducting analyses of randomized, active-control clinical trials**
- **Post hoc identification of covariates should be based on objective criteria in order to minimize biases**

REFERENCES

1. Tangen CM, Koch GG. *Stat Med.* 2000;19(8):1039–1058. **2.** Hall DL et al. *J Am Acad Psychiatry Law.* 2012;40(2):221–231. **3.** Essock SM et al. *Am J Psychiatry.* 2006;163(12):2090–2095. **4.** Alphs L et al. *J Clin Psychiatry.* 2015;76(5):554–561. **5.** Tsiatis AA et al. *Stat Med.* 2008;27(23):4658–4677. **6.** Jiang H et al. *Stat Med.* 2008;27(28):5850–5860.

Disclosures

L. Mao and I. Turkoz are employees of Janssen Research & Development, LLC, and Johnson & Johnson stockholders. L. Alphs is an employee of Janssen Scientific Affairs, LLC, and a Johnson & Johnson stockholder.

Acknowledgments

The authors thank Matthew Grzywacz, PhD, for his writing and editorial assistance, which was supported by Janssen Scientific Affairs, LLC.