Roflumilast or Azithromycin to Prevent COPD Exacerbations (RELIANCE) Trial

Statistical Analysis Plan

03 June 2026

NCT Number: NCT04069312

RofLumilast or Azithromycin to preveNt COPD Exacerbations (RELIANCE)



Statistical Analysis Plan Version 4.0

03 June 2026

Funded by PCORI PCS-1504-30430

Table of Contents

Table of Contents 2

1. List of Abbreviations 3

2. Overview of the Study Design 3

3. Schematic of Design 4

4. Primary Outcome 5

5. Analysis Samples 5

5.1 As randomized (Intention-to-Treat) Sample 6

5.2 Per-protocol Samples 6

6. Randomization 6

7. General Sample Size Considerations 6

8. Interim Efficacy Analysis 8

9. Analysis 8

9.1 Primary outcome 8

9.2 Secondary outcomes 9

9.3. Compliance assessment 11

9.4. Linkage with Medicare Data 12

9.5. Missing data 12

10. References 15

11. Appendix A 15

1. List of Abbreviations

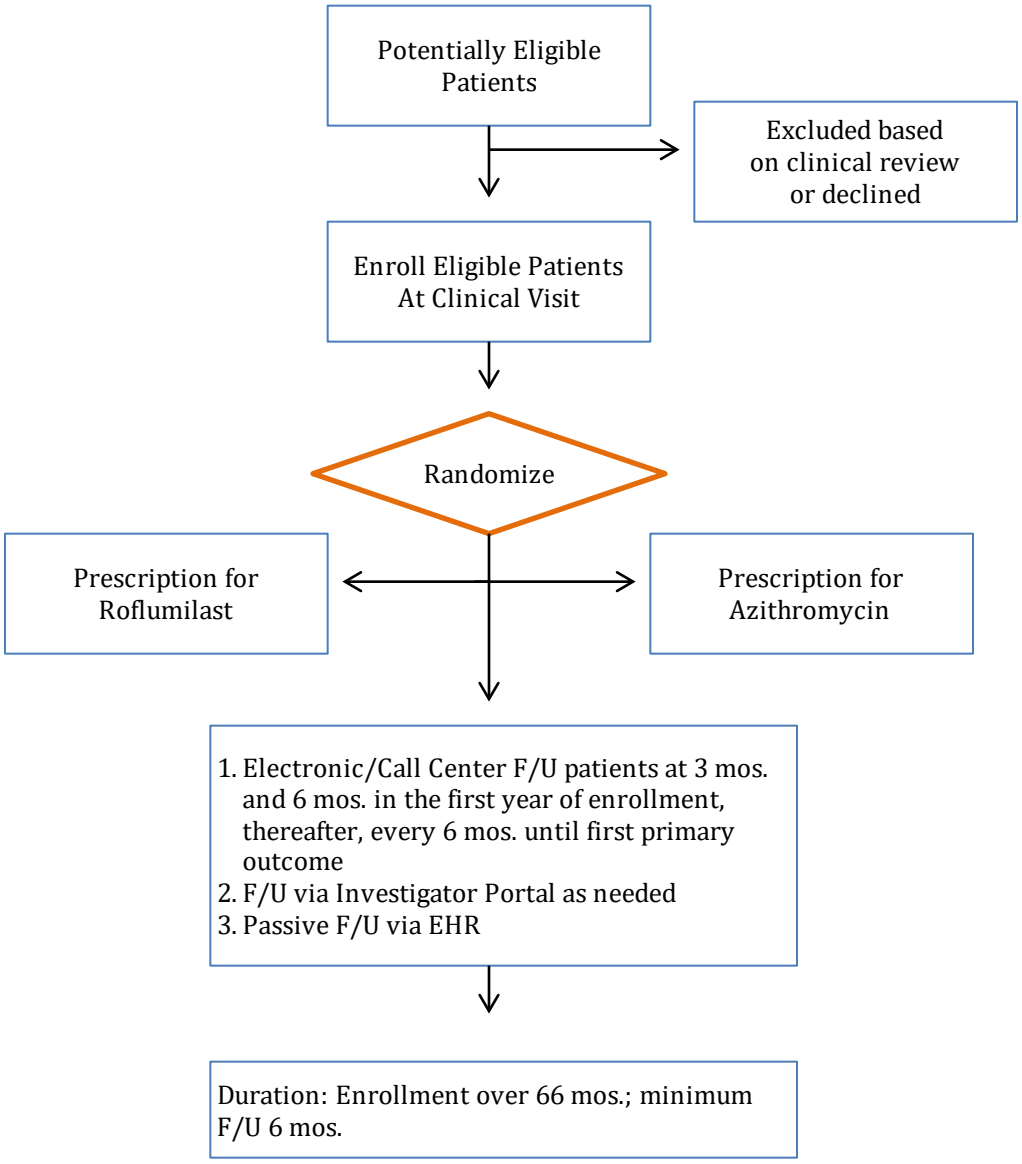
AE	Adverse Event
COPD	Chronic obstructive pulmonary disease
DCC	Data Coordinating Center
DSMB	Data Safety and Monitoring Board
ITT	Intention-to-Treat
PP	Per protocol
RELIANCE	Roflumilast or Azithromycin to prevent COPD Exacerbations
SAE	Serious Adverse Event

2. Overview of the Study Design

RELIANCE is a U.S.-based pragmatic clinical trial funded by the Patient-Centered Outcomes Research Institute (PCORI) to compare long-term use of roflumilast vs. azithromycin in up to 1,250 patients. It is intended to support hospital efforts to reduce the risk of all-cause hospitalization and reduce pre-mature deaths in individuals with chronic obstructive pulmonary disease (COPD). The COPD Patient Powered Research Network (PPRN) and affiliated investigators will conduct the trial in sites in the U.S.

Both roflumilast and azithromycin have been shown to reduce the risk of COPD exacerbations compared to placebo. However, there has not been a head-to-head comparison of the two medications. So, the relative harms and benefits of the two medications are unknown. Eligible patients will be randomized (1:1) to receive either a prescription for roflumilast or a prescription for azithromycin, and will be followed for at least 6 and up to 72 months, depending on the date of enrollment. The primary endpoint is the combined outcome of all-cause hospitalization or death; the secondary endpoints include premature treatment discontinuation, patient-reported adverse effects, and physical, social, and emotional health. Patients will provide data via an online patient portal or via a call center. Baseline and outcome data will also be collected from site medical records.

3. Schematic of Design



4. Primary Outcome

The primary outcome is all-cause hospitalization or death. Individuals who do not experience an event prior to study close-out or are lost to follow-up will be censored at the date of the last known contact.

5. Analysis Samples

RELIANCE is designed to comparative the effectiveness as opposed to pharmacologic efficacy of roflumilast and azithromycin. We will collect data about patient and clinician practices in the study, including the initial and subsequent dosing regimens, adherence, treatment discontinuations, and cross-overs. These practices plus the pharmacologic efficacy of roflumilast and azithromycin in individual patients will contribute to the observed comparative effectiveness of the two treatment options. The primary analysis will therefore be "as randomized" (commonly referred to as intention-to-treat, ITT), secondary sensitive analyses will include per protocol analyses, i.e., the subgroup of patients who adhere to their assigned treatment.

There is considerable debate about the optimal approach to analysis of non-inferiority trials, as randomized or per protocol. The major concern about the as randomized analyses for non-inferiority hypotheses is that if a significant number of patients do not adhere to their assigned treatment, the test treatments may falsely appear to have similar effects thereby introducing a bias favoring a conclusion of non-inferiority. Per protocol analyses, which limits the analyses to the subgroup of patients who adhere to their assigned treatment, are recommended for testing non-inferiority hypotheses primarily to avoid this potential bias. However, per protocol analyses of non-inferiority trials have all the same flaws as they do in superiority trials (1). For example, if patients discontinue study assigned treatments because of side effects, the potential decreased effectiveness of the treatment is correctly attributed to the assigned treatment in an as randomized analysis but would not be in a per protocol analysis. The preferred approach is to maintain as randomized (ITT) as the primary analysis and to institute performance measures that ensure the integrity of the trial and provide essential information for interpretation of the results (2). Per protocol analysis will be a key secondary analysis and will be compared with the as randomized analysis. Data

on protocol performance such as treatment adherence, losses to follow-up and missing data will be used to help understand and interpret differences, if any, in the results of the two methods.

5.1 As randomized (Intention-to-Treat) Sample

The as randomized, or intention-to-treat (ITT), sample includes all patients who are randomized, regardless of the amount of treatment received. Patients will be analyzed according to their assigned treatment. The primary analysis will use the as randomized sample.

5.2 Per-protocol Samples

A variety of per-protocol, or as treated, samples based upon medication adherence will be constructed for use in secondary analyses of the primary outcome (see Section 9.3).

6. Randomization

Patients will be randomized (1:1) to receive a prescription for one of the two treatments (roflumilast or azithromycin). Treatment assignments will be stratified by site and smoking status (current vs former). A permuted block design with multiple block sizes will be used with a separate assignment schedule generated for each stratum (i.e. site and smoking status combination). The order of treatment assignments will be concealed with the next assignment revealed only after a patient has been randomized. Once a patient is assigned to receive a treatment, it will be known to the clinician, site coordinator, and patient (unmasked).

7. General Sample Size Considerations

RELIANCE is a randomized, pragmatic, comparative effectiveness trial designed to evaluate whether azithromycin is non-inferior to roflumilast in high-risk patients with COPD for prevention of hospitalizations or death. Our choice of a non-inferiority margin of 1.2 was based upon recently completed clinical trial in COPD (3), as well as opinions solicited from clinicians and patients. In a survey of 48 clinicians, the minimum clinically important difference was identified to be 1.2 or higher by 75% of

the survey participants and the margin of 1.2 was the most commonly selected (33%). An online survey of 50 patients conducted by the COPD PPRN identified a range of different effect sizes, which seemed to vary based on the history of previous COPD exacerbations, and hospitalizations, and tolerance for adverse effects. A margin of 1.2 is within the range of interest expressed by patients.

The original sample size was 3200 which would have provided 92% power to establish non-inferiority assuming a recruitment period of 30 months, minimum follow-up of 6 months, 1-sided type 1 error of 0.025, non-inferiority margin of 1.2, true hazard ratio of 1, cumulative proportion with an event in 1 year of 0.3, and 5% loss to follow-up per year. The sample size was revised based upon the slow recruitment rate. At the time of the revised calculation, the observed pooled proportion with an event among all participants was 0.6, nearly double the original estimate, and the recruitment period was extended to 66 months. The planned recruitment period was assumed to be 36 months. This allowed for 30 additional months of recruitment from the time of the calculation and condensed the recruitment period that had already occurred into 6 months to maintain a constant recruitment rate across the whole period. This approach was conservative since the 300 participants recruited in the first 30 months would have additional follow-up time in which to have an event. The cumulative loss to follow-up was decreased slightly to 4.7% at 1 year and 13.4% at 3 years based upon available data. All other assumptions were unchanged. A revised sample size of 1,250 was estimated to provide 82% power with a 1- sided type I error of 0.025 to establish non-inferiority.

Post-hoc analyses of trials suggest that while roflumilast appears to benefit both past and current smokers, the benefits of chronic azithromycin may be limited to past smokers (4,5). The sample size of 1,250 participants provide 91% power to detect a heterogeneous effect based upon smoking status (HR = 1 for roflumilast vs HR = 0.65 for azithromycin for past vs current smokers) assuming a reference cumulative incidence of 60% at one year for current smokers in both treatment groups, equal recruitment to all four subgroups, 5% loss to follow-up per year (14.3% cumulatively) and a two-sided type I error rate. These analyses were computed using PASS software and SWOG Statistical tools, respectively (6,7).

8. Interim Efficacy Analysis

A single, formal, interim efficacy analysis will be performed once 50% of the information has been collected (i.e. 467 of the expected 934 events of hospitalization or death have been observed). The trial will be stopped early only if one of the two treatments is demonstrated to be superior, i.e. it will not stop if the non-inferiority criteria are met. Based upon an O'Brien-Flemming type boundary, the type I error rate would be 0.00305 for the interim analysis and 0.04695 at the final analysis to maintain a global 0.05 error rate over the course of the trial for the superiority comparison. If the trial does not stop early for superiority, then at the end of the trial the two arms will first be compared for non-inferiority at the 0.023475 level. With a type I error rate of 0.023475, we have 82% power for the non-inferiority comparison (a loss of < 1% from that described in Section 7). The final superiority assessment will take place at the 0.04695 level only if non-inferiority is established. Additional ad hoc efficacy analyses may be requested by the data safety and monitoring board as needed during the course of the trial.

9. Analysis

9.1 Primary outcome

The primary outcome is the time to first all-cause hospitalization or death. The event time will be defined as the earliest occurrence of either hospitalization or death. Participants will continue to be followed after the occurrence of a hospitalization event in order to complete scheduled 6-month follow-up contact. For time-to-event analyses, individuals without an event will be censored at the date of last contact. For participants who experience an event, the event time will be used in the analysis, regardless of continued follow-up after the event. A Cox proportional hazards model will be used to assess whether azithromycin is non-inferior to roflumilast. Non-inferiority will be achieved if the upper boundary of the confidence interval for the hazard ratio (azithromycin/roflumilast) is less than 1.2, the non-inferiority margin. The level of the confidence interval will be adjusted from 95% to account for type I error spending during interim efficacy analyses and will match the type I error used in the final analysis. If we achieve non-inferiority, then an additional test for

superiority will be performed. The primary analysis will adjust for the stratification variables smoking status (current vs former) through the inclusion of a fixed effect and clinic with a random effect, respectively. A key secondary hypothesis is that the treatment effect will differ for current and former smokers. A test of interaction will be used to assess the heterogeneity of treatment effects (HTE) based upon smoking status (current vs former). Secondary analyses will be performed both adjusting for and investigating the heterogeneity of the treatment effect (HTE) for key baseline characteristics including but not limited 3 pre-specified comparisons: a history of asthma (yes vs no), gender (male vs female) and BMI (obese vs not obese). The Cox proportional hazards model described above will be the primary analysis for the primary outcome. All additional analyses, including alternative modeling strategies or sensitivity analyses, will be considered secondary and will be used to assess the robustness of the primary findings.

9.2 Secondary outcomes

All analyses of secondary outcomes are distinct from the primary outcome analysis and will not replace or modify the primary inference based on the Cox proportional hazards model for the primary endpoint.

Analysis of the secondary composite outcome (moderate COPD exacerbation, all-cause hospitalization, or all-cause death) will employ similar time-to-first-event methods as described for the primary outcome, including Cox proportional hazards models for time-to-first-event analyses. Although participants may continue follow-up after experiencing an event to complete scheduled study assessments, all time-to-event analyses will be based on the time to first occurrence of the event of interest. Moderate COPD exacerbation is defined as treatment with antibiotics or systemic corticosteroids for respiratory exacerbations not associated with hospitalization. Individual components of the primary and secondary composite outcomes (all cause death, all-cause hospitalization, and COPD exacerbation) will also be analyzed using appropriate time-to-event and competing risk methods. Since follow-up was halted after all-cause death or all-cause hospitalization, those two events will be treated as a completing risk for the other outcomes. Cumulative incidence functions (CIFs) will be used to estimate event probabilities in the presence of competing risks, and Fine-Gray subdistribution hazard

regression models will be used to compare treatment groups for selected non-fatal outcomes where death precludes subsequent event occurrence. Cause-specific hazard models may also be explored in sensitivity analyses where appropriate.

For some study outcomes collected through participant self-report during scheduled follow-up contacts, the exact timing of events may not always be completely ascertained and may instead only be known to have occurred within an interval between completed follow-up assessments. Interval-censoring techniques may be employed accordingly for outcomes such as moderate COPD exacerbation, treatment discontinuation, and other secondary outcomes derived from intermittent follow-up assessments when exact event dates are not completely ascertained. Medication adherence and other longitudinal measures collected repeatedly over follow-up will be analyzed using generalized linear and mixed effects modeling approaches, as appropriate. Recurrent exacerbation and hospitalization outcomes may additionally be analyzed using negative binomial regression models, as appropriate.

Additional secondary outcomes include patient-reported outcomes assessed using NIH PROMIS instruments (physical function, sleep disturbance, fatigue, anxiety, and depression); adverse events including hearing decrement, diarrhea, nausea, suicidal ideation, and microbiology cultures identifying macrolide-resistant organisms; crossover to alternate study treatment; treatment discontinuation; out-of-pocket medication costs; and weight. Analysis of secondary outcomes by treatment group will employ logistic regression models for binary outcomes, linear regression models for continuous outcomes, negative binomial regression models for count and recurrent event outcomes, and Cox proportional hazards models for time-to-event outcomes, as appropriate. Mixed effects models will be used to analyze outcomes with repeated measurements over time. The primary time-point of interest is 6 months, the minimum follow-up for all participants, since follow-up varied for participants with participants exiting the study at the later of 6 months or the time of the primary outcome.

Adjusted models including stratification variables (clinic and smoking status) and key baseline covariates (e.g., history of asthma, gender, BMI, medication use) will be used for secondary analyses; unadjusted analyses will be presented descriptively. Summary statistics, model estimates with confidence intervals, and p-values will be used to

present findings. A large number of secondary analyses, including heterogeneity of treatment effect (HTE) analyses, are planned and caution is necessary in interpreting the results, especially since it is generally not possible to pre-specify all potential analyses. The number of secondary analyses explored will be clearly specified, and the expected number of false positives will be considered in the interpretation of findings. Several methods of adjusting p-values for multiple comparisons exist; however, no clear consensus regarding the most appropriate method is available. As recommended by Wang and Lagakos (8), reporting will focus on point estimates, confidence intervals, and p-values, and findings from secondary analyses will be considered hypothesis-generating.

Performance monitoring is a key component of clinical trials. Our performance monitoring will include site-specific and overall metrics on patient screening, enrollment, retention, missing data, and adherence to treatment. These metrics will be used to assess the internal validity of the trial. External validity will be assessed by evaluation of the characteristics of the screened and enrolled population, the ratio of enrolled to screened participants, and comparison of these characteristics to the overall population with COPD. We will rely on published literature and available Medicare-linked data to evaluate the characteristics of U.S. patients with COPD.

9.3. Compliance assessment

Compliance will primarily be assessed based upon patient-reported medication usage collected through scheduled follow-up assessments via the Participant Portal and Call Center. The degree of compliance is an important factor for interpretation of the as-randomized (intention-to-treat, ITT) analyses as well as secondary per-protocol and sensitivity analyses. However, as noted by Wittes (9), analyses relying on post-randomization variables are likely to be biased and/or require complex mathematical models with numerous assumptions. Individuals will be classified according to a number of different methods (e.g. any compliance [at least one usage], minimum compliance [above a certain threshold]) for sensitivity analyses. A variety of causal inference techniques, including inverse probability weighting methods similar to those described by Murray and Hernan (10), maybe considered to assess the potential impact of adherence patterns on treatment comparisons. In addition, medication adherence collected

longitudinally over follow-up may be analyzed using generalized linear and mixed effects modeling approaches, as appropriate.

9.4. Linkage with Medicare Data

Medicare administrative claims data linked through collaborations with the FDA Sentinel Program (see Appendix A) may be used to provide supplemental information regarding hospitalization and mortality outcomes among the subset of RELIANCE participants enrolled in Medicare. Medicare linkage and analyses will be conducted within the CMS Virtual Research Data Center (VRDC) environment by approved collaborating analysts. Only aggregate summary results approved through CMS disclosure review procedures may be exported from the VRDC environment.

Consistent with the revised scope of the CMS linkage activities, Medicare data will primarily be used to provide supportive descriptive assessments regarding completeness of ascertainment for selected primary outcomes in the linked Medicare subset. Analyses involving direct comparison, reconciliation, or pooled integration of participant-level trial and Medicare claims data will not be performed outside of the CMS-approved analytic environment.

Given the limited scope, timing, and availability of linked CMS data, analyses based on Medicare data will be considered supportive and exploratory and will not replace the primary trial analyses based on prospectively collected RELIANCE study data.

9.5. Missing data

As noted in the IOM Guidelines, prevention is the most important method for handling missing data (11). Recommended strategies include careful consideration of potential barriers to data collection during the design phase of the trial, including patient and outcome selection. An advantage of the primary outcome, all-cause hospitalization or death, is that it may be ascertained through multiple complementary sources, including participant follow-up, site coordinator review, and available electronic health record information.

For patient reported outcomes, reaching the patient is essential. Several strategies will

be employed to minimize missing data during trial conduct. These include recording and periodically verifying participant contact information, including secondary contacts such as relatives or caregivers; automated reminders through emails and text messages; telephone follow-up by the Call Center; and assistance from site coordinators in locating participants lost to follow-up. Real-time data performance monitoring reports will also be used to identify trends in missing data and support corrective actions when appropriate.

All randomized participants will be included in the primary intention-to-treat analyses. For time-to-event analyses of the primary outcome (all-cause hospitalization or death), Cox proportional hazards models will be used, with participants censored at the time of the last known follow-up contact if subsequent outcome information is unavailable. Data collected up until the time of withdrawal will be retained from those participants who do not continue on in the study. Potential bias related to informative censoring or missing data mechanisms will be evaluated through sensitivity analyses. The range of potential influence will be established by imputing the 'best' and 'worst' case scenarios. Analytic methods will be used to assess the censoring patterns. Based upon these finding, sensitivity analyses using multiple imputation, inverse probability weighting, or a pattern mixture approach will be used (12,13). For longitudinal and repeated-measures outcomes, the primary analysis will be based upon mixed effects models (for replicate measurements over time) or multiple imputation (for single measurements), both of which are robust to data that is missing at random. 'Best' and 'worst' case analyses similar to those used for time-to-event outcomes will also be used to gauge the potential effect of missing data.

Reasons for missing data will be summarized descriptively when available. For data collected from clinical centers, the reason for missing data will be collected directly from the clinic, e.g., data not assessed, patient refused or no follow- up for patient at the clinical center. The UIC Call Center will record attempts to remind and collect data for patient reported outcomes. These will include patient refusal, inability to contact patient, or death. In comparative effectiveness trials where the hypothesis is non-inferiority, treatment cross-overs, treatment cessation, lack of adherence, and loss to follow-up tend to support the inference of non-inferiority between treatment

groups. To some extent, these deviations from protocol are the rationale for conducting a “real-world” pragmatic trial. Interpretation of these factors can be assisted by comparison of as randomized and per-protocol analyses as well as sensitivity analyses of specified subgroups. However, failure to ascertain the primary outcome, particularly if there is differential data loss between treatment groups remains a potential threat to the validity of the trial and substantial efforts will be made throughout the study to maximize follow-up completeness and outcome ascertainment.

10. References

1. Fleming TR, Odem-Davis K, Rothman MD, Shen YL. Some essential considerations I the design and conduct of non-inferiority trials. Clin Trials 2011;8:432-439
2. FDA 2016. Non-inferiority Clinical Trials to Establish Effectiveness: Guidance for Industry. http://www.fda.gov/downloads/Drugs/.../Guidances/UCM202140.pdf (accessed November 24, 2016).
3. Wise RA, Anzueto A, Cotton D, Dahl R, Devins T, Disse B, Dusser D, Joseph E, Kattenbeck S, Koenen-Bergmann M, Pledger G, Calverley P; TIOSPIR Investigators. Tiotropium Respimat inhaler and the risk of death in COPD. N Engl J Med 2013; 369: 1491-501.
4. Rennard SI, Calverley PM, Goehring UM, Bredenbroker D, Martinez FJ. Reduction of exacerbations by PDE4 inhibitor reflumilast—the importance of defining different subsets of patients with COPD REsp Res, 2011; 12:18.
5. Han MK, Tayob N, Murray S, Dransfield MT, Washko G, et al. Predictors of Chronic Obstructive Pulmonary Disease Exacerbation Reduction in Response to Daily Azithromycin Therapy. Am J Respir Crit Care Med 2014; 189: 1503-8.
6. NCSS Statistical Software. http://www.ncss.com/software/pass/ (accessed July 29, 2015)
7. SWOG Statistical Center. https://www.swogstat.org/statoolsout.html (accessed July 29, 2015).
8. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Reporting of subgroup analyses in clinical trials. N Engl J Med 2007;357:2187-2194. PMID: 15548779.
9. Wittes J. Commentary on Murray and Hernan. Clinical Trials 2016; 13: 379-81.
10. Murry EJ, Hernan MA. Adherence adjustment in the Coronary Drug Project: A call for better per-protocol effect estimates in randomized trials. Clinical Trials 2016; 13: 372-378.
11. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med 2012;367:1355-60.
12. Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2nd ed. New York: John Wiley & Sons, 2002.
13. Daniels MJ, Hogan JW. Missing Data in Longitudinal Studies; Strategies for Bayesian Modeling and Sensitivity Analysis. Boca Raton: Chapman & Hall/CRC Press, 2008.

11. Appendix A

Table. Summary of Medicare Administrative Data Files and Study Applications

Data File Name	Data Period	Included Beneficiaries		Purpose
		Part A/B (Fee-for-service)	Part C (Managed Care)	
Master Beneficiary Summary File (MBSF) Base	2020 - 2022 CY, 2023 Q1 – Q4	X	X	To examine demographics (e.g., age, race, ethnicity), enrollment status, and mortality
FFS Outpatient (OP)	2020 - 2022 CY, 2023 Q1 – Q4	X		To examine resource use, conditions, clinical encounters, and clinical outcomes of interest in the outpatient setting
Encounter OP*	2020 - 2022 CY		X	
FFS Carrier (CARR)	2020 - 2022 CY, 2023 Q1 – Q4	X		To examine resource use, conditions and outcomes of interest, and clinical encounters
Encounter CARR*	2020 - 2022 CY		X	
MedPAR (SS/LS/SNF version)*	2020 - 2022 CY	X	X	To examine resource use, conditions, clinical encounters, and clinical outcomes of interest in the inpatient hospital and skilled nursing facility settings
FFS Inpatient (IP)	2023 Q1 – Q4	X		To examine resource use, conditions, clinical encounters, and clinical outcomes of interest in the inpatient hospital setting
FFS Skilled Nursing Facility (SNF)	2023 Q1 – Q4	X		To examine resource use, conditions, clinical encounters, and clinical outcomes of interest in the skilled nursing facility setting
Part D Event (PDE)*	2020 - 2022 CY	X (subset)	X	To examine utilization patterns, outcomes, and exposure to medical therapies
Part D Drug Characteristics*	2020 - 2022 CY	X (subset)	X	To characterize utilization patterns for the drugs and conditions of interest
Part D Pharmacy Characteristics*	2020 - 2022 CY	X (subset)	X	To identify the pharmacies utilized by our cohort and examine the impact of the pharmacy choices on other variables (e.g., utilization, death or hospitalization, cost)
Plan Characteristics*	2020 - 2022 CY	X (subset)	X	To describe costs associated with the drugs and medical conditions of interest
RELIANCE ID to Bene ID Crosswalk	2020 Q1 – 2023 Q4	X	X	To identify RELIANCE trial participants in the Medicare data

* Only available in an annual format