



Using multiple imputation of real-world data to estimate clinical remission in pediatric inflammatory bowel disease

Journal of **Comparative Effectiveness Research**

Nanhua Zhang^{*,1,2} , Chunyan Liu¹ , Steven J Steiner³ , Richard B Colletti⁴ , Robert Baldassano⁵ , Shiran Chen¹ , Stanley Cohen^{6,7} , Michael D Kappelman⁸ , Shehzad Saeed⁹ , Laurie S Conklin¹⁰ , Richard Strauss¹⁰ , Sheri Volger¹⁰ , Eileen King^{1,2} & Kim Hung Lo¹⁰

¹Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA

²University of Cincinnati, Cincinnati, OH 45229, USA

³Riley Hospital for Children/Indiana University School of Medicine, Indianapolis, IN 46202, USA

⁴University of Vermont College of Medicine, Department of Pediatrics, Burlington, VT 05405, USA

⁵The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

⁶GI Care For Kids, Atlanta, GA 30342, USA

⁷Morehouse School of Medicine, Atlanta, GA 30310, USA

⁸University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

⁹Dayton Children's Hospital/Wright State University, Dayton, OH 45404, USA

¹⁰Janssen Research & Development, LLC, Spring House, PA 19477, USA

*Author for correspondence: Tel.: +1 513 803 9108; nanhua.zhang@cchmc.org

Aim: To evaluate the performance of the multiple imputation (MI) method for estimating clinical effectiveness in pediatric Crohn's disease in the ImproveCareNow registry; to address the analytical challenge of missing data. **Materials & methods:** Simulation studies were performed by creating missing datasets based on fully observed data from patients with moderate-to-severe Crohn's disease treated with non-ustekinumab biologics. MI was used to impute sPCDAI remission statuses in each simulated dataset. **Results:** The true remission rate (75.1% [95% CI: 72.6%, 77.5%]) was underestimated without imputation (72.6% [71.8%, 73.3%]). With MI, the estimate was 74.8% (74.4%, 75.2%). **Conclusion:** MI reduced nonresponse bias and improved the validity, reliability, and efficiency of real-world registry data to estimate remission rate in pediatric patients with Crohn's disease.

Plain language summary: Multiple imputation of registry data to estimate clinical remission in pediatric patients with Crohn's disease.

What is this article about? The purpose of this study was to investigate whether multiple imputation can help researchers address the challenge of missing data when estimating effectiveness or how well a treatment works in patients treated in clinical practice rather than in a clinical trial. Multiple imputation is a statistical technique that fills in a patient's missing data with values based on existing data from other patients with similar characteristics. The ImproveCareNow registry provided a suitable source of "real-world" clinical practice data for studying treatment effectiveness in pediatric patients with Crohn's disease. However, registries often have partial or missing data, which can cause bias, or the tendency to overestimate or underestimate an outcome. In this study, multiple imputation was used to try to reduce bias caused by missing clinical remission outcome for some patients.

What were the results? The true clinical remission rate was 75.1%. Without multiple imputation, the remission rate was underestimated (72.6%). With multiple imputation, the estimated remission rate (74.8%) was closer to the true remission rate.

What do the study results mean? These results show that multiple imputation can reduce bias and improve validity, reliability, and efficiency when using registry data to estimate remission rates in pediatric patients with Crohn's disease when some patients are missing remission data.

Tweetable abstract: In a new study, multiple imputation reduced bias and improved validity, reliability, and efficiency of real-world #registrydata to estimate clinical remission in #pediatricpatients with #Crohnsdisease

First draft submitted: 27 July 2022; Accepted for publication: 30 January 2023; Published online: 17 February 2023

Keywords: clinical remission status • disease-specific patient registries • drug repurposing and relabeling • ImproveCareNow registry • inflammatory bowel disease • missing data • multiple imputation method • pediatric Crohn's disease • real-world evidence • Short Pediatric Crohn's Disease Activity Index (sPCDAI)

The US FDA and European Medicines Agency have both issued guidance encouraging the use of relevant and reliable real-world data (RWD) and scientifically robust methodologies to generate real-world evidence (RWE) to support regulatory decisions [1–3]. RWD collected from pediatric patients by disease-specific registries and quality improvement initiatives in real-world clinical settings have the potential to generate RWE for regulatory decisions, in turn, accelerating access to treatments for pediatric indications [4,5]. Because outcomes data in disease-specific registries are typically collected in clinical practice, registries may have partial or missing data, and the data may be collected at variable timepoints, adding to the challenges of utilizing and evaluating RWD. Furthermore, the risk of bias from partial and missing data needs to be addressed by the statistical methodologies applied.

The multiple imputation (MI) method has been widely used to deal with missing data and for parameter estimation in the presence of incomplete data. The procedure creates multiple imputed datasets that can be analyzed with standard software packages, and the results can be combined using Rubin's rule [6]. The assumption for MI to be valid is that the data are missing at random, which implies that the probability for a data point to be missing depends on data that are observed but not the missing value itself [7,8]. In this study, we used MI to assess the impact of missingness in the ImproveCareNow (ICN) registry. ICN is a large, multicenter pediatric inflammatory bowel disease (IBD) quality improvement and research network [9], and maintains a centralized registry of standardized clinical data routinely collected in clinical practice at the time of diagnosis and at every outpatient clinic visit for more than 50,000 patients with Crohn's disease (CD) or ulcerative colitis [10]. The ICN registry collects clinician-reported outcome measures, such as the Short Pediatric Crohn's Disease Activity Index (sPCDAI) and Physician Global Assessment (PGA) [11], to evaluate disease severity and to monitor therapeutic responses [12–14].

In addition to outcome measures, the ICN registry also includes a variety of patient data such as patient demographics, clinical disease characteristics, laboratory values, medication use, prior surgery and nutrition status [9,10]. These rich patient-level data strengthen the missing at random assumption for MI in this RWE study. Furthermore, the MI method has been successfully applied to ICN registry data for evaluation of the effectiveness of anti-tumor necrosis factor- α (TNF α) agents in the management of pediatric CD [15]. Building on this previous work, we extended the MI from 9 predictors and medication variables to 24, which predicted either the sPCDAI component scores or the missingness of these scores.

ICN data from pediatric patients with CD treated with ustekinumab, a monoclonal antibody to the p40 subunit of interleukin (IL)-12/IL-23 approved for use in adults with CD [16], from 1 January 2007 through 31 July 2019 ($n = 1696$ ICN visits) [17] were examined in a feasibility analysis. Approximately 35.7% of participants had documented moderate-to-severe CD (by PGA or corticosteroid use) at the ustekinumab initiation visit, demonstrating that the registry is a suitable source of RWE to evaluate effectiveness of ustekinumab in children with CD.

To address the problem of partial and missing data in this dataset, we designed the present study to evaluate the performance of the MI method for estimating the clinical remission (defined as sPCDAI ≤ 10) rate from data collected in the ICN registry for pediatric patients with moderate-to-severe CD treated with non-ustekinumab biologics.

Materials & methods

Study design & setting

The study included ICN registry data for pediatric patients who were treated with non-ustekinumab biologics (primarily TNF α blockers, the name of the biologic agents in the records were masked) and who had no previous exposure to ustekinumab from 1 January 2014 to 31 December 2019. Only data from patients not exposed to

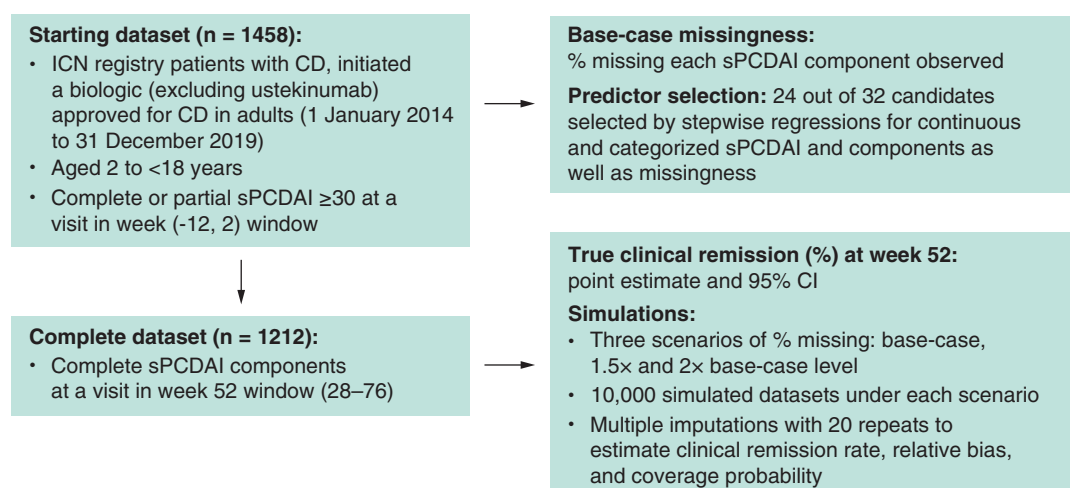


Figure 1. Study design diagram.

CD: Crohn's disease; CI: Confidence interval; ICN: ImproveCareNow; sPCDAI: Short Pediatric Crohn's Disease Activity Index.

Box 1. Inclusion criteria for data simulations.

- Documented Crohn's disease (CD) diagnosis in the baseline window
- Initiation of a biologic agent (i.e., index drug) approved for the adult CD indication in the USA (i.e., infliximab, including the biosimilars, adalimumab, certolizumab, natalizumab, and vedolizumab) between 1 January 2014, and 31 December 2019
- Aged between 2 and <18 years at the time of initiation of the index biologic. If a patient had multiple biologic initiations that met all other inclusion criteria, then the date of the latest biologic initiation was used to determine if the age criteria were met
- At least 1 complete or partial sum of score for sPCDAI ≥ 30 within the baseline window must be available
- No exposure to ustekinumab between 1 January 2014, and 31 December 2019
- At least 1 visit entered in the ICN Registry within the week 52 window
- Provided informed consent or assent for the use of their data in the ICN registry for research purposes

ICN: ImproveCareNow; sPCDAI: Short Pediatric Crohn's Disease Activity Index.

ustekinumab were used to preserve the integrity of the data from patients treated with ustekinumab for use in a separate planned analysis of RWD from the ICN registry. Figure 1 summarizes the study design. Data management and statistical analyses were conducted at Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. The statistical analysis plan was developed and approved prior to study initiation.

Study procedures

Data simulations

To evaluate the performance of MI for handling missing data, a series of simulations were conducted, which included the creation of missingness in the component scores of the sPCDAI based on a subset of fully observed data of ICN patients with moderate-to-severe CD treated with non-ustekinumab biologics. The MI method was used to impute the missing remission statuses for the estimation of remission rates for each simulated dataset. Missingness was defined as the status of a value missing for an outcome of interest (e.g., nonresponse) [18]. The complete dataset used in the simulations was defined as the subset of patients with complete sPCDAI components at 52 weeks following treatment initiation.

A starting dataset was extracted from the ICN registry of patients who met the inclusion criteria (Box 1). Key inclusion criteria included patients 2 to <18 years old who initiated treatment with a biologic agent other than ustekinumab from 1 January 2014 to 31 December 2019, and who had both a baseline and week 52 visit. Baseline and week 52 visits were chosen as primary time points and mimicked the key time points in an ongoing Phase 3 study examining the effectiveness of ustekinumab in pediatric patients with moderate-to-severe CD (ClinicalTrials.gov

Identifier: NCT04673357). Time windows were defined as -12 to +2 weeks around time-zero (T_0) for baseline visit and 28 to 76 weeks from T_0 for week 52 visit, ± 24 weeks around week 52 (T_{52}), where T_0 equals date of first administration of non-ustekinumab biologic, and T_{52} equals 1 year from T_0 (+365 days). The week 52 visit was defined as a visit within the week 52 window for sPCDAI evaluation. If there were multiple visits within this time window, precedence was given to the visit(s) for which remission status could be determined algebraically (i.e., sPCDAI could be determined to be ≤ 10 or > 10 based on all available nonmissing components) and closest to T_{52} ; if the remission status could not be determined algebraically, the visit closest to T_{52} was chosen. The missingness pattern in each sPCDAI component in this starting dataset was summarized (hereafter referred to as the “base case”). The missingness follows a general missing data pattern, i.e., intermittent missingness. The other type of missingness pattern is monotone missingness, which is common in longitudinal studies with dropout. The subset of patients who had fully observed sPCDAI component scores at the week 52 visit was extracted from the starting dataset (hereafter referred to as the “complete dataset”). The true clinical remission rate used in the data simulations was the remission rate calculated for the complete dataset.

For the simulations, missing data were randomly imposed on individual sPCDAI components in the complete dataset that matched the missing pattern, i.e., the missingness in each sPCDAI component identified from the starting dataset, as well as $1.5\times$ and $2.0\times$ the percentage of missingness in the base case. Ten thousand datasets were simulated for each of the 3 scenarios.

Multiple imputation

MI was used to substitute values for the clinical remission status in the simulated datasets. Potential predictors (i.e., the independent variables to be used in the MI) for various representations of the sPCDAI end point were identified based on predictors of sPCDAI components previously reported [15] for another study using the same starting dataset of ICN registry data. Separate regression models were used to identify predictors for sPCDAI score and components as well as for missingness of remission status. All of the data in the starting dataset were utilized to allow for a more robust evaluation of these associations. For categorical covariates, “missing” was treated as a category. All continuous covariates were categorized and treated as categorical predictor variables.

Next, the starting dataset was used to examine the association between the independent variables, the missingness status of clinical remission at week 52, and the following representations of the sPCDAI end point: (1) clinical remission status, i.e., dichotomized sPCDAI > 10 versus ≤ 10 ; (2) complete sPCDAI composite score (range: 0–90 points); (3) categorized variable based on complete sPCDAI composite score (remission = sPCDAI score of 0–10; mild = sPCDAI score of 11–29; or moderate-to-severe = sPCDAI score of 30–90); and (4) 6 individual components of the sPCDAI. Logistic regression modeling was used for the two binary outcomes (clinical remission status and missingness in clinical remission status), multiple linear regression modeling was used for the continuous sPCDAI composite score (range: 0–90 points), and cumulative logistic regression modeling was used for ordinal outcomes (categorized sPCDAI and individual sPCDAI components). Stepwise regression was used to select the variables that were predictive of each end point. The threshold for retaining the predictors was defined as p -value < 0.1 . The union of the selected variables were all included in the MI model as predictors. The SAS procedure PROC GLMSELECT (SAS version 9.4, SAS Institute, Inc., NC, USA) was used for continuous outcomes, and PROC HPGENSELECT was used for binary or ordinal outcomes.

After the artificial missing data were imposed, the clinical remission status was calculated in the complete dataset and in each of the 10,000 simulated datasets. If the absence of clinical remission status could be determined with partial component data, no imputation of the remission status was performed. If the clinical remission status could not be determined, MI was performed.

The SAS PROC MI procedure was used for the MI within each of the 10,000 simulation datasets generated for each scenario. The number of imputations within PROC MI was set to 20. A fully conditional specification logistic regression method was used to impute the sPCDAI components as ordinal outcomes. This fully conditional specification may be applied to missing data with arbitrary missingness patterns. Cumulative logistic regression models were used to impute all the missing sPCDAI components, and complete sPCDAI scores were calculated by adding the individual component values. The missing clinical remission status was then calculated based on the complete sPCDAI score > 10 versus ≤ 10 .

Lastly, the 95% confidence interval (CI) for the clinical remission rate for each simulated dataset where missing data were created was established. From the 20 imputed datasets for each simulation within each scenario (where missing was replaced with imputed values), the clinical remission rates were summarized and aggregated to obtain

Table 1. Missing patterns of sPCDAI components observed at week 52 in the starting dataset.

Starting dataset (n = 1458)	
Missing stool, n (%)	57 (3.9)
Missing abdominal pain, n (%)	8 (0.5)
Missing well-being, n (%)	7 (0.5)
Missing abdominal mass, n (%)	53 (3.6)
Missing EIM, n (%)	4 (0.3)
Missing body weight, n (%)	142 (9.7)

EIM: Extra intestinal manifestation; n: Number of patients; sPCDAI: Short Pediatric Crohn's Disease Activity Index.

the point estimates and 95% CIs. Rubin's rule was used to aggregate the estimated remission rates (SAS procedure PROC MIANALYZE) because this procedure combines across imputations and accounts for statistical uncertainties. The 95% CI estimate for each imputed dataset was calculated using the "Wald" method, which used large-sample normal approximation (the asymptotic standard error) [19].

MI performance was measured by relative empirical bias and coverage probability [20] of the true remission rate derived from the complete dataset, the subset of data including only those patients who had all 6 components of the sPCDAI scores 52 weeks after initiation of therapy (designated using the parameter value p). The absolute empirical bias of the MI was estimated by taking the average of the differences between the estimated clinical remission rate from each simulation and p. Relative empirical bias was estimated as absolute bias/p. Coverage probability was calculated as the proportion of the 95% CIs that cover p out of the 10,000 simulations. As a comparison, we also evaluated the performance of the before imputation results, which used only patients whose remission status could be determined with observed sPCDAI component scores.

Results

A total of 1458 patients met the simulation inclusion criteria and, as such, were included in the starting dataset. The missing patterns of sPCDAI components observed at week 52 in the starting dataset are summarized in Table 1. Of these patients, 83% (1212/1458) had all 6 of the sPCDAI components available at week 52 for calculating the sPCDAI score; therefore, these 1212 patients comprised the complete dataset. The true clinical remission rate calculated for the complete dataset (p) was 75.1% (95% CI: 72.6%, 77.5%), which was used as the benchmark for evaluating the performance of the MI method.

Variable selection yielded a total of 24 candidate predictors for all 10 of the representations of the outcomes at week 52 (Table 2), which were included in the MI model. One variable, PGA, was considered *a priori* to be the most relevant predictor and was kept in all the models. In the 10,000 simulated datasets for the base case, the mean missing percentage in remission status at week 52 was 11.5% (standard deviation [SD] = 1.2%; 95% CI: 7.0%, 16.0%). The distribution of clinical remission rate estimates at week 52, when the missingness was simulated for the base case, 1.5× and 2× that of the starting dataset, was reported (Table 3).

The mean clinical remission rate after MI was 74.8% (SD = 0.2%; 95% CI: 74.4%, 75.2%). Determining the clinical nonremission status for those patients whose partial sum of sPCDAI components was >10, while discarding those whose remission status could not be determined by available sPCDAI components resulted in an underestimation of the clinical remission rate (72.6% [SD = 0.4%; 95% CI: 71.8%, 73.3%]) compared with when MI was used (74.8% vs 72.6%) (Table 3).

In all 3 simulated scenarios (base case, 1.5× and 2.0×), the relative empirical bias in the estimates of the clinical remission rate after MI (-0.38%, -0.60% and -0.84%, respectively) was only approximately 11.5% of that before imputation (-3.3%, -5.16% and -7.18%, respectively) and 28.6% more in coverage probability (100% vs 71.40%) (Table 4). In the base case, where the missingness was simulated at 1.0× that observed in the starting dataset, the MI method showed minimal relative empirical bias (-0.38%) compared with the true clinical remission rate. The 95% CIs included the true clinical remission rate 100% of the time. Accordingly, the findings for the base case were supported by the additional analyses when the missingness was simulated at 1.5× and 2× of the starting dataset. The relative empirical biases remained low (<-1.0% in all 3 simulated scenarios) after MI, while the length of the 95% CI increased slightly with the increase in the amount of missing data, and the coverage probabilities remained 100%.

Table 2. Predictors selected for multiple imputation of missing data.

Description	sPCDAI	sPCDAI categorical	Remission week 52	Stool	Well-being	Abdominal mass	EIM	Abdominal pain	Weight loss	Missing week 52 remission	Times being selected
Age group at induction									1		1
Age group at initial IBD diagnosis	1	1	1	1							4
Albumin categories at week 52	1		1						1		3
Corticosteroid use at baseline									1		1
ESR categories at week 52							1		1		2
CD/IC extent of disease - macroscopic lower GI disease at baseline						1					1
Father's education		1						1			2
Gender	1	1	1					1	1	1	6
Prior GI surgery				1							1
Growth status at week 52	1	1				1			1		4
Hematocrit categories at week 52		1		1							2
Methotrexate use at baseline										1	1
If methotrexate newly started at week 52									1		1
Mother's education								1			1
Nutritional status at week 52	1	1	1	1					1	1	6
PGA at week 52	1	1	1	1	1	1	1	1	1	1	10
Health Insurance			1	1		1					3
Any prior biologic use before induction date	1										1
Number of prior biologics used										1	1
Race				1							1
Since the previous visit, if the patient has been in continuous remission at week 52	1	1	1	1							4
Last known sPCDAI partial score before the week 52 visit date to week 28 after induction date categories	1	1	1	1	1	1		1	1	1	8
Thiopurines (azathioprine, 6-mercaptopurine) use at baseline					1						1
If thiopurines newly started at week 52					1						1

"1" indicates the predictors were selected for predicting the corresponding outcome in each column at p-value <0.1 except PGA, which was considered *a priori* to be the most relevant predictor and was not subject to the selection. CD/IC: Crohn's disease/ileocolonic; EIM: Extra intestinal manifestation; ESR: Erythrocyte sedimentation rate; GI: Gastrointestinal; IBD: Inflammatory bowel disease; PGA: Physician Global Assessment; sPCDAI: Short Pediatric Crohn's Disease Activity Index.

Table 3. Distribution of week 52 clinical remission rate estimates.

True remission rate in the complete dataset: 75.1%		
Extent of simulated missingness of sPCDAI components [†]	Percent remission with no imputation (%)	Percent remission with multiple imputation [‡] (%)
1.0× Base case [§]		
Mean (SD)	72.6 (0.4)	74.8 (0.2)
Median (25th, 75th percentile)	72.6 (72.4, 72.9)	74.8 (74.7, 74.9)
95% CI	71.8, 73.3	74.4, 75.2
Min, Max	71.0, 73.9	74.0, 75.7
1.5× Base case		
Mean (SD)	71.2 (0.5)	74.6 (0.3)
Median (25th, 75th percentile)	71.2 (70.9, 71.6)	74.6 (74.5, 74.8)
95% CI	70.2, 72.2	74.1, 75.1
Min, Max	69.3, 72.9	73.6, 75.6
2× Base case		
Mean (SD)	69.7 (0.6)	74.5 (0.3)
Median (25th, 75th percentile)	69.7 (69.3, 70.1)	74.5 (74.3, 74.6)
95% CI	68.5, 70.8	73.9, 75.0
Min, Max	67.0, 72.1	73.4, 75.5

[†] Statistics based on 10,000 simulated datasets per scenario.
[‡] Twenty imputations were performed per simulated dataset.
[§] Base case determined by pattern of missingness in the starting dataset.
CI: Confidence interval; Max: Maximum; Min: Minimum; SD: Standard deviation; sPCDAI: Short Pediatric Crohn's Disease Activity Index.

Table 4. Empirical biases and coverage probabilities (of the 95% CI) for multiple imputation.

Relative biases [†] (mean with 95% CI)			
Scenario	1.0× Base case	1.5× Base case	2.0× Base case
Before multiple imputation	-3.30% (-4.32%, -2.33%)	-5.16% (-6.52%, -3.90%)	-7.18% (-8.79%, -5.68%)
After multiple imputation	-0.38% (-0.92%, 0.16%)	-0.60% (-1.25%, 0.08%)	-0.84% (-1.61%, -0.08%)
Coverage probabilities [‡]			
Scenario	1.0× Base case	1.5× Base case	2.0× Base case
Before multiple imputation	71.40%	0.80%	0%
After multiple imputation	100%	100%	100%

[†] Relative bias = $\frac{\text{Remission rate from the simulated data} - \text{True remission rate}}{\text{True remission rate}} \times 100\%$. The true clinical remission rate was calculated from the complete dataset (75.1%).
[‡] Coverage probability was the percent of times the clinical remission rate was in the 95% CI for each simulated scenario.
CI: Confidence interval.

Discussion

Our study demonstrated the robustness of using MI as the primary analytical method for estimating the week 52 clinical remission rate in RWD from the ICN registry. The simulations showed that the MI method reduced nonresponse bias in the clinical remission rate observed from only using the available sPCDAI score (i.e., observed before imputation) by selecting a large number of predictors readily available in the ICN registry. Information was borrowed from rich data from other variables to predict the values of the missing sPCDAI components, and information from auxiliary variables was used to improve the estimation of the clinical remission rate by reducing relative empirical bias and improving coverage probability. Additionally, in comparison to the true clinical remission rate of 75.1%, performing a complete-case analysis for the datasets before simulation (i.e., an analysis restricted to patients with complete remission status data) resulted in approximately 3% to 7% relative empirical bias (underestimation) and coverage probabilities of the true clinical remission rate by a 95% CI of 0% to 71.4% across the 3 simulated scenarios. MI reduced nonresponse bias and had good coverage under the simulated missingness of up to 2.0× that observed in the starting dataset. The amount of missingness did not have an impact on the bias for the MI method, while the bias of the complete-case analysis method increased with the amount of missing data. Another imputation method called nonresponder imputation, which is more conservative and assumes those with a missing remission status are not in remission, will further bias the remission status from the true value [21].

This study does have a few limitations. The MI method is based on the assumption that data are missing at random and that all the factors that were predictive of the outcomes and missingness were included in the missing data imputation model, which may not be true in this study; consequently, we included as many relevant variables as possible from the ICN registry to strengthen the assumption. In contrast, as previously mentioned, using MI to substitute values for the missing data outperformed selectively omitting the patients with the missing data (e.g., with unavailable sPCDAI components), as selective omission may result in underestimation. Additionally, categorizing the continuous variables may have resulted in a potential loss of information or may have distorted the relationship among variables; however, it allowed for including more patients in our analysis by including a missing category for these patients. Furthermore, the outcomes were not measured exactly at T_0 and T_{52} but rather within a window of T_0 and T_{52} due to the observational nature of RWD studies; discretizing the time windows into T_0 and T_{52} may have resulted in a potential loss of information. Finally, only data from ICN patients not exposed to ustekinumab were used to avoid making any inference about these patients prior to regulatory agency approval of a future study protocol and statistical analysis plan for a separate, planned analysis of RWD from the ICN registry. Although patients receiving ustekinumab may be different from patients not receiving ustekinumab with regard to disease severity, we do not expect the association between the clinical remission status and the predictors to be different between the 2 patient groups.

Although not an issue for this study, for studies involving complex analytical models, it is important to ensure the compatibility between the analytical model and the imputation model, which can be achieved by using an imputation model that is more general than the analytical model [22]. For example, if the analytical model involves significant interaction between 2 variables or the transformation of a variable, the imputation model should include these as well [23]. White and colleagues presented a nice tutorial of the MI method with a fair discussion of the limitations and pitfalls [22].

Despite these limitations, the use of RWD from disease-specific patient registries, such as ICN, and the use of scientifically robust methodologies, such as data simulation and MI, have the potential to generate relevant and reliable RWE that accelerates regulatory approvals and informs patient care [4]. Our study utilized ICN registry data as well as data simulations and MI to estimate clinical remission rates from the sPCDAI score, an outcome measure commonly used in clinical trials. The overall methodology that was utilized is not specific to the sPCDAI end point and may be applied to deal with any missing outcome in rich data where many relevant variables were measured. In a similar manner, Forrest and colleagues utilized ICN registry data to compute clinical remission rates comparable with those observed in single-group efficacy studies among children and controlled clinical trials among adults [15]. Extracting and analyzing ICN registry data for the period of April 2007 to March 2012, they found that 8.0% of the sPCDAI components were missing across the full dataset (4130 patients with CD) and that 41% of the ICN visits (20,456) had at least 1 missing component; in turn, they used MI to address the missingness observed. Other RWE studies have reached conclusions similar to those from randomized clinical trials, demonstrating the validity of using RWD [24,25].

Pediatric IBD data prospectively collected by the ICN registry are a relevant, reliable, and rich source of RWD for evaluating the effectiveness of IBD treatment strategies in the pediatric population. Although some regulatory authorities have established a framework [26] for the use of RWD to generate RWE that supports the approval of new indications for IBD treatments, the need to establish a consistent approach for minimizing the impact of partial and missing data from RWD sources remains. The MI method that we applied may serve as a model for other RWD registry analyses.

Conclusion

The MI method both optimized and maximized the use of real-world registry data for estimating clinical remission status in pediatric patients with CD, making this scientific methodology a valuable approach for addressing missing data in other studies of real-world registry data where missingness is assumed to occur at random.

Author contributions

All the authors were involved in the conception, preparation, and approval of the manuscript and, as such, met the criteria for authorship recommended by the International Committee of Medical Journal Editors (www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html).

Acknowledgments

K Ruley Sharples of Janssen Scientific Affairs, LLC, provided input and critical review of the manuscript.

Financial & competing interests disclosure

This work was financially supported by Janssen Research & Development, LLC (a subsidiary of Johnson & Johnson). N Zhang, C Liu, S Chen and E King are employees of Cincinnati Children's Hospital Medical Center (CCHMC), and their participation in the present manuscript is funded through a contract between ImproveCareNow (ICN) and CCHMC. R Baldassano serves in a consultancy/advisory role for Janssen Research & Development, LLC. S Cohen serves in a consultancy/advisory role for Janssen Research & Development, LLC and Kate Farms; owns stock in Nutrition4Kid.com; and receives grant support from AbbVie, Arena, Eli Lilly, Janssen, Kate Farms, and Takeda. RB Colletti serves in a consultancy/advisory role for Janssen Research & Development, LLC. MD Kappelman has served in a consultancy/advisory role for AbbVie, Janssen, Pfizer, and Takeda; is a shareholder in Johnson & Johnson; and has received research support from AbbVie, Arenapharm, Boehringer Ingelheim, Bristol Myers Squibb, Celltrion, Eli Lilly, Genentech, Janssen, Pfizer, and Takeda. S Saeed serves on a Speaker's Bureau and Advisory Board for AbbVie. SJ Steiner has no conflicts of interest to disclose. LS Conklin, R Strauss, S Volger, and KH Lo are employees of Janssen Research & Development, LLC, and own Johnson & Johnson stock and/or stock options. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Medical writing support was provided by SC Thompson, ELS, and H Capasso-Harris, of Certara Synchrogenix under the direction of the authors in accordance with Good Publication Practice guidelines (*Ann. Intern. Med.* 175, 1298–1304 [2022]) and was funded by Janssen Research & Development, LLC.

Ethical conduct of research

To protect the patients, the study was conducted in accordance with Good Clinical Practice and Health Insurance Portability and Accountability Act of 1996 requirements. A limited dataset with no patient identifiers was used, and the confidentiality of the patient records was maintained at all times.

Data sharing statement

The data underlying this article are available in the ImproveCareNow Registry and shared in accordance with the ImproveCareNow Data Sharing Policy (www.improvecarenow.org/research-resources).

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Summary points

- Disease-specific patient registries are sources of real-world data for studying treatment effectiveness in pediatric Crohn's disease (CD); however, missing data in these registries presents an analytical challenge.
- The ImproveCareNow registry routinely collects clinical remission data, particularly the Short Pediatric Crohn's Disease Activity Index (sPCDAI).
- A feasibility analysis of ImproveCareNow remission data, from baseline (first ustekinumab dose induction) to week 52, found that approximately one-third of patient visits had incomplete sPCDAI data (29.7% [503/1696]).
- Utilizing ImproveCareNow registry data for pediatric patients with moderate-to-severe CD treated with non-ustekinumab biologics as a case study, we evaluated use of the multiple imputation method to impute clinical remission (defined as $sPCDAI \leq 10$) statuses.
- Missing data patterns were randomly imposed on a completely observed dataset, and various percentages of missing data were simulated.
- Multiple imputation was applied to the simulated datasets to estimate clinical remission rate.
- The true remission rate (75.1% [95% CI: 72.6%, 77.5%]) was underestimated without imputation (72.6% [71.8%, 73.3%]); with multiple imputation of the missing data, the estimate was 74.8% (74.4%, 75.2%).
- The multiple imputation method reduced nonresponse bias and improved the validity, reliability, and efficiency of using real-world data from an inflammatory bowel disease registry to estimate the clinical remission rate in pediatric patients with CD.

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest

1. United States Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Oncology Center of Excellence. Real-world data: assessing registries to support regulatory decision-making for drug and biological products: draft guidance for industry (2021). Available at: www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-registries-support-regulatory-decision-making-drug-and-biological-products (Accessed: 12 July 2022).
2. United States Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Oncology Center of Excellence. Real-world data: assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products: draft guidance for industry (2021). Available at: www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory (Accessed: 12 July 2022).
3. European Medicines Agency. Guideline on registry-based studies (2021). Available at: www.ema.europa.eu/en/documents/scientific-guideline/guideline-registry-based-studies.en-0.pdf (Accessed: 12 July 2022).
4. Crowley E, Ma C, Andic M, Feagan BG, Griffiths AM, Jairath V. Impact of drug approval pathways for paediatric inflammatory bowel disease. *J. Crohns Colitis*. 16(2), 331–335 (2021).
- **Reports that few drugs are currently approved for use in pediatric patients with inflammatory bowel disease (IBD), with long interval delays after adult approval, and that many drugs used in pediatric patients with IBD are used off-label.**
5. Halfvarson J, Cummings F, Grip O, Savoye G. Inflammatory bowel disease registries for collection of patient iron parameters in Europe. *World J. Gastroenterol.* 24(10), 1063–1071 (2018).
6. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons Inc, NY, USA (1987).
- **Lays the foundation for the multiple imputation method in research.**
7. Rubin DB. Inference and missing data. *Biometrika* 63(3), 581–592 (1976).
8. Little RJA, Rubin DB. *Statistical Analysis With Missing Data (3rd Edition)*. John Wiley & Sons, Inc, NJ, USA (2020).
9. Egberg MD, Kappelman MD, Gulati AS. Improving care in pediatric inflammatory bowel disease. *Gastroenterol. Clin. North Am.* 47(4), 909–919 (2018).
10. Steiner SJ, Samson CM, Howe KM, Saeed S, Duganiero TM. ImproveCareNow network 2020 annual data report (2021). Available at: www.improvecarenow.org/ (Accessed: 12 July 2022).
11. Kappelman MD, Crandall WV, Colletti RB et al. Short pediatric Crohn's disease activity index for quality improvement and observational research. *Inflamm. Bowel Dis.* 17(1), 112–117 (2011).
12. Hyams J, Markowitz J, Otley A et al. Evaluation of the pediatric Crohn disease activity index: a prospective multicenter experience. *J. Pediatr. Gastroenterol. Nutr.* 41(4), 416–421 (2005).
13. Loonen HJ, Griffiths AM, Merkus MP, Derkx HH. A critical assessment of items on the pediatric Crohn's disease activity index. *J. Pediatr. Gastroenterol. Nutr.* 36(1), 90–95 (2003).
14. Hyams JS, Ferry GD, Mandel FS et al. Development and validation of a pediatric Crohn's disease activity index. *J. Pediatr. Gastroenterol. Nutr.* 12(4), 439–447 (1991).
15. Forrest CB, Crandall WV, Bailey LC et al. Effectiveness of anti-TNF α for Crohn disease: research in a pediatric learning health system. *Pediatrics* 134(1), 37–44 (2014).
- **Article reporting missingness in 4130 patients with Crohn's disease in the ImproveCareNow registry, and demonstration of the use of the multiple imputation method to address the missingness observed.**
16. STELARA® (ustekinumab) injection, for subcutaneous or intravenous use: prescribing information. Janssen Biotech, Inc, PA, USA (2020).
17. Steiner S, Colletti RB, King E et al. Feasibility of using real-world data from the ImproveCareNow registry to examine the efficacy of ustekinumab for pediatric Crohn's disease. Poster presented at: *Digestive Disease Week (DDW) Medfyle Conference*, Virtual, P581 21–23 May 2021.
18. United States National Research Council. Introduction and Background. In: *The prevention and treatment of missing data in clinical trials. Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education*. The National Academies Press, DC, USA, 7–20 (2010).
19. Wald A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Am. Math. Soc.* 54(3), 426–482 (1943).
20. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat. Med.* 38(11), 2074–2102 (2019).
- **Tutorial that provides a structured approach for planning and reporting simulation studies, which involves defining aims, data-generating mechanisms, estimands, methods, and performance measures (ADEMP).**
21. Chen J, Hunter S, Kisfalvi K, Lirio RA. A hybrid approach of handling missing data under different missing data mechanisms: VISIBLE 1 and VARSITY trials for ulcerative colitis. *Contemp. Clin. Trials* 100, 106226 (2021).

22. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat. Med.* 30(4), 377–399 (2011).
23. Allison PD. Handling missing data by maximum likelihood (paper 313). SAS Global Forum (2012). <https://statisticalhorizons.com/wp-content/uploads/MissingDataByML.pdf> Accessed 17 January 2023.
24. Franklin JM, Patorno E, Desai RJ *et al.* Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative. *Circulation* 143(10), 1002–1013 (2021).
- **Overview of a structured process to design real-world evidence studies emulating randomized, controlled trials and to compare results. The regulatory conclusions were found equivalent in 6 of 10 studies, and the real-world evidence emulations achieved a hazard ratio estimate that was within the 95% confidence interval from the corresponding randomized, controlled trials in 8 of 10 studies.**
25. Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. Feasibility of using real-world data to replicate clinical trial evidence. *JAMA Netw. Open* 2(10), e1912869 (2019).
- **Demonstration that 33/220 (15.0%) clinical trials could be replicated using observational data since the intervention, indication, inclusion and exclusion criteria, and primary end points of these trials could be routinely ascertained from insurance claims and electronic health records data.**
26. United States Department of Health and Human Services, Food and Drug Administration. Framework for FDA's real-world evidence program (2018). Available at: www.fda.gov/media/120060/download (Accessed: 12 July 2022).