U.S. FOOD & DRUG ADMINISTRATION

Center for Biologics Evaluation and Research
Office of Biostatistics and Epidemiology

CBER Surveillance Program Biologics Effectiveness and Safety (BEST) Initiative

Supplementing Administrative Claims COVID-19 Vaccine Data with Immunization Information Systems (IIS) Data: A Feasibility Study Report

Final Report

July 26, 2022

Prepared by: Karen Schneider, PhD;* Elizabeth J. Bell, PhD;* Cindy Ke Zhou, PhD; Grace Yang, MPA, MA; Patricia Lloyd, PhD; Tainya C. Clarke, PhD; Michael Wilkinson, MPH; Emily Myers, MPH; Kandace Amend, PhD; John D. Seeger, DrPH, PharmD; Richard Forshee, PhD; Azadeh Shoaibi, PhD, MHS; Steven Anderson, PhD, MPP; Hui-Lee Wong, PhD

*Co-first authors

Study Team:

US FDA: Hui-Lee Wong, PhD; Cindy Ke Zhou, PhD; Patricia Lloyd, PhD; Tainya C. Clarke, PhD; Richard Forshee, PhD; Azadeh Shoaibi, PhD, MHS; Steven A. Anderson, PhD, MPP

Optum: Karen Schneider, PhD; Elizabeth J. Bell, PhD, MPH; Michael Wilkinson, MPH; Emily Myers, MPH; Holly Phipps, MSW; Rebecca Braun, BA; Kandace Amend, PhD; John D. Seeger, DrPH, PharmD; Grace Yang, MPA, MA

Table of Contents

I. Background

The Food and Drug Administration (FDA) launched the Biologics Effectiveness and Safety (BEST) Initiative in October 2017 to augment the Center for Biologics Evaluation and Research's (CBER) access to better data, methods, tools, expertise, and infrastructure to conduct studies that ensure the safety and effectiveness of biologic products, like vaccines. BEST collaborators use a variety of data sources, including medical and pharmacy claims collected for billing purposes.

Currently, BEST collaborators (e.g., Optum) are working with the FDA to design and conduct safety and effectiveness studies of the coronavirus disease 2019 (COVID-19) vaccines using commercial medical and pharmacy claims. However, a COVID-19 vaccine exposure measure based on health insurance claims will incompletely capture COVID-19 vaccination because COVID-19 vaccines are provided to individuals free of charge and may be administered outside of the health care system. Therefore, certain groups administering the vaccines may not send claims to insurance companies for all individuals who carry health insurance.[1,2] Indeed, based on preliminary analyses, only 37.8% of commercially insured individuals under 65 years of age in a state had at least one dose of a COVID-19 vaccine using administrative claims data alone. During the same time period in this state, an age-standardized estimate of 72.8%[3,4] of 5 – 64 year old Americans (note discrepancy in age bands) had received at least one dose of a COVID-19 vaccine based on the Centers for Disease Control and Prevention's (CDC's) vaccine administration data. Although CDC's estimates of at least one dose tend to be overestimates, discussed more in Section IV, this large discrepancy illustrates the point that a claims-only measure of COVID-19 vaccines is incomplete.

Accurate data on COVID-19 vaccination, including complete person-level information on date, dose, and brand name of vaccine received, is necessary for COVID-19 vaccine studies, especially for studies comparing vaccinated people to unvaccinated people because misclassification of vaccination status would lead to biased effect estimates. For instance, a study with incomplete exposure ascertainment that categorized vaccinated individuals as unvaccinated based on the absence of COVID-19 vaccine claims would underestimate vaccine effectiveness.[5]

Thus, the BEST Initiative sought to evaluate the utility of enhancing claims data with COVID-19 vaccine records from Immunization Information Systems (IIS) for more complete vaccine capture. IIS is defined by the CDC as "confidential, population-based, computerized databases that record all immunization doses administered by participating providers to persons residing within a geopolitical area."[6]

While BEST linkage efforts with IIS are ongoing, this report presents the feasibility results from this linkage and evaluation of COVID-19 vaccine capture with a subset of the linked claims-IIS data.

A. Study Aims

This study evaluates the extent to which IIS data enhance claims-based COVID-19 vaccine capture and estimates the remaining misclassification of vaccinated individuals as unvaccinated.

Aim 1: ***To evaluate the extent to which IIS data increase the completeness of COVID-19 vaccine data in administrative claims databases.***

Aim 2: ***To estimate the magnitude of remaining misclassification of vaccinated individuals as unvaccinated using the vaccination status estimates from Aim 1 (i.e., vaccine records from both IIS and claims data).***

- Aim 2.1: To estimate misclassification of vaccinated individuals as unvaccinated using capture-recapture analyses.[7]
- Aim 2.2: To estimate misclassification of vaccinated individuals as unvaccinated by comparing our estimates of vaccination to those reported by the CDC.
- Aim 2.3: To estimate misclassification of vaccinated individuals as unvaccinated by comparing our estimates of vaccination to those reported by departments of health (DOH).

II. Methods

A. Data Sources

A linked claims-IIS database is the main data source used for this study. This database was developed by enhancing a pre-adjudicated medical and pharmacy claims database with IIS COVID-19 vaccine data for the individuals within the claims database. A description of each data source follows.

1. Pre-Adjudicated Medical and Pharmacy Claims Data

The claims database combines pre-adjudicated medical claims with linked enrollment and pharmacy claims. Hospital and physician claims undergo initial daily processing from a large number of providers across the United States (US) who accept individuals with health insurance. The pre-adjudicated medical claims have an approximately two-month delay for 90% completeness for inpatient claims and over 70% completeness at one-month for outpatient claims. The pre-adjudicated claims database provides the advantage of a shorter time lag as compared to adjudicated claims, for more timely capture of COVID-19 vaccine records. Pharmacy drug claims are adjudicated at the point of sale in real time, by communication between the pharmacy and insurer computer systems.

The pharmacy claims used for this study were refreshed on 18 January 2022 and pre-adjudicated medical claims were refreshed on 15 January 2022.

2. IIS Repositories

IIS data include vaccination data reported by a variety of providers, including physicians, retail pharmacies, hospitals, local public health departments, and mass vaccination campaigns in the jurisdiction.8 There are currently 64 IIS in the US, comprised of all 50 states, the District of Columbia, eight US territories, and five cities. While all of the IIS collect individuals' COVID-19 vaccine administration data, the IIS databases are individually operated and managed by each jurisdiction. A national and consolidated IIS database with complete adoption of functional standards9 and harmonized processes for COVID-19 vaccine administration is not available at this time.

Each IIS captures person-level information data on the vaccination, including (but not limited to): dates of vaccination dose, vaccine manufacturer, vaccine dose number, vaccine dose volume, in addition to identifiable information on the vaccine recipient.

To create IIS repositories, all COVID-19 vaccine data for health plan members were obtained from 14 IIS in 11 states, henceforth referred to as State 1 through State 11. State names were masked because the purpose of this study is to assess the added benefit of enhancing claims data with IIS, not to compare state vaccine capture. Each IIS used an IIS-specific linkage algorithm to match members on a roster to IIS vaccine records administered in their jurisdiction. Linkage algorithms varied by IIS, but typically included a combination of the following variables: member ID, first name, last name, middle initial, date of birth, and zip code.

The IIS repositories are refreshed with data from the IIS at various intervals. The IIS refresh schedule is daily for one state (State 9), weekly for three states (State 5, State 6, and State 7), semi-monthly for one state (State 4), and monthly for six states (State 1, State 2, State 3, State 8, State 10, and State 11). This report is based on COVID-19 IIS repository data as of 10 January

2022. Analytically restricting to the study period (see Section II B), the last observed date of service in each IIS was 21 November 2021 for State 1, 23 November 2021 for State 10, 02 December 2021 for State 2, 03 December 2021 for State 4, 17 December 2021 for State 11, 23 December 2021 for State 8, 30 December 2021 for State 9, and 31 December 2021 for State 3, State 5, State 6, and State 7. These last dates of service give an indication of the data lag for each IIS.

3. Linked Claims-IIS Database

To create the linked claims-IIS database, the pre-adjudicated medical and pharmacy claims database was linked by member ID to the IIS repository data for the individuals within the claims database. The linked claims-IIS database also includes a small number of Medicare claims-derived COVID-19 vaccine records (approximately 1,630 vaccine records after data cleaning; see Section II C for more information) for the commercially insured population who were concurrently eligible for Medicare.

B. Study Period

The study period was 01 December 2020 (month of emergency use authorization by the FDA for the first COVID-19 vaccine) to 31 December 2021. Records in the linked claims-IIS database were restricted to those with dates of service through 31 December 2021.

C. Deduplication of Vaccine Records

Individuals in the linked claims-IIS database may have no vaccine records or they may have a vaccine record from claims only, IIS only, or both. Thus, we conducted deduplication at the individual level to remove redundant vaccine records for the same person. To do this, we first deduplicated the vaccine records within each vaccine data source (i.e., IIS, claims). Specifically, vaccine records with the same member ID, date of service, and vaccine type were deduplicated within each source, with priority on keeping records with a non-missing dose number.

The IIS and claims vaccine records output after this step were utilized to create claims-based and IIS-based vaccine indicators, as further described in Section II E. Records were then deduplicated *across* data sources, keeping only one record with the same date of service and vaccine type per member ID, with priority on keeping records according to the following order: 1) commercial medical or pharmacy claims, 2) IIS records, 3) Medicare claims. For example, if an individual had vaccine records sourced from both commercial claims and IIS with the same date of service and vaccine type, the record from the commercial claims source was retained. If only an IIS-based record was available, the IIS-based record was retained. The supplemental Appendix A includes additional information about the hierarchical data cleaning algorithm.

D. Study Population

The study population was comprised of individuals less than 65 years old who resided in one of the 11 states and were commercially insured by health plans in the claims database at any time during the study period (i.e., between 01 December 2020 [month of emergency use authorization from the FDA for the first COVID-19 vaccine] and 31 December 2021.

E. Defining Vaccination Status

Vaccine records were identified using vaccine administered (CVX) codes, Current Procedural Terminology (CPT) codes, Healthcare Common Procedure Coding System (HCPCS) codes, National Drug Codes (NDC), and International Classification of Diseases, Tenth Revision (ICD-10) procedure codes during the study period.

Dose number was determined using the date of service/vaccine administration. The first dose was defined as the vaccine record with the earliest administration date and the second dose was defined as the second record with an administration date at least three days apart from the first dose.

For this study, vaccination status was defined as:

1. **At least one vaccine dose:** Individuals with at least one dose of a COVID-19 vaccine, regardless of vaccine brand/type.

2. **Completed vaccine series:** The definition of completed vaccine series varied by brand/type of vaccine. If the vaccine was Ad26.COV2.S (Janssen), then only one vaccine record was required to have a completed vaccine series. For the purposes of this study, if the vaccine was BNT162b2 (Pfizer-BioNTech), mRNA-1273 (Moderna), a mixed combination of BNT162b2 or mRNA-1273, or unknown, we defined a completed vaccine series as presence of at least two vaccine records at least three days apart.

3. **At least one additional dose:** The definition of at least on additional dose (e.g., booster or third dose) varied by brand/type of vaccine. If the manufacturer for individuals with a completed vaccine series was Ad26.COV2.S, then at least one additional dose at least three days apart was required (regardless of manufacturer) after the first dose. If the manufacturer was BNT162b2, mRNA-1273, mixed, or unknown, then at least one additional dose at least three days apart was required after the dose that completed the series.

F. Statistical analyses

1. Aim 1

For Aim 1, we sought to evaluate the extent that IIS data increase the completeness of COVID-19 vaccine data in administrative claims databases. We performed descriptive analyses of vaccination estimates for the study population. Specifically, we provided percentages and counts of our study population with at least one vaccine dose, a completed vaccine series, and at least one additional dose using vaccination information from claims only (claims-based vaccine indicator), IIS only (IIS-based vaccine indicator), and the two sources combined (combined claims-IIS vaccine indicator). Vaccination estimates were stratified by key variables (Table 1; supplemental Appendix B includes additional information about variable definitions).

Table 1. Stratification variables

Variable	Categories
State of residence	State 1 – State 11
Age at end of study period	<5 y, 5 – 11 y, 12 – 17 y, 18 – 25 y, 26 – 35 y, 36 – 45 y, 46 – 55 y, 56 – 64 y

Variable	Categories
Sex	Male, female, unknown
Type of vaccine	mRNA-1273, BNT162b2, Ad26.COV2.S, Mixed, Unknown
Dose number	Dose number was determined using the date of service/vaccine administration
Month-Year of dose	Dec-20, Jan-21, Feb-21, Mar-21, Apr-21, May-21, Jun-21, Jul-21, Aug-21, Sep-21, Oct-21, Nov-21, Dec-21

2. Aim 2

To address the second study objective, we estimated the magnitude of misclassification in vaccination estimates using two approaches: 1) comparing our vaccination coverage estimates with a data-driven estimate based on capture-recapture analyses and 2) comparing our vaccination coverage estimates with age-standardized external references (CDC and DOH), as detailed below.

To assess under-recording of vaccination (i.e., the percent of vaccinated individuals misclassified as unvaccinated), we compared CDC, DOH, and capture-recapture adjusted vaccination estimates to the combined claims-IIS vaccination estimates at the state-level. We calculated under-recording of vaccination as the difference between the vaccination estimate from another source and our combined claims-IIS estimate, divided by the vaccination estimate from the other source. It can be interpreted as a percent change from the CDC, DOH, or capture-recapture adjusted estimates.

(Aim 2.1) To estimate the number of individuals in our study population who were vaccinated, but not identified by either of the vaccine data sources (IIS or claims), we used a capture-recapture analysis method overall and by state.[7] This analysis estimated the number of people who would be misclassified as unvaccinated if vaccination status was based on a combination of claims and IIS vaccine records.

Briefly, capture-recapture analysis uses counts from two independent data sources to estimate the number of individuals who were not captured by either of the data sources. The assumptions of capture-recapture are 1) independence of capture by the two data sources, 2) all individuals have the same probability of being captured, 3) no migration in or out of the catchment area, and 4) no loss of identifier. Although these assumptions are for the most part untestable, the likelihood that they are being met and the impact on results if not met is considered in the discussion section. In this analysis, the two data sources were vaccine records from IIS and claims.

Under the assumption of independent samples, the maximum likelihood estimator of the odds ratio, $\frac{ad}{bc}$, should equal 1.0, indicating no relationship. Since the values in cells *a*, *b*, and *c* are known from the source of the two samples (assuming individuals have a unique identifier so those common to the two samples can be identified), the unknown value is *d* (Table 2). Rearranging the terms of the equation to solve for *d* yields: $d = \frac{bc}{a}$. As illustrated in Table 2, by deriving cell d using the factors *a*, *b*, and *c*, we estimated the number of individuals in our study population who were truly vaccinated, but misclassified as unvaccinated because a vaccine record was not captured in claims or IIS. We then re-calculated vaccination estimates, re-classifying the number identified in cell *d* as vaccinated for this analysis.

Table 2. 2x2 table used for capture-recapture analysis

		Vaccine record identified in IIS?	
		Yes	No
Vaccine record identified in claims?	Yes	a	b
	No	c	d

(Aim 2.2) We estimated potential misclassification of vaccinated individuals as unvaccinated by comparing our vaccination estimates to CDC vaccination estimates.[3,4] CDC's estimates are based on vaccine administration, similar to this study. However, CDC estimates include different data sources than this study, including not just IIS, but also territories, tribes, local entities, retail pharmacies, long-term care facilities, dialysis centers participating in the Federal Dialysis Center Program, Federal Emergency Management Agency partner sites, Health Resources and Services Administration partner sites, and other federal entity facilities in a given jurisdiction.[10] However, vaccine records are de-identified and may not be linkable (i.e., a person's dose 1 and dose 2 records might be counted as first doses to different people), hence overestimating the prevalence of vaccine recipients, particularly for first doses.[10-12] We age-standardized[13] (age groupings used: 0 – 4, 5 – 11, 12 – 17, 18 – 64) the CDC estimates for all states. CDC data as of 13 January 2022 were used for all states but State 1 and State 10. CDC data from 03 November 2021 were used for State 1 and State 10 to match the last date of service in the IIS data. Notably, we accessed CDC's data through a Mayo Clinic website,[3] which refreshes the CDC data daily and presents age-stratified vaccination estimates by state.

(Aim 2.3) Additionally, we estimated potential misclassification of vaccinated individuals as unvaccinated by comparing our vaccination estimates to DOH vaccination estimates (age-standardized against the age-distribution of our study population). While DOH estimates are largely based on IIS data, some states also include federal sources. For example, State 6 vaccination estimates include doses identified in IIS and doses administered through Veterans Affairs. States that do not provide age-stratified estimates on their website, such as State 2, were unable to be used in the comparison to DOH estimates. To compare DOH data that closely matched the last date of service within each IIS, we used data from 21 November 2021 for State 1; 11 January 2022 for State 3, State 4, State 5, and State 6; 17 January 2022 for State 7; 14 January 2022 for State 8; 01 January 2022 for State 9; 23 November 2021 for State 10; and 15 December 2021 for State 11. Each state's DOH data were accessed through their respective DOH website.

All analyses were performed in SAS Enterprise Guide version 6.1 (SAS Institute, Inc., Cary NC).

G. Approvals and Data Exchange

Institutional review board approval was not required for this study because it was conducted for FDA BEST public health surveillance. All IIS involved in this study allow the use of their COVID-19 vaccine data for public health surveillance activities.

Data exchange occurred via various secure platforms including, but not limited, to HL7 Web Service call, state website, secure email, flat file, and health information exchange (HIE).

III. Results

A. Findings from Aim 1

Aim 1: To evaluate the extent to which IIS data increases the completeness of COVID-19 vaccine data in administrative claims databases.

A total of 5,112,722 commercially insured individuals less than 65 years old from the 11 states were identified within the study period and included in the analyses (Supplemental Figure 1). Attrition of the vaccine records for the study population is presented in Supplemental Figure 2. Supplemental Tables 1 – 4 describe vaccination patterns among the study population who received at least one dose.

IIS data increased vaccine capture substantially. The percentage of individuals with at least one dose of COVID-19 vaccine was 32.8% using the claims alone and increased to 48.1% with the combined claims-IIS (Table 3; comprehensive results in Supplemental Table 5a). The difference in estimates for at least one vaccine dose between the claims and the combined claims-IIS varied by state, with an increase by 25.1 percentage points (pp) for State 9 and 8.5 pp for State 3. The discrepancy between the percentages with at least one dose of COVID-19 vaccine using the claims indicator and the combined claims-IIS indicator increased with age; for instance, the difference in overall estimates was 1.3 pp for individuals 5 – 11 years old and 20.6 pp for individuals 56 – 64 years old. Results for individuals with a completed vaccine series were similar to what was observed for those who received at least one dose (Table 4; comprehensive results in Supplemental Table 5a). Of those with a completed vaccine series (n= 2,143,556), 1,285,343 (60.0%) received BNT162b2, 613,805 (28.6%) received mRNA-1273, 121,005 (5.6%) received Ad26.COV2.S, 122,972 (5.7%) received a mix of vaccine types, and 431 (0.02%) received doses with unknown type.

Using the claims indicator and restricting the denominator to those with a completed vaccine series, 22.3% (277,945 out of 1,248,637) had at least one additional dose (booster or third dose) . This increased to 27.7% (592,808 out of 2,143,556) when using the combined claims-IIS indicator (Table 5). The small contribution of IIS data to the percent of individuals with an additional dose is likely explained by the timing of eligibility of booster shots for adults. On 21 November 2021, the CDC expanded eligibility of booster shots to all adults who had their second dose more than six months prior;[14] the last date of service for IIS data in this study was in November or December 2021.

Table 3. Among commercially insured individuals aged < 65 years from 11 states, percent with at least one dose[a] of COVID-19 vaccine between 01 December 2020 and 31 December 2021: Total and by state, sex, and age

Characteristic	Total study population	Members with at least one dose in claims[b]		Members with at least one dose in combined claims-IIS[c]		Difference between combined claims-IIS and claims only[d]
	Denominator	Number	Percent	Number	Percent	Percentage points
Total	5,112,722	1,676,235	32.8%	2,458,231	48.1%	15.3
State						
State 1	643,602	201,474	31.3%	316,177	49.1%	17.8
State 2	158,385	47,831	30.2%	76,820	48.5%	18.3
State 3	1,143,375	422,934	37.0%	520,249	45.5%	8.5
State 4	696,305	184,312	26.5%	265,936	38.2%	11.7
State 5	786,234	255,544	32.5%	401,634	51.1%	18.6
State 6	318,060	136,090	42.8%	167,745	52.7%	9.9
State 7	330,165	124,739	37.8%	191,327	58.0%	20.2
State 8	360,267	110,016	30.5%	179,787	49.9%	19.4
State 9	87,663	18,927	21.6%	40,901	46.7%	25.1
State 10	219,939	54,303	24.7%	105,376	47.9%	23.2
State 11	254,098	76,424	30.1%	133,781	52.7%	22.6
Multiple states[e]	114,629	43,641	38.1%	58,498	51.0%	12.9
Sex						
Female	2,618,098	866,261	33.1%	1,275,478	48.7%	15.6
Male	2,491,037	809,051	32.5%	1,179,453	47.4%	14.9
Unknown	3,587	923	25.7%	3,300	92.0%	66.3
Age						
05 – 11 y	443,138	80,538	18.2%	86,600	19.5%	1.3
12 – 17 y	413,499	121,160	29.3%	185,500	44.9%	15.6
18 – 25 y	649,067	195,367	30.1%	283,126	43.6%	13.5
26 – 35 y	941,101	317,135	33.7%	463,163	49.2%	15.5
36 – 45 y	890,290	336,142	37.8%	507,362	57.0%	19.2
46 – 55 y	824,346	329,839	40.0%	496,422	60.2%	20.2
56 – 64 y	676,105	295,635	43.7%	434,418	64.3%	20.6
Missing or < 5 y[f]	275,176	419	0.2%	1,640	0.6%	0.4

[a] Defined as at least one dose of a COVID-19 vaccine, regardless of vaccine type (Ad26.COV2.S, BNT162b2, mRNA-1273, mixed, or unknown)
[b] Prior to hierarchical deduplication of vaccine records across IIS and claims
[c] After hierarchical deduplication of vaccine records across IIS and claims
[d] Calculated as percent combined claims-IIS subtracted by percent claims only
[e] "Multiple states" means there were multiple states listed for a patient, among the eleven states of interest
[f] Missing and < 5 y were combined to mask small cell counts to protect privacy

Table 4. Among commercially insured individuals aged < 65 years from 11 states, percent with a completed COVID-19 vaccine series[a] between 01 December 2020 and 31 December 2021: Total and by state, sex, and age

Characteristic	Total study population	Members with a completed series in claims[b]		Members with a completed series in combined claims-IIS[c]		Difference between combined claims-IIS and claims only[d]
	Denominator	Number	Percent	Number	Percent	Percentage points
Total	5,112,722	1,248,637	24.4%	2,143,556	41.9%	17.5
State						
State 1	643,602	145,137	22.6%	287,198	44.6%	22.0
State 2	158,385	38,294	24.2%	68,478	43.2%	19.0
State 3	1,143,375	310,479	27.2%	404,913	35.4%	8.2
State 4	696,305	135,725	19.5%	228,643	32.8%	13.3
State 5	786,234	193,105	24.6%	366,046	46.6%	22.0
State 6	318,060	102,514	32.2%	144,224	45.3%	13.1
State 7	330,165	101,157	30.6%	180,397	54.6%	24.0
State 8	360,267	83,987	23.3%	159,617	44.3%	21.0
State 9	87,663	12,709	14.5%	36,876	42.1%	27.6
State 10	219,939	39,386	17.9%	95,468	43.4%	25.5
State 11	254,098	54,735	21.5%	122,816	48.3%	26.8
Multiple states[e]	114,629	31,409	27.4%	48,880	42.6%	15.2
Sex						
Female	2,618,098	641,978	24.5%	1,112,255	42.5%	18.0
Male	2,491,037	606,161	24.3%	1,028,218	41.3%	17.0
Unknown	3,587	498	13.9%	3,083	86.0%	72.1
Age						
05 – 11 y	443,138	52,979	12.0%	58,791	13.3%	1.3
12 – 17 y	413,499	100,304	24.3%	164,815	39.9%	15.6
18 – 25 y	649,067	146,150	22.5%	239,174	36.9%	14.4
26 – 35 y	941,101	233,084	24.8%	397,343	42.2%	17.4
36 – 45 y	890,290	249,629	28.0%	447,997	50.3%	22.3
46 – 55 y	824,346	246,537	29.9%	443,184	53.8%	23.9
56 – 64 y	676,105	219,678	32.5%	390,892	57.8%	25.3
Missing or < 5 y[f]	275,176	276	0.1%	1,360	0.5%	0.4

[a] Defined based on type of vaccine and observed number of vaccine records. If the type was Ad26.COV2.S, then only one vaccine record was required to have a completed vaccine series, regardless of what other types of vaccines were present. If the type was BNT162b2, mRNA-1273, a mixed combination that did not include Ad26.COV2.S, or unknown, then a complete vaccine series is defined as presence of at least two vaccine records at least three days apart.
[b] Prior to hierarchical deduplication of vaccine records across IIS and claims
[c] After hierarchical deduplication of vaccine records across IIS and claims
[d] Calculated as percent combined claims-IIS subtracted by percent claims only
[e] "Multiple states" means there were multiple states listed for a patient, among the eleven states of interest
[f] Missing and < 5 y were combined to mask small cell counts to protect privacy

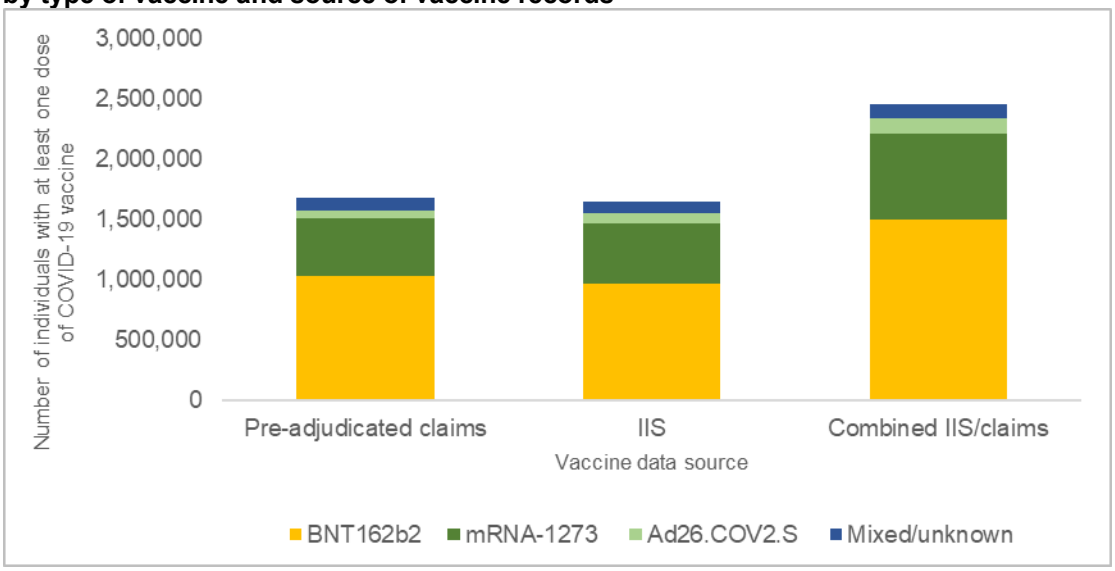
Table 5. Among commercially insured individuals aged < 65 years from 11 states, percent with an additional (e.g., booster or third) COVID-19 dose[a] between 01 December 2020 and 31 December 2021: Total and by state, sex, and age

Characteristic	Members with an additional dose (e.g., booster or third) in claims[b]		Members with an additional dose (e.g., booster or third) in combined claims-IIS[c]		Difference between combined claims-IIS and claims only[d]
	Number	Percent[e]	Number	Percent[e]	Percentage points
Total	277,945	22.3%	592,808	27.7%	5.4
State					
State 1	26,256	18.1%	72,809	25.4%	7.3
State 2	8,087	21.1%	16,458	24.0%	2.9
State 3	76,876	24.8%	101,574	25.1%	0.3
State 4	28,218	20.8%	48,114	21.0%	0.2
State 5	46,419	24.0%	106,673	29.1%	5.1
State 6	22,000	21.5%	60,490	41.9%	20.4
State 7	24,837	24.6%	64,637	35.8%	11.2
State 8	19,824	23.6%	54,051	33.9%	10.3
State 9	2,770	21.8%	13,604	36.9%	15.1
State 10	5,070	12.9%	17,296	18.1%	5.2
State 11	10,134	18.5%	22,539	18.4%	-0.1
Multiple states[f]	7,454	23.7%	14,563	29.8%	6.1
Sex					
Female	146,320	22.8%	317,606	28.6%	5.8
Male	131,551	21.7%	274,295	26.7%	5.0
Unknown	74	14.9%	907	29.4%	14.5
Age					
05 – 11 y	765	1.4%	3,934	6.7%	5.3
12 – 17 y	5,791	5.8%	18,471	11.2%	5.4
18 – 25 y	23,564	16.1%	55,078	23.0%	6.9
26 – 35 y	46,452	19.9%	94,338	23.7%	3.8
36 – 45 y	62,374	25.0%	130,577	29.2%	4.2
46 – 55 y	66,893	27.1%	140,050	31.6%	4.5
56 – 64 y	72,063	32.8%	150,059	38.4%	5.6
Missing or < 5 y[g]	43	15.6%	301	22.1%	6.5

[a] If the type of vaccine for individuals with a completed vaccine series was Ad26.COV2.S, then at least one additional dose was required (regardless of type), at least three days apart from the first dose. If the type of vaccine was BNT162b2, mRNA-1273, mixed, or unknown, then a third dose was required at least three days apart from the dose that completed the series.
[b] Prior to hierarchical deduplication of vaccine records across IIS and claims
[c] After hierarchical deduplication of vaccine records across IIS and claims
[d] Calculated as percent combined claims-IIS subtracted by percent claims only
[e] Denominator for percentages with ≥1 additional dose is members with a completed vaccine series from the same vaccine indicator family (e.g., percentages with ≥1 additional dose using the IIS vaccine indicator were calculated out of the number of members with a completed vaccine series using the IIS vaccine indicator)
[f] "Multiple states" means there were multiple states listed for a patient, among the eleven states of interest
[g] Missing and < 5 y were combined to mask small cell counts to protect privacy

Of those receiving at least one vaccine dose using the combined claims-IIS vaccine indicator (n=2,458,231), BNT162b2 was the most frequently used vaccine with 1,497,191 individuals receiving at least one dose, followed by mRNA-1273 (n=715,821), Ad26.COV2.S (n=121,005), mixed types (n=123,748), and unknown type (n=466). These results are shown in Figure 1, with comprehensive results in Supplemental Table 5a. The distribution of type of vaccine among individuals with at least one COVID-19 vaccine dose was similar across sources of vaccine records.

Figure 1. Number of persons with at least one dose of COVID-19 vaccine by type of vaccine and source of vaccine records



Among those with a completed vaccine series, the greatest number of individuals received their first doses in March and April 2021 (Figure 2) and the largest number of individuals received their last observed doses in April and May 2021 among individuals with a completed vaccine series (Figure 3; comprehensive results in Supplemental Table 5a). Similar patterns were observed across sources of vaccine records.

Figure 2. Count of individuals with a completed vaccine series by month of first observed dose by source of vaccine records

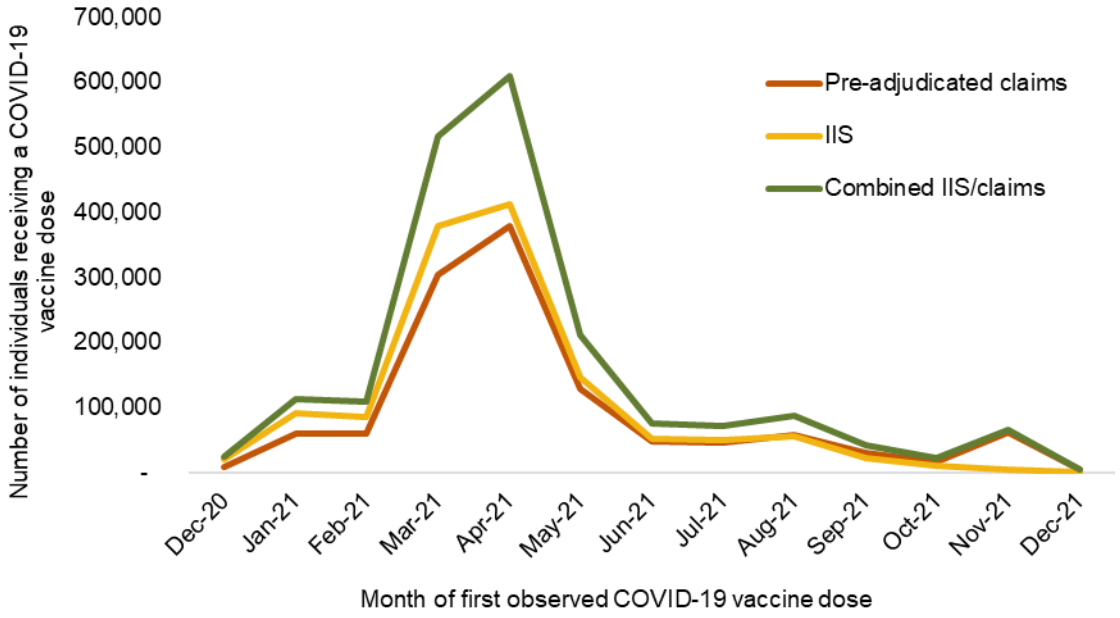
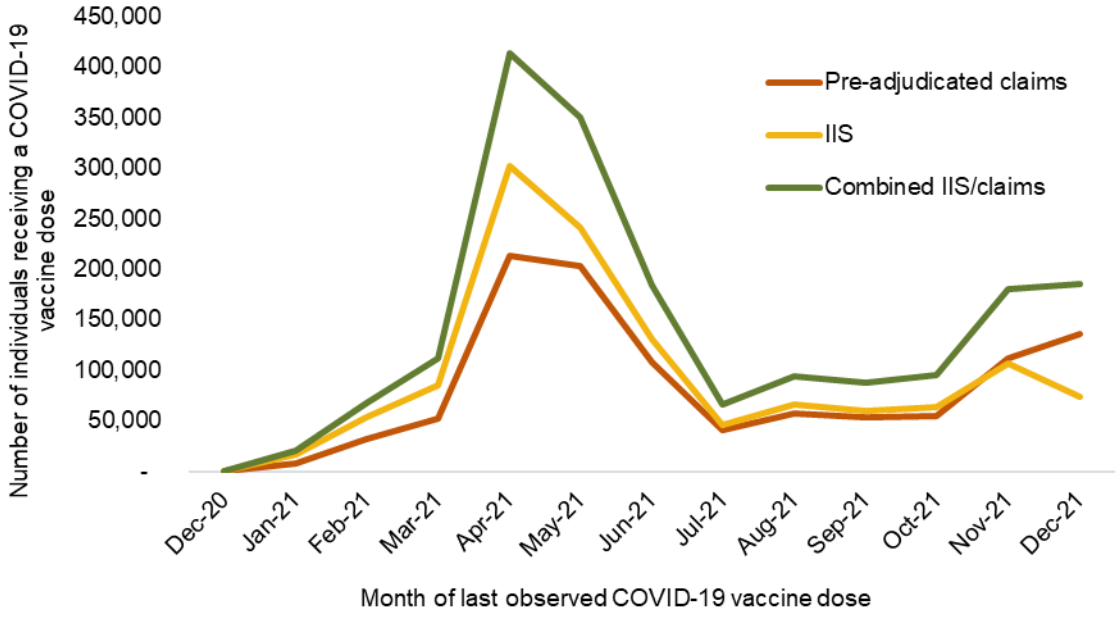


Figure 3. Count of individuals with a completed vaccine series by month of *last* observed dose by source of vaccine records



B Findings from Aim 2

Aim 2.1: To estimate misclassification of vaccinated individuals as unvaccinated using capture-recapture analyses.

Using a capture-recapture analysis, we estimated that 739,128 individuals in total ([781,996 * 814,498] / 861,737) from our study population (14.5%) were misclassified as unvaccinated because they had no vaccine records in the IIS or claims, but truly did have a vaccine dose (derived from Supplemental Table 6). This would reduce the estimated number of unvaccinated individuals in our 11 states from 2,654,491 to 1,915,363 and increase our estimated number of individuals with at least one vaccine dose to 62.5%. State-specific estimates with capture-recapture for at least one dose and completed vaccine series are shown in Tables 6 and 7 (derived from Supplemental Tables 6 and 7). Compared to capture-recapture estimates, the percent under-recording in the combined claims-IIS data varied between 9.2% (State 7) and 38.7% (State 3) for individuals that received at least one dose, and 11.1% (State 7) and 50.9% (State 3) for individuals that received a completed vaccine series (Tables 6 and 7).

Aim 2.2: To estimate misclassification of vaccinated individuals as unvaccinated by comparing our estimates of vaccination to those reported by the CDC

State-specific, age-standardized CDC estimates, in comparison to our claims-IIS combined estimates, are presented in Tables 6 and 7, and the extent of discrepancy varied by state. The CDC estimates were always higher than our estimates. Compared to the CDC estimates, the percent under-recording in our combined claims-IIS data varied between 15.1% (State 7) and 44.2% (State 3) for individuals that received at least one dose, and 12.1% (State 7) and 45.7% (State 3) for individuals that received a completed vaccine series.

Aim 2.3: To estimate misclassification of vaccinated individuals as unvaccinated by comparing our estimates of vaccination to those reported by departments of health (DOH)

The DOH estimates were also always higher than our estimates and varied by state for individuals that received at least one dose and individuals that received a completed vaccine series, although the differences are typically smaller than the magnitude of difference between our estimates and the CDC estimates (Tables 6 and 7). When comparing DOH estimates to the combined claims-IIS estimates, State 7 had the lowest percent of under-recording (10.2% for individuals that received at least one dose and 9.1% for individuals that received a completed vaccine series), and State 3 had the highest percent of under-recording (39.3% for individuals that received at least one dose and 46.9% for individuals that received a completed vaccine series).

Table 6. Comparison of vaccination estimates for at least one dose with capture-recapture adjusted estimates and age-standardized CDC and DOH estimates.

State and Ages	Combined claims-IIS vaccine indicator	Capture-recapture-adjusted estimates[a]	CDC[b]	DOH	Compared to capture-recapture: % under-recording of 1+ dose in our combined claims-IIS data[a]	Compared to CDC: % under-recording of 1+ dose in our combined claims-IIS data	Compared to DOH: % under-recording of 1+ dose in our combined claims-IIS data
State 1, Ages 0 – 64	49.1%	56.2%	59.0%	57.3%	12.6%	16.8%	14.3%
State 2, Ages 0 – 64	48.5%	57.5%	61.5%	NA	15.7%	21.1%	NA
State 3, Ages 5 – 64	48.3%	74.2%	86.5%	79.6%	38.7%	44.2%	39.3%
State 4, Ages 0 – 64	38.2%	50.7%	61.5%	59.8%	24.7%	37.9%	36.1%
State 5, Ages 5 – 64	53.9%	61.7%	74.6%	69.5%	17.2%	27.7%	22.4%
State 6, Ages 0 – 64	52.7%	66.3%	89.0%	82.6%	20.5%	40.8%	36.2%
State 7, Ages 5 – 64	61.8%	63.8%	72.8%	68.8%	9.2%	15.1%	10.2%
State 8, Ages 5 – 64	52.6%	62.7%	78.9%	63.7%	20.4%	33.3%	17.4%
State 9, Ages 0 – 64[c]	46.7%	64.2%	88.3%	72.1%	27.3%	47.1%	35.2%
State 10, Ages 5 – 64[d]	52.3%	59.8%	64.8%	65.0%	19.9%	19.3%	19.5%
State 11, Ages 12 – 64	59.1%	63.8%	85.4%	76.9%	17.5%	30.8%	23.1%

[a] Estimates for this column were for ages 0–64 years. All other columns are restricted to the ages indicated for each row, unless otherwise specified.
[b] CDC data were from 11/3/2021 for State 1 and State 10 to match the last date of service in the IIS data and 1/13/2022 for all other states.
[c] State 9 age-stratified vaccine data cover 0 to 69 years of age.
[d] State 10 age-stratified vaccine data cover 5 to 69 years of age.
Note: Some jurisdictions cover the 5-64 years age range because that was the age range available for the age-specific jurisdiction DOH data. Combined claims-IIS vaccine estimates were calculated to match these same age ranges. Percent under-recording is calculated as the difference between the combined claims-IIS estimate and other external estimate, divided by the other external estimate.
NA = not available on the DOH website.

Table 7. Comparison of our completed series vaccination estimates with capture-recapture adjusted estimates and age-standardized CDC and DOH estimates.

State and Ages	Combined claims-IIS vaccine indicator	Capture-recapture-adjusted estimates[a]	CDC[b]	DOH	Compared to capture-recapture: % under-recording of completed series in our combined claims-IIS data[a]	Compared to CDC: % under-recording of completed series in our combined claims-IIS data	Compared to DOH: % under-recording of completed series in our combined claims-IIS data
State 1, Ages 0 – 64	44.6%	53.9%	51.0%	NA	17.3%	12.5%	NA
State 2, Ages 0 – 64	43.3%	53.3%	49.7%	NA	18.8%	12.9%	NA
State 3, Ages 5 – 64	37.6%	72.1%	69.3%	70.8%	50.9%	45.7%	46.9%
State 4, Ages 0 – 64	32.8%	49.0%	50.7%	NA	33.1%	35.3%	NA
State 5, Ages 5 – 64	49.1%	60.0%	65.4%	62.0%	22.3%	24.9%	20.8%
State 6, Ages 0 – 64	45.4%	63.9%	73.1%	71.7%	29.0%	37.9%	36.7%
State 7, Ages 5 – 64	58.2%	61.4%	66.2%	64.0%	11.1%	12.1%	9.1%
State 8, Ages 5 – 64	46.7%	60.3%	63.3%	52.1%	26.5%	26.2%	10.4%
State 9, Ages 0 – 64[c]	42.1%	63.7%	74.5%	65.2%	33.9%	43.5%	35.4%
State 10, Ages 5 – 64[d]	47.3%	56.9%	56.6%	56.1%	23.7%	16.4%	15.7%
State 11, Ages 12 – 64	54.7%	62.6%	76.6%	70.5%	22.8%	28.6%	22.4%

[a] Estimates for this column were for ages 0–64 years. All other columns are restricted to the ages indicated for each row, unless otherwise specified.
[b] CDC data were from 11/3/2021 for State 1 and State 10 to match the last date of service in the IIS data and 1/13/2022 for all other states.
[c] State 9 age-stratified vaccine data cover 0 to 69 years of age.
[d] State 10 age-stratified vaccine data cover 5 to 69 years of age.
Note: Some jurisdictions cover the 5-64 years age range because that was the age range available for the age-specific jurisdiction DOH data. Combined claims-IIS vaccine estimates were calculated to match these same age ranges. Percent under-recording is calculated as the difference between the combined claims-IIS estimate and other external estimate, divided by the other external estimate. NA = not available on the state website.

IV. Discussion

This study demonstrated the feasibility of obtaining and linking IIS data to claims data for timely and complete capture of COVID-19 vaccine records. We found that the linked claims-IIS data substantially improved the ascertainment of COVID-19 vaccine administration compared to claims alone. However, even with the combined claims-IIS data, the ascertainment of COVID-19 vaccination is incomplete, and estimates of remaining misclassification were obtained from a comparison with CDC, DOH, and capture-recapture estimates.

The percent of individuals with at least one dose of a COVID-19 vaccine increased by 15 percentage points when we augmented claims with IIS data. Without the incorporation of IIS data into insurance claims, considerable misclassification of vaccination status would occur, particularly vaccinated individuals being counted as unvaccinated. This misclassification would make COVID-19 vaccines appear less effective in the context of an effectiveness study, because some individuals in the unvaccinated group could have received the vaccine, reducing the contrast between the vaccinated and unvaccinated groups.[5] Also, this misclassification could limit statistical power for vaccine safety studies because of smaller sample sizes of vaccinees, which could impair the ability to detect rare adverse events.

We hypothesized that IIS data would provide a more complete source for vaccination information and would be inclusive of the majority of the claims-based vaccine records. However, 16% of individuals identified as vaccinated had vaccine records in the claims only and were not present in the IIS data. This may be explained by variation in IIS operations and technology capacity. While functional standards are available,[9] adoption of standards and modernization of technology and data quality may vary by IIS capacity. Further, there is variation across states and jurisdictions in IIS reporting requirements (mandatory vs. optional, ability of individuals to opt out, etc.).[15] Also, jurisdiction matching algorithms may not be the optimal. Each IIS identified members in their data using a variety of matching methods, and the ability and accuracy of matching was driven by the availability of underlying variables that allow for matching. Further, individuals identified in our analyses as present in claims and not IIS could be explained by the individual crossing into another state for his/her vaccination. Heterogeneity across IIS sites was also found in a 2017 CDC Morbidity and Mortality Weekly Report (MMWR) that analyzed data completeness among IIS for childhood vaccination data. They found that “30 of 55 IIS produced 7-vaccine series coverage estimates that were at least 10 percentage points lower than the corresponding National Immunization Survey-Child coverage estimates in 2016, suggesting incompleteness of IIS data.”[16] Regardless of the reason for IIS not capturing some claims-based vaccine records, this finding supports the need for using a variety of data sources to estimate vaccination coverage and conduct data quality checks when linked claims-IIS data are used.

The combined claims-IIS estimates were 9.2% to 50.9% lower than the capture-recapture adjusted estimates, varying by state. One of the assumptions of the capture-recapture analysis is that the capture mechanisms for IIS and claims vaccine records are independent.[7] This assumption is likely not satisfied because providers who submit a claim are also asked to report to IIS, although this is voluntary in some states. Because the assumption of independence between data sources is violated, the calculation will underestimate misclassification of vaccination.

State-level combined claims-IIS estimates were 12.1% to 45.7% lower than the age-standardized CDC estimates. The CDC vaccination data includes vaccination records from not only IIS, but also territory, tribe, and local entities, retail pharmacies; long-term care facilities; dialysis centers participating in the Federal Dialysis Center Program; Federal Emergency Management Agency partner sites; Health Resources and Services Administration partner sites; and other federal entity

facilities (e.g., Department of Defense, Veterans Health Administration, Indian Health Service, and Bureau of Prisons). Because we age-standardized our CDC estimates using wide age groupings (for the adult population), this method may not have fully accounted for age differences between populations, and we did not control for payer and underlying health conditions. Commercially insured individuals under age 65 may be healthier than the underlying population that CDC monitors for public health surveillance. CDC estimates potentially overestimate vaccination because CDC received vaccine records on a rolling basis from multiple sources that could include duplicates.[10-12] Vaccine records submitted to the CDC are de-identified with a unique identification number and are not always linkable, sometimes precluding deduplication. Also contributing to overestimation, CDC's denominator may sometimes be too small, based on outdated census counts.[17]

State-level combined claims-IIS estimates were 9.1% to 46.9% lower than the age-standardized DOH estimates. A strength of the DOH data is that it can more easily link multiple doses to an individual resident of the jurisdiction than the CDC, which makes DOH estimates not as susceptible to duplicates. Also, more granular age stratifications were typically available for DOH estimates, allowing for better age-standardization to our population. However, as the IIS data are one component in our vaccination database and are generally what DOH estimates are based on, the estimates may be vulnerable to the same missingness. Also notable, IIS data presented on DOH websites were not specific to the commercial population and included vaccination of the uninsured or those covered by other payers such as Medicaid, which could explain the differences between the estimates. Beyond informing potential misclassification, the comparison with DOH data serves as an internal validation of our own data, driving conversations with IIS to understand discrepancies.

Finally, we recommend that researchers using IIS data conduct quality checks for data anomalies. Reason being, IIS are data repositories and, ultimately, the vaccine administrators are responsible for the data they enter into IIS. If a researcher identifies a major data anomaly within IIS data, we recommend notifying the IIS because IIS may notify vaccine administrators.

A. Strengths

In this feasibility study, we established a workflow for IIS data engineering, management, integration, and validation and have successfully enhanced a medical and pharmacy claims database with data from 14 IIS, representing 11 states in the US. This study has several strengths. First, we successfully linked IIS data to claims with a short lag time (time between vaccine administration and observed data is typically less than a month) for timely capture of COVID-19 vaccine records. Second, this study established processes that can be leveraged to scale-up to link additional IIS data to the claims databases of the BEST data partners to benefit public health surveillance. Third, the addition of the IIS data significantly improved the capture of COVID-19 vaccine administrations compared to claims alone. Finally, the findings of this study may add rigor to safety and effectiveness studies of COVID-19 vaccines under the BEST Initiative.

B. Limitations

Limitations of this study warrant discussion. First, different IIS implemented different matching protocols to link to members, which may contribute to differences in matching quality. We are learning more about the linkage practices of each IIS (probabilistic vs. deterministic matching, variables used, how to handle multiple matches) and assessing these practices in relation to matching quality. Second, this study required only one day of medical/pharmacy coverage during

the study period, which would miss vaccinations in claims data for a member who disenrolled from health insurance during the study period and subsequently received a vaccine. To investigate the impact of this on results, a post-hoc secondary analysis restricted to individuals with continuous enrollment during the study period (unless a shorter period due to death). With this restriction in place, the vaccination coverage estimates increased from 48% to 56% for at least one dose and from 42% to 49% for completed series (data not shown). Thus, our estimates of under-recording are likely overestimates for studies that restrict to continuously enrolled individuals and would result in larger bounds for quantitative bias analyses than may be the truth. Finally, vaccines received outside of an individual's state that did not result in a claim would be missed, with the exception of states that have agreements to share IIS data with other, often neighboring, states. To assess the impact of this on results, a post-hoc secondary analysis restricted to individuals with claims from a single state was conducted. The rationale for this was that out-of-state claims could be a surrogate for someone who was likely to receive out-of-state medical care, like a COVID-19 vaccine. With this restriction in place, the vaccination coverage estimates increased from 48% to 49% for at least one dose and from 42% to 43% for completed series (data not shown). Thus, this provides some evidence that the impact of cross-state migration on vaccination coverage estimates in this study is minimal.

C. Conclusions

Our findings support the need for using a variety of data sources for estimating vaccination coverage. Additionally, this report quantifies the capture of vaccine information in the linked claims-IIS databases and informs strategies to account for incomplete vaccine data in safety and effectiveness studies. We are currently expanding this feasibility study to claims linkage with other IIS jurisdictions to better reflect the US commercially insured population. We continue to collaborate with IIS and federal stakeholders to develop a modernized, interoperable immunization data infrastructure to improve surveillance of vaccines.

Appendix A: Steps for IIS and Claims Vaccine Record Cleaning

Step #	Data Cleaning Step Description
1	**Linked IIS vaccination data to claims**
2	**Transformed claims vaccine data (CPT®, NDC, HCPCS codes, and ICD-10 procedure codes) to CVX codes and dose number for the purposes of deduplication, but kept the original vaccine codes for analyses**
3	**Restricted data to COVID-19 vaccine codes within appropriate time frame:** 1) Restricted to COVID-19 vaccine codes: CVX codes 207 (mRNA-1273), 208 (BNT162b2), 212 (Ad26.COV2.S), 213 (unknown vaccine type), 217 (BNT162b2 tris-sucrose formulation for ages 12+), 218 (BNT162b2 tris-sucrose formulation for ages 5 to <12), and 219 (BNT162b2 tris-sucrose formulation for ages 2 to <5) 2) Removed records with blank member ID or date of service 3) Removed records with dates before vaccine approval/authorization: If CVX Code = 207, then removed records prior to 18 December 2020 If CVX Code = 208, then removed records prior to 11 December 2020 If CVX Code = 212, then removed records prior to 27 February 2021 If CVX Code = 213, then removed records prior to 11 December 2020 If CVX Code = 217, then removed records prior to 03 September 2021 If CVX Code = 218, then removed records prior to 29 October 2021 If CVX Code = 219, then removed records because it was not yet authorized
4	**Deduplicated vaccine records within each data source,* prioritized non-missing dose number** 1) Within claims and IIS data sources, records were grouped by member ID, date of vaccine, CVX vaccine code, and dose number. Observations with duplicates of all four variables were removed. 2) Records with the same member ID, date of service, and CVX code were then further deduplicated, with priority on keeping records with non-missing dose number. In other words, we only removed records that had a missing dose number when there was another record with a non-missing dose number, with the same CVX and record date for that member. But if there were two records on the same date, same CVX, with different non-missing doses (e.g., dose 3 and dose 2), both were kept.
5	**Deduplicated vaccine records, prioritized by data source** 1) Set up data source hierarchy rules: If the source was commercial pharmacy claims, then priority was set to '4' If the source was commercial medical claims, then priority was set to '3' If the source was IIS, then priority was set to '2' If the source was Medicare, then priority was set to '1' 2) Records were grouped by member ID, date of service, and CVX vaccine code. Then, these records were deduplicated, keeping only one record with the same date of service and CVX vaccine code per member ID, with priority on keeping the record highest in the data source hierarchy (4 being highest).

*IIS data was treated as a single source

Appendix B: Variable Definitions

Variable	Description
State of residence	Based on pre-adjudicated claims data. Claims were searched between 01 January 2017 and 31 December 2021 for all associated values. State was set to "Multiple States" if there was more than one state of interest per patient.
Age	Based on pre-adjudicated claims and IIS data. Age calculated on 31 December 2021. When member records indicated differing dates of birth or, when both were missing, it was set to unknown.
Sex	Based on pre-adjudicated claims and IIS data. When member records indicated differing sex or, when both were missing, it was set to unknown.
Type of vaccine	If a member had only mRNA-1273 vaccine records, they are in the mRNA-1273 category. Only BNT162b2, they are in the BNT162b2 category. Only Ad26.COV2.S, they are in the Ad26.COV2.S category. If someone had only unknown and another type of vaccine, they went in the known type of vaccine category (e.g., BNT162b2 and unknown would be considered BNT162b2). If someone had at least two different known types of vaccine (e.g., Ad26.COV2.S and BNT162b2, BNT162b2 and BNT162b2, or Ad26.COV2.S and BNT162b2) they were classified as "mixed" type of vaccine.
Dose number	Dose number was determined using the date of service/vaccine administration. The first dose was defined as the vaccine record with the earliest administration date and the second dose was defined as the second record with an administration date at least three days apart from the first dose.
Month of first dose	The month-year the first vaccine dose was administered. Those who did not have a vaccine record went into unvaccinated. Those whose data sources (IIS, claims) had conflicting dates had a 'Conflicting dates' value.
Month of last observed dose	The month-year the last observed vaccine dose was administered. Those who did not have a vaccine record went into unvaccinated. Those whose data sources (IIS, claims) had conflicting dates had a 'Conflicting dates' value.

References

1. Goralnick E, Kaufmann C, Gawande AA. Mass-Vaccination Sites - An Essential Innovation to Curb the Covid-19 Pandemic. *N Engl J Med.* 2021;384(18):e67.
2. Franklin JM, Lin KJ, Gatto NM, Rassen JA, Glynn RJ, Schneeweiss S. Real-World Evidence for Assessing Pharmaceutical Treatments in the Context of COVID-19. *Clin Pharmacol Ther.* 2021;109(4):816-828.
3. Mayo Clinic. U.S. COVID-19 vaccine tracker: See your state's progress. https://www.mayoclinic.org/coronavirus-covid-19/vaccine-tracker.
4. Centers for Disease Control and Prevention (CDC). COVID Data Tracker: COVID-19 Vaccinations in the United States. https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-rate-total.
5. De Smedt T, Merrall E, Macina D, Perez-Vilar S, Andrews N, Bollaerts K. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. *PLoS One.* 2018;13(6):e0199180.
6. Centers for Disease Control and Prevention. About Immunization Information Systems. https://www.cdc.gov/vaccines/programs/iis/about.html. Accessed March 15, 2022.
7. Seeger JD, Schumock GT, Kong SX. Estimating the rate of adverse drug reactions with capture-recapture analysis. *Am J Health Syst Pharm.* 1996;53(2):178-181.
8. American Immunization Registry Association. *An Environmental Scan on Progress, Challenges, and Opportunities: Expanding Immunization Information Systems for Adults in the United States.* July 2020.
9. Centers for Disease Control and Prevention. Immunization Information System (IIS) Functional Standards, v4.1. https://www.cdc.gov/vaccines/programs/iis/functional-standards/func-stds-v4-1.html. Accessed March 17, 2022.
10. Centers for Disease Control and Prevention (CDC). Reporting COVID-19 Vaccinations in the United States.
11. Wingrove J. Vaccine Data Gaps Point to Millions More in U.S. Who Lack Shots. *Bloomberg.* December, 18, 2021.
12. AP News. Data correction drops COVID vaccine rates in Pennsylvania. 2021.
13. Rothman K, Greenland S, Lash T. *Modern Epidemiology.* Third ed: Lippincott Williams & Wilkins; 2008.
14. Centers for Disease Control and Prevention. CDC Expands Eligibility for COVID-19 Booster Shots to All Adults. https://www.cdc.gov/media/releases/2021/s1119-booster-shots.html. Accessed March 18, 2022.
15. Martin DW, Lowery NE, Brand B, Gold R, Horlick G. Immunization information systems: a decade of progress in law and policy. *J Public Health Manag Pract.* 2015;21(3):296-303.
16. Murthy N, Rodgers L, Pabst L, Parker Fiebelkorn A, Ng T. *Morbidity and Mortality Weekly Report. Progress in Childhood Vaccination Data in Immunization Information Systems — United States, 2013–2016.* US Department of Health and Human Services/Centers for Disease Control and Prevention;2017.
17. Bump P. It's probable that fewer seniors have been vaccinated than the government claims. *The Washington Post*.