

Methodology on Creating the Linked Retail Health Clinic (LiRHC) Database



DISCLAIMER. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the National Center for Health Statistics, the Centers for Disease Control and Prevention, or the U.S. Census Bureau . The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 7529814: CBDRB-FY23-CES019-007).

CENSUS BUREAU

ALICE ZAWACKI

PRESENTER

JOEY MARSHALL



United States[®]
Census
Bureau

NATIONAL CENTER FOR
HEALTH STATISTICS

DONALD CHERRY

XIANGHUA YIN

BRIAN W. WARD



Federal Committee on
Statistical Methodology

October 25, 2023

What are
RHCs and
why do we
care about
them?

**“An in-store clinic with
healthcare professional(s) who
provides medical care”**

(2017 Census of Retail Trade, Census Bureau)

Services

- Minor illnesses/injuries
- Health screenings
- Medication management
- Immunizations

Consumer appeal

- Convenient locations and hours
- Walk-in accessibility
- Short wait times
- Transparent pricing

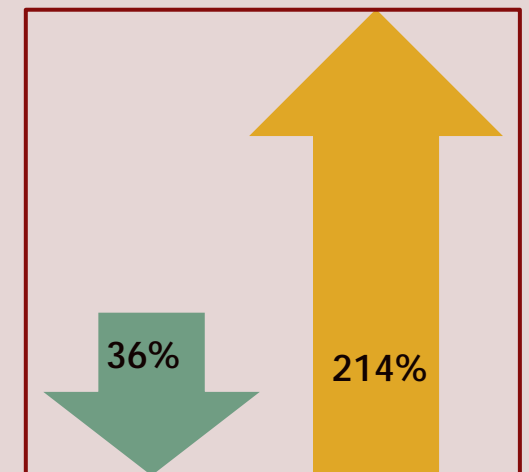


What are **RHCs** and why do we care about them?

Role of RHCs?

- Population health
- Healthcare access
- Healthcare provision

2008 → 2015
Change in visit rate for
low-acuity conditions



Hospital
Emergency
Departments

RHCs

Poon, Schuur, and Mehrotra (2018)



Interagency Collaboration

CENSUS BUREAU & NCHS COLLABORATION

Agencies jointly saw benefits in creating a database of RHCs and their locations throughout the United States

Benefits

- Create new dataset
 - Nationally representative sample of RHCs
 - Support analysis of this relatively “new” health care setting
- Produce new estimates describing areas & populations near RHCs
- Potential use
 - Solicit EHR data from RHCs
 - Collect National Ambulatory Medical Care Survey data

2018-2020 LiRHC Database

2,000

Number of RHCs operating

97%

RHC share in MSAs

47%

Share in the South region

Methods

Data

1. **Convenient Care Association (CCA)**
 - Members name and address
 - 2018-2020, multiple pulls each year
 - RHCs and other walk-in health centers
2. **Census Bureau's Business Register (BR)**
 - Universe of U.S. businesses
 - Name, location, payroll, employment size, industry, etc.
3. **National Plan and Provider Enumeration System (NPPES)**
 - Centers for Medicare and Medicaid Services
 - Includes individual health care providers, organizations
 - Evaluate match quality, distinguish clinic type



Methods

Matching

Probabilistic matching: Similarity of addresses

STEPS

1. Preprocess original records
2. Link CCA + BR
3. Match NPPES
4. Postprocess matched records

Methods

Matching

STEPS

1. Preprocess original records: CCA, BR, NPPES
2. Link CCA + BR
3. Match NPPES
4. Postprocess matched records

Methods

Matching

1. **Preprocess original records: CCA, BR, NPPES**
 - Standardize basic address text
 - Remove address-related “stop words”
 - Concatenate address and city names
 - Deduplicate CCA → 1 Observation/Year
2. Link CCA + BR
3. Match NPPES
4. Postprocess matched records

Methods

Matching

1. Preprocess original records
2. Link CCA + BR
3. Match NPPES
4. Postprocess matched records



Methods

Matching

1. Preprocess original records
2. **Link CCA + BR**
 - Use Fellegi-Sunter probabilistic linkage model
 - Enamorado, Fifield, and Imai, 2019
 - *fastLink* R library
 - Match records within each state-year block
 - Conduct 8 sequential matching passes
 - Address “next-door neighbor” problem
3. Match NPPES
4. Postprocess matched records

Methods

Matching

- Preprocess original records
 1. **Link CCA + BR**
 - Use Fellegi-Sunter probabilistic linkage model
 - Match records within each state-year block
 - Conduct 8 sequential matching passes
 - Address “next-door neighbor” problem
 2. Match NPPES
 3. Postprocess matched records

Methods

Matching

- Preprocess original records
 1. **Link CCA + BR**
 - Use Fellegi-Sunter probabilistic linkage model
 - Match records within each state-year block
 - Conduct 8 sequential matching passes
 - Address “next-door neighbor” problem
 2. Match NPPES
 3. Postprocess matched records

Methods

Matching

1. Preprocess original records
2. **Link CCA + BR**
 - Use Fellegi-Sunter probabilistic linkage model
 - Match records within each state-year block
 - Conduct 8 sequential matching passes
 - Same address including exact number match
 - Mixed-use retail settings
 - Accept matches with quality score > 14
 - Remove year and state blocking
 - Address “next-door neighbor” problem
3. Match NPPES
4. Postprocess matched records

Methods

Matching

- Preprocess original records
 1. **Link CCA + BR**
 - Use Fellegi-Sunter probabilistic linkage model
 - Match records within each state-year block
 - Conduct 8 sequential matching passes
 - Address “next-door neighbor” problem
 2. Match NPPES
 3. Postprocess matched records

Methods

Scoring Matches

Match Quality Score >14 → Accept match

Score	Criteria
23	Maximum possible score

Methods

Matching

Match Quality Score >14 → Accept match

Score	Criteria
+5	BR name references a known RHC or contains "clinic" or "health"
+5	All numbers in address match ("123 main, Suite 500" = 123500)
23	Maximum possible score



Methods

Matching

Match Quality Score **>14** → Accept match

Score	Criteria
+5	CBPBR name references a known RHC or contains "clinic" or "health"
+5	All numbers in address match ("123 main, Suite 500" = 123500)
+3	Address number match
+3	BR name = Name of known grocer or major retailer
+3	Industry code = Healthcare provider, retail trade
23	Maximum possible score



Methods

Matching

Match Quality Score **>14** → Accept match

Score	Criteria
+5	CBPBR name references a known RHC or contains "clinic" or "health"
+5	All numbers in address match ("123 main, Suite 500" = 123500)
+3	Address number match ("123 Main, Suite 500" = 123)
+3	CBPBR name = Name of known grocer or major retailer
+3	Industry code = Healthcare provider, retail trade
+1	Positive payroll (BR)
+1	Positive employment (BR)
+1	Active business (BR)
+1	Included in Census Bureau tabulations
23	Maximum possible score

Methods

Matching

- Preprocess original records
- Link CCA + BR
- **Match 2018 NPPES to CCA-BR matches**
 - Probabilistic street address match with(out) state block
- Postprocess matched records

Methods

Matching

1. Preprocess original records
2. Match CCA + BR
3. Match 2018 NPPES to CCA-BR matches
4. **Postprocess records**
 1. Deduplicated RHCs operating in multiple years
 2. Remove unmatched CCA records

Methods

Quality Assessments

1. Benchmarking
2. Match statistics
3. Manual review



Methods

Quality Assessments

2,008 RHCs
(Merchant
Medicine)

1,949 RHCs
(Statista)



2018-2020 LiRHC

2,000 RHCs operating



Methods

Quality Assessments

1. Benchmark
2. Review match statistics

95% CCA records match BR

Match Quality Score (0-100)	Percent
High (75-100)	81
Moderate to low (0-74)	19

Pass Number	Percent
1 or 2	92
3-8	8

Source. Based on authors' merging of the 2018-2020 Convenient Care Association data (CCA) with the Census Bureau's 2017-2010 Business Register (BR).

Firm name assigned to CCA record vs. firm name in BR

3. Manual review

Methods

Quality Assessments

1. Benchmarking
2. Match statistics
3. **Manual review**

500 randomly selected matches

- 1 CCA record → > 1 BR record
- NPPES taxonomy vs. BR industry
- NPPES taxonomy → Urgent care centers?
- Geographic distribution by match status

LiRHC

97%
Metropolitan Statistical Areas



Linked Retail Health Clinic (LiRHC) Locations by U.S. Census Region and Division: 2018-2020



Note: Division totals exceed 2,000 due to rounding, which is required under the Census Bureau's disclosure avoidance rules. Alaska and Hawaii are in the Pacific division.

Source: Based on authors' merging of the 2018-2020 Convenient Care Association data, the 2017-2020 County Business Patterns Business Register, and the 2018 National Plan and Provider Enumeration System data from the Centers for Medicare and Medicaid Services.



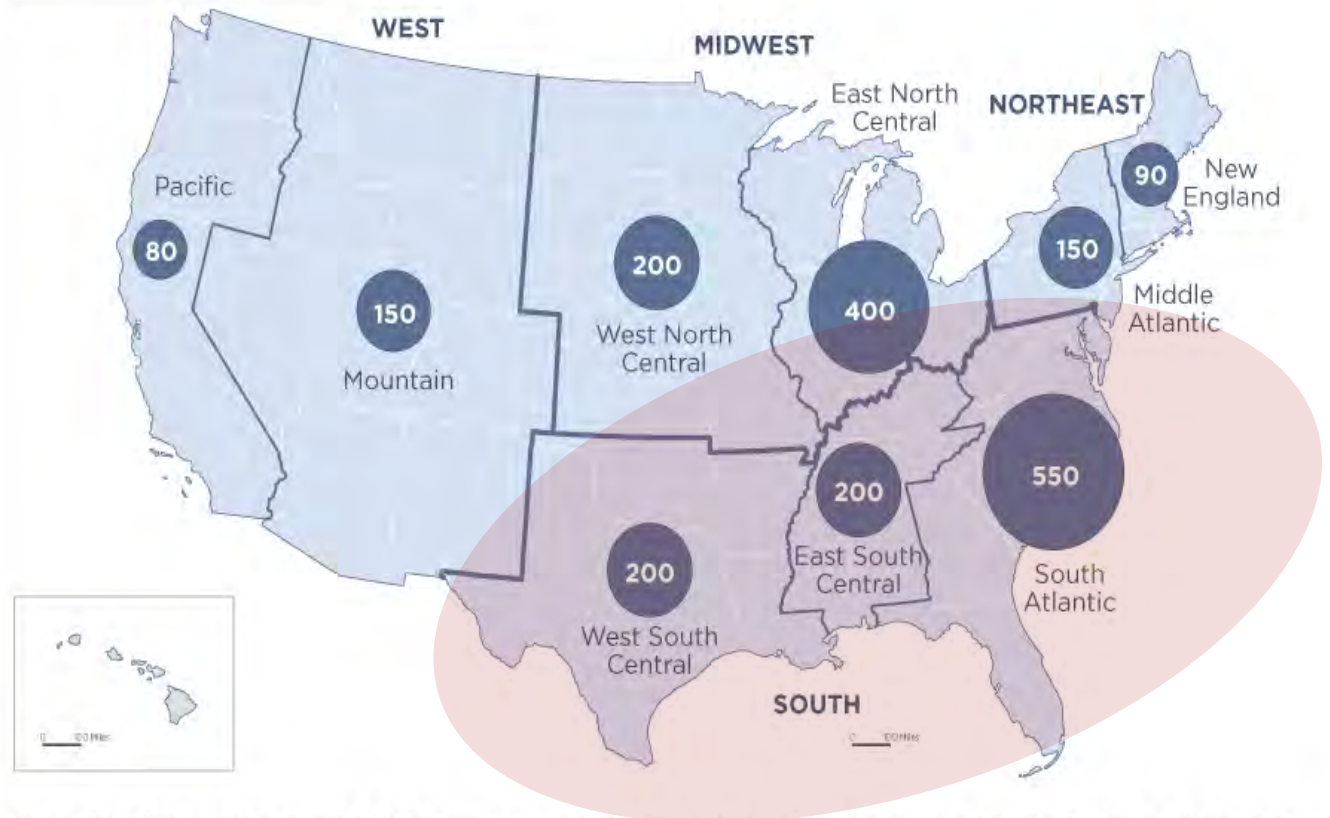
LiRHC

47% South region

27% South Atlantic division



Linked Retail Health Clinic (LiRHC) Locations by U.S. Census Region and Division: 2018-2020



Note: Division totals exceed 2,000 due to rounding, which is required under the Census Bureau's disclosure avoidance rules. Alaska and Hawaii are in the Pacific division.

Source: Based on authors' merging of the 2018-2020 Convenient Care Association data, the 2017-2020 County Business Patterns Business Register, and the 2018 National Plan and Provider Enumeration System data from the Centers for Medicare and Medicaid Services.



REGION	SHARE	COUNT
South	47%	950
Midwest	30%	600
Northeast	12%	240
West	11%	230
TOTAL	100%	2,020



Linked Retail Health Clinic (LiRHC) Locations by U.S. Census Region and Division: 2018-2020



Note: Division totals exceed 2,000 due to rounding, which is required under the Census Bureau's disclosure avoidance rules. Alaska and Hawaii are in the Pacific division.

Source: Based on authors' merging of the 2018-2020 Convenient Care Association data, the 2017-2020 County Business Patterns Business Register, and the 2018 National Plan and Provider Enumeration System data from the Centers for Medicare and Medicaid Services.

Next Steps

Using LiRHC

SECOND OBJECTIVE

Use LiRHC to improve understanding on how RHCs might affect population health and access and provision of healthcare services



Next Steps

Using LiRHC

SECOND OBJECTIVE

Use LiRHC to improve understanding on how RHCs might affect population health and access and provision of healthcare services

Characteristics of nearby populations?

Demographic ~ Socioeconomic
Insurance coverage ~ Presence of children

Characteristics of other healthcare providers?

Types ~ Distribution ~ Visits
Revenues ~ Payroll ~ Employment size



United States Census Bureau

Partners Researchers Educators Survey Respondents News NAICS Codes Jobs About Us Contact Us Help

Topics Data & Maps Surveys & Programs Resource Library Search data, events, resources, and more

Is There a Retail Health Clinic Near You?

March 28, 2023
Written by: Alice Zawacki, Joey Marshall, Donald Cherry, Xianghua Yin and Brian W. Ward

Two thousand Retail Health Clinics Provided Walk-In Health Services Mostly in Metropolitan Areas

March 28, 2023
Written by: Alice Zawacki, Joey Marshall, Donald Cherry, Xianghua Yin and Brian W. Ward

Facebook Twitter LinkedIn

[Is There a Retail Health Clinic Near You? \(census.gov\)](https://www.census.gov/newsroom/2023-03-28-retail-health-clinics.html)

United States Census Bureau

Partners Researchers Educators Survey Respondents News NAICS Codes Jobs About Us Contact Us Help

Topics Data & Maps Surveys & Programs Resource Library Search data, events, resources, and more

Census.gov / Census (History) Pages / Methodology on Creating the U.S. Linked Retail Health Clinic (LiRHC) Database

Methodology on Creating the U.S. Linked Retail Health Clinic (LiRHC) Database

March 2023
Written by: ALICE ZAWACKI, JOEY MARSHALL, DONALD CHERRY, XIANGHUA YIN, AND BRIAN W. WARD
Working Paper Number CEP-23-19

Facebook Twitter LinkedIn

Abstract

Retail health clinics (RHCs) are a relatively new type of health care setting and understanding the role they play as a source of ambulatory care in the United States is important. To better understand these settings, a joint project by the Census Bureau and National Center for Health Statistics used data science techniques to link together data on RHCs from Convenient Care Association, County Business Patterns Business Register, and National Plan and Provider Enumeration System to create the Linked RHC (LiRHC, pronounced "lync") database of locations throughout the United States during the

Download Methodology on Creating the U.S. Linked Retail Health Clinic (LiRHC) Database (PDF - 1MB)

Related Information
Center for Economic Studies (CES) Working Paper Series
PUBLICATION
Discussion Paper Series

[Methodology on Creating the U.S. Linked Retail Health Clinic \(LiRHC\) Database \(census.gov\)](https://www.census.gov/newsroom/2023-03-28-retail-health-clinics.html)

CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

National Center for Health Statistics

NCHS and the U.S. Census Bureau are partnering to learn more about retail health clinics as their reach and popularity grow. Read about the project.

Thank you

Alice.M.Zawacki@Census.gov

