

# **Integrated Data for Population Health Research: The New Jersey iPHD**

**Institute for Health Brown Bag Seminar**

**March 16, 2023**

**Joel C. Cantor, ScD**

**Jose Nova, MS**

# Outline

---

- About the iPHD
  - Establishment and governance
  - Research priorities
  - Data sources
  - Data linkage strategy
- 2022 Cycle I Application Progress Report
  - Approved projects
- 2023 Cycle II Application Opportunities
- Discussion: Building iPHD Value and Sustainability

# Acknowledgements

---

CSHP colleagues Manisha Agrawal, Oliver Lontok, Kate Scotto, Jolene Chou, Joe Brecht, and Margaret Koller have made major contributions to iPHD development and operations

The iPHD Governing Board members have worked diligently to guide and support iPHD implementation, we especially acknowledge the dedication of Board Chair Rachel Hammond from NJDOH and thank the agency's data stewards for their essential contributions

Rick Hall, Julian Rodriguez and Todd Rossi from ChoiceMaker LLC have provided their expertise in the initial iPHD data linkage process

Funding for the iPHD has been provided by the NJDOH with additional resources from the Robert Wood Johnson Foundation

# iPHD Overview

---

- The **Integrated Population Data Health Data (iPHD) Project** was established under NJ P.L. 2015, c. 193, enacted Sept. 11, 2016
- Establishes a process to integrate data from publicly supported programs for population health research for the purpose of:
  - Improving **public health, safety, security and well-being** of NJ residents
  - Improving **cost-efficiency of government assistance programs**
- Establishes a Governing Board and process
- Authorizes operations within Rutgers Center for State Health Policy (CSHP)
- Start-up support from NJDOH, self-supporting with fees and grants over time

More information: <https://iphd.rutgers.edu/>

# iPHD Governing Board Members

<p><b>Rachel Hammond, CIPP/G, HCISPP</b>  Governing Board Chair  Designee, NJ Commissioner of Health  <i>Ex Officio</i></p>	<p><b>Joel C. Cantor, ScD</b>  Rutgers Center for State Health Policy  <i>Ex Officio</i>, Non-voting</p>	<p><b>Rashmi Jain, PhD</b>  Montclair State University  Gubernatorial Appointment  <b>Large Data Systems &amp; Data Security Expert</b></p>
<p><b>Greg Woods, MPA</b>  Designee, NJ Commissioner of Human Services  <i>Ex Officio</i></p>	<p><b>Elizabeth G. Litten, JD</b>  Fox Rothschild LLP  Gubernatorial Appointment  <b>Privacy &amp; Security Expert</b></p>	<p><b>VACANT</b>  Designee, NJ State Treasurer  <i>Ex Officio</i></p>
<p><b>Francis Baker, JD</b>  Designee, NJ Attorney General  <i>Ex Officio</i></p>	<p><b>Michele L. Norin, MEd</b>  Rutgers University, Office of Information Technology  <i>Ex Officio</i></p>	<p><b>Kathleen Noonan, JD</b>  CEO, Camden Coalition of Healthcare Providers  Senate President Appointment  <b>Advocate on behalf of persons whose data may be in the iPHD</b></p>
<p><b>Janet Currie, PhD</b>  Professor of Economics and Public Affairs, Princeton University  <b>Expert in Human Subjects Research from a NJ Research University</b></p>	<p><b>Mark McNally, JD</b>  Deputy Attorney General  NJ Office of the Attorney General  <i>Board Counsel</i></p>	

# iPHD Research Priorities

---

## Opioid Epidemic



Informing New Jersey's integrated approach to addressing the opioid epidemic.

## Maternal and Infant Health



Improving maternal and infant health.

## Social Determinants of Health



Assuring access to physical and behavioral health services and addressing social determinants of health.

## COVID-19 and Other Public Health Emergencies



Conducting analyses to support NJ's response to COVID-19 and other public health emergencies

More information: <https://iphd.rutgers.edu/research-priorities>

# iPHD Data Sources

Source	Description	Years
NJ mortality records	Cause of death and demographic attributes of NJ decedents	2000-2020
NJ birth records	Health and demographic data for newborns and mothers for NJ deliveries	2000-2020
NJ All-payer hospital uniform billing (UB) records	Utilization, diagnostic, and billing information for inpatient and ED visits at NJ acute care hospitals	2010-2020
NJ COVID-19 surveillance data	COVID-19 diagnostic lab test results (NJ CDRSS)	2020-2021
NJ Emergency Medical Services data	Medical and incident information from NJ EMS transports	2017-2022*

\*Subject to confirmation

More information: <https://iphd.rutgers.edu/data>

# iPHD Data Sources

---

## Additional Sources Under Discussion by the iPHD Governing Board

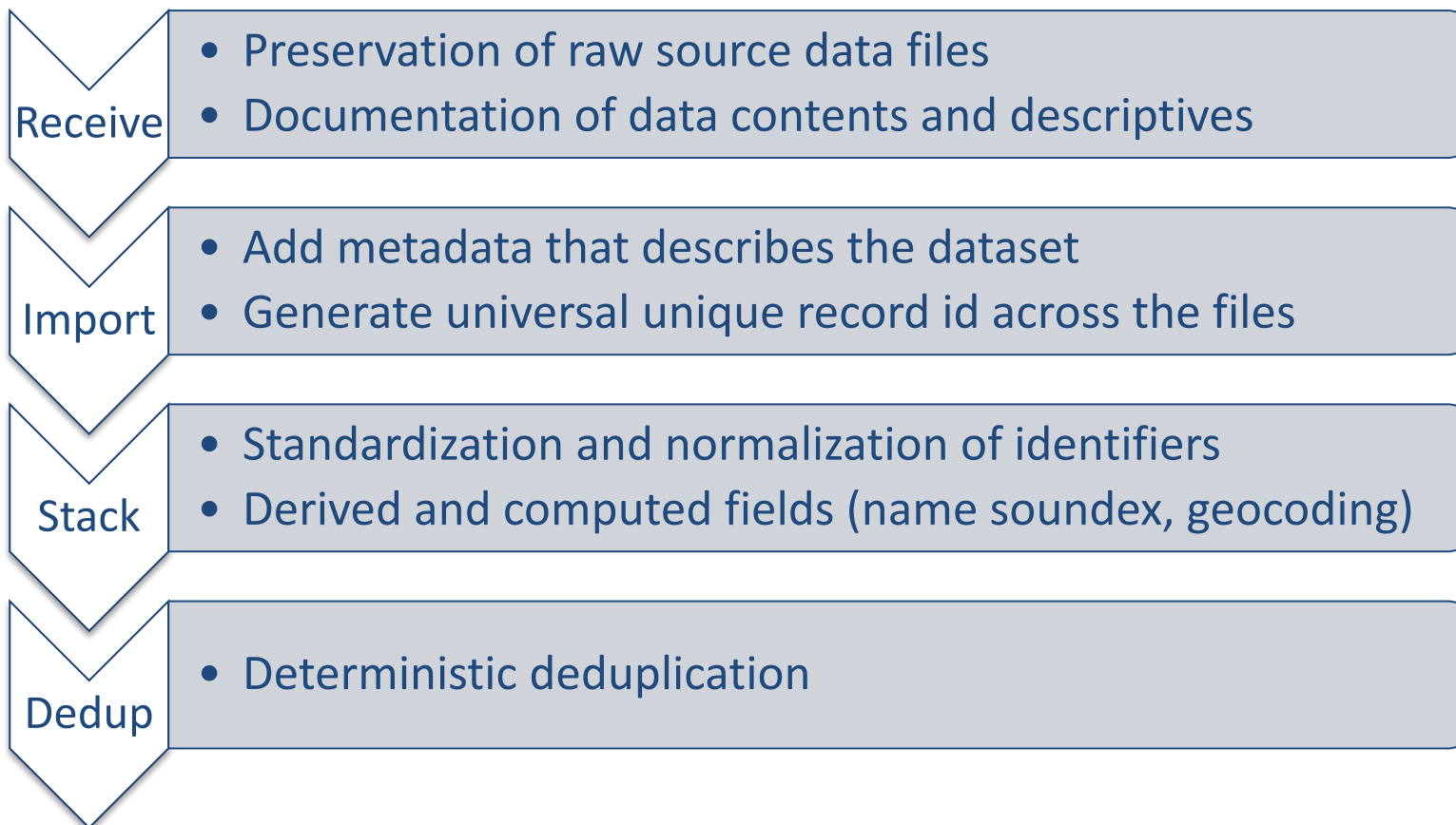
- COVID-19 vaccination registry data (NJDOH)
- NJ Medicaid enrollment, claims and encounter data (NJ DHS)



# Data Linkage

---

ETL (Extract, Transform, Load) process for data matching

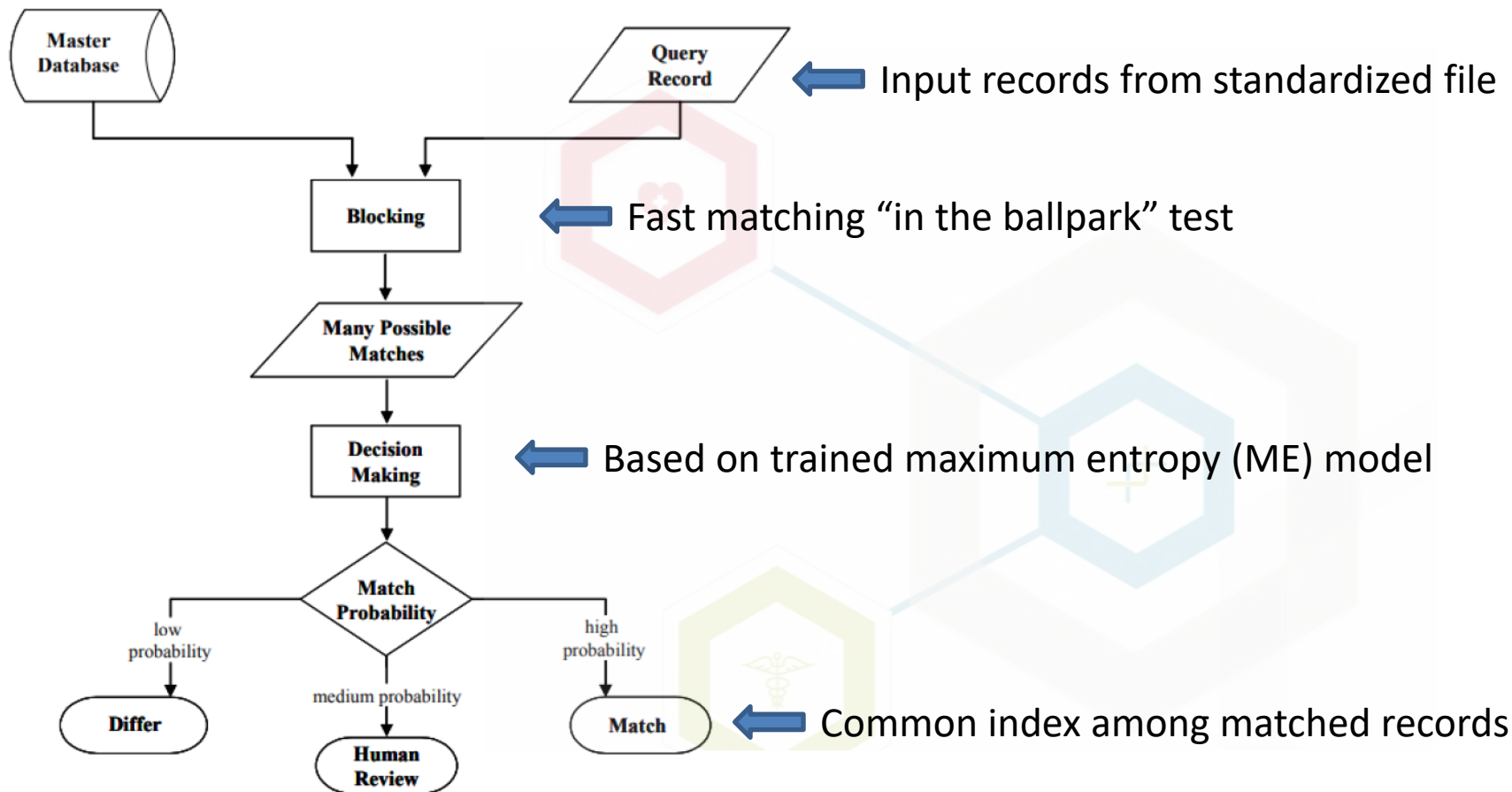


# Data Linkage

Record Matching with



**CHOICEMAKER**



# Data Linkage

---

Record Matching with  **CHOICEMAKER**

## Machine Learning and Model Development

- Built around a set of 50 – 200 “clues” or “features” indicating match or differ decisions
- The model is trained on a set of pairs of records that have been tagged as a “match”, “differ”, or “hold (unsure)”.
- The maximum entropy training process then assigns each clue a weight, which is a positive real number indicating the relative predictive strength of the clue.

$$\text{Probability} = \frac{\text{MatchProduct}}{\text{MatchProduct} + \text{DifferProduct}}$$

# Data Linkage

Record Matching with  CHOICEMAKER

## ME Example:

	Field	Query Record	Match Record	Clue name	Prediction	Weight
1	First name	Jim	Jim	First names match	Match	1.5
2	Last name	Conner	Connor	Last names differ	Differ	2.2
3	Soundex last name	C560	C560	Soundex last names match	Match	5.5

MatchProduct =  $1.5 * 5.5 = 8.25$

DifferProduct = 2.2

Probability =  $8.25 / (8.25 + 2.2) = 79\%$

Given this simple example the formula gives us a probability of 79% that “Jim Conner” and “Jim Connor” are the same person.

Weights are determined during the training by the ME engine to produce a decision that is consistent with human markings.

# Data Linkage

---

## Initial Matching and Next Steps

- Reviewing match results among about 58 million records including about 2.4 million records for birth mothers
- Creating a master person ID and relevant pseudonymous identifiers for limited datasets for initial 4 study projects
  - Constructing crosswalks for birth record child and birth mother
- Continuing to improve matching models as new data is introduced



# Outline

---

- About the iPHD
  - Establishment and governance
  - Research priorities
  - Data sources
  - Data linkage strategy
- 2022 Cycle I Application Progress Report
  - Approved projects
- 2023 Cycle II Application Opportunities
- Discussion: Building iPHD Value and Sustainability

# iPHD Cycle I Application Process & Timeline

---

- Request for applications: pilot funding and “data only” (with applicable fees) – June 1, 2022
- Applications submitted – July 15, 2022
- Review by 3 Research Advisory Committee (RAC) members
- Selection by Governing Board – October 21, 2022
- Execute subawards, data use agreements, etc. – in progress
- Data transfer – March-April 2023

# Cycle I (2022) iPHD Projects

---

## Trends in Adverse Birth Outcomes: Variations by Race and COVID-19 Exposure

**Institution:** Central Jersey Family Health Consortium

**PI:** Cheryl A. S. McFarland, PhD

**Study goal:** Compare pre-pandemic and pandemic periods adverse birth outcomes to better understand the impact COVID-19 and the state's response had on the population of pregnant women.

**iPHD data:** Birth (2019-2020), Mortality (2019-2020), Hospital UB (2018-2020), and COVID-19 surveillance data (2020)



# Cycle I (2022) iPHD Projects

---

## **Perinatal Depression and Emergency Department Visits in the Postpartum Period: A Quasi-Experimental Analysis**

**Institution:** Rutgers School of Public Health

**PIs:** Slawa Rokicki, PhD & Mark McGovern, PhD

**Study goal:** Examine demographic, social determinants, obstetric health, and mental health factors associated with overall and psychiatric ED use, with a particular focus on the impact of depression symptoms at delivery, in the postpartum period; and estimate disparities in ED use by maternal characteristics.

**iPHD data:** Birth (2010-2020) and Hospital UB (2010-2020)

# Cycle I (2022) iPHD Projects

---

## **Ensuring Programmatic Exposure and Efficacy in Areas of Greatest Need: A Geographical Study of Mental Health Outcomes and Provision of Behavioral Health Services by the New Jersey Pediatric Psychiatry Collaborative**

**Institution:** Hackensack Meridian School of Medicine

**PI:** Morgan Peltier, PhD

**Study goals:** Identify trends in hospitalizations and deaths associated with mental health disorders, investigate histories of hospitalizations prior to death, and understand how programmatic services provided by the NJ Pediatric Psychiatry Collaborative (NJPPC) have addressed the needs in communities, and where services need to improve to support communities at risk.

**iPHD data:** Mortality (2010-2020) and Hospital UB (2010-2020)

# Cycle I (2022) iPHD Projects

---

## **Social Vulnerability, Disparities, and the Health Impacts of the Intersecting COVID-19 and Opioid Epidemics on New Jersey Communities**

**Institution:** Rutgers Institute for Health

**PI:** Steve Crystal, PhD

**Study goals:** Better understand the effects of the dual epidemics of COVID-19 and opioid use disorder (OUD) on NJ communities, with a focus on incidence of selected acute healthcare events indicative of the impact of the two epidemics.

**iPHD data:** Mortality (2010-2020), Hospital UB (2010-2020), and COVID surveillance data (2020-2021)

# Outline

---

- About the iPHD
  - Establishment and governance
  - Research priorities
  - Data sources
  - Data linkage strategy
- 2022 Cycle I Application Progress Report
  - Approved projects
- 2023 Cycle II Application Opportunities
- Discussion: Building iPHD Value and Sustainability

# iPHD Cycle II Application Process & Timeline

---

- Request for applications released – March 1, 2023
  - Pilot funding (up to 2), “fee waivers” (up to 2), and “data only” (fees apply)
- Letters of intent due (mandatory, non-binding), application portal opens – March 22, 2023
- Full applications due, application portal closes – April 26, 2023
- Notification of applicable data fees – May 2023
- Awards announced – November 2023

RFA and instructions: <https://iphd.rutgers.edu/instructions-forms>



Center for State Health Policy  
*Institute for Health, Health Care Policy and Aging Research*

# Potential Additional 2023 Funding Opportunity!

---

- Discussions underway with the Robert Wood Johnson Foundation Health Data for Action (HD4A) Program
- National competitive call for proposals for research with novel state-based data, including iPHD
  - RWJF to fund grants for successful applicants
  - RWJF will fund CSHP to cover data preparation
  - Proposals due late summer or fall 2023 (TBD)
- Projects also subject to approval of the iPHD Governing Board
  - Must serve the purpose of the iPHD and address iPHD research priorities

Watch these spaces: <https://academyhealth.org/about/programs/health-data-action> and <https://iphd.rutgers.edu>

# Outline

---

- About the iPHD
  - Establishment and governance
  - Research priorities
  - Data sources
  - Data linkage strategy
- 2022 Cycle I Application Progress Report
  - Approved projects
- 2023 Cycle II Application Opportunities
- Discussion: Building iPHD Value and Sustainability



# Implementation Challenges

---

- Achieving sustainable funding
  - Operating the iPHD is costly
    - In part due to high RU fringe rate and IFH Data Core passthrough
  - NJDOH funding short-term and less than annual operating cost
  - Creating an affordable **fee schedule** is difficult
    - “Cart and horse” problem – fees come down with increased # of projects, but current fees deter potential applicants
    - Building volume through active outreach and marketing will help
  - Finding **grant support** is also challenging

# Fee Schedule Example

---

## **Social Vulnerability, Disparities, and the Health Impacts of the Intersecting COVID-19 and Opioid Epidemics on New Jersey Communities**

**Institution:** Rutgers Institute for Health

**PI:** Steve Crystal, PhD

**Study goals:** Better understand the effects of the dual epidemics of COVID-19 and opioid use disorder (OUD) on NJ communities, with a focus on incidence of selected acute healthcare events indicative of the impact of the two epidemics.

**iPHD data:** Mortality (2010-2020), Hospital UB (2010-2020), and COVID surveillance data (2020-2021)

**iPHD data fee:** \$34,250 (supported by NJDOH, in this case)

# Implementation Challenges (continued)

---

- Adding new data sources
  - On the critical path for building the value of the iPHD
  - Requires complex negotiations and extensive DUA requirements

# Questions and Discussion

---

- What additional NJ administrative datasets would enhance the **value** of iPHD for addressing its research priorities?
- What **grant sources** should we pursue to help enhance the iPHD and promote its use?