



*Resultant*

**Connecting Datasets to  
Deepen Analysis and Reduce  
Indiana's Infant Mortality Rate**



## SETTING THE STAGE

Indiana's infant mortality rate—the number of infants who die before one year of age—had surpassed the national average for several decades, despite significant efforts to reduce it. In 2012, that rate put Indiana in the bottom 20% of states, standing at 7.7 deaths per 1,000 live births, compared with 6.1 nationally—a rate that also is significantly higher than peer nations like Canada (4.8) and United Kingdom (4.1).

With forward-looking health leadership, Indiana had improved its infant mortality rate over the preceding decades but had reached a plateau. Indiana Department of Health (IDOH) undertook a significant analysis, comparing Indiana to similarly situated states with disparate infant mortality rates, and found that the health data fell short of explaining the differences. Traditional epidemiological analysis couldn't adequately address a problem so closely tied to social factors like economic status.

In 2013, the state commissioned a data-driven analysis that unified information from previously unlinked sources across agencies, and Resultant utilized sophisticated machine learning techniques to build predictive models that estimate risk for adverse birth outcomes.

Based on new insight from this refocused data analysis effort, Indiana's leadership designed programs that brought the state to its lowest-ever infant mortality rate in 2020.

**In 2012, Indiana was in the bottom 20% of states, standing at 7.7 deaths per 1,000 live births, compared with 6.1 nationally.**



## THE PROBLEM

Infant mortality is a complex social problem that cannot be isolated from education, economic factors, county of birth, and myriad other contributing elements—social determinants of health. Epidemiologists at work on the issue were getting only health data, which comprises a sliver of the full picture of infant mortality. Agencies were well aware of the infant mortality rate but lacked context such as mean age of death by county and number of prenatal visits.

State officials recognized the complex web of health and other factors that contributed to the problem and began working toward an entirely new solution. To achieve a fuller analysis that would enable more effective programs, IDOH would:

- Compile massive amounts of data from disparate sources to provide broader perspective
- Build a high-security environment to keep personally identifiable information safe
- Ensure accuracy with reliable record-matching across a wide range of agencies
- Perform advanced mathematics on joined datasets to uncover new insights
- Give the data its greatest reach and power for policymakers and analysts through user-friendly interaction and reporting tools

## THE APPROACH

Moving from narrow data to an overly broad approach would create an impenetrable mass of data that did little to clarify solutions, and so the first step was evaluating which datasets would prove most useful and address gaps in existing research.

### 1. DISCOVERY SESSIONS

Our team worked across relevant State agencies in close collaboration with leadership, subject matter experts, program directors, and system administrators. Various contributors to the project, from those who'd help analyze the data to policymakers who'd eventually apply it to programming and even leaders working in previously untapped resource agencies, worked with our team through guided discovery sessions to compile relevant details and design a way to achieve real change in Indiana.

We designed these discovery sessions to obtain relevant system and data documentation, and to gain a high-level understanding of the data collected by the agencies, collection methodologies, source data systems, potential data quality issues. This initial information-gathering process enabled us to determine the breadth of data available, system interconnectivity, and areas in which information contained in these systems overlapped.

### 2. DATASET SELECTION

From more than 100 datasets, the project team identified 22 that would best provide information and insight related to underlying causes of infant mortality and help the State obtain a fuller understanding of how to reduce its rate. The team next selected variables for integrated (horizontal) analysis and determined that nine datasets held the predictive power to fuel individualized interventions.

### 3. ANALYTICS PLATFORM

To support the analytic endeavor to come, Resultant developed an Advanced Data Analytics (ADA) platform, which provided a secure data repository where analysts could connect and collaborate. This platform foundation has since powered other research, collaboration, and analysis efforts across agencies in- and outside State government.





## FINDINGS

Agency and sector knowledge combined with data science yielded an entirely new way of considering infant mortality in Indiana—one that provided actionable information that now informs outreach and treatment programs. Through machine learning and logistic regression techniques, the project team identified models for estimating the risk of adverse birth outcomes based on maternal age, Medicaid status, and prenatal care stratified by region and demographics at the ZIP-code level.

**The project unlocked a critical factor:** Nearly 50% of infant deaths occurred within a tiny fraction (1.6%) of Indiana's population. The State could identify this high-risk subpopulation through data inference and launch proactive interventions where they would produce the greatest change and by acting on other key risk factors the study identified:

- **Inadequate prenatal care:** Of all factors studied, access to prenatal care was the most important predictor of adverse outcomes. The study showed that the highest risk of infant deaths is to mothers with less than 10 prenatal visits.
- **Medicaid enrollment:** Significant disparities exist in infants born to mothers enrolled in Medicaid, with increased risk for having low birth weight infants.
- **Young maternal age:** 15- to 20-year old mothers with fewer than 10 prenatal visits were at the highest level of risk for adverse health outcomes. These mothers are most at risk because they are more likely to have inadequate prenatal care and be enrolled in Medicaid.

Health data alone shows the number of 15- to 20-year-old women who give birth in a given year. Logistic regression modeling from disparate datasets reflecting social determinants of health can identify women who are at risk for infant mortality before they even become pregnant—and intervene.

Accurate identification of controllable factors created opportunities for the State to address risk and reduce the infant mortality rate.



Nearly 50% of infant deaths occurred within a tiny fraction (1.6%) of Indiana's population and could be identified before pregnancy.



## USING DATA TO EFFECT CHANGE

Working with the State of Indiana, the Resultant team developed a birth outcome risk quantification tool. This dynamic tool enables public health experts and policymakers to identify variables like age, ZIP code, and number of prenatal visits to calculate the risk of adverse outcomes. Evaluating the factors contributing to infant mortality at a granular level enables precise intervention where it can have the most profound effect.

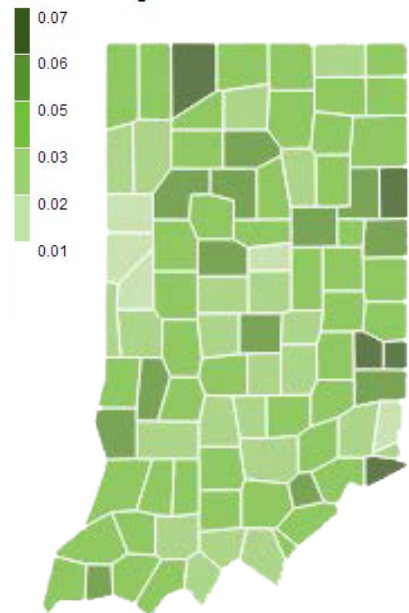
The tool includes dynamic dashboards that make results clear so that end users easily understand the information that feeds their decision-making. Because they can group individuals into subpopulations, analysts can more accurately infer causal relationships among individuals and subpopulations, and they have a powerful tool for identifying trends over time to examine program efficacy.

Applying early findings made possible by the birth outcome risk quantification tool, the State immediately adjusted a federally funded outreach campaign to more effectively identify high-risk citizens and put a number of focused education and outreach initiatives in place, including a mobile app that provides important health information to expectant and new mothers.

The State of Indiana can utilize the tool to perform continuous analysis and find new ways to connect at-risk mothers with the resources that will support a positive birth outcome. As a result of the findings, the State immediately adjusted a federally funded outreach campaign to more effectively identify those who were identified as high risk.

Because the birth outcome risk quantification tool is dynamic, analysts can utilize it to explore developing questions related to infant mortality by including new variables or focusing on specific subpopulations. Indiana citizens will receive the State's best efforts at providing effective resources even as factors contributing to infant mortality shift.

Low Birth Weight Risk



**CALCULATED  
RISK OF LOW  
BIRTH WEIGHT  
BY COUNTY**





## TOOLS AND METHODS

Diverse data elements and sophisticated machine learning algorithms enabled discovery of small subsets of the population that accounted for a highly disproportionate percentage of infant deaths. Building from these results, Resultant and the state collaboratively performed additional analyses to form actionable recommendations tailored to these at-risk subpopulations. The project moved forward through logistic regression modeling, self-organizing maps, and analysis of high-risk cluster populations.

## IMPACT

Relying on a fuller picture of infant mortality, the State of Indiana developed focused, flexible programs to reach the citizens who most needed them. And they achieved incredible results: After years of stagnation in infant mortality, Indiana has seen progress for the first time since delving into the data in 2014.

**Indiana's infant mortality rate has fallen every year since 2017, and in October 2020, the governor announced an all-time low rate: 6.5 deaths per 1,000 live births.**