Benchmarking INPC Using Birth Certificates

Mallory Dreyer, Kosali Simon, Coady Wing

Executive Summary —The Indiana Network for Patient Care (INPC) is a large health information exchange (HIE) containing over 16 billion clinical data elements on more than 25 million individuals. Despite its scale, little is known what share of health care provided in Indiana is captured in the INPC, or how coverage varies across geography, patient groups, and types of care. Assessing coverage is essential for judging the external validity and generalizability of research using INPC data. In this study, we examine coverage for one important service — childbirth and delivery — by comparing births recorded in the INPC with the full universe of Indiana births from the National Vital Statistics System (NVSS) in 2022.

Data and Methods

- **INPC Cohort**: Constructed linked-mother infant dyads (N=55,519) for births occurring in 2022, with demographic characteristics merged from maternal records and implausible records excluded.
- **NVSS Benchmark**: All births occurring in Indiana in 2022 (N=79,950) with maternal demographics, residence, and birth setting.
- **Comparison**: We assessed statewide coverage, demographic representativeness, and county-level coverage.

Comparison of Births: NVSS vs. INPC (2022)

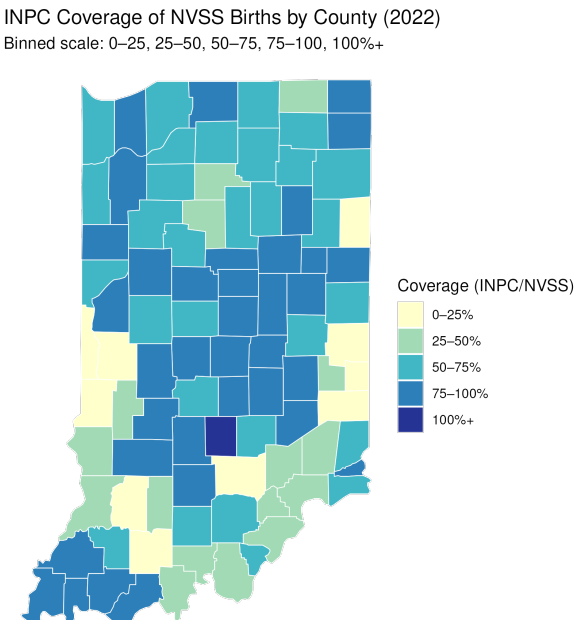
Characteristic	NVSS Vital Stats	INPC
Total births	79,950	55,519
Average age	28.39	28.96
White	81.1% (64,825)	72.5% (40,246)
Black	13.2% (10,525)	12.5% (6,945)
Asian / Pacific Isl.	3.3% (2,668)	2.8% (1,575)
AIAN	0.09% (70)	0.15% (81)
Multiracial	2.3% (1,862)	5.2% (2,903)
Hispanic	12.5% (9,967)	10.9% (6,029)
Unknown race	—	6.8% (3,769)
Hospital birth	98.5% (77,239)	—
Out-of-state res.	3.5% (2,769)	2.5% (1,364)

Findings

- INPC captures approximately 70% of Indiana births.
- Maternal age distributions are similar.
- Maternal race/ethnicity distributions differ, perhaps due to differences in how EHRs and vital records capture race/ethnicity (e.g., INPC includes “unknown,” NVSS does not).
- Out-of-state residents account for 3.5% of NVSS births and 2.5% of INPC births.
- Coverage is highest in populous counties (e.g. Marion, Allen, Lake, Hamilton, St. Joseph, Elkhart). Some border counties show low-coverage which may reflect cross-border care.

Limitations

- Maternal residence is determined from the most recent ZIP code in INPC, which may not align with residence at delivery.
- Some ZIP codes map to multiple counties, so some maternal residence counties in INPC may be misclassified.
- Some mother-infant pairs in INPC may reflect families who enter the system post-birth.



Conclusion —INPC captures a substantial share of Indiana births, representing about 70% of statewide deliveries in 2022 with strong coverage in populous counties but some gaps in border counties and differences in race and ethnicity. Future work can benchmark INPC to other state or national databases such as HCUP’s state inpatient database or Medicare claims.

1 — Introduction

The Indiana Network for Patient Care (INPC) is a large database of electronic health records from multiple health care providers and systems that operate in different parts of Indiana. For most types of care, it is reasonable to assume that the database includes records of all of the care delivered by contributing providers. However, it is not clear what fraction of care received by patients in Indiana is captured. The coverage rate likely varies across types of care. It may have changed over time. It is also possible that coverage differs by patient sub-population.

According to a recent case study, INPC has information on more than 16 billion clinical data elements, including encounters, lab results, pharmacy data, radiology reports, and text notes, among others (Williams et al. 2025). The Indiana Health Information Exchange (IHIE) manages the INPC, drawing on more than 123 hospitals, 19,000 practices, and 54,000 providers across Indiana and some neighrboring states. IHIE's coverage recently expanded to include two additional health information exchanges to include more of Indiana and portions of southern Michigan. As of December 2023, INPC had information on almost 25 million patients, encompassing Indiana and non-Indiana residents, living and deceased individuals, and those with comprehensive care histories to those with only one clinical observation or encounter.

Despite the large volume of records in INPC, one way to improve the the value INPC data is to clarify what is known about coverage from both an organizational/contractual point of view, and from benchmarking studies that compare INPC to "ground truth" estimates from outside data sources. In this short paper, we present a very simple benchmarking study based on vital statistics data. In the United States, birth data is collected and compiled federally into the National Vital Statistics System (NVSS). We compare the births in INPC to births in the NVSS in 2022. Similar exercises for benchmarking can be done with other statewide and national data sources, such as HCUP inpatient data or Medicare claims.

2 — Data

INPCIn INPC, our sample was constructed using two files: (1) the first file links a given mom's unique identifier ($STUDYID$) to a unique birth identifier ($BIRTHID$) and (2) the second file links a given infant's unique identifier ($STUDYID$) to a unique birth identifier ($BIRTHID$). These files can be linked based on the unique birth identifier which creates mother-infant pairs. The infant file also includes the date of birth. For this project, we restrict to infants born in 2022.

Demographic information was merged on using the demographics file. Age at delivery for the mother was constructed using the mother's date of birth and the infant's date of birth. The mother's race and ethnicity were also linked. We dropped implausible mothers from the sample–those who were younger than 12 or older than 50 and those who had a non-female sex (male or unknown) listed. The first ZIP code to occur in the demographics

file was kept for each mother and county codes were merged on using a HUD crosswalk. The resulting cohort represents all linked mother–infant dyads captured in INPC during 2022 with demographic information coming from the mother. This is what is most directly comparable to the NVSS birth certificates because the NVSS also captures maternal demographic characteristics such as county of residence, age, race, etc.

In INPC, there is no variable that defines the location of the birth (hospital or elsewhere). The INPC cohort is intended to approximate the population of Indiana births captured in NVSS, though it represents only those births occurring in hospitals and health systems participating in the health information exchange.

NVSS Birth Certificates Using the 2022 NVSS data, we limited the sample to births that occurred in the state of Indiana. There were a total of 79,950 births in Indiana in 2022. The NVSS data include the mother's age, race/ethnicity, and residence county, as well as a variable that defines the birth location as being in a hospital or not. NVSS caputres all births in Indiana, which makes it the benchmark against which INPC is compared.

3 — Benchmarking Results

3.1. Statewide Coverage

NVSS records a total of 79,950 births in Indiana during 2022, while INPC captures approximately 55,519 linked mother–infant dyads. Taken together, this implies that INPC represents roughly 70% of all births statewide. Table 2 compares counts and shares of births across age–race subgroups.

Overall, INPC coverage by age and race groups vary significantly. Some differences reflect variations in how race/ethnicity are captured across EHRs and NVSS, as NVSS does not include an unknown race category.

3.2. Representativeness of INPC Births

Table 3 compares the characteristics of the births in INPC to the NVSS births. While the average age is similar in the two systems, there are differents in the racial composition, which again is partially driven by differences in how race and ethnicity are captured in EHRs vs. birth certificates. Most births in NVSS occurring in Indiana in 2022 took place in a hospital (98.5%). In NVSS, approximately 3.5% of births are to out-of-state residents whereas 2.5% of births in INPC are to out-of-state residents.

3.3. Coverage by County

Coverage also varies substantially across counties. In NVSS, county of occurrence defines where a birth took place, while county of residence reflects the mother's home address. To facilitate comparison, we use county of residence in both datasets, since INPC includes out-of-state mothers who deliver at Indiana facilities and excludes Indiana residents who may deliver in non-INPC hospitals.

In Figure 1, the total number of births in INPC and NVSS are plotted by county of residence. The county with the most births is Marion County, which reflects the fact that it is the largest county by population. Allen, Lake, Hamilton, St.Joseph, and Elkhart counties also have higher total numbers of births, also reflecting that they are counties with higher populations.

Figure 2 shows the coverage of INPC births, computed by taking the number of INPC births over the number of NVSS births by county of residence. Some counties have low coverage, shown by the yellow in the figure. One county has more than 100% coverage, implying that there are more births to residents in that county in INPC than there were in NVSS. Some border counties to Illinois, Kentucky, amd Ohio have low coverage, which may imply that there may be some cross-border care utilization where Indiana residents receive care at facilities out-of-state that do not participate in IHIE. However, there is substantial variation in coverage across counties. Marion County falls in the 50-75% coverage category.

Figure 3 shows the total number of births in INPC and NVSS by county of residence. The pink lines show the count of births in INPC and the navy lines show the count for NVSS.

3.4. Limitations

While INPC provides a valuable source of linked mother-infant pairs for Indiana, several limitations should be noted when interpreting these results.

First, we do not directly verify that each INPC mother-infant pair corresponds to a documented birth event in the clinical records. A more rigorous approach would involve linking the encounter-level ICD-10 codes from both the mother and infant to confirm delivery. Without this step, it is possible that some pairs represent families who began receiving care at INPC-contributing facilities after the birth, in which case their deliveyr records would not appear in INPC even though their demographic information and subsequent encounters do. If the birth certificate information is available for all births in Indiana, it is possible that someone may show up in INPC as a linked pair well after their actual birth.

Second, the measure of maternal residence relies on the most recent address available in the demographics file. Because we do not observe historical address changes, we cannot confirm that the ZIP code reflects the mother's residence at the time of birth. The limitation likely contributes to the case where the county-level coverage exceeded 100%, as families may move counties after delivery. Furthermore, the INPC geographic information has the ZIP code of residence, and we report at the county level. A single ZIP code may map to more than one county, but we assigned a ZIP code to the county with the highest share of residential population.

4 — Next Steps

In addition to births, INPC can be benchmarked against other statewide and national data sources to assess its representativeness for broader populations and healthcare utilization. Future work could use sources like AHRQ's HCUP state inpatient data, state ambulatory surgery and services data, or state emergency department data for Indiana to benchmark. Indiana Medicaid or the 100% sample of Medicare claims could also provide additional useful data sources for benchmarking exercises.

We currently have the 100% sample of Medicare claims for Part A and Part B (not Medicare Advantage). These claims could provide a useful benchmark for older adults. We could compare hospitalization rates for conditions like heart failure or pneumonia for older adults. Medicaid claims provide a complementary benchmark for lower-income populations. They could be used to assess how well INPC captures pediatric care, preventive services, and maternal care among Medicaid enrollees. Because INPC includes a payer type variable (at the encounter level), it is possible to directly compare the share of encounters attributable to Medicare or Medicaid with those observed in the claims data, offering a way to assess representativeness by payer.

HCUP data could also provide another useful comparison point. HCUP's State Inpatient Database (SID) captures all hospital discharges and would allow us to compare inpatient utilization in INPC against the universe of hospitalizations in Indiana and examine coverage stratified by age, payer, or diagnosis. The State Ambulatory Surgery and Services Database (SASD) includes data on outpatient surgeries and services which could serve as a benchmark for INPC outpatient encounters to assess coverage of same-day surgical procedures like joint replacements or cataract surgeries. The State Emergency Department Database (SEDD) includes all ED visits that do not result in an admission and would allow us to examine how well INPC captures emergency department utilization, particularly for conditions like asthma, injury, or substance use–related visits.

5 — Figures and Tables

Table 2 — Age-Race Comparison of Birth Counts and Shares (NVSS vs. INPC)

Age Bin	Race	NVSS	INPC	Share (%)
<20	AIAN	<10	<10	–
<20	Asian PI	32	<10	–
<20	Black	343	170	49.6
<20	Multi	109	43	39.4
<20	Unknown	–	110	–
<20	White	1433	548	38.2
20–29	AIAN	37	30	81.1
20–29	Asian PI	724	395	54.6
20–29	Black	5354	3433	64.1
20–29	Multi	1222	1494	122.3
20–29	Unknown	–	1814	–
20–29	White	31942	18296	57.3
30–39	AIAN	29	46	158.6
30–39	Asian PI	1696	1024	60.4
30–39	Black	4339	2964	68.3
30–39	Multi	498	1247	250.4
30–39	Unknown	–	1649	–
30–39	White	29015	19637	67.7
40+	AIAN	<10	<10	–
40+	Asian PI	216	142	65.7
40+	Black	489	378	77.3
40+	Multi	33	119	360.6
40+	Unknown	–	196	–
40+	White	2435	1765	72.5

Figure 1 — INPC (top) and NVSS (bottom) number of births by mother's county of residence in 2022.

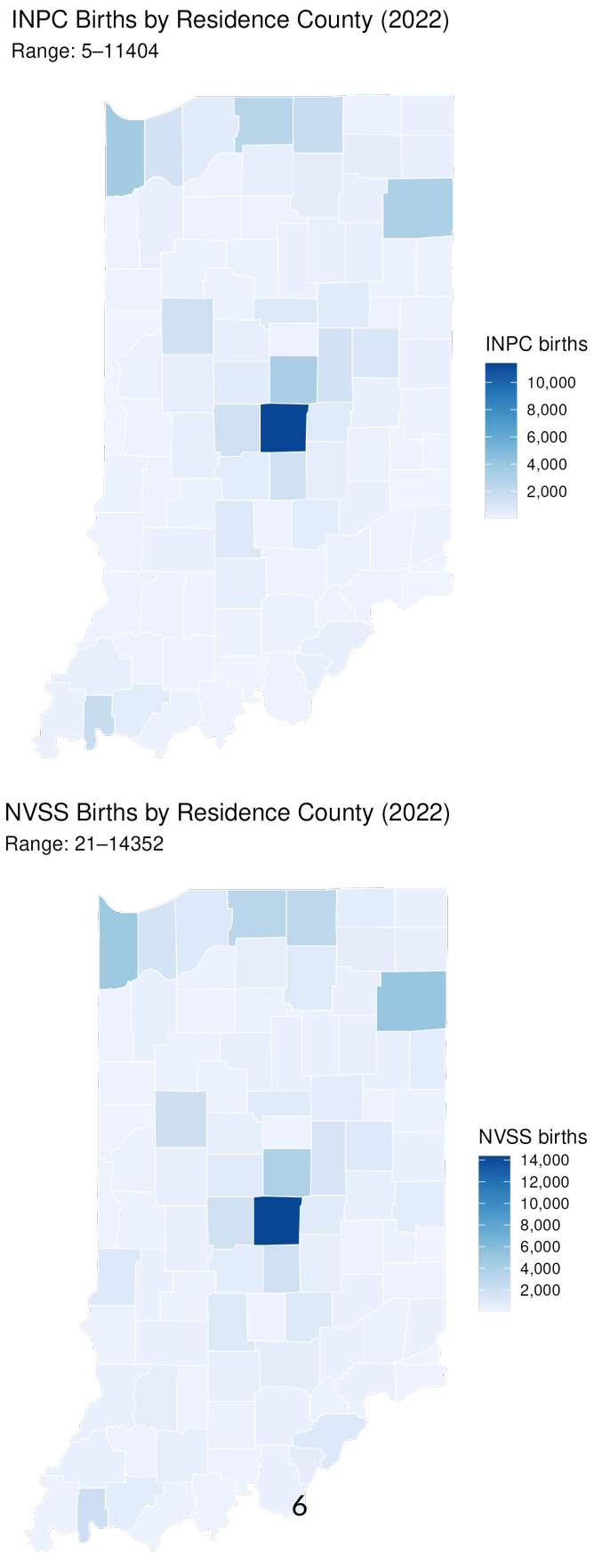


Table 3 — Comparison of Births: NVSS vs. INPC (2022)

Characteristic	**NVSS Vital Stats**	**INPC**
Total births	79,950	55,519
Average age	28.39	28.96
White	81.1% (64,825)	72.5% (40,246)
Black	13.2% (10,525)	12.5% (6,945)
Asian / Pacific Isl.	3.3% (2,668)	2.8% (1,575)
AIAN	0.09% (70)	0.15% (81)
Multiracial	2.3% (1,862)	5.2% (2,903)
Hispanic	12.5% (9,967)	10.9% (6,029)
Unknown race	—	6.8% (3,769)
Hospital birth	98.5% (77,239)	—
Out-of-state res.	3.5% (2,769)	2.5% (1,364)

Figure 2 — INPC Coverage (%) of NVSS Births by County (2022)

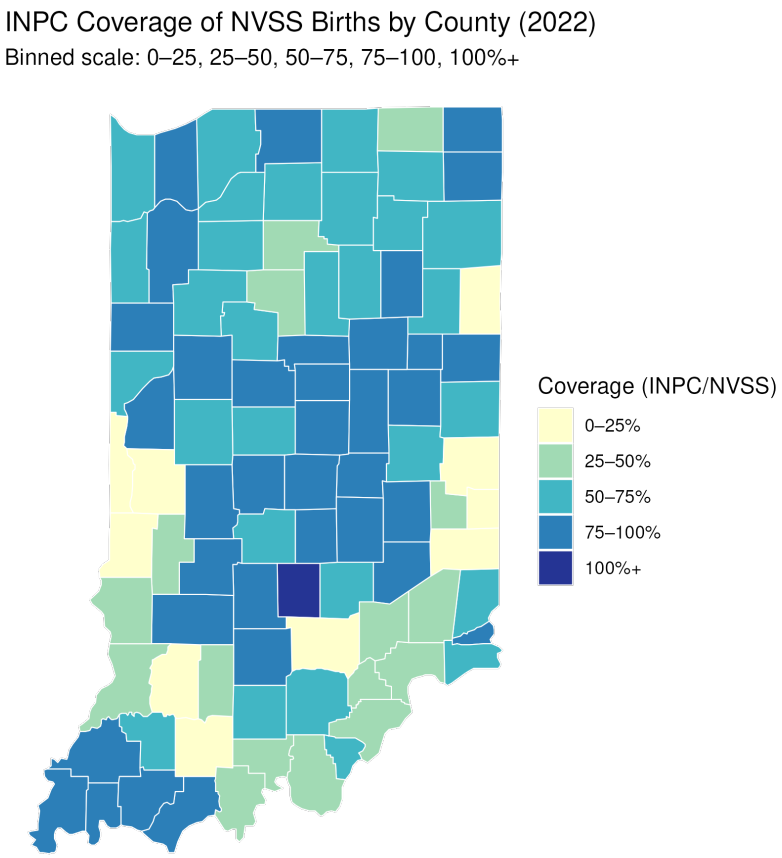
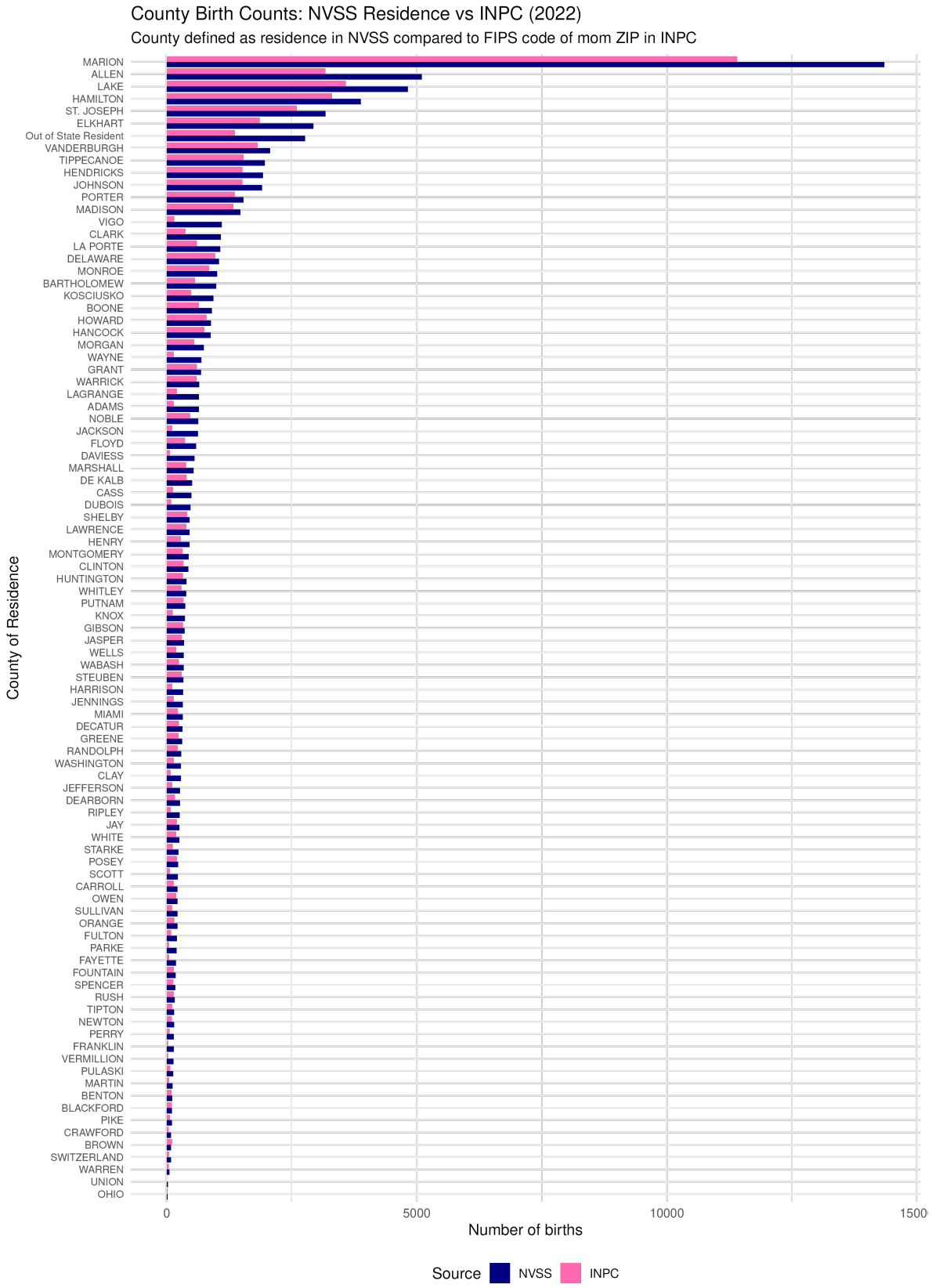


Figure 3 — County Birth Counts by Mother County of Residence in 2022: Comparison of INPC to NVSS



References

Williams, Karmen S, Saurabh Rahurkar, Shaun J Grannis, Titus K Schleyer, and Brian E Dixon. 2025. “Evolution of clinical Health Information Exchanges to population health resources: a case study of the Indiana network for patient care.” *BMC medical informatics and decision making* 25 (1): 97. (Cited on page 1).