OSIIS Data Quality

Martin Lansdale, OSIIS Data Quality Manager

martinl@health.ok.gov

OKLAHOMA
State Department of Health

OSIIS DATA QUALITY REVIEW

- Today we will discuss:

    -What is Data Quality

    -IIS Data Quality Blue Print

    -Why Does Data Quality Matter

    -OSIIS Data Quality Plan

    -OSIIS Data Quality Review

    -OSIIS Data Quality Improvement

DATA QUALITY: WHAT IS IT?

- Data Quality is an integral part of data governance that ensures your organizations data (i.e. IIS registry) is fit for its purpose.

- Generally refers to the overall utility of a dataset and includes 6 core concepts: Accuracy, Completeness, Consistency, Timeliness, Validity, and Uniqueness.

- All core concepts are important but an organization can focus on specific ones depending on need.

- The CDC, MIROW, and AIRA have created a 'best practices" guide that includes a IIS Data Quality Blue Print which focuses on 4 specific core concepts.

DATA QUALITY: WHAT IS IT?

DATA QUALITY CHARACTERISTICS

As mentioned, the best practice recommendations made in the guide support the Immunization Information Systems (IIS) Data Quality Blueprint.[7]

MAIN CHARACTERISTICS

- **Availability:** IIS data can be readily used by authorized users to inform IIS and immunization program activities.

- **Completeness:** The IIS contains all (historic and new) vaccination and demographic records for persons residing in the jurisdiction, and all vaccination and demographic records contain complete data fields.

- **Timeliness:** Patient records are established, and vaccination events are recorded
in the IIS within specified time frames.

- **Validity:** Vaccination and demographic records in the IIS conform with generally accepted standards (e.g., Advisory Committee on Immunization Practices recommendations, United States Postal Service standards).

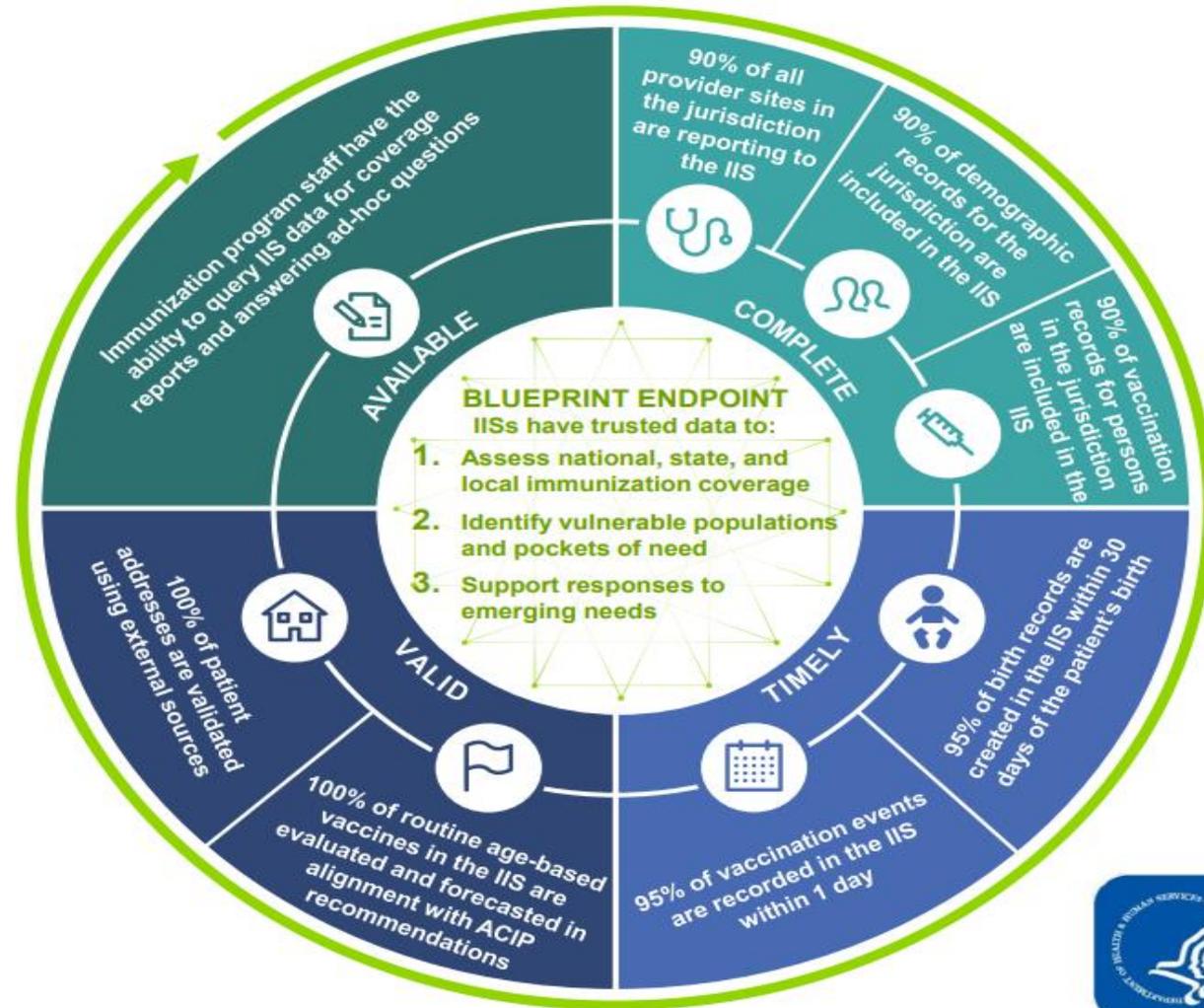## Immunization Information Systems (IIS) Data Quality Blueprint

The IIS Data Quality Blueprint is a guide to help immunization programs address and advance data quality within IISs. Its measures focus on improving IIS data throughout the lifespan.

The blueprint focuses on driving IIS data that are:
- **AVAILABLE**
- **COMPLETE**
- **TIMELY**
- **VALID**

Other enabling characteristics (i.e., ACCURATE, CONSISTENT, and UNIQUE) and functional processes (e.g., interoperability) are inherent to high quality IIS data and complement the blueprint data quality characteristics.

Sample activities for achieving high quality data and the associated recommended measures are found in the IPOM Sample Activities 2023 document.

BLUEPRINT ENDPOINT
IISs have trusted data to:
1. Assess national, state, and local immunization coverage
2. Identify vulnerable populations and pockets of need
3. Support responses to emerging needs

AVAILABLE
Immunization program staff have the ability to query IIS data for coverage reports and answering ad-hoc questions

COMPLETE
90% of all provider sites in the jurisdiction are reporting to the IIS

90% of demographic records for the jurisdiction are included in the IIS

90% of vaccination records for persons in the jurisdiction are included in the IIS

TIMELY
95% of birth records are created in the IIS within 30 days of the patient's birth

95% of vaccination events are recorded in the IIS within 1 day

VALID
100% of patient addresses are validated using external sources

100% of routine age-based vaccines in the IIS are evaluated and forecasted in alignment with ACIP recommendations

Data Quality: Why Does It Matter?

"You can have all of the fancy tools, but if [your] data quality is not good, you're nowhere." — Veda Bawo, director of data governance, Raymond James

"Without clean data, or clean enough data, your data science is worthless." — Michael Stonebraker, adjunct professor, MIT

DATA QUALITY: WHY DOES IT MATTER?

- Data Quality ensures that data is fit for its purpose and can accurately and reliably be used for analysis.

- Poor data quality can lead to erroneous associations causing bad policy/decision making.

- Measuring data quality can help organizations identify data gaps/errors that need to be resolved.

- Mantra: Crap in, Crap out. If your data is unreliable then your analysis is suspect or at least reduced in scope.

OSIIS Data Quality Plan

OSIIS Data Quality Plan

- Completed in June 2022 in conjunction with OSDH and AIRA staff.

- Based on IIS Data Quality Blueprint and MIROW guidelines.

- SCOPE: Improving the quality of data at rest in OSIIS.

- The scope includes assessment of patient demographic and vaccination information but not a review of incoming data.

- Plan includes 8 Objectives and strategies to improve data quality.

- Objectives include 6 high priority issues.

OSIIS DATA QUALITY PLAN: HIGH PRIORITY

- <u>Lack of data exchange with Vital Records:</u> Need Vital Records feed for birth and death records.

- <u>Missing jurisdiction-level patient status:</u> Jurisdiction-level patient status is important in order to appropriately include or exclude individuals from vaccination coverage assessments.

- Cost of maintaining storage space for legacy OSIIS data.

- Incomplete phone number/email data in patient records.

- Incomplete race and ethnicity data in patient records.

- Errors in vaccine funding source in vaccination records.

OSIIS Data Quality Review

OKLAHOMA
State Department
of Health

OSIIS Data Quality Review

- Assessed 2021 and 2022 data for all data elements in the Oklahoma Data Quality Plan.

- A total of 9,610,838 records reviewed, core concept: Completeness.

- Current administrations and history shots separated.

- Manual Entry and HL7 submission compared along with patient age.

- Veterans Health Administration (VHA) and IZ Gateway data compared.

- Analysis was completed with SAS v9.4 and Excel.

OSIIS Data Quality: History vs Current Administration
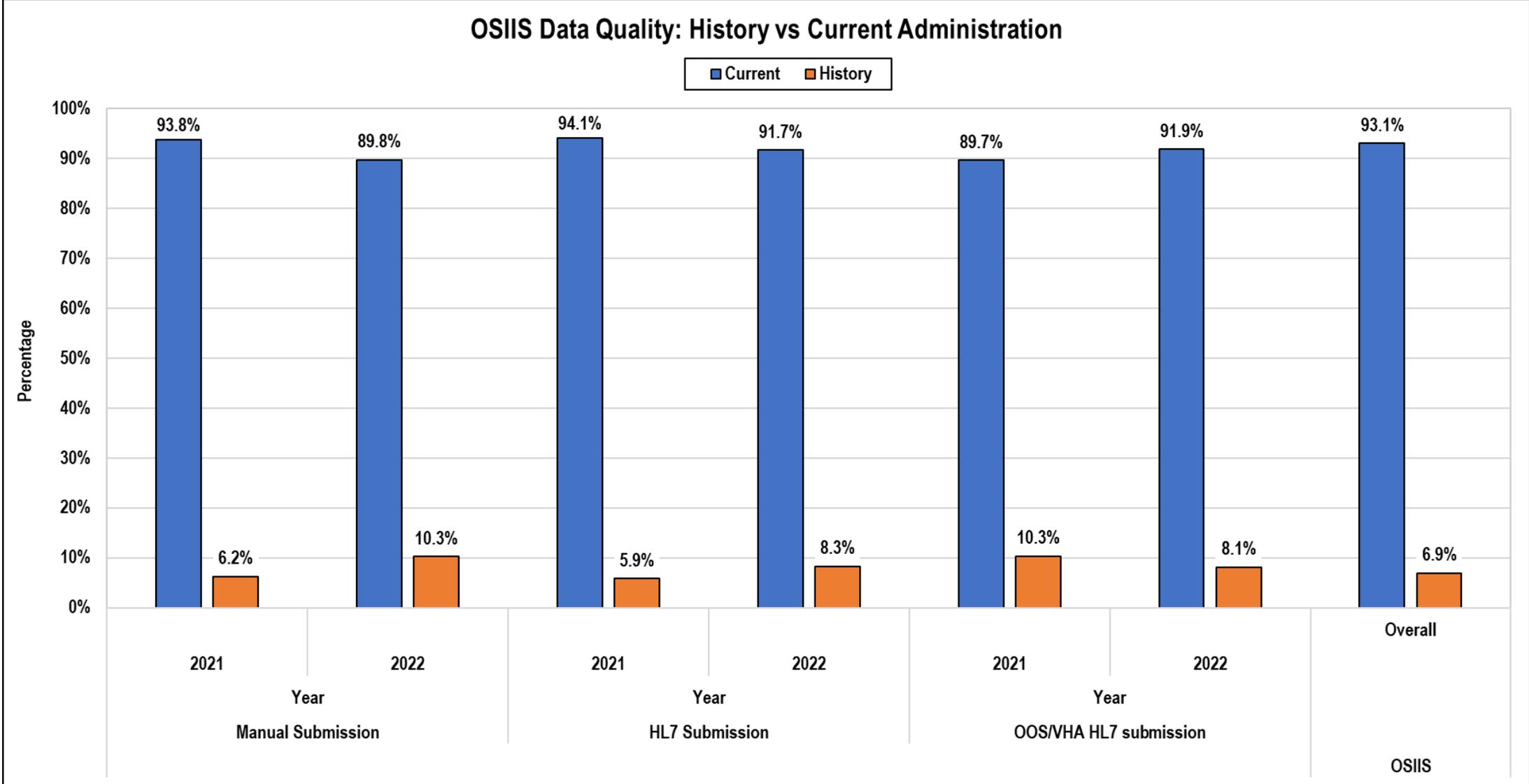
OSIIS Data Quality by Submission Type, 2021-2022
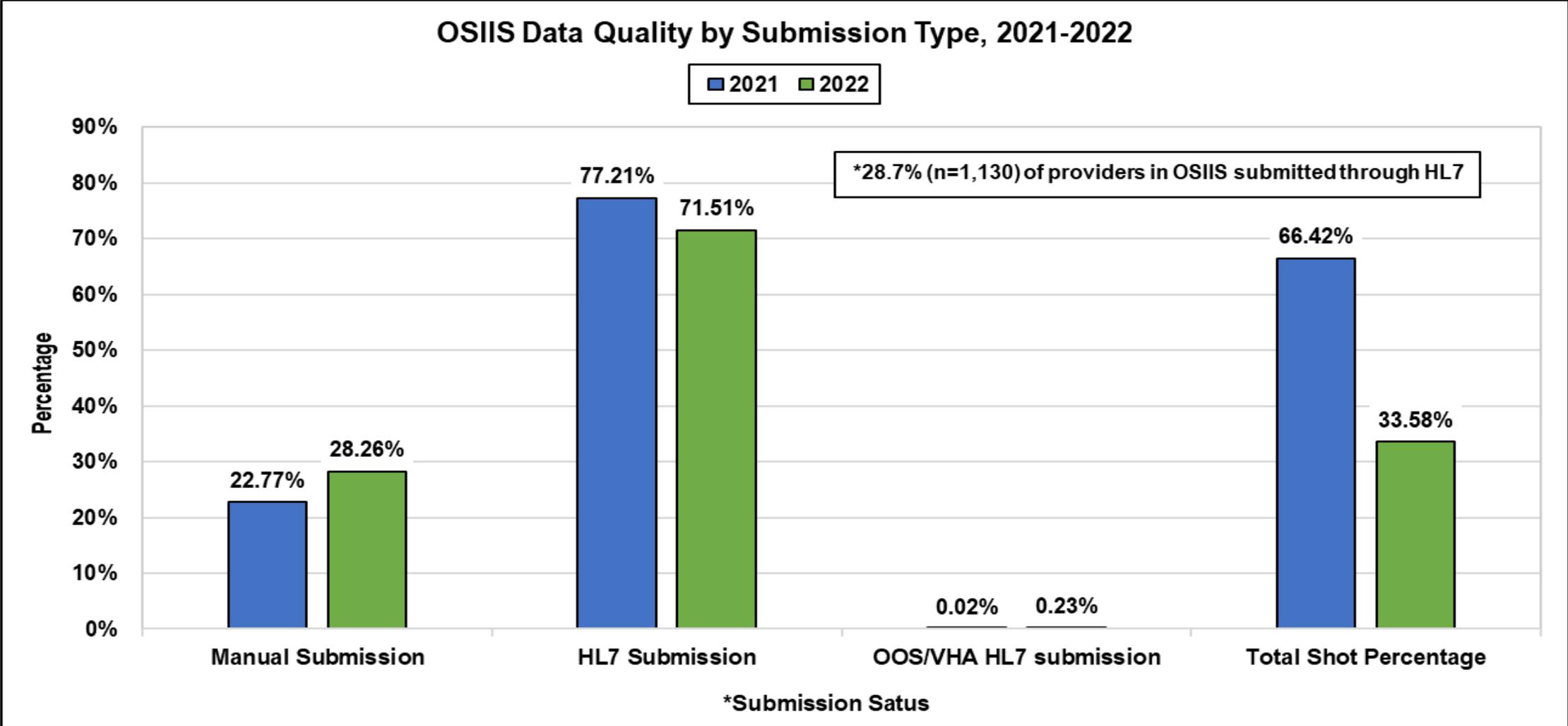
OSIIS Shot Administrations by Funding Source, 2019-2022

Legend: 2019 | 2020 | 2021* | 2022 | Total (4 Year)

* Covid Vaccine Roll-Out. PAN funding source used for covid shots only.
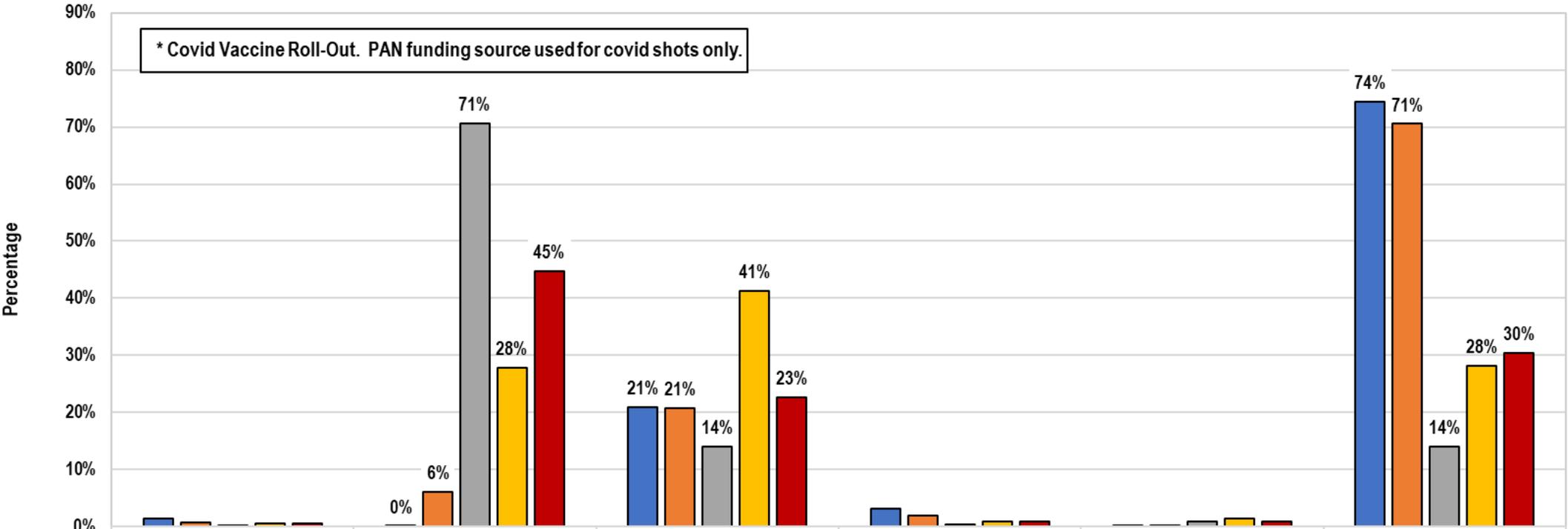
	317	PAN	PRIVATE	STATE	UNKNOWN	VFC
2019	1.3%	0.0%	20.9%	3.1%	0.2%	74.5%
2020	0.8%	6.0%	20.7%	1.8%	0.2%	70.5%
2021*	0.2%	70.7%	14.0%	0.4%	0.8%	14.0%
2022	0.5%	27.8%	41.3%	0.8%	1.5%	28.1%
Total (4 Year)	0.4%	44.8%	22.6%	0.9%	0.9%	30.4%

OSIIS Data Quality Review: Lowest Missing Percentage

OSIIS Data Elements with Lowest Missing Percentage			
Data Element	**OSIIS Overall % missing**	**MIROW Guideline**	**Meets?**
Patient First Name	0%	0%	Yes
Patient Last Name	0%	0%	Yes
Patient Date of Birth	0%	0%	Yes
Patient Gender	0%	5%	Yes
Patient Address	1.08%	15%	Yes
Patient City	1.10%	15%	Yes
Patient County	2.02%	15%	Yes
Patient State	1.08%	15%	Yes
Patient Zip Code	1.10%	15%	Yes
Shot Manufacturer	5.78%	1%	No
Lot Number	1.15%	1%	Yes
Site	6.69%	10%	Yes
Route	5.71%	10%	Yes
Expiration date	5.72%	10%	Yes
Eligibility	5.69%	1%	No
Funding Source	6.75%	0%	No

Data Quality: Racial Category by Age and Year

Data Quality: Racial Category by Submission Status and Year

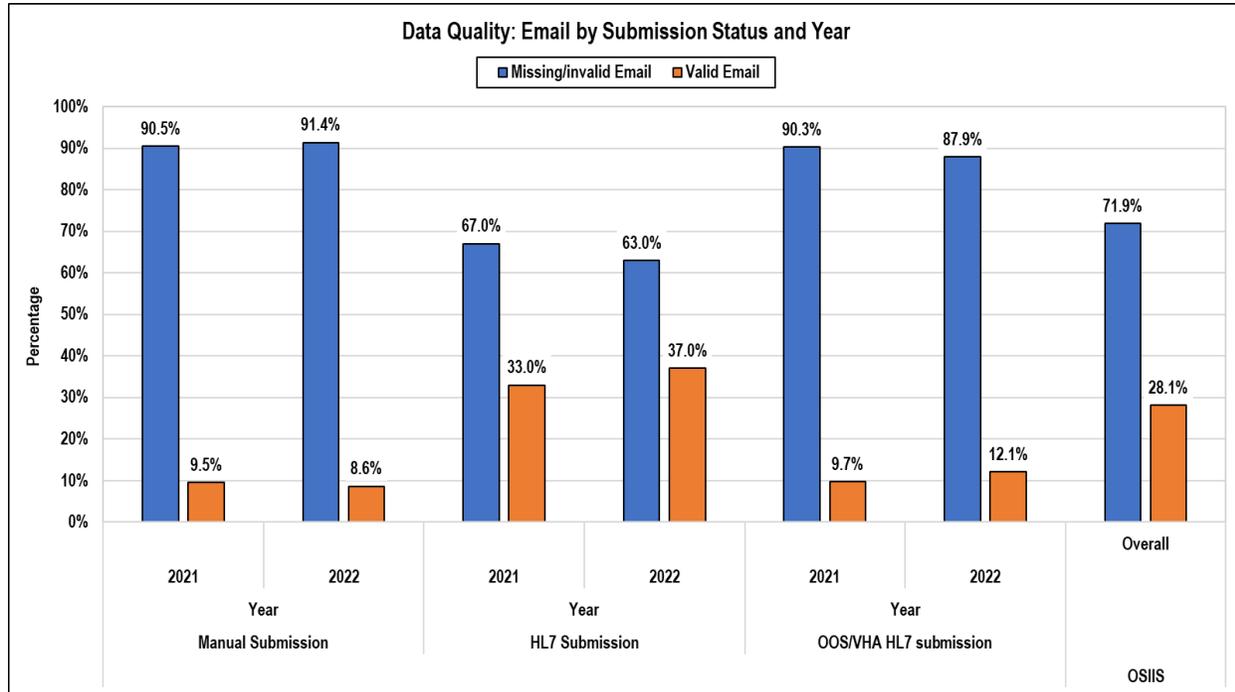*Valid Race means an actual racial category was picked (not missing/blank or Other).
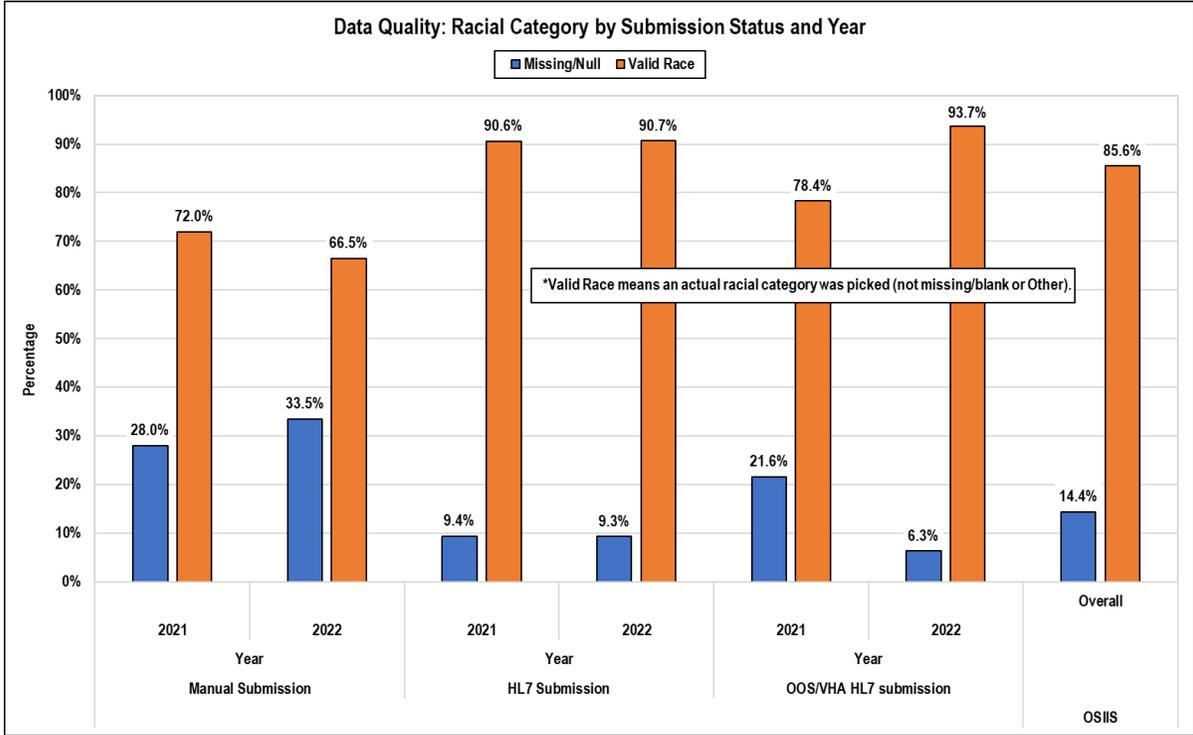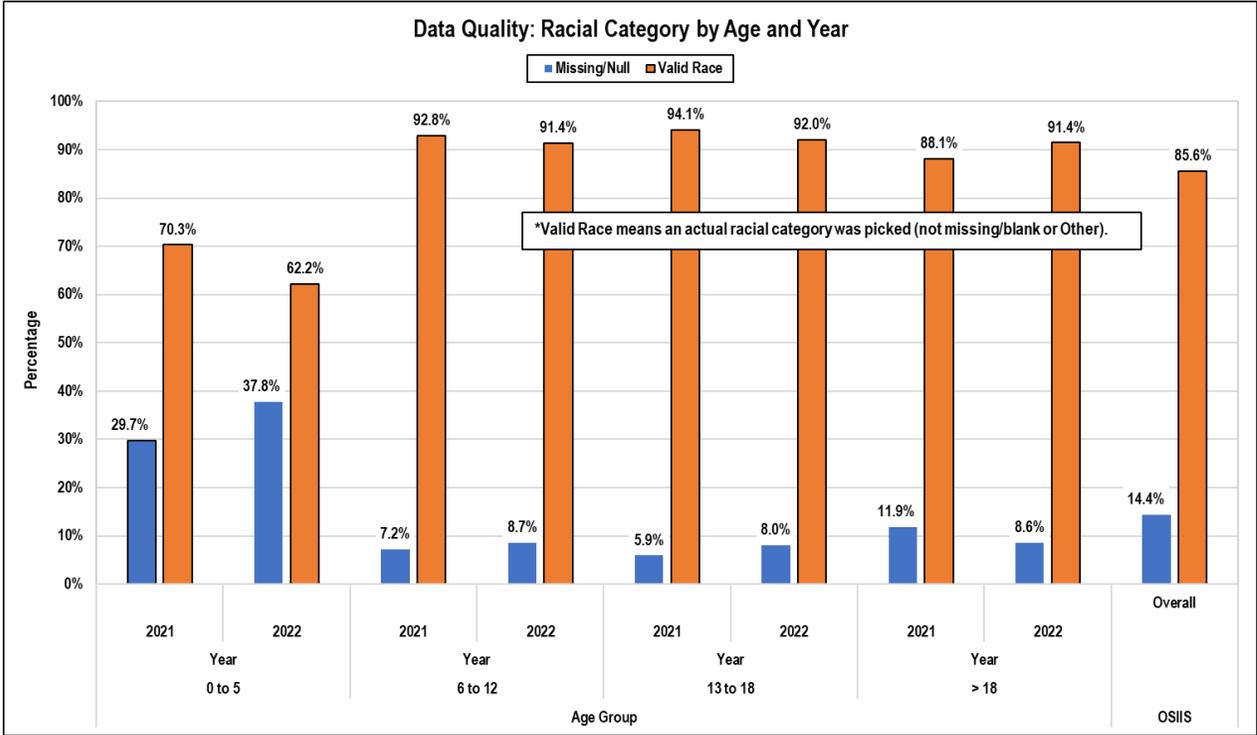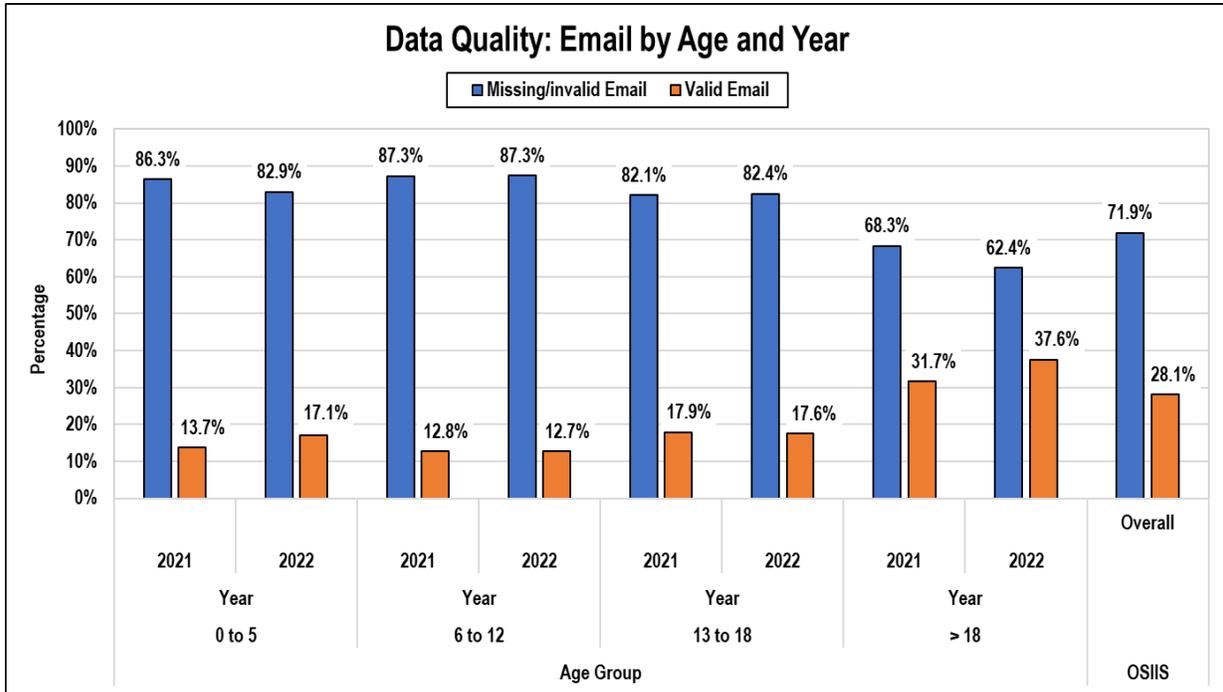
Data Quality: Ethnicity by Age and Year
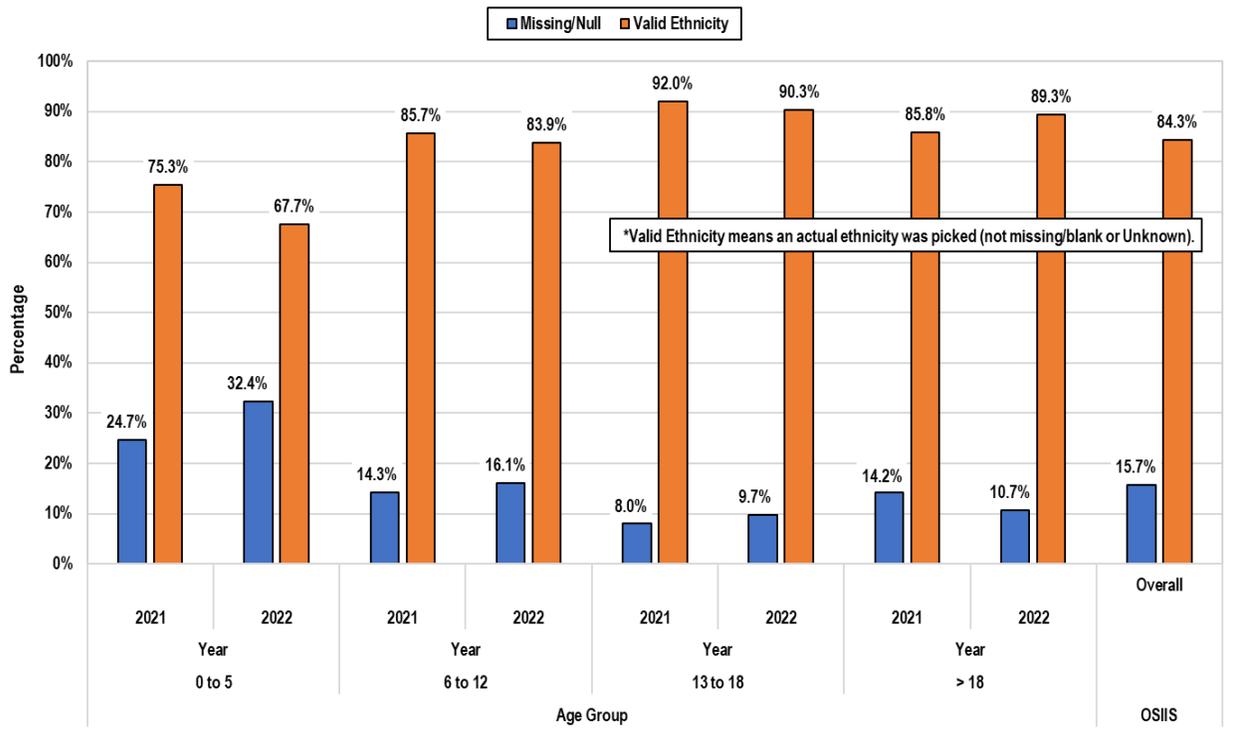
Data Quality: Ethnicity by Submission Status and Year
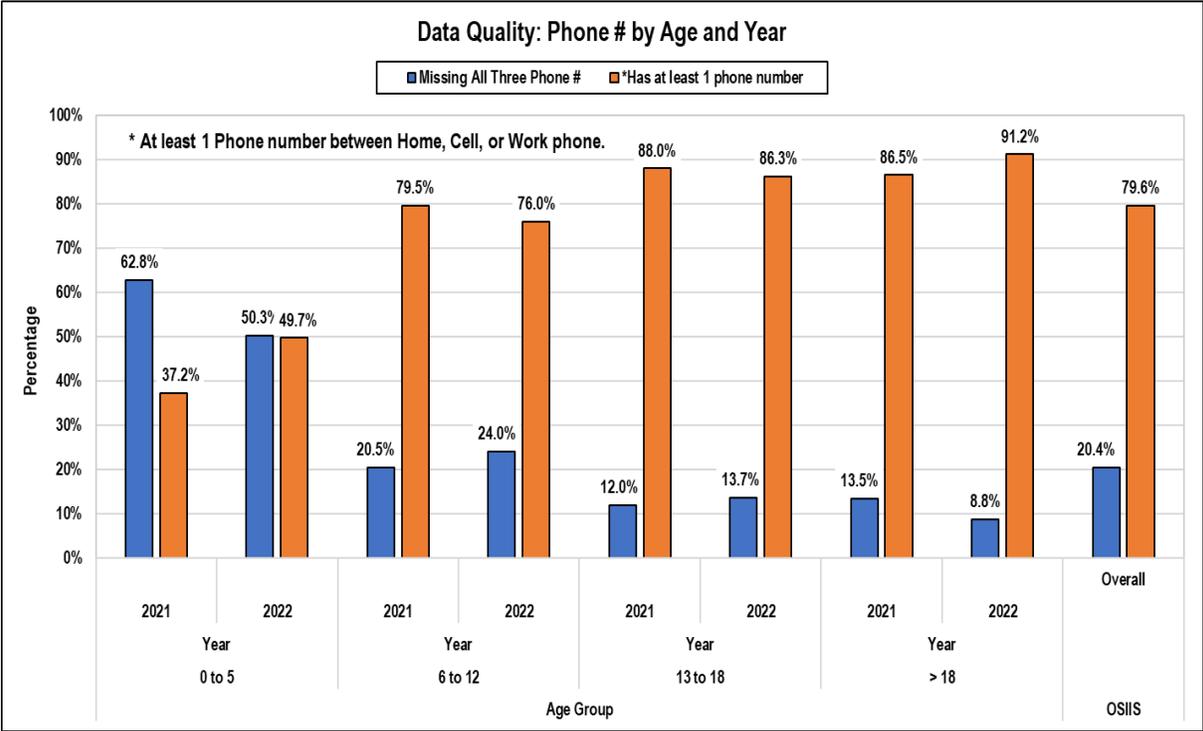
Data Quality: Email by Age and Year
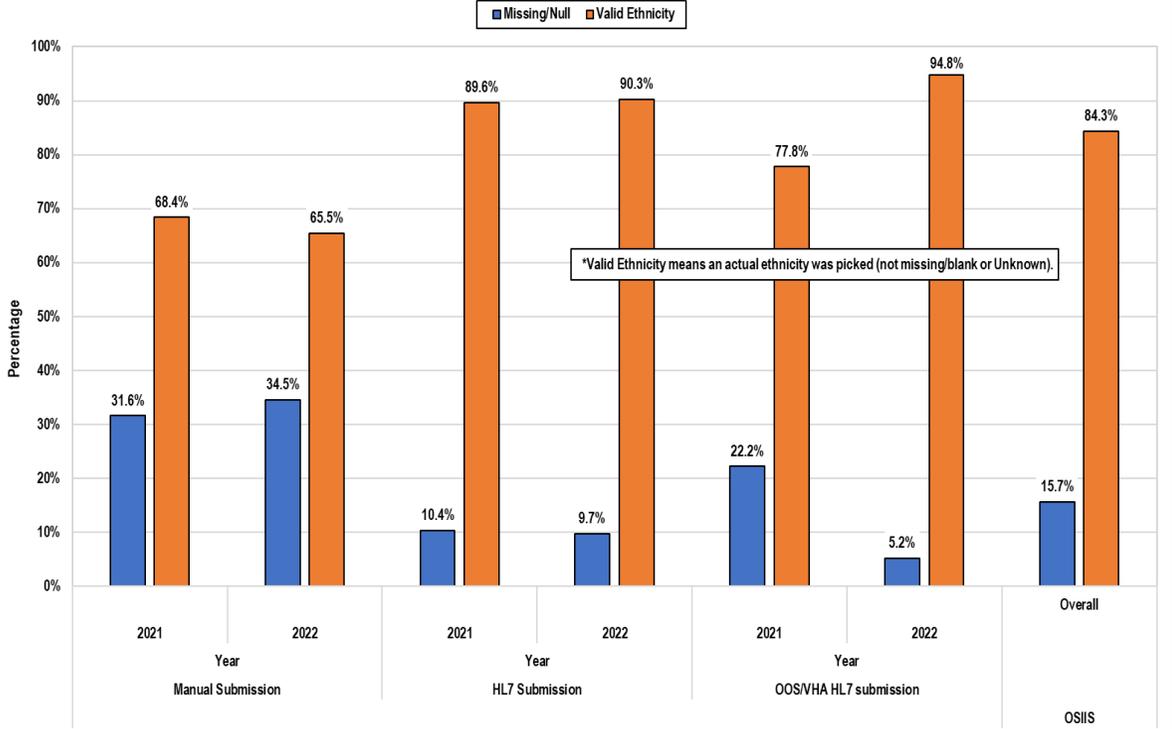
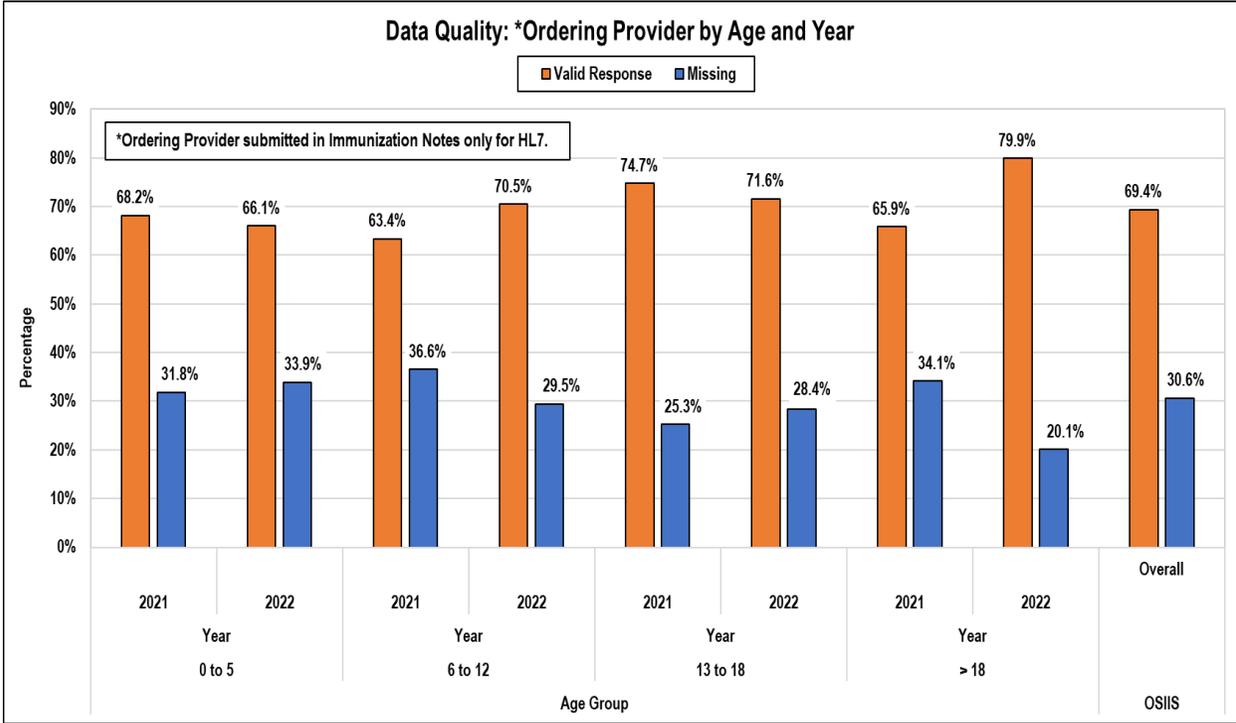Data Quality: Email by Submission Status and Year

OSIIS Public Portal

OSIIS Data Quality Review: Phone Number
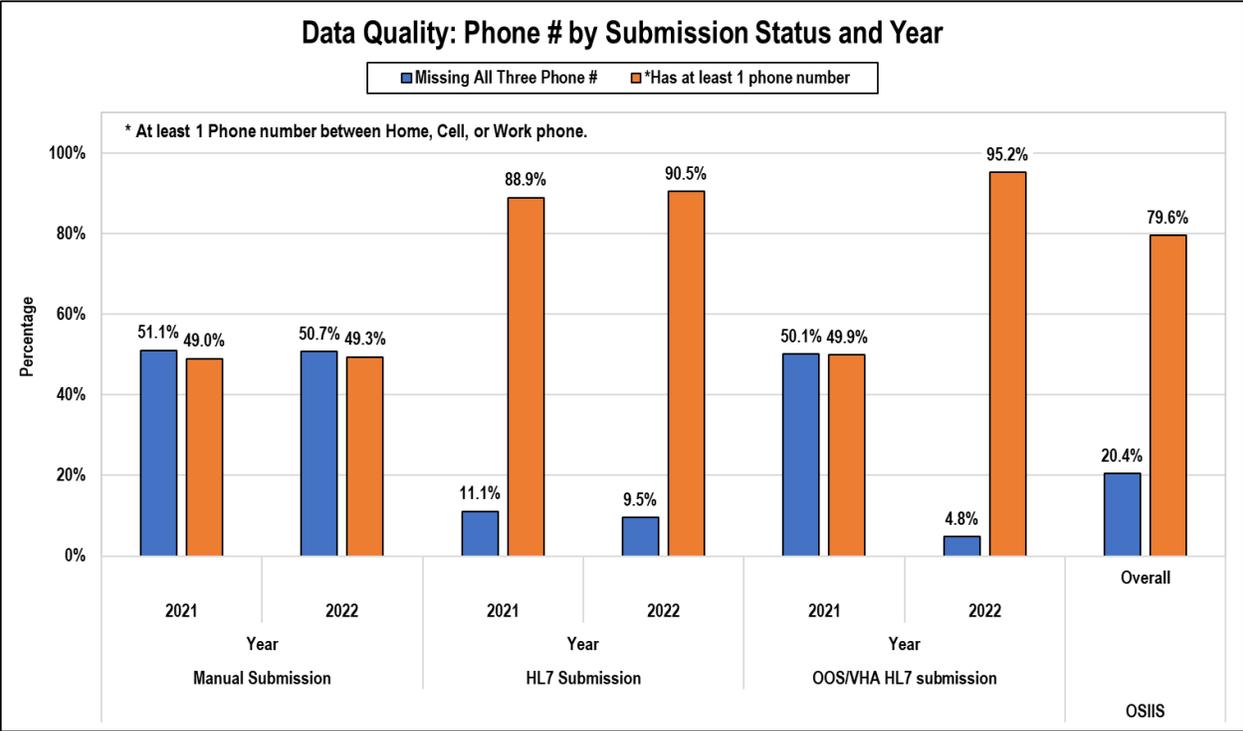
Data Quality: Phone # by Age and Year

Data Quality: Phone # by Submission Status and Year

Data Quality: *Ordering Provider by Age and Year

*Ordering Provider submitted in Immunization Notes only for HL7.

Data Quality: *Ordering Provider by Submission Status

*Ordering Provider submitted in Immunization Notes only for HL7.
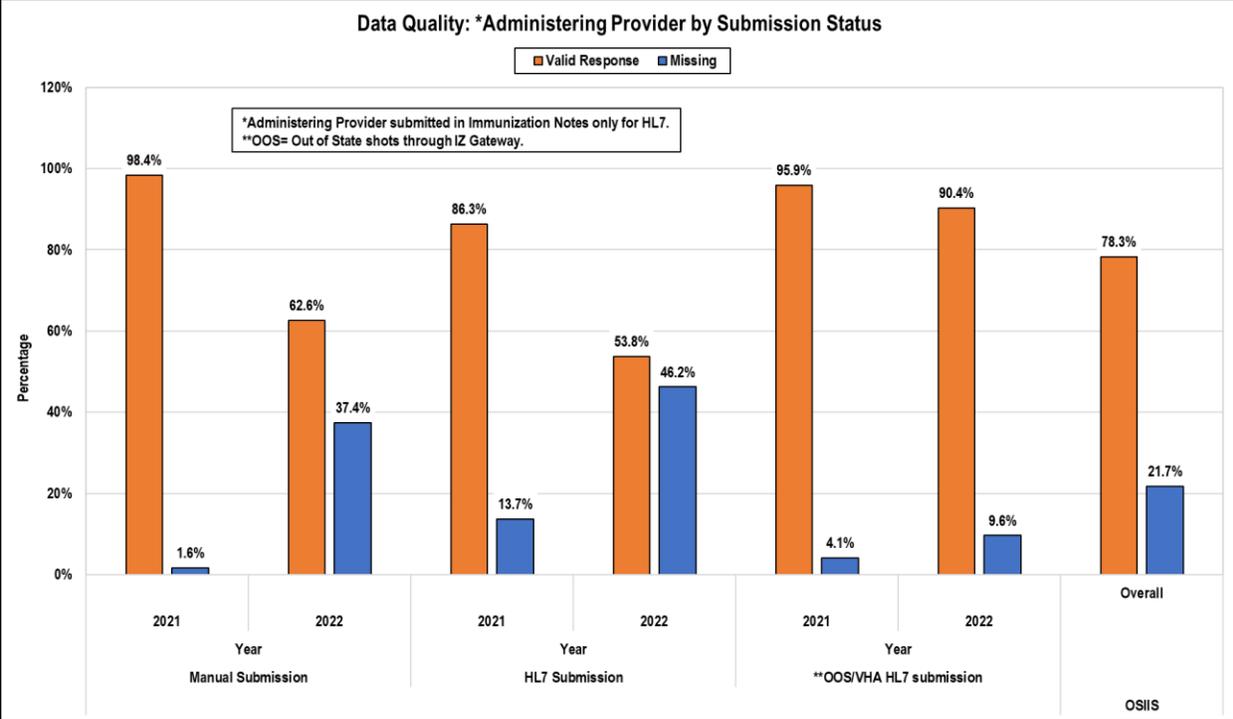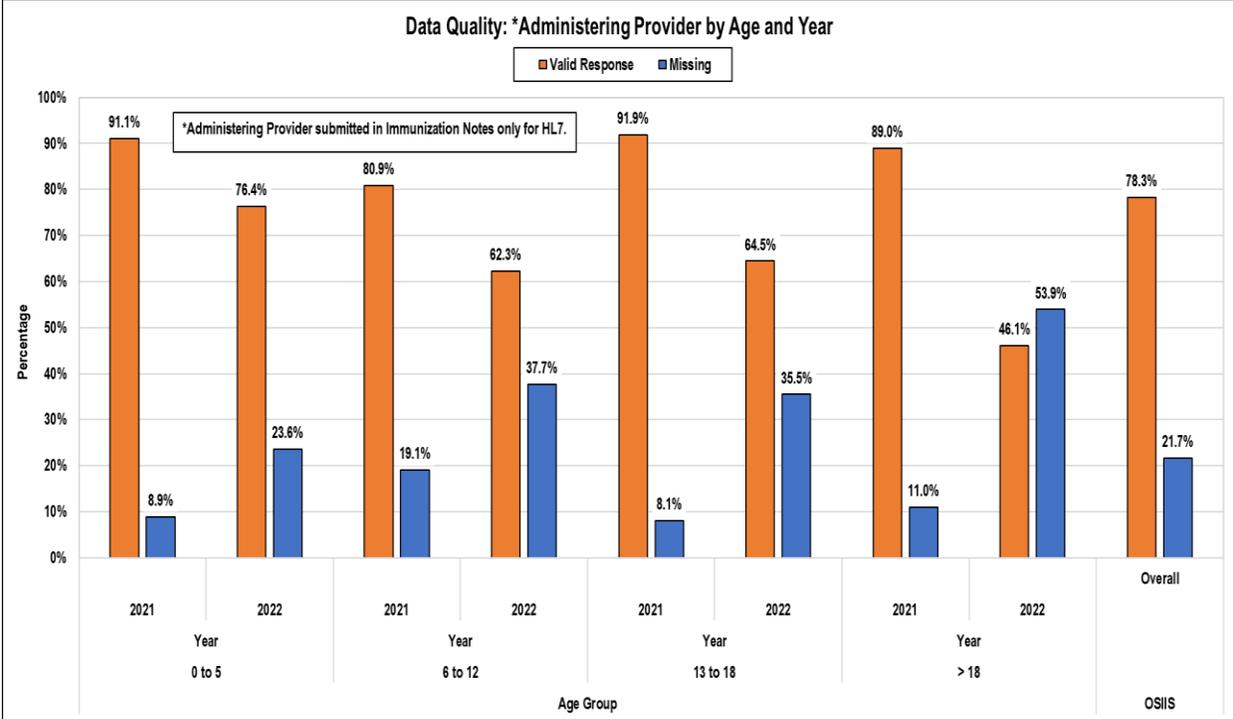**OOS= Out of State shots through IZ Gateway.

Data Quality: *Administering Provider by Age and Year

Data Quality: *Administering Provider by Submission Status
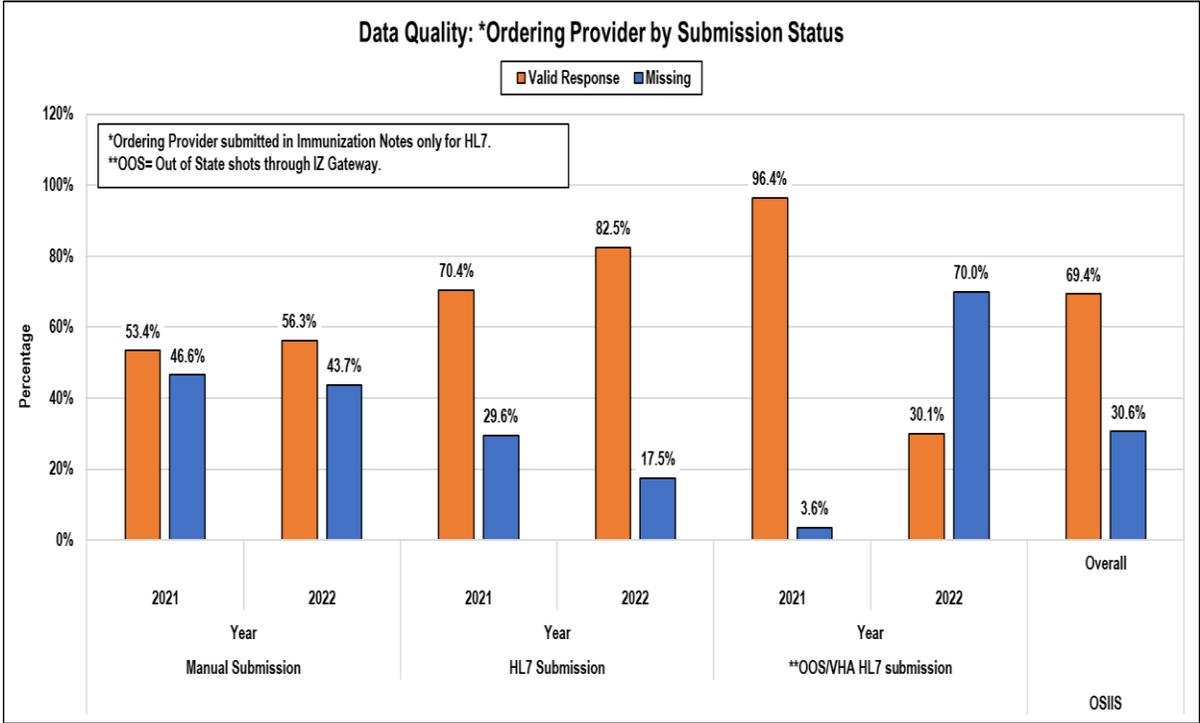
OSIIS DATA QUALITY REVIEW: TRENDS

- Data Quality for certain demographic data elements is low in the younger age groups(0 to 5).

- Data Quality for providers entering manually is lower overall compared to providers submitting through HL7.

- Missing shots (VFC vs. Privately funded shots).

- Data for the actual shots given has better overall data quality than demographic elements.

- Email submission is extremely low (affects public portal) for all age groups and submission type.

OSIIS Data Quality Improvement

- Ensure high quality data through HL7 submission (onboarding thresholds, monthly data quality reports).

- Completing high priority objectives and strategies in data quality plan.

- Assessing data quality at the health district level and contacting providers about areas of low data quality.

- Utilizing CDC and OSDH Immunization Service staff to work with providers to fix various data quality issues.

- Engaging state OSIIS vendor to automate data quality reports for all submission types.

Data Quality Improvement: Future

- To ensure high data quality, multiple approaches have to be used including programmatic and technical resources.

- **Educate provider organizations about data quality (make it visible).**

- Prioritize data quality for data elements that have the most impact for public health and vaccine accountability.

- Assess data quality throughout the process of receiving and managing data (onboarding through evaluating data at rest) <u>for all submission types (manual and HL7).</u>

- Adoption of uniform business rules across jurisdictions that allow for some flexibility.

Questions?

Immunization/OSIIS Information