

Research Letter

# Explainable Machine Learning for Public Health Informatics in HEDIS Childhood Immunization Status Combination 10

---

JiaLi Ruan, MSHealth Care Service Corporation, Naperville, IL, United States

---

**Corresponding Author:**

JiaLi Ruan, MS  
Health Care Service Corporation  
1100 E Warrenville Rd  
Naperville, IL 60563  
United States  
Phone: 1 630-701-4247  
Email: [jruan4529@gmail.com](mailto:jruan4529@gmail.com)

## Abstract

---

Completion of the HEDIS (Healthcare Effectiveness Data and Information Set) Childhood Immunization Status Combination 10 (Combo 10) measure among US children aged 24-35 months declined from 53.7% in 2021 to 44.6% in 2023, with a statistically significant survey-weighted annual trend. An explainable machine learning approach identified influenza vaccination and rotavirus series completion as the strongest component-level drivers of Combo 10 completion, supporting targeted public health quality improvement.

*Online J Public Health Inform 2026;18:e90241*; doi: [10.2196/90241](https://doi.org/10.2196/90241)

---

**Keywords:** public health; childhood immunization; HEDIS; CIS Combo 10; National Immunization Survey; machine learning; random forest; explainable artificial intelligence; SHAP; public health informatics; quality measurement; Shapley Additive Explanations; Childhood Immunization Status Combination 10; Healthcare Effectiveness Data and Information Set

## Introduction

---

Immunizations have substantially reduced vaccine-preventable morbidity and mortality worldwide [1]. The Childhood Immunization Status (CIS) Combination 10 (Combo 10) measure is a core HEDIS (Healthcare Effectiveness Data and Information Set) pediatric quality indicator [2]. Combo 10 assesses whether children receive the complete series of 10 recommended vaccines by 24 months of age; however, national performance remains below desired benchmarks [3]. Identifying component-level drivers of Combo 10 failure is a public health informatics problem: it converts population-level quality data into actionable information for health plans, immunization programs, and policy stakeholders.

Traditional Combo 10 reporting summarizes overall completion but gives limited insight into component-level drivers. Public health informatics can extend routine quality reporting by linking population data, transparent analytics, and actionable interpretation. Explainable artificial intelligence (XAI) can show which inputs drive model output [4-9]. This study applied survey-weighted regression, random forest classification, and SHAP (Shapley Additive Explanations) to

national survey data to identify components most predictive of HEDIS CIS Combo 10 completion.

## Methods

---

### Study Design

The 2021-2023 National Immunization Survey–Child (NIS-Child) public-use files, a nationally representative CDC (Centers for Disease Control and Prevention) survey with provider-verified vaccination records, were analyzed. The age-eligible cohort included 32,997 children aged 24-35 months, aligning with HEDIS CIS specifications. Weighted modeling was restricted to records with valid provider-adjusted child weights (PROVWT\_C) and complete outcome information, yielding 16,021 children. Analytic features were 10 CDC-defined up-to-date (UTD) binary indicators corresponding to Combo 10 components: DTaP, IPV, MMR, Hib, hepatitis B, varicella, PCV, hepatitis A, rotavirus, and influenza. Indicators were harmonized across years; missing or insufficient provider records were coded as not UTD. Survey-weighted logistic regression estimated national trends, odds ratios (ORs), 95% CIs, and *P* values. For

XAI analysis, a child-level random forest classifier modeled Combo 10 completion from the 10 UTD indicators, incorporating provider weights; 5-fold cross-validation was used as a model-checking step. Because Combo 10 is defined by these indicators, the classifier was used for SHAP-based decomposition rather than independent outcome prediction. SHAP values estimated each component's contribution to Combo 10 classification; mean absolute SHAP values were used for global ranking.

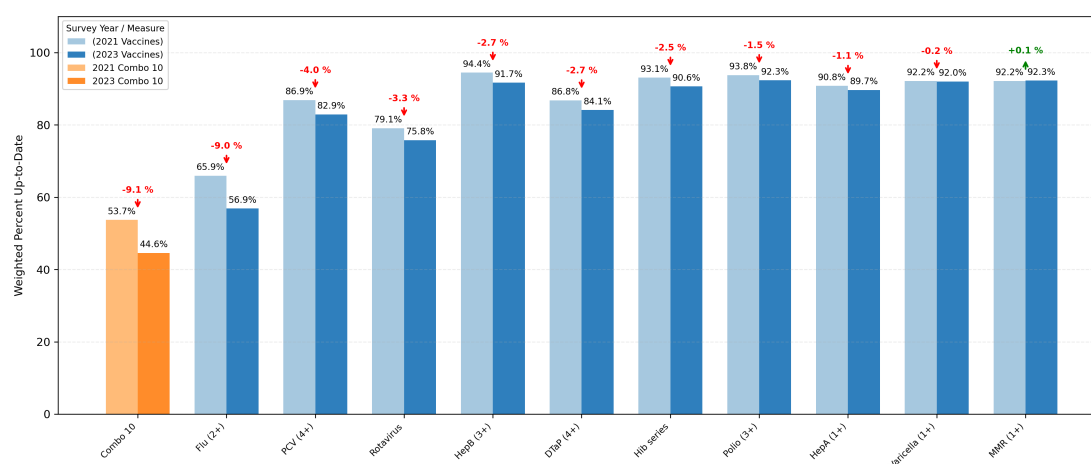
## Ethical Considerations

This study used publicly available, deidentified NIS-Child public-use data and did not involve human subjects research as defined by federal regulations; therefore, institutional review board approval was not required.

## Results

The age-eligible cohort included 32,997 children; 16,021 had valid provider weights and complete outcome information. National HEDIS CIS Combo 10 completion declined from 53.7% in 2021 to 49.6% in 2022 and 44.6% in 2023, a 9.1-percentage-point decrease. The survey-weighted regression model was statistically significant for the annual trend (annual OR 0.83, 95% CI 0.831-0.834;  $P < .001$ ). Figure 1 shows nonuniform component declines, supporting component-level informatics analysis rather than aggregate rates alone.

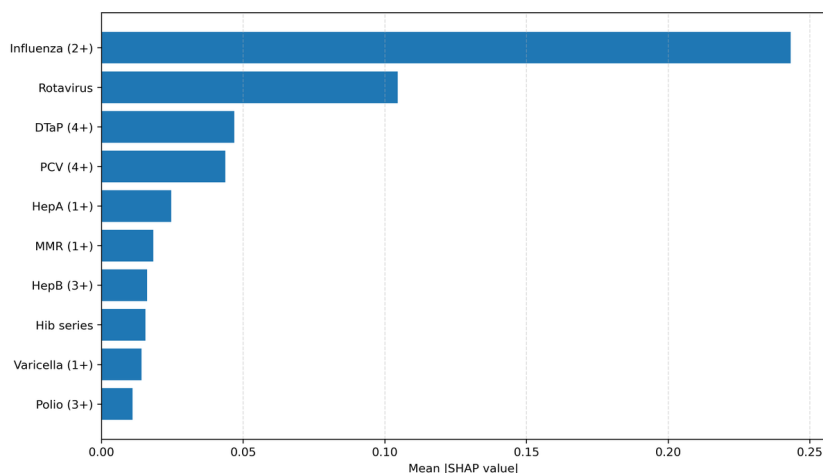
**Figure 1.** National vaccine coverage changes from 2021 to 2023, sorted by decline, with the Childhood Immunization Status (CIS) Combination 10 (Combo 10) measure highlighted.



Five-fold cross-validation was used to check model stability. Because CIS Combo 10 completion is defined by the 10 vaccine component indicators, classifier performance was expected to be near-perfect and was not interpreted as independent predictive validity. The random forest model was used for SHAP-based decomposition. SHAP rankings showed that influenza vaccination and rotavirus series completion had

the largest mean absolute SHAP values and were therefore the strongest drivers of predicted Combo 10 status. DTaP and PCV had moderate importance, while MMR, varicella, IPV, Hib, and hepatitis B had lower SHAP influence. Figure 2 displays SHAP-based feature importance rather than native impurity-based importance measures.

**Figure 2.** Mean absolute Shapley Additive Explanations (SHAP) values for CIS Combination 10 (Combo 10) prediction.



## Discussion

The main finding was that incomplete influenza and rotavirus vaccination disproportionately explained HEDIS CIS Combo 10 underperformance in this national sample. This pattern is actionable because practices and health plans can contact patients who are due or overdue for immunizations, use visit-based prompts to reduce missed vaccines, and train staff to discuss immunization with families. Future motivational interviewing programs could pair these outreach workflows with family-centered communication and could be adapted to other HEDIS measures to improve public health quality.

This study illustrates how XAI can complement quality-measure reporting within public health informatics. Rather

than treating machine learning as a black box, SHAP linked each prediction to vaccine-specific inputs and distinguished components merely included in HEDIS from those driving noncompletion. The approach can support transparent, data-driven quality improvement for clinicians, health plans, public health agencies, and policy stakeholders [5,6,8,9].

Limitations include use of public-use NIS-Child data rather than plan-level HEDIS claims and reliance on a global SHAP bar plot. SHAP values indicate model-based component contributions rather than causal effects; future public health informatics work should validate findings in plan-level HEDIS data and examine subgroup patterns with beeswarm or dependence plots.

## Funding

The author declared no financial support was received for this work.

## Conflicts of Interest

None declared.

## References

1. Zhou F, Jatlaoui TC, Leidner AJ, et al. Health and economic benefits of routine childhood immunizations in the era of the Vaccines for Children program - United States, 1994-2023. *MMWR Morb Mortal Wkly Rep*. Aug 8, 2024;73(31):682-685. [doi: [10.15585/mmwr.mm7331a2](https://doi.org/10.15585/mmwr.mm7331a2)] [Medline: [39116024](https://pubmed.ncbi.nlm.nih.gov/39116024/)]
2. Office of Disease Prevention and Health Promotion. Healthcare Effectiveness Data and Information Set (HEDIS). Healthy People 2030. URL: <https://odphp.health.gov/healthypeople/objectives-and-data/data-sources-and-methods/data-sources/healthcare-effectiveness-data-and-information-set-hedis> [Accessed 2026-06-20]
3. Hill HA, Yankey D, Elam-Evans LD, et al. Decline in vaccination coverage by age 24 months and vaccination inequities among children born in 2020 and 2021 - National Immunization Survey-Child, United States, 2021-2023. *MMWR Morb Mortal Wkly Rep*. Sep 26, 2024;73(38):844-853. [doi: [10.15585/mmwr.mm7338a3](https://doi.org/10.15585/mmwr.mm7338a3)] [Medline: [39325676](https://pubmed.ncbi.nlm.nih.gov/39325676/)]
4. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Presented at: 31st Conference on Neural Information Processing Systems (NIPS 2017); Dec 4-9, 2017; Long Beach, CA, United States. URL: <https://papers.neurips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions> [Accessed 2026-07-28]
5. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med*. Sep 2019;25(9):1337-1340. [doi: [10.1038/s41591-019-0548-6](https://doi.org/10.1038/s41591-019-0548-6)] [Medline: [31427808](https://pubmed.ncbi.nlm.nih.gov/31427808/)]
6. Paneru B, Paneru B, Sapkota SC, Poudyal R. Enhancing healthcare with AI: sustainable AI and IoT-powered ecosystem for patient aid and interpretability analysis using SHAP. *Measurement: Sensors*. Dec 2024;36:101305. [doi: [10.1016/j.measen.2024.101305](https://doi.org/10.1016/j.measen.2024.101305)]
7. Arya V, Bellamy RKE, Chen PY, et al. One explanation does not fit all: a toolkit and taxonomy of AI explainability techniques. *arXiv*. Preprint posted online on Sep 6, 2019. [doi: [10.48550/arXiv.1909.03012](https://doi.org/10.48550/arXiv.1909.03012)]
8. Salih A, Raisi-Estabragh Z, Boscolo Galazzo I, et al. A perspective on explainable artificial intelligence methods: SHAP and LIME. *arXiv*. Preprint posted online on May 3, 2023. [doi: [10.1002/aisy.202400304](https://doi.org/10.1002/aisy.202400304)]
9. Ghasemi A, Hashtarkhani S, Schwartz DL, Shaban-Nejad A. Explainable artificial intelligence in breast cancer detection and risk prediction: a systematic scoping review. *arXiv*. Preprint posted online on Jul 12, 2024. [doi: [10.48550/arXiv.2407.12058](https://doi.org/10.48550/arXiv.2407.12058)]

## Abbreviations

**CDC:** Centers for Disease Control and Prevention  
**CIS:** Childhood Immunization Status  
**Combo 10:** Combination 10  
**HEDIS:** Healthcare Effectiveness Data and Information Set  
**NIS-Child:** National Immunization Survey-Child  
**OR:** odds ratio  
**SHAP:** Shapley Additive Explanations  
**UTD:** up to date  
**XAI:** explainable artificial intelligence

*Edited by Edward Mensah; peer-reviewed by Manop Kaewmoracharoen, Sonnia Oliva; submitted 29.Dec.2025; final revised version received 23.Jun.2026; accepted 21.Jul.2026; published 31.Jul.2026*

Please cite as:

Ruan J

*Explainable Machine Learning for Public Health Informatics in HEDIS Childhood Immunization Status Combination 10*

*Online J Public Health Inform 2026;18:e90241*

URL: <https://ojphi.jmir.org/2026/1/e90241>

doi: [10.2196/90241](https://doi.org/10.2196/90241)

© JiaLi Ruan. Originally published in the Online Journal of Public Health Informatics (<https://ojphi.jmir.org/>), 31.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Online Journal of Public Health Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://ojphi.jmir.org/>, as well as this copyright and license information must be included.