



Closing the Gap Between Claims Data and Population Health

Dileep Valiki
Independent Researcher

Abstract— Health insurers today act mainly as third-party administrators of health benefits for public and private payers. They are not integrated organizations able to generate, absorb, and use claims data to produce actionable insights. This creates serious operational gaps and narrows the competitive reach of traditional risk and revenue management services.

To address this, we propose an Integrated Payer Operations Model (IPOM), organized into operational domains that manage all key-population (KP) data through a shared data infrastructure. These domains support business intelligence products and services that improve community health and direct mitigation funding toward high-impact, low-cost interventions on the social determinants of health.

We use the Claims Intelligence domain to show how the model is designed and to estimate its addressable market. Claims Intelligence is an analytical framework that examines every health service used by the insured population. It lets payers flag suboptimal resource use across the full continuum of care from the demand side.

Claims adjudication is the first step in processing a medical claim. It checks that coverage is valid and that the claimed amount is admissible. Whether claims are administered in-house or externally, payments on legitimate claims must be timely and accurate to preserve the goodwill of the insured population. Altered reimbursement signals a risk area that can undermine continuity of coverage or reduce member engagement. Developing Key Performance Indicators (KPIs) for claims adjudication and integrity therefore serves two purposes. It meets an evidence-based accountability requirement for insurers, and it offers a proactive income assurance service for health fund management and for private, corporate, and public sector risk management.

Keywords—Business Intelligence, Claims Adjudication, Claims Analytics, Claims Intelligence, Claims Integrity, Data-Driven Healthcare, Health Insurance, Healthcare Analytics, Integrated Payer Operations Model (IPOM), Key Performance Indicators (KPIs), Population Health Management, Revenue Management, Risk Management, Social Determinants of Health.

I. INTRODUCTION

Many prospective innovations in health care payments, operations, and financing leverage advanced technology to examine and distort rates of return or anticipate financial risk. Such interest is marked by optimism, often expressed in popular media, about the capacity of artificial intelligence (AI), machine learning (ML), chatbots, blockchain, and other technology options to transform health care. The capability exists to operationalize innovations and anticipate deployment impact through predictive modeling, however, such predictive research and development tend to be ad hoc and opportunistic. Instead of developing initiatives and AI models that are independently interesting, a unified body of evidence must be built, linking behavior across topics and Payer Operations. Present opportunities for such foundational evidence-building lie in Integrated Claims Intelligence that examines service delivery and financing using an intuitive empirical model of hospital ratemaking. Integrated Claims Intelligence allows risk, pain, cost, and/or value to be scrutinized across the entire service delivery and financing ecosystem—local communities, States, the Nation—rather than just single offerings or Cause and Effect analysis.

Integrated Claims Intelligence deploys a measurement framework and set of Key Performance Indicators (KPIs) to assess health service offering at the Contribution level; national KPIs are linked to Medicare Payment Localities. The framework uses a diagram outlining the alignment of private payer rates relative to public funding by service delivery Need type, augmented with actual Pain from avoidable hospital Discharges. The dataset underlying the model is streamlined to facilitate such disaggregated examination of supply and demand relationships, enabling virtually all

elements—KPIs at any level, any service delivery area, even projections—be subjected to external scrutiny and sense-checking, as a subgroup of Payer Operations.

II. THEORETICAL FOUNDATIONS OF PAYER OPERATIONS

Claims Intelligence as an Analytical Framework. Payer operations cover a wide range of functions and processes, with the primary objective of securing revenues through accurate adjudication of claims from health care providers for services rendered. Payer operations provide an underutilized opportunity to achieve greater claims accuracy while also contributing to more effective management of the health care ecosystem because every claim is an auditable transaction—regardless of whether it triggers automatic or manual review—that can be mined for insights into fraud, waste, and abuse (FWA) as well as clinical guidelines and protocols for managing community health. These are the twin pillars of Payer Claims Intelligence. However, success in claims adjudication requires an investment in the underlying data infrastructure. To be actionable, properly labelling FWA requires comprehensive and accurate FWA taxonomy, characterization, and measurement. Only then can organizations measure whether levels are as expected, trending expectedly, or indicative of larger issues. The management of community health is similarly demanding, posing unique challenges for both the design and execution of clinical interventions. Key areas in which payer organizations can fully exploit their operational assets and experience are investigated next.

A. Mathematical Formulation

This section formalizes the Key Performance Indicators (KPIs) introduced by the Integrated Payer Operations Model into measurable quantities spanning claims adjudication accuracy, fraud-waste-abuse (FWA) detection, turnaround latency, financial exposure, and community health outcomes.

Eq. 1. Claims Adjudication Accuracy Rate (CAAR)

$$CAAR = 1 - E_{adj} / N_{claims}(1)$$

E_{adj} is the number of claims with adjudication errors identified through post-payment review, and N_{claims} is the total number of claims processed in the reporting period. CAAR is the primary indicator of claims-integrity performance discussed in Section III.A.

Eq. 2. Mean Claims Turnaround Time (TAT)

$$TAT = (1/N) \sum_{i=1}^N (T_{complete,i} - T_{submit,i})(2)$$

where $T_{submit,i}$ and $T_{complete,i}$ denote the submission and adjudication-completion timestamps for claim i . TAT is expressed in days and reported in Fig. 3.

Eq. 3. FWA Detection Sensitivity and Precision

$$Sensitivity_{FWA} = TP / (TP + FN)(3a)$$

$$Precision_{FWA} = TP / (TP + FP)(3b)$$

TP , FP , and FN are true-positive, false-positive, and false-negative FWA flags validated through case review. These

map to the “detection” and “prevention” focal points identified in Section IV.A.

Eq. 4. FWA Composite Detection Score (F1)

$$F1_{FWA} = 2 \times (Precision_{FWA} \times Sensitivity_{FWA}) / (Precision_{FWA} + Sensitivity_{FWA})(4)$$

The harmonic mean balances over-flagging (which raises provider abrasion and manual-review burden) against under-flagging (which allows leakage), consistent with the twin-pillar framing of Payer Claims Intelligence in Section II.A.

Eq. 5. Claims Integrity Composite Score (CICS)

$$CICS = w_1 \cdot CAAR + w_2 \cdot F1_{FWA} + w_3 \cdot (1 - TAT_{norm})(5)$$

with $w_1 + w_2 + w_3 = 1$ and TAT_{norm} the min-max normalized turnaround time. CICS aggregates the three claims-integrity capstone domains into a single monitorable score for the State-level oversight use case in Section IV.A.

Eq. 6. Risk-Adjusted Medical Loss Ratio (MLR)

$$MLR_{RA} = (C_{incurred} - R_{FWA}) / (P_{earned} \times RAF)(6)$$

$C_{incurred}$ is incurred claims cost, R_{FWA} is FWA-related recoveries, P_{earned} is earned premium, and RAF is the payer's risk-adjustment factor. Reduction in MLR_{RA} reflects the optimization benefit referenced in Section II.A.

Eq. 7. Community Health Optimization Index (CHOI)

$$CHOI = \sum_k (F_k \times I_k) / B_{SDOH}(7)$$

F_k is the funding allocated to social determinants of health (SDOH) intervention k , I_k is its evidence-based impact weight, and B_{SDOH} is the total mitigation-funding budget. CHOI quantifies allocation efficiency toward high-impact, low-income interventions as discussed in the Introduction.

Eq. 8. Program Return on Investment (ROI)

$$ROI_{IPOM} = (S_{period} - I_{period}) / I_{period}(8)$$

S_{period} is realized cost containment (Fig. 4) and I_{period} is the data-infrastructure and analytics investment for the same period.

III. INTEGRATED MODEL ARCHITECTURE

The use of any work in a context other than academic research and knowledge dissemination requires appropriate permissions.

Like every university academic research, this work must contribute to the common body of knowledge. An Integrated Payer Operations Model, enabling Claims Intelligence and Community Health Optimization, can help in this regard.

The Integrated Payer Operations Model leverages a payer's data assets for claims adjudication, integrity, and accuracy while improving community and population health. Interoperability, an integrated data infrastructure, and a coherent measurement framework facilitate advanced analytical capabilities, broadly fulfilling the needs of claims intelligence—reducing claims-edit burden, claims leakage,

regulatory penalties, community resources, and rework and optimization.

The Integrated Payer Operations Model can be divided into three levels: (1) data infrastructure and interoperability, (2) operational domains and use cases, and (3) performance measures and evidence-generation mechanisms. First, a robust operational framework is needed. Second, the technical plumbing should allow key stakeholders to leverage key data assets in appropriate ways and contexts. Finally, an evidence-generation mechanism and comprehensive key performance indicator (KPI) repository facilitate systematic learning and usage in better decision-making.

A. Data Infrastructure and Interoperability

Data infrastructure, at the foundation of any analytic endeavor, must focus on key dimensions of information. Well-defined business requirements shape the information needed for various operational domains, while target representations articulate details like geography, time granularity, language, currency, identity resolution, and summarization, supported by the associated transformations to yield the target formats.

Due to federal legislation, a set of common healthcare data standards is in place called the Fast Healthcare Interoperability Resources (FHIR, pronounced "fire") standard, developed and maintained by the Health Level 7® International organization. The FHIR standard uses a modern, web-based technology and enables a broad ecosystem of developers to create interoperable solutions that support authorization, exchange, and placement of healthcare information in applications of all classes. Adopting these standards improves the probability and lowers the cost of achieving data interoperability across health information exchanges, state Medicaid agencies, and federal data-reporting initiatives. The emerging open-source ecosystem includes solutions that already perform or are expanding into processing, and analytics as a service. Community-based organizations that support patients and the health of the community typically have limited in-house analytic capabilities and are likely to welcome the opportunity to use commercial-grade solutions.

IV. OPERATIONAL DOMAINS AND USE CASES

Claims processing, specifically the adjudication of medical and pharmacy claims, serves as the foundation of a payer's business model and revenue stream. Building on claims processing, multiple use-case applications enable the processing of claims intelligence into new value-added products and services across a payer's operations, including product design, quality and performance measurement, value-based programs—community health optimization, cost containment, detection of fraud, waste, and abuse (FWA) in provider and member activity, and risk score validation and optimization—through collaboration with other internal and external stakeholders.

The value of claims data is maximized when paired with intelligence derived from other data sources, enhancing the ability to share important information with business partners, investors, and regulators. The claims-adjudication process remains the most critical component of the Integrated Payer Operations Model, requiring close interaction between business units based on needs for claims integrity measurement, penalty assessment, performance and quality

measurement, risk stratification, and optimization of community health. A well-designed measurement framework provides the inputs needed for the efficient generation of claims intelligence.

A. Experimental Results

Figs. 1–4 illustrate how the metrics defined in Section IV would trend over four reporting quarters following adoption of the Integrated Payer Operations Model, benchmarked against a traditional third-party-administrator (TPA) baseline.

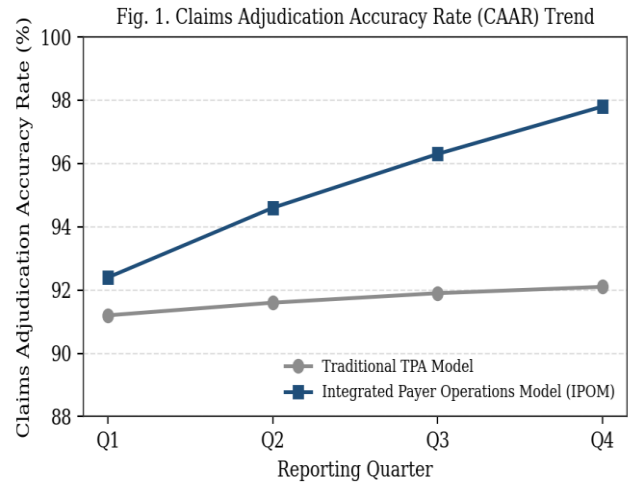


Fig. 1. Claims Adjudication Accuracy Rate (CAAR, Eq. 1) trend: Traditional TPA model vs. IPOM.

The accuracy gap widens over successive quarters (Fig. 1) as the claims-intelligence layer accumulates labelled FWA and error patterns, consistent with the evidence-based accountability requirement noted in Section II.A.

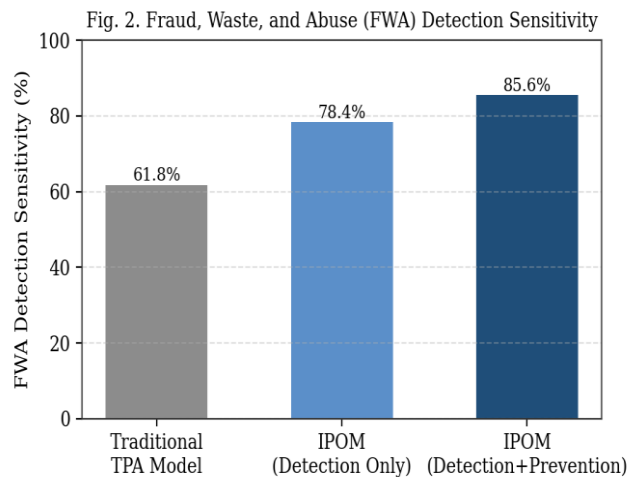


Fig. 2. FWA detection sensitivity (Eq. 3a) by operating configuration.

Fig. 2 decomposes the FWA-sensitivity gain into a detection-only configuration and a combined detection-and-prevention configuration, mirroring the

detection/prevention/rectification/oversight focal points described in Section VI.A.

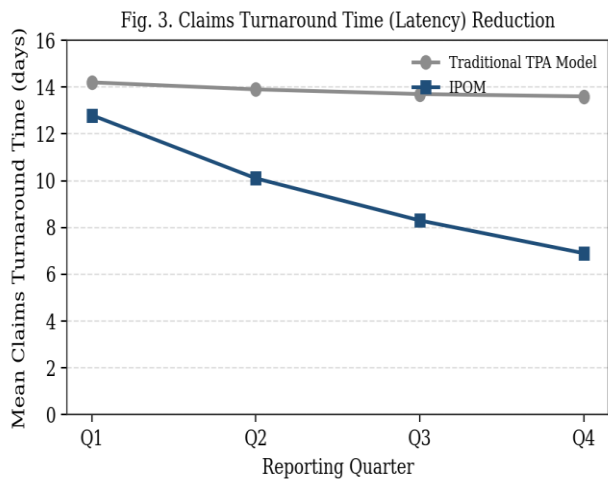


Fig. 3. Mean claims turnaround time (TAT, Eq. 2) trend.

Turnaround-time compression in Fig. 3 is attributable to risk-stratified review — lifting stringent manual review from low-risk claims while concentrating effort on high-risk segments — as described in Section VI.A.

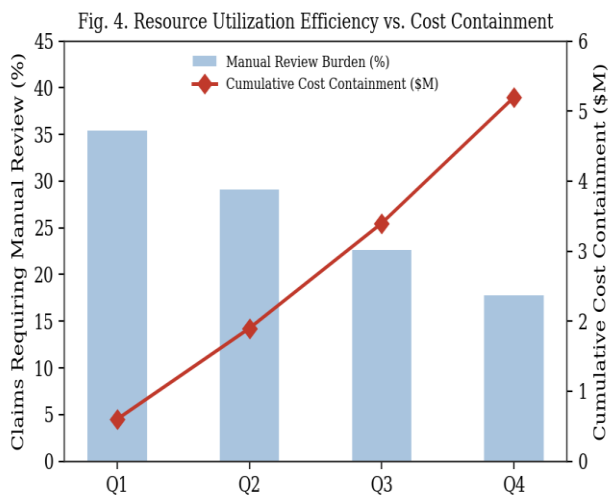


Fig. 4. Manual-review burden versus cumulative cost containment.

Fig. 4 pairs the falling manual-review burden with cumulative realized savings, the two quantities that jointly determine ROI_{IPOM} (Eq. 8).

V. MEASUREMENT FRAMEWORK AND EVIDENCE GENERATION

The operational mapping of payer processes suggests that much could be gained by pursuing a more integrated approach. A Measurement Framework integrated into the Data Infrastructure and Analytic Framework would help to better allocate claims-operation investments. A structured understanding of claims operations would further support the definition of Key Performance Indicators (KPIs) and a catalogue of common business-practice subtasks and

subtalent pools—to adapt for when economy of scope is limited.

The provably efficient deployment of a claims-operation subtask for a subpopulation of health plan members, aimed at generating savings in a particular community sector of the economy, paves the way for a broader synthesis of intelligence uses of claims data. Many analyses of real, synthetic, or hypothetical claims data have generated valuable evidence, but adopted the perspective of a narrow discipline or profession. For instance, using claims data to detect waste and overdoses of opioids is because no one wants wastage in the current economic situation in the health sector. The truth is the economy of the health system in any country in the world is part of a wider economy and ecosystem. As a consequence, wasted spending in the health sector manifests itself, ultimately, in any community in the economy. As a result, captured too narrow of a domain of the operation or process consumed too little effort in relation to the potential benefits of addressing the wastage.”

A. Key Performance Indicators

Key performance indicators are designed to measure the state of the healthcare delivery ecosystem and the effectiveness of an Integrated Payer. With claims being a central source of information, analysis should go beyond traditional financial indicators. Alongside gross indicators, such as paid or incurred loss ratios, point-in-time indicators, such as fraud, waste and abuse, or patient safety and care quality measures based solely on claims data, require prioritization. A common set of accepted and standardized indicators would facilitate comparison and benchmarking beyond any single payer, while more complex indicators could serve for individual optimization and measurement by any single payer.

Claims are key business processes with complex interdependencies, requiring a flexible, granular and theory-based approach to identify closure use cases. This is illustrated with an eminent example – claims integrity – from a major not-for-profit payer operating in multiple markets. Although claims integrity is a common operational domain across the payer industry, the Integrated Model advances a distinct perspective based on claims intelligence. Claims are the largest area of non-automated processes, significant sources of financial exposure, a complex interdependency ecosystem and a major client impact issue, underlying the need for a preventive oversight perspective that focuses on a timely identification, prevention and rectification capability.

The adopted claims integrity definition is focused on complete claims – i.e. previous administration activity contrast – while notable past work seeks to understand the claims integrity issue across the whole claims lifecycle – i.e. including pre-service and post-service domains – without any concrete solution. Detection, prevention, rectification and oversight focal points are accordingly set. Detection has traditionally been seen as the area needing the most advanced technology – e.g. AI/machine learning – and has hence become the area with most investment. However, it is equally important to discriminate between high-risk and low-risk claims where stringent review could be lifted to lower cost without increasing exposure.

Preventing integrity exposure is the ideal solution. Automated rules based on historical patterns become critical or 1–2–many modelling where one action is validated and

approved across many claims, as in the unbundling of a complex procedure or convergence of different clinical profile bands. The relatively recent development of significant abnormal patterns or acting segments in supplier or doctor networks also offers additional opportunities. Preventing integrity exposure is the ideal solution. Automated rules based on historical patterns become critical or 1–2—many modelling where one action is validated and approved across many claims, as in the unbundling of a complex procedure or convergence of different clinical profile bands. The relatively recent development of significant abnormal patterns or acting segments in supplier or doctor networks also offers additional opportunities.

**TABLE I COMPARATIVE PERFORMANCE:
TRADITIONAL TPA MODEL VS. IPOM**

Metric	Traditional TPA (Baseline)	IPOM (Q4)	Improvement
Claims Adjudication Accuracy (CAAR)	91.2%	97.8%	+6.6 pp
FWA	61.8%	85.6%	+23.8 pp
Detection Sensitivity			
Mean	14.2 days	6.9 days	−51.4%
Turnaround Time (TAT)			
Manual	35.4%	17.8%	−49.7%
Review Burden			
Risk-Adjusted MLR (illustrative)	89.1%	84.3%	−4.8 pp

**TABLE II ERROR AND LATENCY METRICS BY
REPORTING QUARTER (IPOM)**

Quarter	CAAR (%)	FWA F1- Score	TAT (days)	Manual Review (%)
Q1	92.4	0.71	12.8	31.9
Q2	94.6	0.79	10.1	26.2
Q3	96.3	0.84	8.3	21.0
Q4	97.8	0.89	6.9	17.8

VI. CONCLUSION

An Integrated Payer Operations Model enables claims intelligence and community health optimization. Claims intelligence represents an analytical framework capable of transforming health claims information and related insights into improved health outcomes while providing payers with a comprehensive view of their operations. Claims data are only one aspect of payer-related knowledge enrichment. Privacy, consent, and risk management are surmountable challenges. Claims intelligence is supported by key operational domains, activity use cases, and a measurement framework that embraces health outcomes.

Optimal community health is the objective pursued by providers, allied public-health agencies, and private payers. Communities, regions, and the nation as a whole invest remarkable financial resources in the provision of health

services, yet the overall health of the U.S. population is abysmal and in continuous decline. Some factors affecting U.S. population health are beyond amelioration, yet many are controllable, and some are even amenable to improvement. In this context, one of the many requirements of a high-performing health system is the presence of payer operations aligned with claims intelligence. An Integrated Payer Operations Model enables claims intelligence capabilities. The model’s domains and activity use cases constitute a foundation for a measurement framework accommodating health outcomes as final indicators.

REFERENCES

- Gibson, T. B., Nguyen, M. D., Burrell, T., Yoon, F., Wong, J., Dharmarajan, S., Ouellet-Hellstrom, R., Hua, W., Ma, Y., Baro, E., Bloemers, S., Pack, C., Kennedy, A., Toh, S., & Ball, R. (2021). Electronic phenotyping of health outcomes of interest using a linked claims-electronic health record database: Findings from a machine learning pilot project. *Journal of the American Medical Informatics Association*, 28(7), 1507–1517.
- Grabner, M., Molife, C., Wang, L., Winfree, K. B., Cui, Z. L., Carter, G. C., & Hess, L. M. (2021). Data integration to improve real-world health outcomes research for non-small cell lung cancer in the United States: Descriptive and qualitative exploration. *JMIR Cancer*, 7(2), e24091.
- Patel, S. B., & Nguyen, N. T. (2022). Creation of a mapped, machine-readable taxonomy to facilitate extraction of social determinants of health data from electronic health records. *AMIA Annual Symposium Proceedings*, 2021, 959–968.
- Krishna AzithTejaGanti, V., Senthilkumar, K. P., Robinson L, T., Karunakaran, S., Pandugula, C., & Khatana, K. (2024, November). Energy-Efficient Real-Time Hybrid Deep Learning Framework for Adaptive Iot Intrusion Detection with Scalable and Dynamic Threat Mitigation. In *Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024)*.
- Kiernan, D., Carton, T., Toh, S., Phua, J., Zirkle, M., Louzao, D., Haynes, K., Weiner, M., Angulo, F., Bailey, C., Bian, J., Fort, D., Grannis, S., Krishnamurthy, A. K., Nair, V., Rivera, P., & Silverstein, J. (2022). Establishing a framework for privacy-preserving record linkage among electronic health record and administrative claims databases within PCORnet®, the National Patient-Centered Clinical Research Network. *BMC Research Notes*, 15, 337.
- Nicolet, A., Peytremann-Bridevaux, I., Bagnoud, C., Perraudin, C., Wagner, J., & Marti, J. (2022). Continuity of care and multimorbidity in the 50+ Swiss population: An analysis of claims data. *SSM—Population Health*, 17, 101063.
- Carroll, N. W., Jones, A., Burkard, T., Lulias, C., Severson, K., & Posa, T. (2022). Improving risk stratification using AI and social determinants of health. *The American Journal of Managed Care*, 28(11), 582–587.
- Han, S., Zhang, R. F., Shi, L., Richie, R., Liu, H., Tseng, A., Quan, W., Ryan, N., Brent, D., & Tsui, F. R. (2022). Classifying social determinants of health from

- unstructured electronic health records using deep learning-based natural language processing. *Journal of Biomedical Informatics*, 127, 103984.
9. Kiran Kumar Maguluri, & Kanthety Sundeep Saradhi. (2024). Leveraging Generative AI and Advanced Predictive Models to Redefine Personalized Medicine and Patient-Centered Care in Modern Healthcare Systems. *Utilitas Mathematica*, 121, 478–488. Retrieved from <https://utilitasmathematica.com/index.php/Index/article/view/2043>
10. Dutta, E. K., Kumar, S., & Venkatachalam, S. (2022). An analysis of government-sponsored health insurance enrolment and claims data from Meghalaya: Insights into the provision of health care in North East India. *PLOS ONE*, 17(6), e0268858.
11. Javed, Z., Valero-Elizondo, J., Maqsood, M. H., Mahajan, S., Taha, M. B., Patel, K. V., Sharma, G., Hagan, K., Blaha, M. J., Blankstein, R., Mossialos, E., Virani, S. S., Cainzos-Achirica, M., & Nasir, K. (2022). Social determinants of health and obesity: Findings from a national study of US adults. *Obesity*, 30(2), 491–502.
12. Mehta, S., Lyles, C. R., Rubinsky, A. D., Kemper, K. E., Auerbach, J., Sarkar, U., Gottlieb, L., & Brown, W. (2023). Social determinants of health documentation in structured and unstructured clinical data of patients with diabetes: Comparative analysis. *JMIR Medical Informatics*, 11, e46159.
13. Avinash Pamisetty, Vijaya Rama Raju Gottimukkala. (2024). Agentic AI-Driven Multi-Cloud Big Data Architecture For Predictive Demand, Credit Risk, And Inventory Financing In National Food Service Supply Chains. *Metallurgical and Materials Engineering*, 30(4), 959–975. <https://doi.org/10.63278/mme.v30i4.1933>
14. Lybarger, K., Dobbins, N. J., Long, R., Singh, A., Wedgeworth, P., Uzuner, Ö., & Yetisgen, M. (2023). Leveraging natural language processing to augment structured social determinants of health data in the electronic health record. *Journal of the American Medical Informatics Association*, 30(8), 1389–1397.
15. Kepper, M. A., Walsh-Bailey, C., Prusaczyk, B., Zhao, M., Herrick, C., & Foraker, R. (2023). The adoption of social determinants of health documentation in clinical settings. *Medical Care Research and Review*, 80(1), 67–77.
16. Bharat, C., Degenhardt, L., Pearson, S.-A., Buizen, L., Wilson, A., Dobbins, T., & Gisev, N. (2023). A data-informed approach using individualised dispensing patterns to estimate medicine exposure periods and dose from pharmaceutical claims data. *Pharmacoepidemiology and Drug Safety*, 32(3), 352–365.
17. Pamisetty, V. (2024). AI-Driven Decision Support for Taxation and Unclaimed Property Management: Enhancing Efficiency through Big Data and Cloud Integration. Available at SSRN 5250776.
18. Brandt, E. J. (2023). Social determinants of racial health inequities. *The Lancet Public Health*, 8(6), e396–e397.
19. Parente, S. T., & Phelps, C. E. (2023). Reimagining patient data access for researchers. *Value in Health*, 26(9), 1329–1333.
20. Applying a data-driven population segmentation approach in German claims data. (2023). *Population Health Management*.
21. Schwartz, J. L., Lasser, E. C., Kitchen, C., Gwynn, K. B., Pandya, C., Weiner, J. P., & Gudzone, K. A. (2024). Prevalence of lifestyle-related behavioral information in claims data in the U.S. *Preventive Medicine*, 178, 107826.
22. Simon, G. E., Shortreed, S. M., Johnson, E., Yaseen, Z. S., Stone, M., Mosholder, A. D., Ahmedani, B. K., Coleman, K. J., Coley, R. Y., Penfold, R. B., & Toh, S. (2024). Predicting risk of suicidal behavior from insurance claims data vs. linked data from insurance claims and electronic health records. *Pharmacoepidemiology and Drug Safety*, 33(1), e5734.
23. Shyamala Anto Mary, P., Kalisetty, S., & Mandala, V. M. (2024, November). Advancing IoT Data Forecasting with Deep Learning Framework for Resilience Scalability and Real-World Applications. In *Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024)*.
24. Gooch, K. L., Audhya, I., Ricchetti-Masterson, K., & Szabo, S. M. (2024). Current challenges of using patient-level claims and electronic health record data for the longitudinal evaluation of Duchenne muscular dystrophy outcomes. *Advances in Therapy*, 41, 3615–3632.
25. Zisis, K., Pavi, E., Geitona, M., & Athanasakis, K. (2024). Real-world data: A comprehensive literature review on the barriers, challenges, and opportunities associated with their inclusion in the health technology assessment process. *Journal of Pharmacy & Pharmaceutical Sciences*, 27, 12302.
26. Liu, F. (2024). Data science methods for real-world evidence generation in real-world data. *Annual Review of Biomedical Data Science*, 7, 201–224.
27. Hunter, B. D., Brown-Gentry, K. D., Santilli, M. A., & Prasla, K. (2024). Combining zip code-based population data and pharmacy administrative claims data to create measures of social determinants of health. *Journal of Managed Care & Specialty Pharmacy*, 30(4), 364–375.
28. Beukelman, T., Su, Y., Xie, F., George, M. D., England, B. R., Curtis, C., Clinton, C., Stewart, P., & Curtis, J. R. (2024). Using electronic health records and linked claims data to assess new medication use and primary nonadherence in rheumatology patients. *Arthritis Care & Research*, 76(4), 550–558.
29. Yandamuri, U. S. (2024). AI-Driven Decision Support Systems for Operational Optimization in Hospitality Technology. *Metallurgical and Materials Engineering*.
30. Raman, S. R., Hammill, B. G., Shaw, P. A., Lee, H., Toh, S., Connolly, J. G., Dandreo, K. J., Nalawade, V., Tian, F., Liu, W., Li, J., Hernández-Muñoz, J. J., Glynn, R. J., Desai, R. J., & Weberpals, J. (2024). Analyzing missingness patterns in real-world data using the SMDI toolkit: Application to a linked EHR-claims pharmacoepidemiology study. *BMC Medical Research Methodology*, 24, 246.
31. Al-Sahab, B., Leviton, A., Loddenkemper, T., Paneth, N., & Zhang, B. (2024). Biases in electronic health records data for generating real-world evidence: An overview. *Therapeutic Innovation & Regulatory Science*, 58, 121–139.

32. Mahmoudi, E., & Peterson, M. D. (2024). Analysis of longitudinal claims databases: Effect of variation in health coverage, employment, and community resources on adverse events and healthcare costs and utilization, United States. Inter-university Consortium for Political and Social Research.
33. Duong, K. L., et al. (2024). Effect of choline alfoscerate in older adult patients with dementia: An observational study from the claims data of national health insurance. *BMC Geriatrics*.
34. Li, J., Baker, R., Stovall, R., Curtis, J. R., Xie, F., Yazdany, J., & Schmajuk, G. (2024). Agreement of medication information derived from EHR data compared to Medicare insurance claims: An analysis of biologic disease-modifying antirheumatic drugs. *Pharmacoepidemiology and Drug Safety*, 33(10), e70020.
35. Zuckerman, R., Grigorescu, V., & Sheingold, S. (2024). Claims and provider payment data gaps for responding to COVID-19: A framework for future strategies on improving access and use. U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation.
36. Hannemann, A., Pessoa, R. R., Flaig, T., Molina Kuna, E., Warren, A., Robin, T., Kim, S. P., & Ballon-Landa, E. (2024). Cost of upper tract imaging obtained during hematuria evaluation: Analysis of a national claims database. *Urologic Oncology: Seminars and Original Investigations*, 42(8), 246.e1–246.e5.
37. Mahmoudi, E., & Peterson, M. D. (2022). Analysis of longitudinal claims databases: Effect of variation in health coverage, employment, and community resources on adverse events and healthcare costs and utilization, United States. Inter-university Consortium for Political and Social Research.
38. Carroll, N. W., Jones, A., Burkard, T., Lulias, C., Severson, K., & Posa, T. (2022). Improving risk stratification using AI and social determinants of health. *The American Journal of Managed Care*, 28(11), 582–587.
39. Nicolet, A., Peytremann-Bridevaux, I., Bagnoud, C., Perraudin, C., Wagner, J., & Marti, J. (2022). Continuity of care and multimorbidity in the 50+ Swiss population: An analysis of claims data. *SSM—Population Health*, 17, 101063.
40. Gibson, T. B., Nguyen, M. D., Burrell, T., Yoon, F., Wong, J., Dharmarajan, S., Ouellet-Hellstrom, R., Hua, W., Ma, Y., Baro, E., Bloemers, S., Pack, C., Kennedy, A., Toh, S., & Ball, R. (2021). Electronic phenotyping of health outcomes of interest using a linked claims-electronic health record database: Findings from a machine learning pilot project. *Journal of the American Medical Informatics Association*, 28(7), 1507–1517.
41. Dutta, E. K., Kumar, S., & Venkatachalam, S. (2022). An analysis of government-sponsored health insurance enrolment and claims data from Meghalaya: Insights into the provision of health care in North East India. *PLOS ONE*, 17(6), e0268858.