# Exploring the Potential of AI-Assisted Analysis for Advanced Utilization of Medical Claims Data

Ryoma Seto[1)], Masatoshi Ishikawa[1)]

1)Tokyo Healthcare University

**Abstract**

Medical claims (receipt) data represent one of the most comprehensive administrative data sources in healthcare systems, covering diagnoses, procedures, prescriptions, and associated reimbursement information. Despite their extensive coverage and routine availability, claims data have traditionally been used primarily for billing and administrative purposes, with secondary analytical use often constrained by data complexity and analytical workload. Recent advances in generative artificial intelligence (AI), particularly large language models, have raised expectations for new forms of analytical support beyond conventional statistical processing.

This study explores the feasibility and implications of an AI-assisted analysis workflow applied to real-world outpatient claims data. Using aggregated procedure-level data derived from routine clinical practice, we examine how generative AI can support descriptive analysis, concentration analysis, and exploratory interpretation without replacing established quantitative methods. The results suggest that AI-assisted workflows may enhance analytical efficiency and interpretability in the exploratory phase of claims data analysis. Ethical and legal considerations related to secondary use of medical claims data are also discussed.

## 1. Background and Objectives

Medical claims data have long been recognized as a valuable resource for evaluating healthcare utilization, cost structures, and patterns of service provision. In many countries, including Japan, claims data cover nearly all insured medical services, making them uniquely suited for population-level health services research. Previous studies have demonstrated the utility of claims data in analyzing disease prevalence, treatment patterns, healthcare costs, and policy impacts (Iezzoni, 1997; Jollis et al., 1993).

However, the practical use of claims data remains challenging. Claims datasets are typically large-scale, highly coded, and primarily structured for reimbursement rather than research purposes. As a result, researchers must devote substantial effort to data preprocessing, aggregation, and interpretation before meaningful analysis can begin. These challenges have limited the routine use of claims data, particularly in exploratory or hypothesis-generating research.

In recent years, generative AI technologies have emerged as potential tools to support data-intensive analytical workflows. Large language models have demonstrated the ability to summarize complex information, identify patterns in structured and semi-structured data, and generate human-readable explanations. Rather than replacing statistical analysis, such models may function as cognitive support tools, assisting researchers in organizing results and articulating analytical narratives.

The objective of this study is to explore the applicability of an AI-assisted analysis workflow for the advanced utilization of medical claims data. Specifically, we aim to examine whether generative AI can support descriptive and exploratory analysis of claims data while maintaining methodological transparency and ethical integrity.

## 2. Materials and Methods

### 2.1 Data Source

This study used anonymized outpatient claims data collected from a single urban clinic over a two-month period. The dataset consisted of procedure-level aggregated records, including procedure names and corresponding frequencies. All records were aggregated prior to analysis, and no individual patient-level identifiers were included.

Claims data were originally generated for routine reimbursement purposes under the Japanese health insurance system. For the purpose of this study, the data were restructured into a tabular format suitable for descriptive analysis and visualization.

### 2.2 Ethical and Legal Considerations

The secondary use of medical claims data raises important ethical and legal considerations, particularly with respect to patient privacy and data governance. In this study, all data were fully anonymized before analysis, and no attempt was made to re-identify individuals or link records across datasets.

According to Japanese guidelines on medical research involving human subjects, studies using fully anonymized secondary data without individual linkage may be exempt from formal ethical review. Nevertheless, data handling procedures were designed in accordance with the principles of data minimization and responsible use, consistent with international discussions on health data ethics (OECD, 2015).

### 2.3 Analytical Workflow and Role of AI

Initial data aggregation and frequency counting were performed using standard spreadsheet-based methods. Subsequently, generative AI was introduced as an auxiliary analytical tool. Importantly, AI was not used to compute numerical results or perform statistical estimation.

Instead, its role was limited to supporting interpretation by summarizing observed distributions, highlighting concentration patterns, and assisting in the formulation of analytical explanations. This approach aligns with previous discussions on human–AI collaboration in data analysis, in which AI functions as a decision-support or sensemaking tool rather than an autonomous analyst (Shneiderman, 2020).

## 3. Results

### 3.1 Distribution of Medical Procedures

Analysis of procedure frequencies revealed a highly skewed distribution. A small number of procedures accounted for a disproportionately large share of total claims. As illustrated in Fig.1, procedures related to infectious disease testing and routine outpatient management appeared most frequently, reflecting the clinical context of the study period.

Such skewed distributions are consistent with previous findings in claims data research, where healthcare utilization often follows a long-tail pattern (Newhouse et al., 2013). The visualization of procedure frequencies enabled rapid identification of dominant service categories within the dataset.

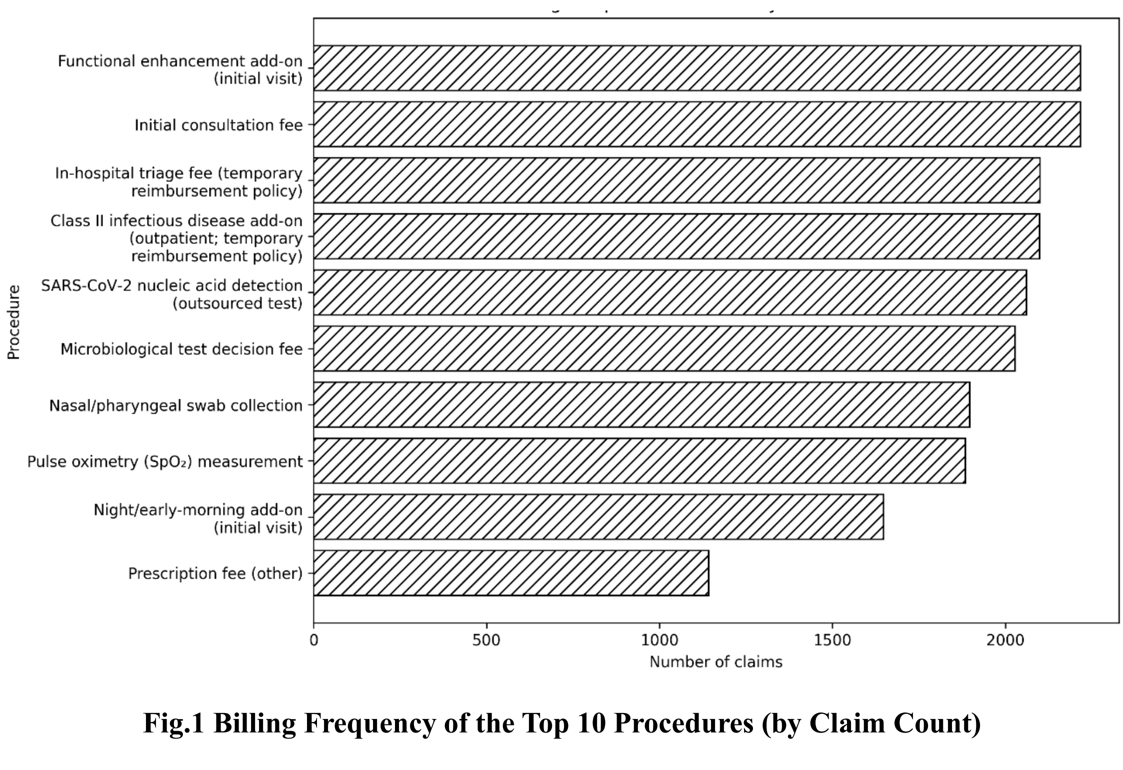


**Fig.1 Billing Frequency of the Top 10 Procedures (by Claim Count)**

### 3.2 Concentration of Claims

To further examine the structure of the data, cumulative concentration analysis was performed. Fig.2 shows the cumulative share of total claims accounted for by procedures ranked in descending order of frequency. The curve rises steeply in the early ranks, indicating that a limited number of procedures explain the majority of observed claims.

This concentration pattern suggests that exploratory analysis can often focus on a relatively small subset of procedures without substantial loss of explanatory coverage. Generative AI was used to assist in articulating this finding in natural language, translating numerical concentration metrics into an interpretable narrative.

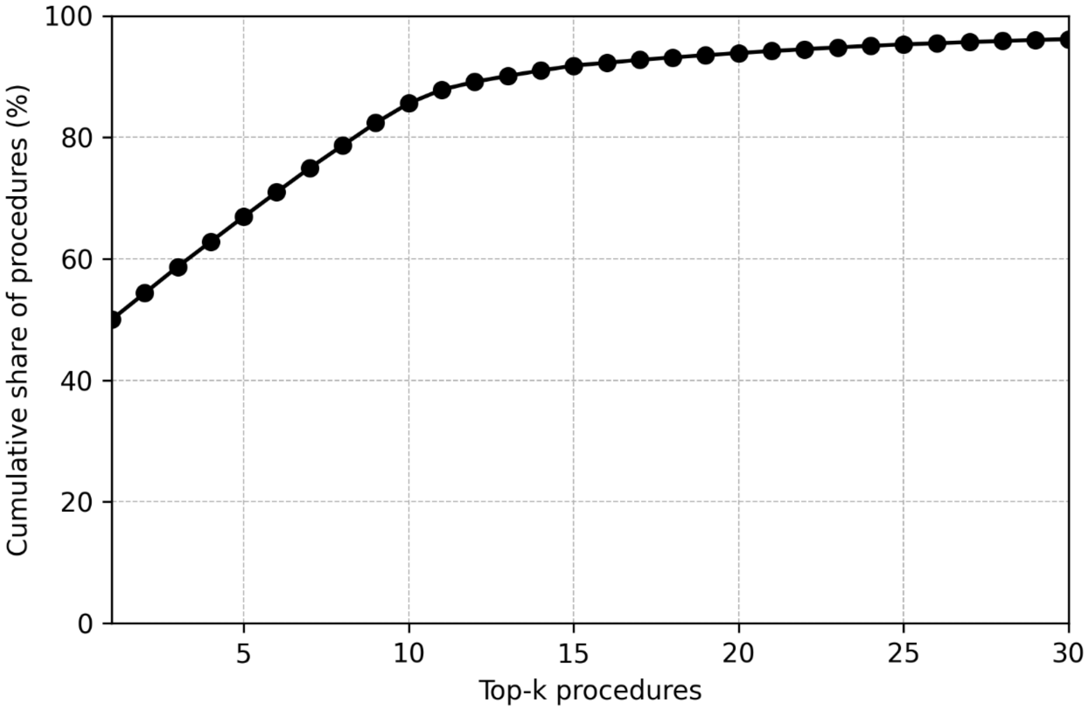


**Fig.2 Concentration Ratio (CRk) of Billed Procedures**

**4. Discussion**

Generative AI may support the interpretation of concentrated claim patterns by introducing an algorithmic, human-in-the-loop workflow. As illustrated in Fig. 3, frequency-based analyses such as CRk first identify dominant procedure combinations, after which generative AI translates these statistically derived patterns into natural language labels and provisional workflow descriptions. Importantly, this process includes explicit decision points and iterative human review, rather than fully automated judgment.

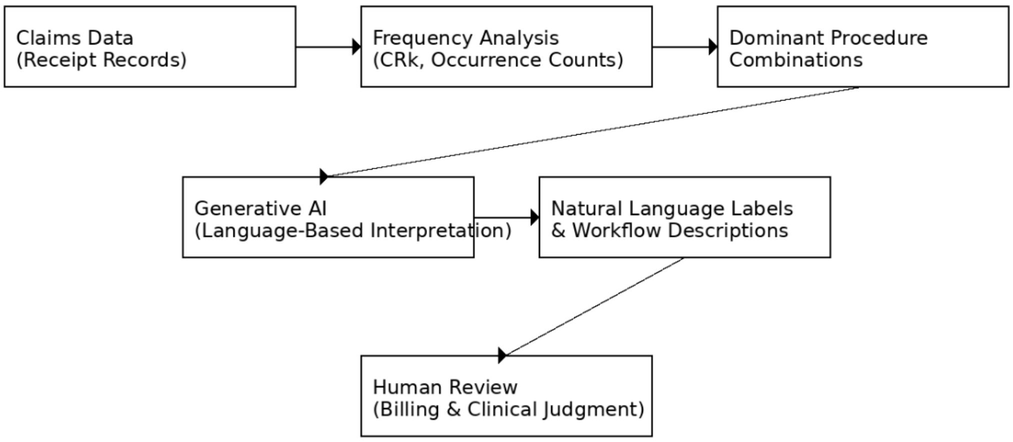


**Fig.3 Conceptual workflow of AI-assisted analysis for claims data interpretation**

From a methodological perspective, the AI-assisted analysis workflow demonstrated potential benefits in the exploratory phase of claims data research. By supporting the interpretation and explanation of descriptive results, generative AI may reduce the cognitive burden associated with large-scale data analysis. This is particularly relevant in early-stage research, where the primary goal is to understand data structure and generate hypotheses rather than test predefined models.

At the same time, the limitations of AI-assisted interpretation must be acknowledged. Generative AI does not inherently verify factual correctness and must therefore be used under human supervision. Transparent reporting of analytical steps and clear separation between computed results and AI-assisted interpretation are essential to maintain scientific rigor.

This study has several limitations. First, the dataset was derived from a single outpatient clinic and reflects local practice patterns, which may limit generalizability to other settings. Validation using multi-institutional or larger-scale datasets will be necessary in future research.

Second, the analysis was based on aggregated procedure-level data rather than patient-level longitudinal records. While suitable for exploratory analysis, this approach does not allow assessment of care trajectories or outcomes at the individual level.

In addition, generative AI was used solely as an assistive tool under expert supervision. AI-generated interpretations may be plausible but inaccurate if applied without domain knowledge, underscoring the importance of human oversight.

Future studies should integrate AI-assisted exploratory workflows with formal statistical methods and robust governance frameworks to support broader and more reliable applications of claims data analysis.

## 5. Conclusions

This study suggests that AI-assisted analysis workflows can serve as a useful complement to conventional methods in the advanced utilization of medical claims data. When applied appropriately, generative AI may enhance efficiency and interpretability in descriptive and exploratory analysis without compromising ethical or methodological standards. Future research should further examine best practices for integrating AI into health services research workflows.

**References**

Iezzoni LI. *Risk Adjustment for Measuring Health Care Outcomes*. Health Administration Press; 1997.

Jollis JG, Ancukiewicz M, DeLong ER, et al. Discordance of databases designed for claims payment versus clinical information systems. *Ann Intern Med*. 1993;119(8):844–850.

Newhouse JP, et al. Variation in health care spending: target decision making, not geography. *Health Aff*. 2013;32(7):1223–1229.

OECD. *Health Data Governance: Privacy, Monitoring and Research*. OECD Publishing; 2015.

Shneiderman B. Human-centered artificial intelligence: Reliable, safe & trustworthy. *Int J Hum-Comput Interact*. 2020;36(6):495–504.