AI Works Better with the ACG® System

Population Health Analytics in an Era of Machine Learning

Authored by: The Johns Hopkins University (JHU) Center for Population Health IT (CPHIT) and The Johns Hopkins HealthCare Solutions (JHHCS) ACG System Team

This report explores the opportunities and challenges of applying artificial intelligence (AI) and machine learning (ML) to health care and population health analytics. It presents a framework for how the Johns Hopkins ACG System — developed at the Johns Hopkins Bloomberg School of Public Health — serves as an ideal "feature layer" for AI/ML applications. The paper also defines key data science terms, highlights proven AI-ACG System use cases, outlines common implementation pitfalls, and provides practical guidance for safe and equitable adoption of AI and ML.

JOHNS HOPKINS MEDICINE

Table of Contents

Introduction – AI, ML and Population Health Analytics

Health care organizations are increasingly integrating artificial intelligence (AI) with established population health analytic systems like the ACG System. This report examines the opportunities, challenges and practical applications of this integration, and provides guidance for organizations seeking to leverage these technologies to maintain and improve patient care quality, efficiency and equity.

Artificial intelligence often makes use of statistical — machine learning (ML) — approaches to drive population health analytics. When combined with the clinically validated ACG System, organizations gain transparent and scalable tools for identifying risk, personalizing outreach and allocating resources. The integration of AI with the ACG System has the potential to substantially advance population health management by providing a more comprehensive and accurate understanding of patient risk.

This report provides:

- An overview of key AI trends in health care with an emphasis on population health;
- Clear working definitions of key AI and ML concepts in the context of health care;
- Practical examples of AI-ACG System integration;
- A critical analysis of key challenges and opportunities related to AI; and
- Actionable recommendations for implementation of AI models.

Towards an Understanding of the AI Landscape in Health Care

The health care AI field is evolving rapidly. Terms, concepts and labels can be confusing, and at times applied differentially by tech companies competing in the crowded AI marketplace. Moreover, what is termed "AI" today may have been labeled something else as recently as a few years ago. In the following section, we provide a series of observations and operational definitions intended to help professionals better understand and work within this swiftly moving field.

AI applications in health care can be categorized into four domains:

1. Chatbots and Knowledge Management:
Systems that can interact with users, answer questions and provide information from existing knowledge bases. These range from simple rules-based systems to sophisticated conversational AI that can engage in complex medical discussions. Innovative tools to manage large, unstructured knowledge bases are usually central to the functioning of chatbots.

2. Automation and Digitalization:
Tools that streamline both administrative and clinical workflows — such as automated coding systems and robotic process automation (RPA) for claims processing — help reduce the manual workload on providers, allowing them to focus more on patient care.

3. Robotics and Sensing:
Physical systems that can perform tasks or interact with the real-world and gather visual, audio or other information. These include surgical robots, diagnostic imaging systems, wearable health monitors and ambient listening devices.

4. Big Data Analytics:
Focuses on analyzing raw data to create knowledge. Includes predictive modeling, machine learning and natural language processing. This domain currently is having a significant impact on health care analytics and decision-making.

While there is not full consensus regarding all terms in the evolving data science field, based on the current technical and popular literature, the below section offers some working definitions of the AI tools and functions introduced above:

Machine Learning (ML) is generally considered a subset of AI and concerns algorithms that learn from data and include supervised and unsupervised methods. ML systems — such as neural networks — can be tailored to improve their performance over time as they are exposed to more data and adapt to a changing health care landscape. When applied appropriately, ML is especially valuable in use cases where patterns are complex, non-linear, and continually evolving, such as risk prediction, patient stratification and resource allocation.

Big Data refers to datasets that are too large or complex for traditional data processing methods, and they are often characterized by the "three Vs": volume (large amounts of data), velocity (rapid data generation), and variety (diverse, often non-standard data types). Health care generates enormous amounts of big data daily, from clinical records to genomic sequences to population health surveillance data. Two big data challenges in health care are the lack of full interoperability and the large amounts of unstructured free text such as clinician notes. AI, ML and other data science techniques have helped address these challenges.

Predictive Analytics uses human-guided statistical techniques — such as linear and logistic regression — as well as machine learning to analyze current and historical data in order to predict future events or behaviors. In health care, these methods can be used to forecast disease outbreaks, identify patients at risk for certain events, like hospital readmission, or anticipate patterns in resource utilization.

Natural Language Processing (NLP) is the encoding of key linguistic information into a format that can be readily used by ML and analytic tools. This is particularly important in health care, where much valuable information exists in unstructured text format — clinical notes found in electronic health records (EHRs), radiology reports, discharge summaries and patient communications. The recent adoption of Ambient AI tools, where voice recognition devices can listen to patient/provider interactions to help document spoken text during a visit, is one type of NLP technology.

Large Language Models (LLMs) are advanced NLP systems that — while used for many purposes — are currently most widely applied to understand and generate human-like text, as exemplified by chatbot systems like ChatGPT®. These systems represent a significant advancement in AI capability, especially when applied to knowledge management and when generating new information (sometimes termed "generative AI"). Though the evaluative evidence is still being collected, LLMs may have many potential applications ranging from clinical documentation to case summarization.

The discussion in this report is relevant to applications across many types of health care. However, the focus will be on the application of machine learning to predictive modeling using big data in the population health context. This discussion will be especially relevant to private or public health care organizations responsible for delivery, management and financing of public health and health care services for communities or large enrolled groups of consumers/patients.

The ACG System as an AI-Ready Feature Layer

For more than three decades, the ACG System has been used to categorize diagnoses, pharmacy and other digital data into over 200+ clinical markers and 15+ risk models that fully describe current morbidity burden and medical/social risk factors. The ACG System algorithms use these risk factors and benchmark data against a reference population to predict future health care utilization for each individual in a population. Its robust mapping logic is based on clinical cogency, predictive accuracy and years of real-world validation (backed up by over 1,000 peer reviewed academic papers). The ACG System is used in 20+ nations and is applied to hundreds of millions of patients daily in hundreds of organizations. It is the most widely used population health risk measurement and population health analysis methodology in use globally.

The ACG System is used in 20+ nations and is applied to hundreds of millions of patients daily in hundreds of organizations.

The development process that created the ACG System's measures and categories made use of the most advanced multivariate statistical approaches available at the time, including techniques that are now considered to be part of the machine learning toolkit (such as recursive partitioning and penalized regression). Under the guidance of panels of Johns Hopkins practicing clinicians, data scientists used these computer-based analytic techniques to create and test rules-based algorithms and forecasting models. When users apply the ACG System's predictive models they either do so: 1) using reference population weights or coefficients derived from large national populations embedded in the ACG System software, or 2) making use of their own site-specific learning data to recalibrate model weights and predictive scores to better fit their local data.

While research suggests that AI or ML are not always beneficial in improving the accuracy of predictions, in the cases where they are, the comprehensive and innovative ACG System measurement framework serves as an ideal feature layer for machine learning models. Rather than relying on dynamic or black-box ML approaches that learn directly from raw, often noisy, data — which attempt to discover new patterns or hierarchies from scratch — the ACG System uses a rules-based methodology grounded in decades of clinical and scientific research. This curated, stable structure provides a clinically meaningful and transparent foundation for further model development.

By applying machine learning within the ACG System framework, organizations can accelerate development timelines, increase model portability across settings, and enhance interpretability within both clinical and administrative contexts. Especially in high-stakes health care applications where explainability, reproducibility and trust are paramount, leveraging a system with such clinical rigor is not only beneficial — it is essential.

Strategic and Practical Applications of AI and Data Science

We draw from extensive experience in health informatics, population health analytics and the data sciences to expertly support organizations considering AI adoption. Importantly, we also provide some suggestions for how the well-tested ACG System analytic framework on its own — and in conjunction with additional AI/ML techniques — can be applied to population health analytics and other related health care applications.

If applied in the right manner and in the right context, the convergence of AI technologies and population health analytics present unprecedented opportunities to improve patient outcomes, optimize resource allocation and enhance care delivery. By combining AI capabilities with the proven ACG System, health care organizations can achieve more precise risk stratification, personalized interventions and scalable population management.

Addressing Some Challenges and Limitations of AI Application to Population Health Analytics

While the potential of AI in population health analytics is significant, it is crucial to maintain realistic expectations about its current capabilities and limitations. The health care AI field has sometimes been characterized by inflated claims and unrealistic expectations, leading to disappointment when systems fail to deliver promised benefits, especially when the comparison starting point is a standard advanced analytic infrastructure, rather than no infrastructure at all.

A landmark systematic review by Christodoulou et al. (2019)[1] analyzed 282 studies comparing machine learning approaches to traditional advanced statistical approaches (such as logistic regression) for clinical prediction problems. Their findings showed that machine learning methods were not uniformly superior to traditional statistical approaches, particularly when study bias was taken into consideration. More recent studies (e.g., Jing et al., 2022) have demonstrated similar findings.[2]

This systematic review underscores the importance of:

- Carefully evaluating whether AI approaches provide meaningful improvements over existing methods, especially when biases and analytic balances are fully considered;
- Considering the additional complexity, cost and resource requirements of AI implementation;
- Ensuring that AI models are generalizable, scalable, parsimonious and easy to interpret;
- Maintaining expertise in traditional statistical, clinical and epidemiological approaches, even when moving towards AI tools; and
- Focusing on the problem being solved, rather than the technology being used.

[1] Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* 2019;110:12-22. doi:10.1016/j.jclinepi.2019.02.004 https://pubmed.ncbi.nlm.nih.gov/30763612/

[2] Jing B, Boscardin WJ, Deardorff WJ, et al. Comparing Machine Learning to Regression Methods for Mortality Prediction Using Veterans Affairs Electronic Health Record Clinical Data. *Med Care.* 2022;60(6):470-479. doi:10.1097/MLR.0000000000001720 https://pubmed.ncbi.nlm.nih.gov/35352701/

Navigating AI Pitfalls: How the ACG System Helps Overcome Them

Despite benefits when applied appropriately and in the right context, AI implementation in the field of population health analytics can falter when organizations underestimate four recurring challenges:

- Algorithmic bias;
- Data quality and fragmentation;
- Transparency or clinician skepticism;
- Overfitting/lack of portability outside of a single context.

Further detail is outlined below:

Challenge 1 – Algorithmic Bias

Under-representation of certain patient groups when creating the measures and models can lead to inequitable predictions when applying the model to patients from these marginalized populations. An ACG-centric "bias-audit" can help provide a structured evaluation method, where stratified calibration by age, sex, race, ethnicity and insurance coverage type helps ensure equitable performance across the full population.

Challenge 2 – Data Fragmentation

Missing or conflicting clinical codes reduce ML signal strength. The ACG System's clinically validated aggregation logic maximizes usable information from incomplete data across multiple sources and can help promote and monitor data-sharing agreements and other interventions that can improve data completeness and interoperability.

Challenge 3 – Black-box Opacity

Clinician distrust and regulatory pushback arise from opaque models. Combining the ACG System measures as explainable AI overlays (e.g., making use of explanatory SHAP/ICE plots as appropriate — which help visualize how individual features influence a model's prediction) provides transparency and builds trust. Transparency is not optional in health care. Each model moved into production must be accompanied by feature-importance summaries, case-based rationale and graphical dashboards that reveal how risk factors drive predictions. The ACG System, with 200+ clinically, administratively and real-world data tested risk markers; 15+ risk models; academic underpinnings; and comprehensive technical user documentation, is second to none in terms of transparency.

Challenge 4 – Over-fitting & Portability

AI models fine-tuned to a single site may not be readily applicable to other contexts. Beginning with transportable ACG System features, monitoring out-of-sample drift, and retraining on large sets of claims data from multiple settings, ensures broader applicability and will substantially increase the likelihood that locally developed advanced AI models will be generalizable and scalable. The ACG System out-of-the-box will provide both reference scores from large external reference populations, as well as advanced multivariate, locally calibrated advanced statistical models. Both of these reference points can serve as gold standard grounded comparators to locally created and weighted AI/ML models. This will allow users to assess the degree to which AI-innovations may (or may not) offer improved predictive abilities while maintaining adequate validity and transparency.

ACG Supported AI Applications in Population Health

These examples of actual applications from ACG System users are briefly outlined below:

Case Study 1: Targeted Care Management

Challenge: Pressures to deliver personalized, timely care with the right resources.

ACG System Supported Solution: The ACG System's population segmentation tool (Patient Need Groups) is used to create targeted care management lists based on clinical profiles that are overlayed with AI for personalized messaging and SMART alerts.

Result: SMART alerts connect the right patients to the right resources at the right time – the result is more efficient resource allocation and increased patient engagement.

Case Study 2: Intelligent Care Recommendations

Challenge: Care managers are overwhelmed by patient complexity when designing interventions that meet individual patient needs.

ACG System Supported Solution: AI generates tailored recommendations and well-being tips via the patient app, informed by ACG System risk markers, identification of comorbidities, predicted health costs and medication adherence, combined with patient personality type and defined health goals.

Result: Significant reduction in care planning time while improving the consistency and quality of interventions, based on clinically defined population health needs.

Case Study 3: Model Selection Improvement

Challenge: Identifying which ACG System risk markers and models are the best fit for analytic needs.

ACG System Supported Solution: ACG System outputs (e.g., risk model scores, PNG segmentation categories, Resource Utilization Bands, etc.) can be input to AI models, improving their predictive accuracy and clinical explainability.

Result: Augmented information and explainable ACG System outputs support the delivery of tailored patient care.

Why AI Works Better with the ACG System

The ACG System provides a robust foundation that helps address many inherent AI limitations. Its efficiency in processing digital data from large populations stems from decades of human and machine optimization, allowing organizations to risk stratify, categorize and predict future utilization for millions of patients more effectively and meaningfully than when using machine learning techniques alone. The ACG System's interpretability and its global gold standard status means providers can understand and trust the clinical rationale behind each model prediction or risk score, fostering adoption and appropriate use.

The ACG System provides a robust foundation that helps address many inherent AI limitations.

From an analytic perspective, a major advantage of incorporating the top-down ACG System feature layer for AI applications is that the final use case will not be dragged down by bottom-up spurious associations, systemic bias or missingness of the data. This tandem use of AI and the ACG System means that such concerns can be mitigated, and users will still benefit from the individualized pattern discovery and automated efficiency that many AI techniques can offer.

The Johns Hopkins Bloomberg School of Public Health faculty and the ACG System team at Johns Hopkins HealthCare Solutions are actively exploring the next generation of AI-enabled ACG tools, along with suggested frameworks for applying the current and expanding ACG System feature layer to a range of AI and ML use cases. As this work progresses, we plan to continue to engage users and gather input to help shape the direction and relevance of these developments to real-world applications. See the additional details at the end of this white paper for more information on some of these activities.

JHU faculty are hard at work developing new AI/ML/NLP oriented ACG System applications and tools

The JHU ACG research team CPHIT, made up of data and computer scientists, as well as health services and epidemiologic researchers and clinicians, have an ongoing robust R&D portfolio focusing on the adoption and evaluation of AI, ML, NLP and other techniques within the ACG System's context. The intent is to explore ways these approaches can be applied to current ACG System risk markers and models and also to improve ACG System analytic approaches using these tools in a number of ways.Working with JHHCS and other key stakeholders, we are also developing ongoing guidance for users on how best to make the ACG System the cornerstone of their current and future AI analytics in support of a wide range of clinical and administrative goals. In the box below are some recent publications — from both ACG System users and JHU researchers and developers — that have applied AI/ML techniques using the ACG System feature layer. Also see more AI/ML/NLP publications by JHU faculty at: https://publichealth.jhu.edu/center-for-population-health-information-technology/publications

To Learn More

To learn more about how the ACG System can support your organization, or to request the latest bias-assessment report, implementation playbook or collaborative pilot discussion, contact us at acginfo@jh.edu.

To stay informed from the population health analytics experts at Johns Hopkins, join our mailing list https://www.hopkinsacg.org/blog-signup/.

More information on ACGs at: www.hopkinsacg.org

Additional Resources

ACG System webinar on AI applied to ACG analytics in a South African Health Plan/Scheme - https://www.hopkinsacg.org/ukwebinar16/

YouTube Grand Rounds Lecture by JHU Professor Jonathan Weiner (ACG System Co-Developer and Scientific Director) Lecture on Pragmatic Application of AI to Population Health Analytics - https://www.youtube.com/watch?v=rDij3K3N1Hk

Tan JK, Zhang X, Cheng D, Leong IYO, Wong CS, Tey J, Loh SC, Soh EF, Lim WY. Using the Johns Hopkins ACG Case-Mix System for population segmentation in a hospital-based adult patient population in Singapore. *BMJ Open.* 2023;13(3):e062786. https://bmjopen.bmj.com/content/13/3/e062786

Jiang Y, Huang YL, Watral A, Blocker RC, Rushlow DR. Predicting Provider Workload Using Predicted Patient Risk Score and Social Determinants of Health in Primary Care Setting. *Appl Clin Inform.* 2024;15(3):511-527. https://pubmed.ncbi.nlm.nih.gov/38960376/

Kan HJ, Kharrazi H, Chang HY, Bodycombe D, Lemke K, Weiner JP. Exploring the use of machine learning fo r risk adjustment: A comparison of standard and penalized linear regression models in predicting health care costs in older adults. *PLoS One.* 2019;14(3):e0213258. https://pubmed.ncbi.nlm.nih.gov/30840682/

Hatef E, Rouhizadeh M, Nau C, Xie F, Rouillard C, Abu-Nasser M, Padilla A, Lyons LJ, Kharrazi H, Weiner JP, Roblin D. Development and assessment of a natural language processing model to identify residential instability in electronic health records' unstructured data: a comparison of 3 integrated healthcare delivery systems. *JAMIA Open.* 2022;5(1): https://academic.oup.com/jamiaopen/article/5/1/ooac006/6529555

Ramachandran R, McShea MJ, Howson SN, Burkom HS, Chang HY, Weiner JP, Kharrazi H. Assessing the Value of Unsupervised Clustering in Predicting Persistent High Health Care Utilizers: Retrospective Analysis of Insurance Claims Data. *JMIR Med Inform.* 2021;9(11): https://medinform.jmir.org/2021/11/e31442

Pandya CJ, Wu J, Hatef E, Kharrazi H. Latent Class Analysis of Social Needs in Medicaid Population and Its Impact on Risk Adjustment Models. *Med Care.* 2024;62(11):724-731. https://journals.lww.com/lww-medicalcare/abstract/2024/11000/latent_class_analysis_of_social_needs_in_medicaid.6.aspx