



Closing Risk Gaps with AI

The art of the possible today and its impact on care and economics for government-sponsored programs

CONTENTS



EXECUTIVE SUMMARY



RISK CODING

The Stakes and Market Dynamics



**CURRENT APPROACHES AND UNMET
OPPORTUNITIES**



**AI FOR CREATING DIAGNOSTIC
RECOMMENDATIONS**



**AI FOR OPTIMIZING CLINICIAN
ENGAGEMENT**



CONCLUSION

ABOUT THE AUTHOR

Jim Manzi

*Partner,
Foundry.ai*

Jim is co-founder of Curia.ai. He was co-founder, CEO and Chairman of Applied Predictive Technologies, which became the world's largest cloud-based AI software company and the dominant platform for rapid, iterative randomized trials of business programs. Previously, Jim developed pattern recognition software at AT&T Laboratories, and worked as a corporate strategy consultant.

Jim is the author of several software patents for the automation of randomized experiments, as well as the 2014 Harvard Business Review article "The Discipline of Business Experimentation." His 2012 book Uncontrolled on the development and application of randomized trials was widely reviewed in the New York Times, Wall Street Journal and other national publications.

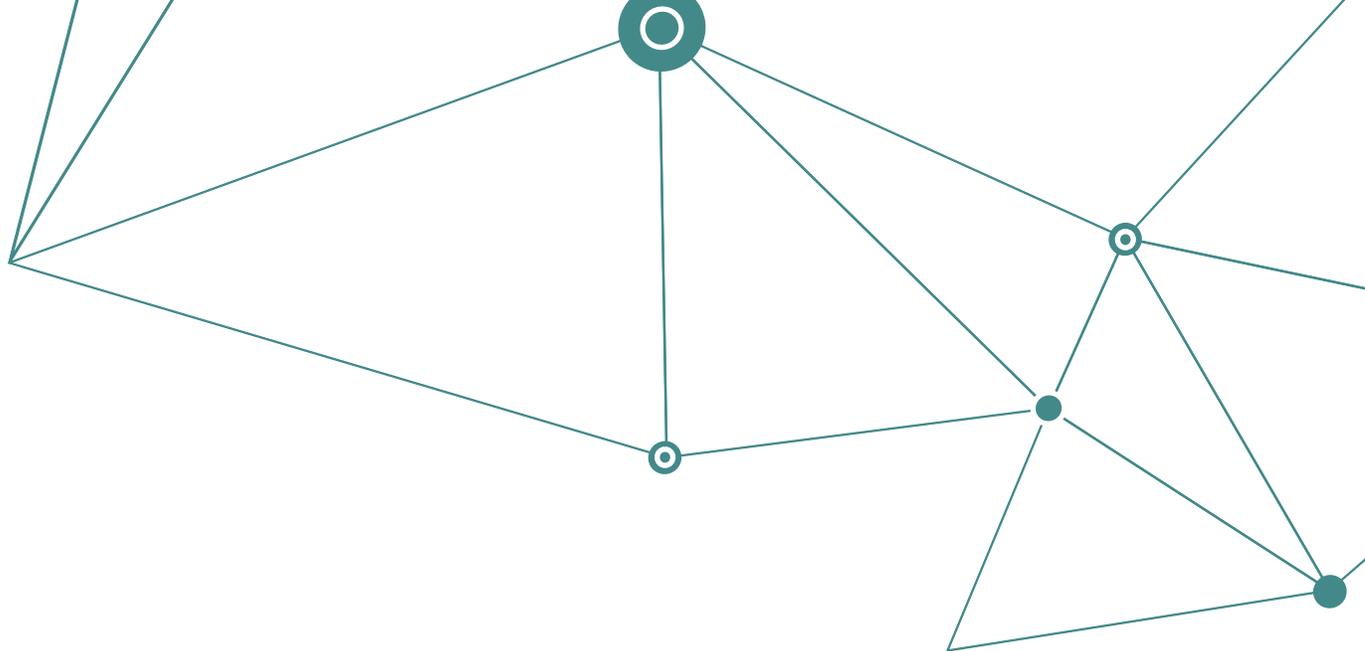
Jim received an SB in mathematics from MIT, and was subsequently awarded a Dean's Fellowship in statistics to the doctoral program at the Wharton School of the University of Pennsylvania. He serves on the Board of Directors of Aledade, an innovative ACO network backed by Venrock and Google Ventures.

ABOUT CURIA.AI

Curia.ai

*A Foundry.ai
Subsidiary*

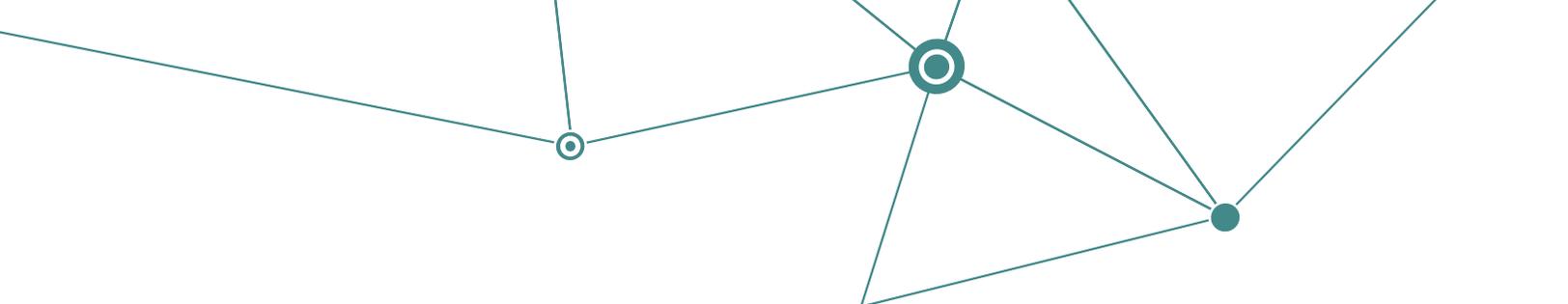
Curia.ai works with leading providers, health plans and integrated care networks to implement leading-edge AI capabilities to close risk gaps through a combination of improved diagnostic coding prediction and optimization of clinician engagement.



EXECUTIVE SUMMARY

Closing risk gaps through correct risk coding has enormous quality of care and economic implications in government-sponsored health plans, and current market dynamics are heightening this issue as they create downward pressure on risk scores and upward pressure on actual risk. Recognizing the critical importance of this issue, most relevant entities have put in place numerous operational and analytical systems to identify and close risk gaps, but the current state of the art still leaves very large risk gaps unaddressed.

Recent technology advances now make it practical to apply AI-driven analytical systems to further close risk gaps by: (1) increasing the relevance and score weight of coding recommendations; and (2), empirically optimizing the communications of these recommendations to clinicians to maximize appropriate conversion of recommendations to coded diagnoses. Accomplishing this requires overcoming some substantial data and analytics challenges. The benefits are significant: better care management through more accurate diagnosing, and millions of dollars of incremental annual margin per fifty thousand members via increased reimbursements.



RISK CODING

The Stakes and Market Dynamics

Establishing correct diagnostic coding improves health outcomes by placing members on the correct care pathway. But for many members served by government-sponsored programs — including Medicare Advantage (MA), Medicare Shared Savings Program (MSSP), capitated Medicaid plans and Commercial ACA plans — these diagnostic coding decisions also have enormous economic consequences.

For Medicare members, the Centers for Medicare & Medicaid Service (CMS) applies a relatively complex approach, termed the Hierarchical Condition Category (HCC) method, to convert the list of chronic conditions reported at the ICD 10 code level in a given year for a given member to that member's Risk Adjustment Factor ("RAF score"). The RAF score in turn determines the reimbursement that will be provided for that member's care. Risk adjustment for Medicaid members is similar and, though details vary by state, typically employs the hierarchical Chronic Illness and Disability Payment System (CDPS) to generate a RAF score and resulting reimbursement level for each capitated Medicaid member.

Capturing the correct diagnostic codes for each member served under MA, MSSP, or capitated Medicaid plans dramatically affects the overall revenue captured by the plan or provider. In 2019, the research firm IDC estimated that companies lose 20 to 30 percent of reimbursements every year due to gaps in the process of capturing all relevant diagnostic codes for all members.

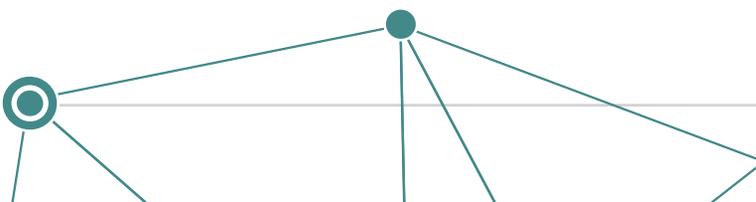
Market dynamics in 2020 have made this problem much more severe, due to changes brought about by the COVID pandemic.

First, claims volume has decreased dramatically since March. This creates some short-term benefits for risk-bearing entities but stores up trouble that will begin to manifest in 2021. A reduction in encounters puts downward pressure on risk scores (because of fewer opportunities to observe and code conditions) and upward pressure on actual risk (because of reductions in diagnosis and preventive care). RAF scores have likely been suppressed at least 5% already, and further reduction is expected. Plus, the increase in actual risk coupled with pent-up demand for unmet medical needs will surely put pressure on utilization once the situation begins to return to normal, presumably sometime in 2021.

Second, the economic disruption caused by the pandemic's lockdowns has changed the healthcare landscape. The unprecedented explosion in unemployment in 2020 has driven large increases in Medicaid and ACA enrollment. The needs of MA and MSSP members are likely to increase as well because of rising unemployment and related economic stress.

Third, the telehealth channel has expanded considerably during the pandemic. While some percentage of visits is likely to return to in-office during 2021, it is probable that certain categories of primary care, mental health and other appointment types will see a permanent partial shift to telehealth. To date, accurate and complete coding has been more difficult for most physicians in telehealth settings, and this will therefore create further incremental risk gaps.

Taken together, the fundamental importance of risk coding, heightened by current market dynamics, has moved this issue to the top of the agenda for healthcare executives. In a Change Healthcare 2020 survey of senior executives at 175 health plans, by far the top two concerns were "Getting members to the doctor" (39%) and "Ensuring accurate capture of all diagnostic codes" (34%).



CURRENT APPROACHES AND UNMET OPPORTUNITIES



Most MA plans, Accountable Care Organizations (ACOs) and Managed Care Organizations (MCOs) have put in place numerous procedures to improve capture of accurate diagnostic codes, often under the heading of “risk gap management.” These procedures include operational methods, such as hiring coding specialists to perform detailed chart reviews, conducting coding training sessions with clinicians and support staff, and member engagement programs to increase the incidence of annual wellness visits with associated coding opportunities. Many organizations also deploy related technology, including simplified clinician user interfaces linked to the underlying medical record, as well as automated identification of obvious missed coding opportunities such as a chronic condition that has been diagnosed in previous years, but which has not yet been diagnosed in the current year.

Cooperatively, these approaches play an important role in ensuring that members are on the correct care pathway and that the organization receives correct compensation. However, many successful MA plans, ACOs and MCOs recognize that there are still substantial opportunities to improve coding accuracy that can provide enormous benefits to the organization and its member group. They are working with sophisticated technology leaders to advance this analytical capability much further. They are constructing systems that employ artificial intelligence and machine learning (AI/ML) techniques to substantially improve the accuracy and completeness of coding recommendations by member. These systems are creating significant and measurable gains to both care delivery and reimbursements due to properly adjusted RAF scores. Beyond improvements in patient care, the economic payoff is typically millions of dollars of identifiable incremental annual margin per fifty thousand members.

Systematically achieving best-in-class risk coding requires integrating two mutually reinforcing capabilities:

- 1 | Automated identification of currently uncoded conditions with maximum impact, and
- 2 | Analytical identification of the specific messaging methods, modes and content that will maximize appropriate conversion of recommendations to coded conditions for each clinician.

The balance of this document will be an explanation of these capabilities.

AI FOR CREATING DIAGNOSTIC RECOMMENDATIONS

The fundamental role of AI in the identification of missed coding opportunities is to increase the precision and scope of the recommendations to clinicians. Accomplishing this requires access to a significant data and analytics infrastructure.

The starting point is a large pool of data on millions of relevant Medicare / Medicaid members, which should include encounters, diagnoses, procedures, prescriptions and lab values. In addition, data on social determinants of health (SDOH) at the member level, plus detailed data by clinician linked to each member encounter, must be incorporated to achieve useful results.

Next, a network of several thousand predictive models tied to individual diagnostic codes must be built and maintained. These models are then integrated with an economic engine that calculates the marginal RAF score impact of each potential incremental diagnosis for each member based on their conditions, attributes, and other factors.



Building from this base network, analogous models should be built for a specific organization's member / patient base, and hybridized with the national models based on the trade-off between the greater specificity at the organization level versus the larger sample size at the national level. In practice, the organization-specific model will tend to dominate for very common conditions and the national model for rarer conditions. This combined network must be updated and maintained as data and practices change.

Operationally, the output from this analytical process is a weekly targeted list of diagnostic recommendations by member that have high probability of validity and significant impact on RAF scores, if accepted. In general, recommendations should include quantitative measures of confidence as well as specific data on the exact diagnoses, procedures and labs that drive the recommendations, along with data provenance for each of these drivers. These recommendations are made available to clinicians via existing clinician engagement portals and pipelines.

AI FOR OPTIMIZING CLINICIAN ENGAGEMENT

Automated identification of granular risk gaps is critical, but it leaves open a key dimension of the problem: the propensity of each clinician to be positively impacted by the coding recommendation (i.e., 'impactability'). For example, some clinicians are much more likely to change coding decisions within targeted clinical areas. Certain physicians are more likely to respond to digital engagement, others to direct contact. Some will respond positively to multiple coding recommendations for each member while others will respond better to at most a single recommendation. The most effective organizations don't make *a priori* assertions about this complex response profile — they learn empirically what recommendations work with what clinicians, delivered how and through what channels.

Typically, the net RAF score impact of a specific coding recommendation if accepted and the probability that it will be accepted by a specific clinician are very weakly correlated and are in practice independent effects, which can be depicted in the following 2x2 matrix. For clarity, the horizontal axis represents a standard *risk* question, which reduces to ‘will the clinician make the optimized coding change without a recommendation,’ whereas the vertical axis represents an *impactability* question, which is ‘will the clinician accept the optimized coding change with a recommendation.’

Clinician Targeting Matrix for Coding Recommendations

With Recommendation	Yes	Persuadables 	Sure Things 
	No	Lost Causes 	Sleeping Dogs 
		No	Yes

Without Recommendation



Ideally, the analytical software utilized will be able to draw this picture accurately and reliably across the member and clinician base on a regular cadence, and all coding recommendations will be provided solely to the clinicians who are 'Persuadables' with respect to that recommendation (prioritizing within this group using the incremental projected RAF score impact). Making recommendations to any other group of clinicians will be counterproductive -- e.g., 'Sure Things' were already going to make the optimal coding change without any recommendation, and 'Lost Causes' were not but will not be influenced by the recommendation.

CONCLUSION

Closing risk gaps through correct diagnostic coding continues to have enormous quality of care and economic implications for government-sponsored health plans, with current market conditions only amplifying the importance of this capability.

While nearly all relevant organizations have put in place some operational processes to identify and close risk gaps, most have not taken advantage of new AI-driven analytical systems that are creating significant additional lift by: (1) increasing the relevance and score weight of coding recommendations; and (2), empirically optimizing the communications of these recommendations to clinicians to maximize appropriate conversion of recommendations to coded diagnoses.

Curia.ai is pleased to be leading the development of new technology in this area and we would welcome the opportunity to share more about our work, its empirical impact on both care and economics, and possibilities for teaming.



Curia.ai

1920 L St NW, Suite 800
Washington, DC 20036

<https://curia.ai>