



**SOCIETY OF
ACTUARIES®**

Article from

Risk Management Newsletter

August 2018

Issue 41

Validating a Novel Health Insurance Fraud Detection Method

By Rebecca Elias and Robert Lieberthal

Editor's Note: The 52nd Actuarial Research Conference (ARC) was held in Atlanta, GA in July 2017, with the theme "Actuarial Research at the Crossroads: Transcending Disciplines." Actuarial educators, practitioners and researchers gathered together to discuss the latest developments and to exchange ideas. In the March issue of Risk Management, Dr. Shaun Wang shared his research on "Modeling of Optimal Spending and Allocation on Cybersecurity." Continued in this issue, we have Dr. Lieberthal and his team to discussing their research on a health insurance fraud detection method.

In a country where health insurance is the main way to pay for health totaling to \$3.3 trillion dollars, it is important to know how the money is appropriated. It is estimated that over 50 percent of health insurer spending on health care comes from social insurance programs such as Medicaid and Medicare.¹ Coupled with the increasing costs of health care in the United States, lowering health care expenditures and assuring the integrity of these payments is vital. This is what led Dr. Robert Lieberthal, assistant professor in Public Health at The University of Tennessee in Knoxville, along with Dr. Jing Ai of The University of Hawaii at Manoa and Dr. Patrick Brockett of The University of Texas at Austin to their most recent collaboration. The aims of their ongoing study are to devise a theoretical basis for predicting health care fraud and to apply PRIDIT, a method validated in other lines of insurance, to health insurance claims.

The benefits of a predictive modeling approach with health care claims data include cost savings, promptness of claim analysis which could lead to faster payment, and potential to enhance fraud detection techniques. Understanding the components of health care fraud informs public policymaking. Detecting health care fraud has the potential to improve value in health care, especially in social insurance programs such as Medicaid and Medicare. The savings from health care fraud detection can be redirected for expansion of other types of care. Furthermore, detection of fraud can reduce pressure and the cost of compliance for law abiding providers, facilities, and their patients with potential to spill over into private insurance.

This health care fraud study aims to develop a predictive modeling approach to fraud management in Medicare claims. The methodology for this study is to apply PRIDIT to determine suspicious scores for each claim and to determine the most important red flags for fraud. PRIDIT is a fraud detection technique producing a rank-ordered score for the intensity of a latent variable by identifying a relationship between this variable and a set of ranked predictive variables. Redit scoring is a method to relate the value of proxy variables such as claim size, patient characteristics, and provider characteristics to the suspiciousness of a given claim.² Then, principal components analysis (PCA) is applied to the Redit score to determine an overall score of claim suspicion based on the underlying variables.³ This PRIDIT approach has been previously validated in multiple lines of insurance including consumer fraud in automobile claims.⁴

The savings from health care fraud detection can be redirected for expansion of other types of care. Furthermore, detection of fraud can reduce pressure and the cost of compliance for law abiding providers, facilities, and their patients with potential to spill over into private insurance.

Through applying and validating PRIDIT as a predictive method for detecting fraud in insured health care claims, the research team used 2009 calendar year claims from a 5% sample file of Medicare beneficiaries to test this approach. The 2009 calendar year claims data comes from the Medicare 5% sample file, a de-identified data file of a random 5% sample of traditional Medicare beneficiaries and their complete Medicare claims history (Parts A, B, and D). The identified fraud predictors included patient age, diagnosis count, claim payment amount, and total charges. The 5% sample is then split between Medicare-only and dual eligible individuals. The study sample included 453,941 Medicare-only beneficiaries with an average age of 76.8 years and 184,178 dual eligible beneficiaries with an average age of 78.4 years.

The identified fraud predictors included patient age, diagnosis count, claim payment amount, and total charges. The 5% sample



allowed for the research team to study specific populations that may be at risk for fraud. For example, they were able to split the study between Medicare-only and dual eligible individuals who are receiving both Medicare and Medicaid benefits. Potentially important predictor variables included the claim payment amount, total PPS capital amount, claim PPS capital DRG weight number, National Claims History (NCH) claim provider payment amount, and NCH carrier claim allowed charge amount. It is then up to analysts to decide the threshold for identifying suspicious claims for further investigation by human experts such as special investigative units (SIU) or law enforcement personnel. This was a research project, so no claims were forwarded to outside parties for further investigation.

In analyzing the data, a small number of highly suspicious inpatient claims and provider claims were identified, and validation variables and certain predictor variables were found to be more

useful for some diagnoses. While the researchers had access to two validation variables—claim processing time and distance from provider to patient—we are pursuing an independent expert evaluation of claims, comparing a validated sample to the results of any analytic method. This is as close as we can get to the gold standard of comparing adjudicated claims to court decisions on whether claims are truly fraudulent. Currently, this study is in the validation phase, where a medical doctor with clinical experience in the inpatient and outpatient settings with expertise in family medicine is analyzing the claims and determining fraud independently of the PRIDIT method. A forthcoming comparison of the results of the data mining method and the human expert approach is likely to inform both the validity of the PRIDIT methods and future approaches to examining health insurance fraud. ■



Rebecca Elias is a graduate research assistant at The University of Tennessee. She can be reached at relias@vols.utk.edu.



Robert Lieberthal, Ph.D., is a health economist and an assistant professor at the University of Tennessee, Knoxville. He can be reached at rob@lieberthal.us.

ENDNOTES

- 1 National Health Expenditures 2016 Highlights. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/downloads/highlights.pdf>
- 2 Bross, I. D. (1958). How to Use Redit Analysis. *Biometrics*, 18–38.
- 3 Brockett, P. L., Derrig, L. L. Golden, A. Levine, and M. Alpert (2002). Fraud Classification Using Principal Component Analysis of RIDITs. *Journal of Risk and Insurance*, 69(3), 341–371.
- 4 Ai, J., P. L. Brockett, and L. L. Golden (2009). Assessing Consumer Fraud Risk in Insurance Claims: An Unsupervised Learning Technique Using Discrete and Continuous Predictor Variables. *North American Actuarial Journal*, 13(4), 438–458.