Predicting Your Next Big Workers' Compensation Claim—

WHILE YOU CAN STILL DO SOMETHING ABOUT IT

Integrating AI and machine learning into workers' compensation is a paradigm shift.

BY KENNY GRIFNO PH.D., ELIZABETH KEIFFER PH.D., SCOTT ROLOFF, BILL MCCALLUM, AND MARK BARTA

TOM'S KNEE BOUNCED UP AND DOWN as he sat waiting for the Director of Human Resources. He turned toward the door as Sally walked in, followed by Bob, the CFO. Tom swallowed—hard.

Sally half-smiled, half-grimaced. Bob didn't even try. "Tom, we need to make cuts in your workers' comp team," he said. "You have three out of control claims that busted through our stop-loss threshold. Even though the insurer will pay them from here on out, what we had to pay before reaching the stop-loss busted the budget. And that's before counting the overtime and temp expenses to cover for those injured employees."

Bob shook his head before continuing. "Plus, the carrier said that next year we're getting a big premium increase because of those three claims."

Sally shifted in her chair. "What kind of cuts are we talking about?"

Bob looked down at a piece of paper, and then up. "Two FTEs."

Tom wrapped his fists around the arms of his chair. "If I have to let two of my folks go, things will get even worse."

Sally took a deep breath. "What's done is done. As long as we don't have any more big claims, can't we make do?"

Bob shook his head. "Not unless there is a way to spot these claims early and head them off before they get out of control. Otherwise, I need to start baking a couple of big claims into each year's budget, and Tom, that money needs to come from somewhere else in your budget."

Sally turned to Tom. "Is there a way to identify these claims when the injury occurs so that we can intervene right away?"

Tom shook his head. "These three claims came out of the blue. When I got the first medical expense reports on them, they were already out of control. There was no way to spot them before that."

Bob stood. "Then you know what you need to do."

Tom was right. There was no way to spot those claims sooner—until now. First, let's lay the groundwork.

COSTS THAT MATTER

In workers' compensation, you track a lot of metrics.[1] One of the most important is the number of days to return an injured employee to work. Return to work and other metrics, however, are merely proxies for what matters most—the cost of a good outcome, and getting the injured employee back to work and keeping them there.

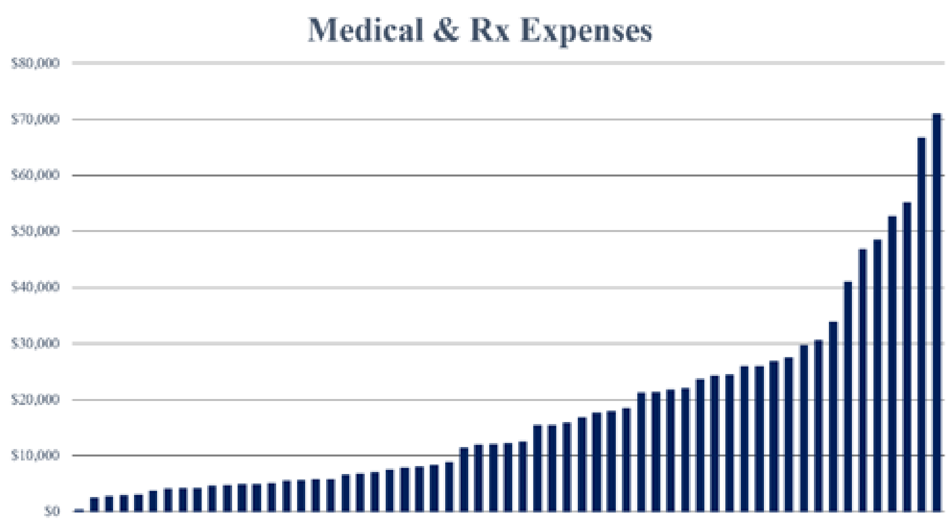
This "outcome" cost has two controllable components:

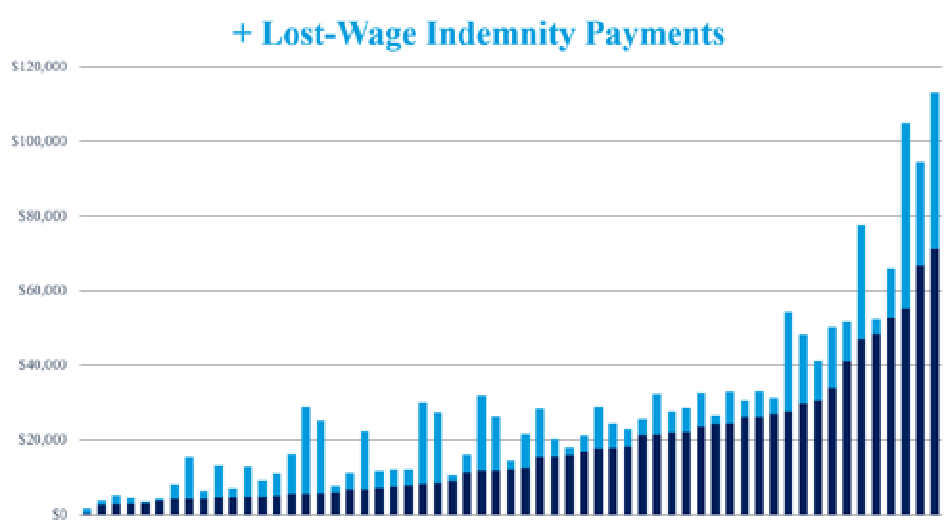
- Medical and pharmacy expenses; and
- Lost-wage indemnity payments to the injured employee.

You may have other indemnity payments too, such as fixed payments when employees lose limbs. You cannot do anything about these, however, because they accrue at the moment of the injury. You need to focus on the costs that you can control.

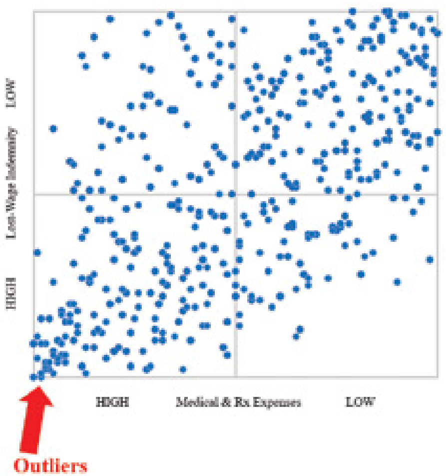
These two components do not move in tandem. Some claims may have high medical and pharmacy expenses, but low lost-wage indemnity payments—and others the reverse. Minor injuries may have only medical and pharmacy expenses and no lost-wage indemnity. The first chart below shows the medical and pharmacy expenses on shoulder injuries that had both medical and pharmacy expenses and lost-wage indemnity payments.

The second chart adds the lost-wage indemnity payments on top of the medical and pharmacy expenses–giving you the complete picture.[2]





Here is another way of looking at it, this time with knee injuries. In the quadrant graph below, each claim is a bubble, as opposed to a bar in the graphs on the previous page. The claims have been placed along the horizontal axis according to their medical and pharmacy expenses—high on the left and low on the right. Likewise, the claims are located along the vertical axis according to their lost-wage indemnity payments—high on the bottom and low on the top. The out-of-control claims that you have to worry about—the outliers—are in the lower-left corner.

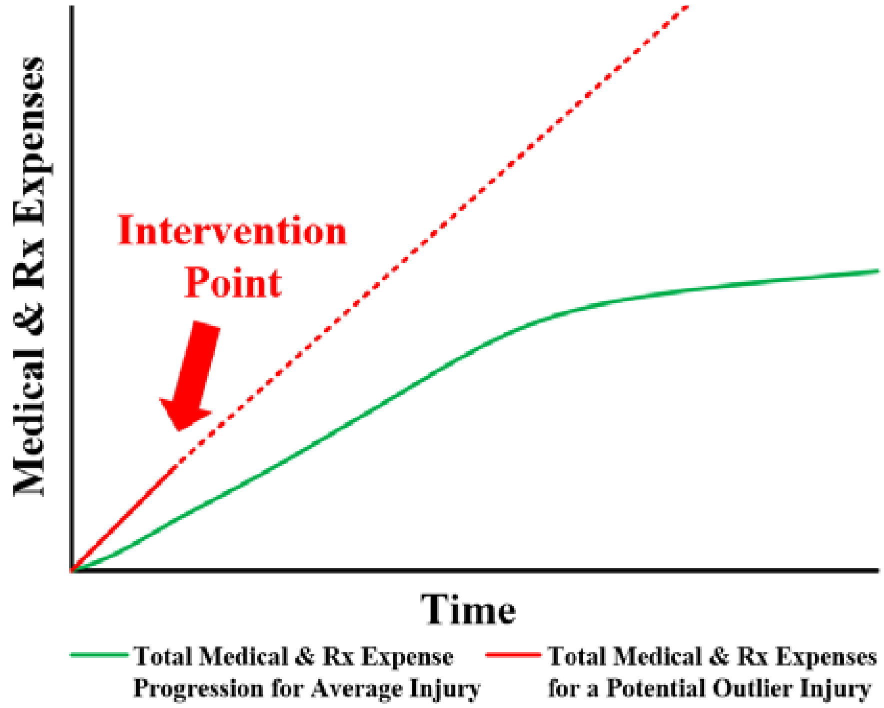


What are outliers? They are claims far outside the averages and norms, often defined as three or more standard deviations above the mean. Perhaps the patient has comorbidities that complicate their recovery? An obese employee suffering from depression with a broken leg will cost more, and take longer to get back to work, than an employee without those comorbidities. Maybe the leg was not merely broken, but shattered? Or maybe the doctor did a bad job?

HOW COSTS MOVE

The lost-wage indemnity payments move in a steady proportional fashion. Two months of indemnity payments are twice as much as one; four months are twice as much as two. That's not where you need to look.

The medical and pharmacy expenses follow a different pattern. They are front loaded–and those out-of-control outliers even more so. You need to intervene quickly to have any chance of controlling them.



But Tom was right. There could be a two-month lag between when the procedure or examination generating a medical expense occurs and when the provider sends it to the third-party administrator (TPA) and it is processed. So you learn about those out-of-control claims from Months #1 & #2 in Months #3 & #4. You need to intervene in the first week or two to make a difference.

You need a real-time data source, and in your workers' compensation data you have one—the adjuster's notes.

ADJUSTER NOTES

The workers' compensation adjuster handles the claim. In many states, the adjuster can even direct the injured employee's care—tell the employee which doctor they have to see.

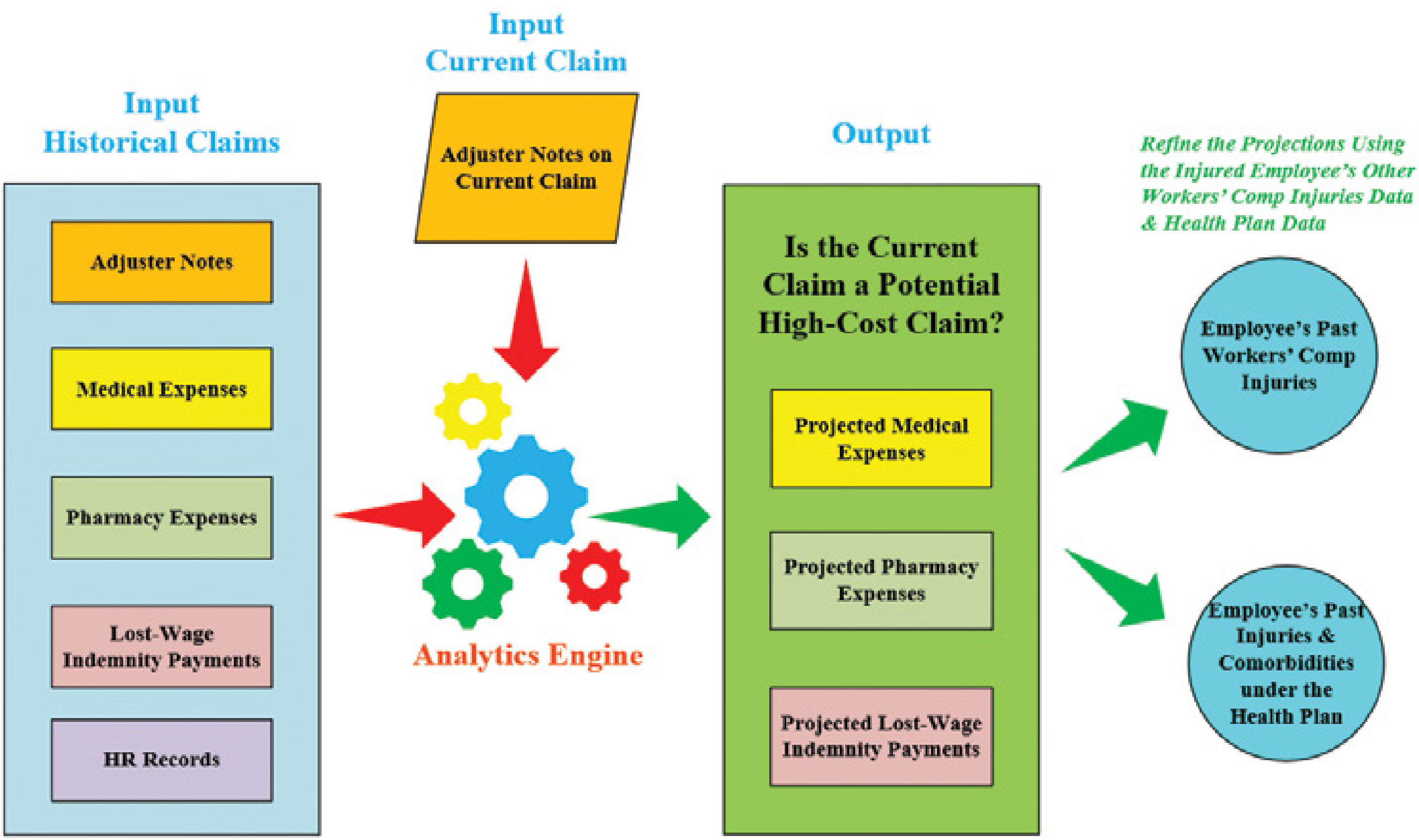
As soon as an injury occurs, the adjuster opens a file and documents what is going on with the case from that moment forward. Using your trove of past adjuster notes and their associated medical and pharmacy expenses, you can text mine the notes and use AI (artificial intelligence) and machine learning to correlate keywords and phrases in the early entries with surging medical and pharmacy costs a few days later. You can also do the same thing with return-to-work periods and lost-wage indemnity payments, although as discussed above you have more time to act here because they progress proportionately, not drastically. You can then program the system into which the adjusters input their notes to flag any claim on which an adjuster uses the trigger words as soon as the adjuster types them.

What you do once you identify a potential outlier will vary with each claim. You may want to direct the employee to the best possible surgeon for that type of injury, even if that surgeon is out of state. You may want to enroll the employee in a one-on-one support program with a counselor specializing in helping employees recover from that type of injury.[3] Whatever you do, you will want to monitor the claim closely, adjusting your intervention strategy if the claim begins to wobble.

PREDICTING OUTCOMES

Integrating AI and machine learning into workers' compensation is a paradigm shift that will enable you to identify, predict, and mitigate costly injuries while you can still do something about them.[4] With this approach, your mindset shifts from reactive to proactive.

When predicting these high-cost claims, time is crucial. You must therefore prioritize the adjuster notes on a current injury over the ensuing



medical and pharmacy expenses because of the 30-to-60-day lag in reporting them.

Your analytics engine will input the past workers' compensation data to learn how to identify potentially high-cost claims based on the adjuster notes, and then as the adjuster types the notes on a current claim, the engine will flag that claim if it could become a problem. Using the past costs on similar claims, the engine will then project the costs on the current one if you do not intervene. Supplementing this data will be your Human Resources (HR) data on the injured employees (e.g., position, age, sex, hours of overtime worked prior to injury, etc.).

REFINING THE PREDICTION

The initial cost projection will be for a generic injury—for example, a lower back injury that the adjuster describes using the keywords "severe," "extreme," and "permanent." You can refine this projection with the information on the employee's past workers' compensation injuries. If this is the employee's second or third back injury, it will likely cost more than their first.

If you use an outside analytics shop to do the analysis, you can also enrich the projection with the injured employee's health plan data. This data source will contain information on prior injuries not in the workers' compensation data set, as well as comorbidities such as diabetes, obesity, and depression that could complicate the employee's recovery. An employee with a back injury who is also battling a comorbidity will cost more and take longer to get back to work than an employee without that complication. If you are doing the analysis yourself, however, the Health Insurance Portability and Accountability Act (HIPAA) will preclude you from using the health plan data. As a general proposition, HIPAA prohibits an employer from accessing its employees' health plan information, but does not prohibit a third party that the employer engages from doing so.[5]

DETAILS YOU CAN LIVE WITH

From this point forward, we will discuss specifics on how to identify a potential high-cost claim based on the initial entries in the adjuster notes. We will do so on a high level—enough so that you can discuss what needs to be done with your data analytics experts, but not nearly enough to tell you how to do it.

DATA

Your data—the adjuster notes, medical and pharmacy expenses and lost-wage indemnity payments under the workers' compensation program, HR records, and maybe even health plan data—will probably not be "ready for prime time." Data acquisition, cleaning, and preprocessing are the foundation of any reliable analytics engine. Only after this stage can data exploration, text analysis, and predictive modeling complete the analytic lifecycle. And understand that this is not a rigid or linear process. New insights may lead to refinements along the way.

Your data will consist of both structured data—think rows and columns on a spreadsheet—and unstructured data like text. You can use the unstructured data from the adjuster notes in two ways. First, you can mine this data to gain insights into the words and phrases linked to past outliers. Second, you can codify those insights as structured data in terms of those outliers' costs, both the medical and pharmacy expenses and the lost-wage indemnity payments, to use when predicting the current claim's costs.

NATURAL LANGUAGE PROCESSING

With the rise of AI, natural language processing (NLP) has transformed the way that we extract, interpret, and comprehend complex unstructured data from text documents. These advances have been particularly insightful in the field of healthcare analytics, where the extraction of nuanced information is critical for making informed decisions. Natural language processing uses sophisticated algorithms to understand context, extract entities, and discern sentiment embedded in a document. This goes far beyond mere keyword spotting to include syntactic and semantic[6] analysis, permitting an in-depth understanding of the text.

Machine learning models detect patterns and correlations in data that humans miss. NLP techniques, such as named entity recognition (NER), can identify and categorize specific body parts or medical conditions, while sentiment algorithms assess the overall tone and emotion of the communication, in our case providing insights into the adjuster's evaluation. Feature extraction wrings out the frequency of high-risk keywords (e.g., fracture, severe, permanent), sentiment analysis (positive, neutral, or negative), length (word or character count), presence of specific body parts or affected areas (e.g., knee, back, shoulder), and severity indicators (e.g., serious injury, minor incident).

BAG OF WORDS

To structure and decrease the size of large data sets, you can treat the adjuster notes and other text data like a "bag of words,"[7] focusing on the presence or frequency of individual words without considering their context. This approach is surprisingly effective when deciding which category or cluster a document falls into for later use. When using this technique, the first step is tokenization, which is breaking the text down into smaller pieces (called tokens) and identifying delimiters, which are the special characters or symbols in a data set that separate different fields (e.g., commas, semicolons, etc.). The next step is stemming, which reduces multiple variants of a word to its root. An example would be converting "claims," "claimed" and "claiming" to "claim." You can then use frequency filters to ignore common words that appear in nearly all documents.

Once you have a reduced the bag of words, you can parse, filter, and transform the text. Using latent semantic indexing (LSI), you can analyze the relationships of the words in the bag and map those terms to a small set of concepts, creating a clean and structured data set. This ultimately sets up the data to be clustered or categorized into meaningful groups that you can use to predict future outcomes.

BUILDING THE ANALYTICS ENGINE

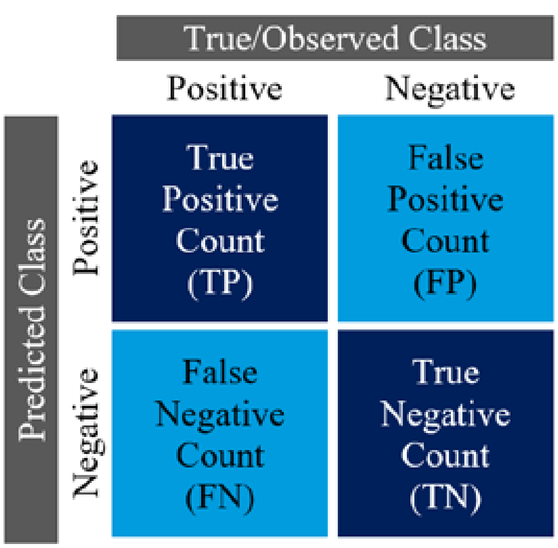
You can build your analytics engine using a variety of tools, including Python, KNIME (Konstanz Information Miner), Alteryx, and SAS EM (SAS Institute's Enterprise Miner). Once designed, you train the engine on prior claims data to distinguish between high-cost and normal-cost claims; and then apply the engine to your current claims as they occur to predict which are likely to be high cost.

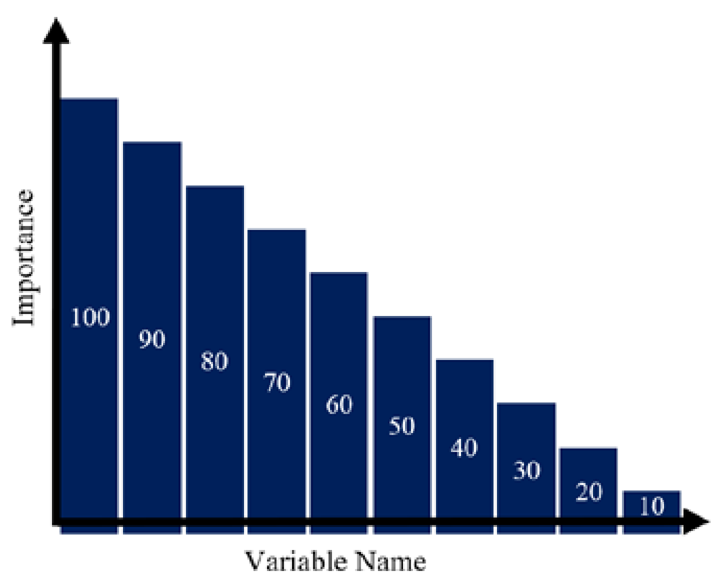
There are several techniques that you can employ to build the engine. Two particularly effective ensemble models for imbalanced data sets, where there are fewer high-cost claims compared to normal-cost ones, are:

- Random Forest—This model builds a series of decision trees from the data; and
- Gradient Boosting—This model constructs an ensemble of sequential models, with each successive model correcting the errors of the previous one.

After you have built the model, you will want to use hyperparameter tuning to tweak the engine's settings to improve its performance. If the data set on which you trained the model only has a few examples of certain groups of claims, you can explore resampling techniques such as oversampling or undersampling to address this class imbalance.

Finally, you test the results using various performance metrics derived from the "confusion matrix."[8]





Machine learning models detect patterns and correlations in data that humans miss. NLP techniques, such as named entity recognition (NER), can identify and categorize specific body parts or medical conditions, while sentiment algorithms assess the overall tone and emotion of the communication, in our case providing insights into the adjuster's evaluation.

These performance metrics include:

- Precision—Measures the accuracy of the positive predictions (i.e., of all the claims that the model identified as high-cost how many really would have been);
- Recall—Measures the completeness of the positive predictions (i.e., of all the potentially high-cost claims, how many did the model identify, and how many did it miss);
- F1-Score—Combines the above precision and recall measures into a single score; and
- AUC-ROC (Area Under the Receiver Operating Characteristic Curve)—Measures the tradeoff between the false positive rate (i.e., identifying claims as high-cost that aren't) versus the false negative rate (i.e., not identifying claims as high-cost that are).

CONCLUSION

Using AI and machine learning to mine your adjuster notes to identify workers' compensation claims likely to spiral out of control—in time for you to do something about them—is a game-changer. Sifting through the adjuster notes for past claims, AI spots words and combinations of words for claims that spun out of control not found in the notes for normal claims.

When an adjuster types those words and phrases into the notes for a current case, the program capturing the notes raises a red flag as soon as the adjuster types them. Algorithms then project the medical and pharmacy expenses and lost-wage indemnity payments on the current case based on the costs from those past cases—if you do not intervene.

Now you can allocate your resources more efficiently and effectively, concentrating on the employee injuries that could bust your budget, and making sure that they don't.

Kenny Grifno, Ph.D., is the chief of analytics of Integer Health Technologies, as well as a teaching assistant professor of information systems at the Sam M. Walton College of Business at the University of Arkansas, where he teaches healthcare business analytics courses.

Elizabeth Keiffer, Ph.D., is an associate director for graduate programs and teaching assistant professor of information systems at the Sam M. Walton College of Business at the University of Arkansas, where she teaches analytics, decision support, and data mining courses.

Scott Roloff is the president of IntegerHealth, which combines advanced analytics with medical expertise to quantify healthcare outcomes for insurers, health systems, and employers, driving down the costs while improving the care for health plans, wellness programs, and workers' compensation. He is a CPA and lawyer.

Bill McCallum is the chief information officer of IntegerHealth and leads its workers' compensation practice. He also holds a patent on merging disparate data sets.

Mark Barta is the City of Fort Worth's assistant director of human resources and its director of risk management. He leads the city's workers' compensation program, as well as its city-run health plan clinics.

1 See Annarino, John, Deborah Kroninger, Bliss Dickerson, Freddie Johnson, Scott Roloff & Kenny Grifno, "Ohio Pilot Program Steps Up Workers' Compensation Effectiveness," *The Journal of Total Rewards,* 1st Quarter 2023, pp. 22-38.

2 Using these outcome analytics, the City of Fort Worth decreased its workers' compensation costs 23% by constructing a provider network of the best providers for each type of injury. Roloff, Scott, Bill McCallum, Mark Barta & Jody Moses, "How Fort Worth Drove Down Workers' Compensation Costs While Getting Injured Employees Better Care," *Public Risk*, March / April 2021, pp. 10-13. You can also use these outcome analytics on your employee health plan and wellness programs. Roloff, Scott, "How to Decrease Your Health Plan Costs," *The Self-Insurer*, September 2020, pp. 42-52; Roloff, Scott & Kenny Grifno, "How to Calculate the ROI on Your Wellness Program," *The Journal of Total Rewards*, 3rd Quarter 2022, pp. 22-32.

3 See Warren, Harvey, *The Optimized Patient 2.0: How to Prepare for, Survive, and Recover from Any Surgery and Major Injury* (2nd ed., 2021).

4 So what is the difference between artificial intelligence (AI) and machine learning? AI is a broad field focused on building computers and software to mirror and eventually surpass human cognitive functions, including understanding language and analyzing data. Machine learning (ML) is a subset of AI that uses algorithms to recognize patterns in data sets.

5 See Roloff, Scott, Kenny Grifno, Bill McCallum & Mark Barta, "Integrating Health Plan, Wellness Program, Workers' Compensation, and Human Resources Data," *The Journal of Total Rewards*, 2nd Quarter 2024.

6 Syntactic analysis concerns the grammatic structure of a sentence, while semantic analysis focuses on the sentence's meaning.

7 See Zhang, Yin, Rong Jin & Zhi-Hua Zhou, "Understanding Bag-of-Words Model: A Statistical Framework," *International Journal of Machine Learning & Cybernetics*, Volume 1, pp. 43-52 (2010); HaCohen-Kerner, Yaakov, Daniel Miller & Yair Yigal, "The Influence of Preprocessing on Text Classification Using a Bag-of-Words Representation," PLoS ONE 15(5): e0232525, https://doi.org/10.1371/journal.pone.0232525.

8 Delen, Dursun, Predictive Analytics: Data Mining, Machine Learning and Data Science for Practitioners, Pearson FT Press (2nd ed., 2020).