A visual story of data management

The scene: A beautiful spring day outside the UTEP Student Learning Center.

Richard: Kendra. Hey. Kendra is over here. Hey, how are you doing?

Kendra: Hello, Richard. How are you? I am doing fine.

Richard: Are you looking forward to summer break?

Kendra: I am. However, I am working on my dissertation over the summer. I am handling additional faculty duties since we are short-staffed.

Richard: That's right. Hey, do you know who is across the street? I have seen him more often lately.

Kendra: Yeah that's John Q. Veteran. He hangs out by the library on nice days. He was a student, but he wasn't doing well. I heard he served in Afghanistan. He seems to be drunk all the time. It's sad.

Richard: It is terrible. But it's not as uncommon as you think.

Kendra: What do you mean?

Richard: Well, did you know that 80% of the service members consume alcohol, and 40% of them binge drink, compared only to 55% and 27% in the civilian population, respectively?

Kendra: How do you know that?

Richard: The Department of Defense surveyed alcohol behaviors over the prior 30-day period.

Kendra: I bet you already know what factors contribute to these behaviors? They are probably related to having served in combat or being injured in war. So, what is the answer?

Richard: Let's visit my office, and I will show you.

Richard: In the Army, we have this acronym – the BLUFf.

Kendra: BLUFf, what's that?

Richard: B-L-U-F. It means bottom line up front—the BLUFf. So, I'll give you the answer straight up—daily stress.

Kendra: Daily stress? Really? Not combat? Not injury? I didn't expect that.

Richard: I did not expect that either. But my research showed that it is daily stress.

Now that you know the answer let me show you how I reached that conclusion. Think of a three-legged stool. You can see the picture has an extra fourth leg, but don't worry about it; I couldn't erase it. The three legs represent the supporting evidence for my conclusion. These are the approach, the analyses, and the result with confirmatory testing. Together, each will build upon the other, providing a step-by-step methodology resulting in a rigorous and evidence-based conclusion. Let's start at the beginning. I used the dataset “Military Well-being Project 2020) from the Inter-university Consortium for Political and Social Research at the University of Michigan. It is an open-source, non-restricted repository. I perform a secondary data analysis.

My project aims to evaluate the association between mediating factors and the risk of alcohol and substance abuse, with the primary research question: How do these mediating factors contribute to the risk of alcohol and substance abuse?

My secondary questions are: What factor contributes most to risk? Finally, is there any buffering effect of protective factors versus risk factors?

Kendra: What do you mean by mediating behaviors? Protective? Risk factors?

Richard: I am glad you asked. Mediating behaviors can be grouped as protective or risk factors. Protective factors include calling and purpose, community engagement, help-seeking behavior, military identity, quality of life, public service motivation, and overall wellness.

Remember, my population consists of veterans, so the factors are related to the qualities and virtues associated with military service.

The risk factors include combat exposure, daily stress, loneliness, moral injury, self-stigma, suicidal ideation, and a history of traumatic brain injury.

So, let's go through the process. To recap, I concluded that daily risk contributes the most. Let me turn on the projector so we can take a closer look.

As taught in biostats class, I did my due diligence with my data management. First, I conducted a descriptive statistical analysis to understand the dataset better. Then, I

performed univariate analyses to understand each variable better. I did several bivariate analyses to identify associations and possible collinearity. I really like the look of the dendrograms. I even confirmed the low likelihood of collinearity with a variance inflation analysis.

Now, for the real statistical analysis. I used stepwise linear regression modeling to answer my overarching hypothesis: Protective factors do not buffer risk factors for alcohol and substance abuse. Oh my, this AI is literal. I said a stepwise modeling and got a model on a step. Goodness. That is better. Now, we have the regression model equation. Don't worry; I won't show you all 122 slides to make my point.

This table shows the stepwise progression of my analysis. So, the models included seven protective and seven risk factors. I also included the known confounders for alcohol abuse risk, which include age, gender, race, and socioeconomic status. I used income as a proxy. I started with reduced models, adding in variables and removing variables. Noticeably, the beta coefficients for the protective factors are small and unchanging across the models. I will explain the importance later.

Now, let's focus only on the final, full model. As you can see, daily stress contributed the most to the outcome. It was significantly significant at a p-value <0.05. This was true across all models. My thought was, how clinically relevant is this finding?

I performed a Cohen-F^2 analysis on my data. Daily stress, moral injury, and suicide risk all had a medium to large. Now, I have both statistical significance and clinical relevance.

But is it really true? This was not as I expected. Further, the regression modeling showed no buffering effect. Remember, the beta coefficients for the protective factors did not change across the models.

So, I did one last analysis: a causal interference analysis of the data. As I showed in the table of models, daily had the largest beta coefficient across all models, and I was the strongest contributor. Further, it suggested that daily stress is not plausibly caused by alcohol and substance abuse risk. Therefore, it was the only variable that directly predicts the outcome. It also showed that both a history of a traumatic brain injury and combat exposure may be confounders in addition to the other known ones. Thus, my conclusion is that **daily stress** contributes the most to the risk of alcohol and substance abuse in veterans—my bluff.

Kendra: So, what are you going to do with this finding?

Richard: Let's recap to set the stage for possible implications, next steps, and the future direction. Remember, I hypothesized that the protective factors would buffer the risk factors, but the data did not show that. The linear regression modeling showed that daily stress was the largest contributor, which was both statistically and clinically significant. It had a medium effect size and was confirmed by a causal inference analysis to be a direct predictor of the outcome.

Potentially, this finding has a great impact on clinical practice in the prevention of alcohol and substance abuse for veterans.

As practitioners, we often ask patients about the negative aspects of their lives. The surveys are worded in the negative": “I feel sad" instead of “I have been feeling optimistic lately." Even though the protective factors did not buffer the risk factors, they did not contribute to the risk. So, questions about these aspects must be included in the patient encounter. Sadly, they rarely are.

Now, getting back to daily stress. It is well known that stress has significant impacts on health. All too often, stress is not discussed in the clinical setting. Instead, the focus is on depression, anxiety, and PTSD. This research shows that there needs to be a shift in the approach to preventive medicine.

The focus has been on combat exposure, the history of TBIs, and the sense of moral injury, especially from those veterans who have served in combat. Not stress and especially not daily stress. Yet our modeling showed that daily stress is the largest contributor to the risk of abuse.

Therefore, this research can help to form policy and clinical practice to educate providers to develop a provider-patient relationship that places greater emphasis on daily stress. Stress is a typical personal occurrence. An open discussion in the clinical encounter will establish a healthy therapeutic alliance between the patient and the provider. This could lead to a decrease in the prevalence of alcohol use and bingeing through prevention.

Kendra: That's fascinating. I would not have guessed that daily stress would be the predictor for the risk of alcohol and substance abuse. So, what are you ultimately going to do with your project?

Richard: I plan to publish the findings because they were so interesting and unexpected. Hopefully, the VA leadership will allow me to educate their providers about daily stress so that they can discuss it during the clinic visit.

Kendra: That's fantastic. I have a meeting to go to, but this was great. Have a great summer, and I will see you next semester.

Richard: You as well.