



Understanding the Ethics and Privacy Concerns with Suicide Risk Prediction Algorithms

Sarah Logan (Siena College) and Dr. Michael Zimmer (Marquette University)



Introduction

- Suicide is one of the leading causes of death in the United States
- A model has been developed to predict suicide risk using elements from electronic health records
- This model could be improved through the use and incorporation of publicly available socioeconomic data
- This poses a problem as there may be potential privacy concerns regarding the use of this data

Methods



We created a survey to measure public opinion on the use of personal data in various contexts, with a specific focus on the context of a suicide prediction model. Participants were presented with 10 vignettes (short scenarios) to which they rated their concern. There were 420 total respondents.

Concern Over Data Use in Specific Scenarios

- Most concerning vignette: tracking adherence to stay-at-home orders during the COVID-19 pandemic using GPS data
- Least concerning vignette: restaurant manager conducting surveys to improve the restaurant's quality of service
- Factors that contributed most to concern over vignettes: the purpose of data collection and the potential future use of the data

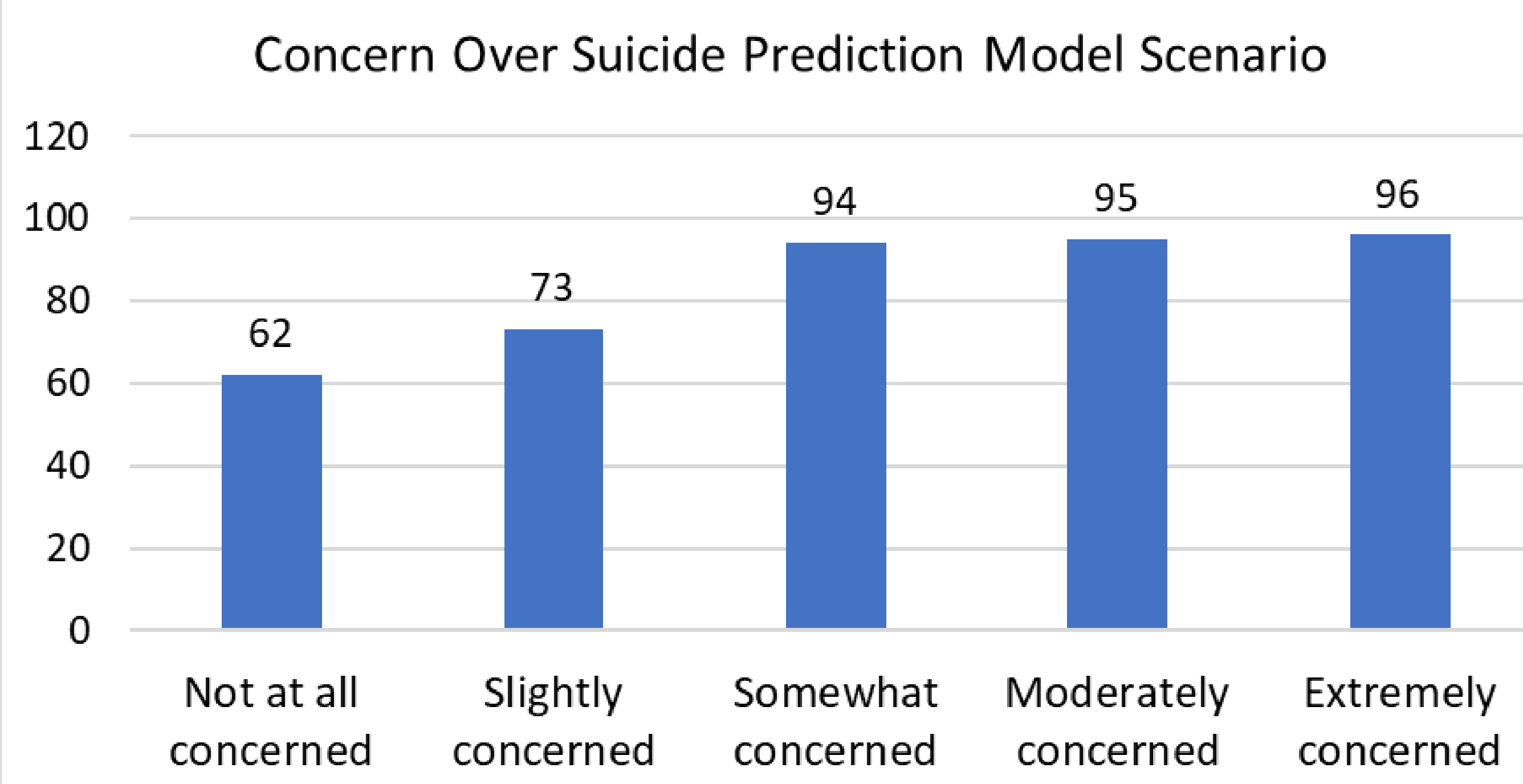
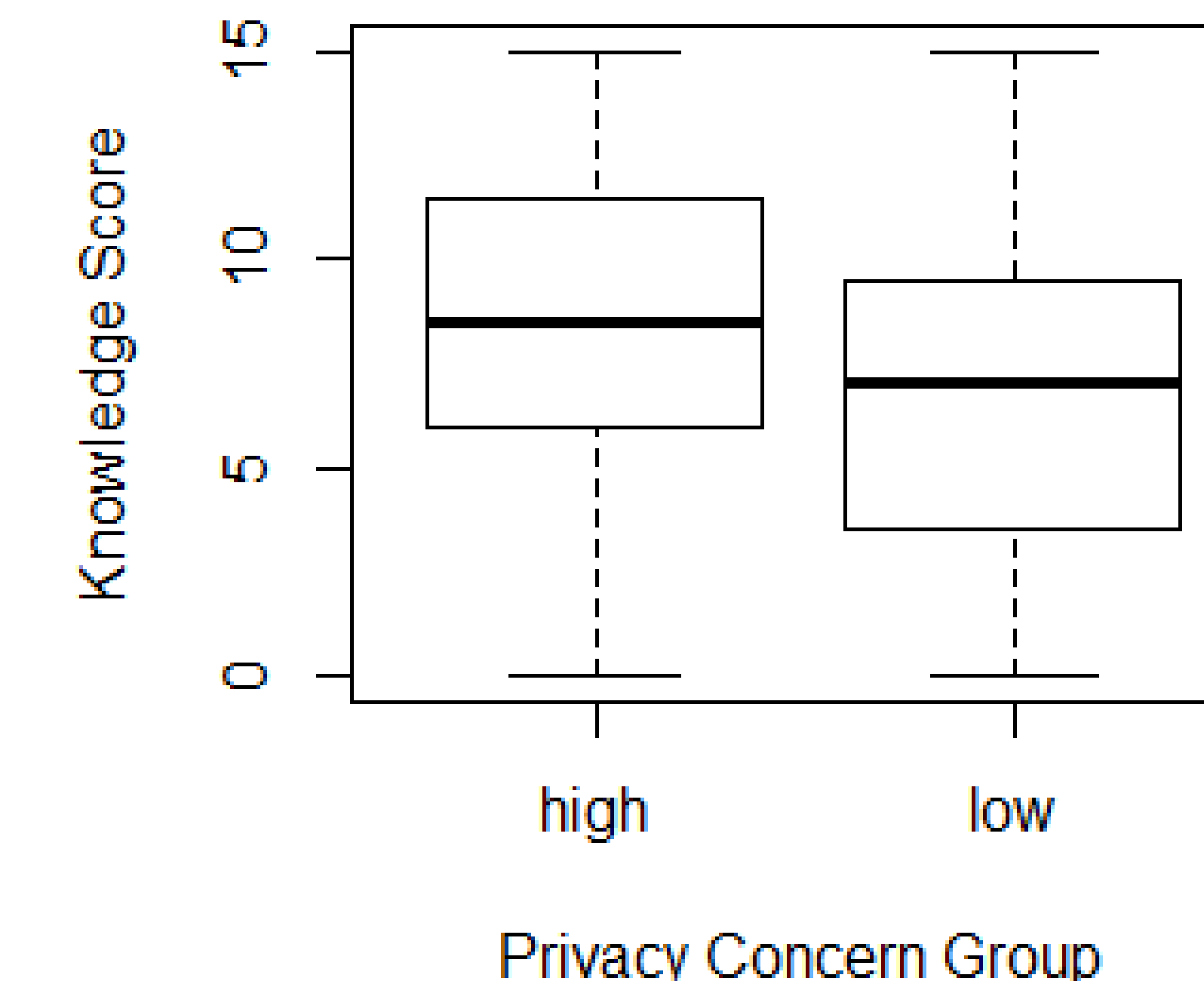


Figure 2. 67.9% of participants said they were somewhat, moderately, or extremely concerned over the general use of publicly available data in a suicide risk prediction model.

Relation Between Knowledge & Concern

Figure 4. According to assigned knowledge scores, participants with high privacy concern generally had greater privacy knowledge than participants with low privacy concern.



Discussion

- The most concerning factors (purpose of data collection and potential future use of data) indicate that participants weren't as concerned about the sensitivity of the data itself but were rather concerned over how the data would or could be used.
- Some data attributes elicited concern from both the high and low privacy concern groups: income, assets, and court records.
- Do individuals in the low privacy concern group have a genuine disregard for their privacy or do they feel a sense of apathy in trying to protect their information?

Knowledge of Publicness of Data Attributes

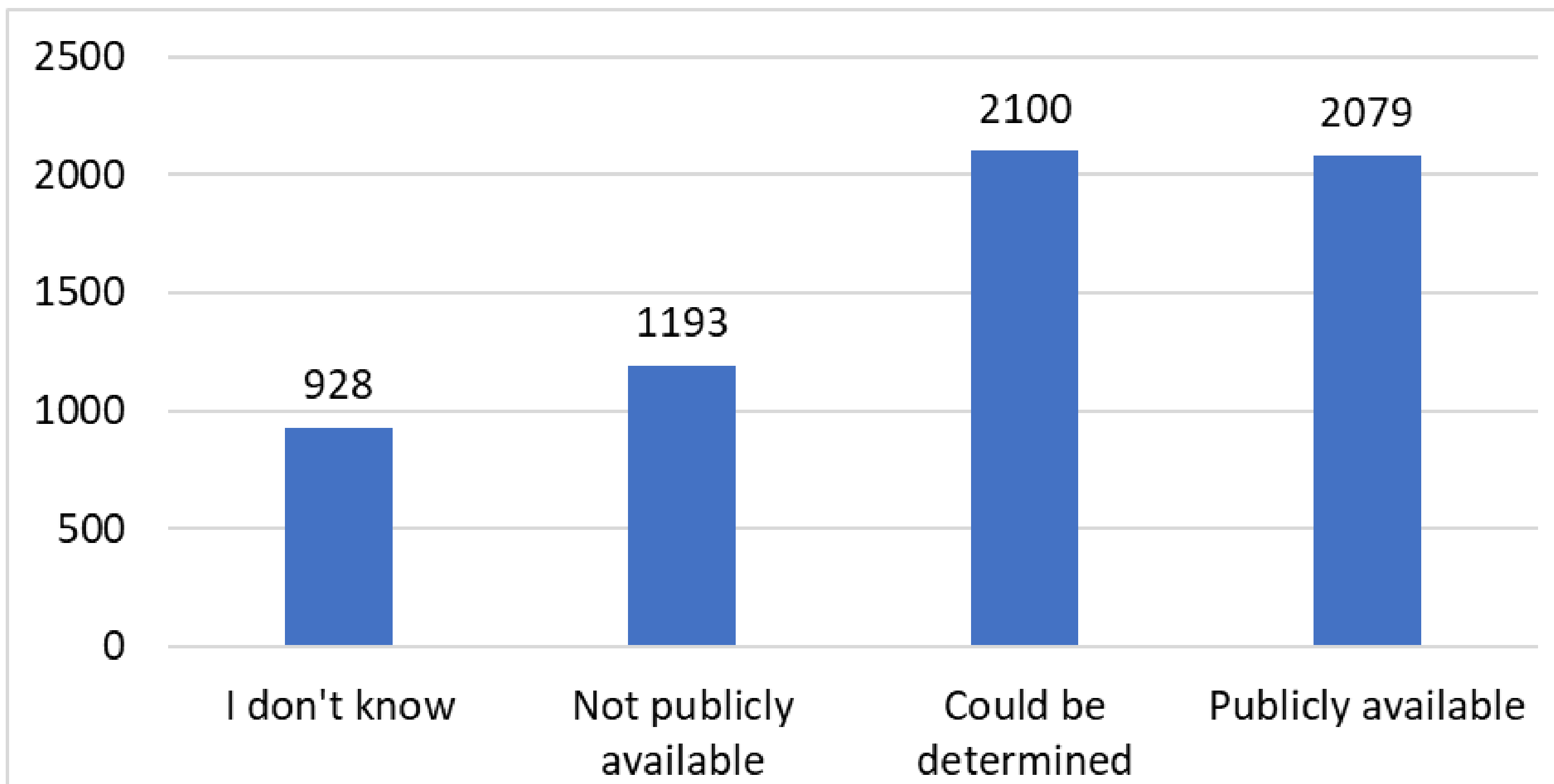


Figure 1. Total responses to all knowledge questions. 66.3% of responses to the knowledge questions were correct (selected "Could be determined" or "Publicly available").

High vs. Low Privacy Concern Participants

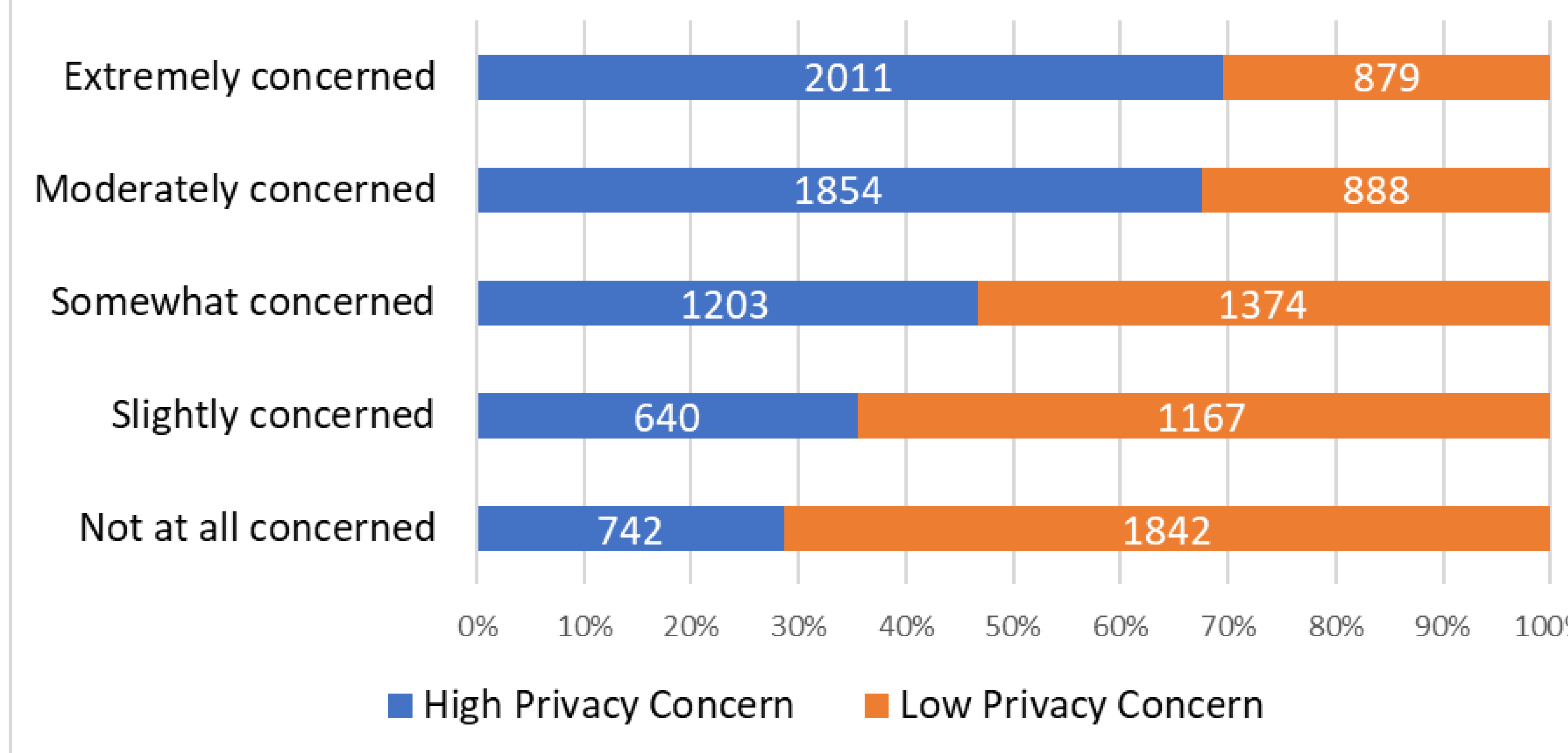


Figure 3. Total responses to concern over the use of specific data attributes for a suicide prediction model. General privacy concern was indicative of participants' concern level over specific data attributes.

Conclusions & Future Work

- Overall, participants had a good amount of knowledge regarding the publicness of various data attributes.
- Privacy concern is contextually dependent, as different vignettes received different levels of concern.
- General privacy concern corresponds to specific concern over the use of various data attributes in the context of a suicide prediction model.
- There is a positive relationship between privacy knowledge and privacy concern.
- It may be beneficial to further investigate what factors (ie: privacy protecting behaviors) can serve as predictors of privacy concern.