## **Peer Review File**

Article information: https://dx.doi.org/10.21037/atm-23-888

## **Review Comments**

This paper focuses on a very interesting topic. The paper is well-written structured. However, some issues must be improved:

1. The Authors don't explain what is in justifying the originality of this paper.

Response: Thanks for the suggestion, we added the originality in introduction section.

Changes in the text: line 96-97, page 4.

2. The description of sampling and data collected in the survey is not clear, the authors do not include information on the analyses conducted. Ideally, there would be more in-depth information on the data collection, reliability of the questionnaire, and its validity.

Response: Thanks for the suggestion, we added the sampling process and questionnaire reliability and validity in methods section.

**Changes in the text:** line 123-124 and 127-130, page 5.

3. In the work of Podsakoff, MacKenzie, Lee and Podsakoff (2003) particular attention is paid to the fact that variance methods may intensify or underestimate the observed relations between constructs, which leads to both type I and type II errors. CMB was not presented in this paper. **Response:** Thanks for the suggestion, we added the common methods bias in limitation section.

Changes in the text: line 266-268, page 9.

4. Next, as shown in Tab. 6 some of your sub-scales give Alphas below the tolerance scale - it is acknowledged that alpha coefficient should exceed 0.7 (Nunnally & Bernstein 1993; Tavakol & Dennick, 2011). It suggest that some of sub-scales are not reliable.

**Response:** Thanks for the suggestion, it is true that 0.7 could be the standard for reliability, but generally, 0.6 is also commonly accepted for the criteria of reliability.

Changes in the text: None.

5. Tab. 7 don't contain Cronbach's α?

**Response:** Thanks for the suggestion, the CR value is the Cronbach's  $\alpha$ , we corrected it in the Table 7.

Changes in the text: line 1, page 19.

6. Why later in the charts do we have z-scores for all constructs (not real value) except selfanchoring work engagement?

Response: Thanks for the comment, it is because we should display clearly the influence of all the variables.

Changes in the text: None.

7. The min for CR is 0.7 and for AVE is 0.5. Table 7 shows that these criteria are not met.

Response: Thanks for the comment, we revised the standard of CR to 0.6. In CR section, communication and clinical skills have value more than 0.6 and in AVE information management had value more than 0.5, which could be a screener for eligible variables.

Changes in the text: None.

8. What does the content of Table 8 show? The results are shown to meet the condition Discriminant validity – Fornell-Larcker Criterion?

Response: Thanks for the comment, we revised the table caption and the matrix is needed to evaluate the validity of all the variables.

Changes in the text: line 1, page 20.

9. The substantive aspects of using SEM are not met, see the paper: Kosiek, K., Staniec, I., Godycki-Cwirko, M., Adam Depta & Anna Kowalczyk Structural equation modeling for identification of patient safety antecedents in primary care. BMC Family Pracice 22, 183 (2021). https://doi.org/10.1186/s12875-021-01533-6

Response: Thanks for the comment, we think Kosiek's article is different form ours in both study design and technique details. In our research more technique details of questionnaire reliability and validity, and competency modelling are provided.

Changes in the text: None.