Dear Reviewers,

Thank you very much for your deep and thorough analysis of our manuscript. Below are the answers and explanations of the raised doubts. In the main text of the article, the changes are marked in red for better detection. We hope that the article in its current form can be accepted for publication.

Kind regards,

On behalf of the team – corresponding author.

Rev 1, ID no : 06298726

"This paper is of high quality and highly innovative, which can be used to screen for peripheral neuropathy in Poland to improve the diagnostic process of peripheral neuropathy. However, scale translation and factor analysis should be supplemented in the research results."

Response:

-The scale translation with the instruction how to perform it were added as a supplementary material no 1 and 2. We hope that it was available for the reviewers. As a supplementary material no 4 the Factor analysis have been added (Suppl no 4 – Table 5 and 6). Results and information about are now available in the main text.

Rev 2, ID no : **06269601**

"This study deals with the validation of the Michigan (…). The innovation is clear, but the manuscript requires large improvements. I have pointed out some grammar issues, but there are several others. The methods and results sections require improvements. Figures and tables also need adjustments in terms of visualization. Methods -Third paragraph, last but one line – exclude "and" -Fourth paragraph. Insert "," after "Thus" Population -Improve the presentation of inclusion/exclusion criteria. Present them in the same sentence, and not in topics. Use a) or i) to distinguish them. Furthermore, explain in the manuscript why these criteria were adopted. -Change "weight" to "body mass" -It is not clear what the following sentence means "A nerve conduction study (NCS) was performed on the same day as the first MNSI" Statistical analysis -Change "The basis from" to "The basis for". -The information regarding NCS as the gold standard should appear before. Results -This section presents information that should appear in the methods. Limit this section to the results, without discussing it or presenting new information on the adopted tools. "

Response:

Suggested changes were made.

We understand that reading tables and figures can be readable to a greater or lesser extent for different audiences. Nevertheless, they have been prepared in accordance with the Journal guidelines. We hope that their interpretation contained in the main text is clearer and helpful. We are open for the suggestions from the Editor**.**

The inclusion and exclusion criteria were adopted based on the previous study (Turkish and Portugal versions – please see references no 20 and 28) and were in line with the neurologist's suggestion on how to avoid misdiagnosis**.** Although, in everyday practice, the MNSI can be applied to the whole population with diabetes (also these patients with exclusion criteria used in our study) very strict criteria should be applied during the validation process.

According to the sentence: „Nerve conduction study (NCF) was performed on the same day as the first MNSI" – we would like to explain that e.g. if the patient N.N. had MNSI assessment on 01st May 2022 on the same day, the NCS was done (after MNSI).

Rev 3, ID no : **06261624**

Dear Reviewer and Editor - Only points that raised some concerns are addressed below.

"4 Background. Does the manuscript adequately describe the background, present status and significance of the study? No. I note that the authors give information about MNSI in the "Methods" section. However, I would also suggest that the authors give a brief background on the MNSI in the "Introduction" section, such as reliability, number of items, and a focus on the current status of the MNSI study.

Response:

A brief description of the method has been added. Due to the length of the article, we did not want to repeat obvious information in several chapters.

"*5. Methods. Does the manuscript describe methods (e.g., experiments, data analysis, surveys, and clinical trials, etc.) in adequate detail? Yes. There are two suggestions that I hope the authors will seriously consider.*

1) The authors mentioned that 80 participants were eventually recruited to take part in the survey. Although the authors described that the minimum sample size was the analysis of the reliability of the test with Cronbach's alpha coefficient. However, small sample size may seriously affect the reliability of the study results, especially in scale revisions."

Response:

The information about the sample size calculation is within the Statistical analysis chapter, please see below:

“ ….Based on the accepted assumptions and literature information, the minimum sample size was set at 80. “

The mathematical structure of most statistics, including those calculated when performing the Cronbach's alpha and reliability analysis, means that the increase in the size of the sample N itself increases the statistical significance, therefore it is highly probable that a larger sample would only increase the statistical significance obtained by us. In our opinion, the surveyed group was not so small that the statistical significance we showed could be considered accidental.

2) In the "Statistical analysis" section, Cronbach's alpha, split-half reliability and other tests and related literature should be given. I notice that the authors present the relevant content in the "3. Reliability and stability results" section. However, some indicators are not given in the corresponding literature. It is suggested to move the relevant content to the "Statistical analysis" section.”

Response:

The suggested part of the text has been moved and suggested changes were made.

“6 Results. Are the research objectives achieved by the experiments used in this study? What are the contributions that the study has made for research progress in this field?

1)The author mentioned that "legs" was changed to "legs and feet" and also gave an explanation while pointing out that the same problem was reported by the team from Portugal. I suggest that the explanation should be put in the discussion section. The results section should just present what the study found. As for the basis of the revision, it should be written in the discussion section.”

Response:

The suggested part of the text has been moved.

2) In general, when the p-value is less than 0.001, it is reported as $p < 0.001$ without going into detail. I note that the authors report results to 6 decimal places, is this necessary?”

Response:

The suggested changes have been made.

“7. Discussion. Does the manuscript interpret the findings adequately and appropriately, highlighting the key points concisely, clearly, and logically? Are the findings and their applicability/relevance to the literature stated in a clear and definite manner? Is the discussion accurate and does it discuss the paper’s scientific significance and/or relevance to clinical practice sufficiently?

The manuscript has explained the study results more adequately and appropriately, however, the discussion of the theoretical significance and clinical value of the paper is partially missing.

1) *To facilitate the reader's reading, the writer should explain the differences between part A and part B.*

Response:

The short description of the tool (MNSI) is now included in the introduction (please see our answer to your comment no 4 Background) and was previously incorporated in the Methodology chapter. Very detailed information was also provided in the Supplements. We are very sorry, but the length of the article does not allow us to refer to this topic again. The validated scale is intended for people who care for patients with diabetes, so there is no need to worry that they will not understand the difference between parts A and B.

2) *The authors mentioned that "we achieved a lower specificity but higher sensitivity and the sensitivity and specificity from our validation study were closer to the results found in the Portuguese [20] and Turkish [23] analyses". So, is it possible for the author to give a relevant explanation?*

Response: Please see below after the next issue

3) Besides, the authors also mentioned that "In our study, we found the optimal cut-off point for section B to be 2 with a sensitivity of 81–84% and specificity of 60–70% which was similar to the Portuguese study (86% sensitive and 61% specific using the same cut-off point) [20]". It is suggested here that the authors should provide an appropriate discussion around the similarities and differences with previous studies.

Response:

Dear Reviewer we also mentioned in the Discussion section: "…the same cut-off points used in our study for sections A and B were proposed by the authors of the Portuguese version [20] with similar levels of sensitivity and specificity and in one of the Turkish validation studies [28]…."

As you cited us above, we summarized in the Discussion section that it was: " similar levels of sensitivity and specificity" with the Portuguese and Turkish studies. I did not find an analysis of the differences between studies about sensitivity and specificity in mentioned articles dedicated to the validation of MNSI, only the values were put with short comment like in our study. We do not want you to find our comment as ignorance but an analysis of such details may be a good theme for a statistical article but is not clinically relevant. Below you can find the answer from our Statistician Team, but (I am sorry) I see no reason to include it in the Discussion, because the study itself will then cease to be the goal and its reading will be illegible :

Due to the method of determining sensitivity and specificity from two-way frequency tables, it is practically impossible to obtain identical values of these parameters in two different studies. In order to answer the question of what difference between the values of these parameters obtained in various studies should be considered statistically significant, the

easiest way is to determine the 95% significance intervals for these values and assess whether they overlap.

The ROC analysis was the basis for determining the 95% significance intervals for the statistically significant values of sensitivity and specificity calculated in the study.

According to the ROC analysis, the statistically significant sensitivity and specificity values discussed in the publication:

PE1) For the cut-off point 2, and the Youden Index value = 0.514, the calculated sensitivity and specificity values were respectively:

Sensitivity=81.43% and Specificity=70.00%.

The 95% significance intervals for these values are:

95% 〖CI〗_Sensitivity={70.34%;89.72%} and 95% 〖CI〗_Specificity={34.75%;93.33%}

PE2) For the cut-off point 2, and the Youden Index value = 0.443, the calculated sensitivity and specificity values were respectively:

Sensitivity=84.29% and Specificity=60.00%

The 95% significance intervals for these values are:

95% 〖CI〗_Sensitivity={73.62%;91.89%} and 95% 〖CI〗_Specificity={26.24%;87.84%}.

A comparative analysis of the calculated 95% significance intervals for the sensitivity and specificity parameters characterizing our study with the discussed literature (Portuguese and Turkish) values clearly shows their similarity. With the assumed confidence level of α=0.05, there is no reason to consider the difference between them as statistically significant.

3) *It is hoped that the authors can write clearly about the theoretical significance and clinical practice of this study.*

The following explanation has been added to the discussion section:

"The Recommendations of the Polish Diabetes Association do not propose the use of appropriate scales for assessing peripheral neuropathy [33]. This is due to the lack of adaptation of scales such as the MNSI to the Polish system and may contribute to a lack of uniformity in assessing the patient's condition. Thus, some centers create their own informal evaluation criteria and rules (even though they use the same tools and tests). The introduction of a validated scale will help standardize screening for neuropathy in

diabetic patients. Similar problems arise in the design of clinical trials, as only the English-language scale can be used, which is not valid."

5) I suggest that the authors could have discussed the study section more carefully, for example, as the sample size was small, the representativeness of the sample and the way it was sampled, etc., could be potential shortcomings.

Response:

Please see the answer for your point 5.1 about the criterium of the sample size calculation.

11. References. Does the manuscript appropriately cite the latest, important and authoritative references in the Introduction and Discussion sections? Does the author self-cite, omit, incorrectly cite and/or over-cite references?

Some of the references are formatted incorrectly, and it is expected that the authors will check them carefully. In addition, it is suggested to add relevant literature from the last 3 years appropriately.

Response:

We checked References - we hope they are now formatted correctly. We have supplemented the literature with the latest publications.

Rev 4, ID no : 05486871

"The manuscript should be revised regarding to the COSMIN guidelines. Also, international translation procedures should be considered."

Response:

Although other reviewers confirmed that the study was conducted correctly and in accordance with the recommendations, we have checked the recommendations you mentioned (https://www.cosmin.nl/). These points were reached in our study indirectly as were a part of the published instruction, we followed, on how to validate (references- please see below).

Ad 1. The subsequent points related to COSMIN which are typical and necessary for translation and validation were described within the text in the Method section, Statistical analysis, as well as in the Supplementary materials. The study was preparing according to available guidelines (Beaton D.E. et al., 2000; Cam and Baysan-Arabaci, 2010) dedicated to validation standards and its construction was approved by the Bioethics Committee. Please

find that we did not validate a new tool, but we adopted a commonly used one for the Polish market. To minimize the risk of bias associated with transmission we also used the gold standard.

Ad 2. We followed the translation procedure, we found in the recent literature and this procedure was described with details in the main text – page number 4 (section Methods). Forward and backward translation as well as cultural adaptation had been applied together with recipients' (patients and practitioners) opinions (described in detail- as an example please see Results section point 1). Therefore, as you can find, we do not avoid statements about the difficulties that arise in the translation process.

The aim of our study was to validate a widely known and used by specialists, partially subjective scale, assessing the presence of peripheral, sensorimotor neuropathy in patients with diabetes. The usefulness of this scale has been confirmed in previous studies. Like any screening tool, mainly if even partially contains a subjective element (patient report- part A in MNSI), we are aware that the MNSI scale is not perfect either. Nevertheless, undermining its usefulness in everyday practice was not the aim of our study. Due to the previously mentioned potential weaknesses of the MNSI, the results of this scale were referred to the so-called gold standard (NCS), which confirmed its usefulness as a screening in this as well as in previous studies.