Agreement between Original and Rasch-Approved Neck Disability Index

ZE LU (luze66269271@gmail.com)
McMaster University School of Rehabilitation Science https://orcid.org/0000-0002-7398-1968

Joy C. MacDermid
Western University

Goris Nazari
Western University

Research article

Keywords: Original, Rasch-Approved, Neck Disability Index

Posted Date: April 25th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-18446/v1

Version of Record: A version of this preprint was published on July 3rd, 2020. See the published version at https://doi.org/10.1186/s12874-020-01069-w.

Abstract

Background: The neck disability index (NDI) has been used widely to monitor the treatment outcomes from patients' perspective, given that neck pain is a notable social burden with a high prevalence (33%) within the adult population. Previous researchers have examined its psychometric properties and latent construct. Uptake of modified versions of NDI have been limited, however, the traditional/original NDI is still commonly administered. Examining the extent of agreement between traditional and Rasch-based versions using Bland-Altman (B&A) plots will inform our understanding of score differences. Therefore, the objective of current study was to describe the extent of agreement between different versions of NDI.

Methods: The current study was a secondary data analysis where the study data was compiled from two prospectively collected data source. We performed a comprehensive literature search to identify Rasch analyses of the NDI within four databases including Embase, Medline, PubMed, and Google Scholar. Modified version with associate score transformation solution was obtained if they meet the predefined criteria. We established B&A plot by the mean difference and the 95% limits of agreement (LoA; ±1.96 times the standard deviation).

Results: Two Rasch approved versions (8- and 5- item) that met the study criteria were identified from 303 publications. We analyzed 201 (43 males and 158 females) patients attending community clinics for neck pain. We found that the mean difference was approximately 10% of the total score between the 10-item and 5-item (-4.6 points), whereas the 10-item versus 8-item and 8-item versus 5-item had similar mean differences that were about half (-2.3 points). The B&A plots displayed wider 95% LoA for the agreement between 10-item and 8-item (LoA: -12.0, 7.4) and 5-item (LoA: -14.9, 5.8) compared with the agreement of 8-item and 5-item (LoA: -7.8, 3.3).

Conclusion: Our study is the first to examine the agreement between each versions of NDI. Due to the unpredictable bias and wide 95% LoA, the traditional 10-item version should not be used interchangeably with other two Rasch-approved versions even with a small magnitude of bias. Different coverage of content should be considered during the decision of 8- and 5-item NDI.

Background

Neck pain is considered a notable social burden and has a high prevalence (33%) within the adult population, and nearly 70% of patients will experience a life-span neck pain (Bovim, Schrader, & Sand, 1994; Covic, Pallant, Conaghan, & Tennant, 2007; Croft et al., 2001; Hogg-Johnson et al., 2008; MacDermid et al., 2009). Clinical decision-making requires monitoring the treatment effect (improvement or deterioration) from both clinician and patient perspectives. Since the first patient-reported outcome measure (PROM) assessing pain and disability in participants with neck pain was published in 1991 – the neck disability index (NDI), it has been used widely in surgical treatment, injection therapies, physiotherapies, as well as the exercise and research context (Iyer et al., 2019; MacDermid et al., 2009; Vernon, 2008). Evidence from systematic reviews and individual studies suggest that NDI has sufficient

level of psychometric properties across various populations (Bobos, Macdermid, Walton, Gross, & Santaguida, 2018; MacDermid et al., 2009). The NDI has been cited and applied in more than 300 publications as the outcome measure to assess for neck-related conditions. In addition, the original version has been translated into 22 languages versions (Evans et al., 2012; Vernon, 2008).

The NDI was developed as a unidimensional instrument assessing neck disability which assumes unidimensionality, a fundamental requirement for using a summary score (Packham & Macdermid, 2013; Pallant & Tennant, 2007; Van Der Velde, Beaton, Hogg-Johnston, Hurwitz, & Tennant, 2009). Previous researchers have examined original NDI using factor analysis, qualitative interview, and construct analysis under the classical test theory (CTT) (Hung et al., 2015). Gabel et al (2014) concluded that the NDI is a one-factor model confirmed by confirmatory factor analysis in a homogenous population with neck pain. However, the other team identified 2 factors by a principal component analysis (Wlodyka-Demaille et al., 2004). MacDermid et al (2009) found that the NDI moderately relates to both the physical and mental component scores of the Short Form 36 health questionnaire. Through the problem elicitation technique (PET), others have concluded that the NDI is a multidimensional scale that measures symptoms, impairments, and disabilities (work, recreation) (Hoving, O' Leary, Niere, Sally, & Buchbinder, 2003). Given the uncertainty identified in latent construct, the Rasch analysis based on item response theory (IRT) and Rasch modelling, was used to evaluate these performance characteristics of PROM (Jerosch-Herold, Chester, Shepstone, Vincent, & MacDermid, 2018; Packham & Macdermid, 2013; Pallant & Tennant, 2007; Smith, Rush, Fallowfield, Velikova, & Sharpe, 2008). It enables examination of unidimensionality and interval level of scaling, and provides the transformation strategy to convert an ordinal score to interval scaling, which can validate the use of a total sum score (Cano, Barrett, Zajicek, & Hobart, 2011). Where outcome measures are not developed using Rasch modelling, they can retrospectively be evaluated for fit to the Rasch model which often results in suggested modifications needed to obtain fit.

Several studies have inspected the NDI using Rasch analysis and found violations of Rasch basic assumptions (Gabel, Cuesta-Vargas, Osborne, Burkett, & Melloh, 2014; Van Der Velde et al., 2009; Walton & MacDermid, 2013). They offered solutions which included exclusion of misfit items and new coding algorithms. Although modified versions of NDI have been constructed that are conceptually and statistically sound, uptake has been limited and the traditional NDI is still commonly used. Studies to date have focused on defining modified versions with better measurement properties but have not defined the extent to which these new versions differ from the traditional NDI scoring outside of the development data set. Examining the amount of agreement between traditional and Rasch-based versions of the NDI using Bland-Altman (B&A) plots will inform our understanding of how these scores might differ (Bland & Altman, 1999, 2010; Nazari, MacDermid, Sinden, Richardson, & Tang, 2019)

Therefore, the objective of current study was to describe the extent of agreement between different versions of NDI in a sample of patients attending community clinics for neck pain.

Methods

Study design

The current study was a secondary data analysis where the study data was compiled from two prospectively collected data source. Both studies received ethical approvals (McMaster REB #03-145 and HiREB #13-300) and all participants provided written, signed consent. Participants were recruited from community clinics presenting for neck pain in Hamilton, ON Canada through paper and online based survey.

Information source

We performed a comprehensive literature search to identify Rasch analyses of the NDI within four databases including Embase, Medline, PubMed, and Google Scholar. Search keywords were set as neck disability index, NDI, Rasch analysis, structural validity, construct validity. The search year range was limited until January 2020. Details of search strategies were presented in Appendix 1.

Eligibility Criteria

We included studies that applied the Rasch modelling to report the structural validity of NDI. According to assumption of the Rasch theory, we defined the acceptable fit of the Rasch model as followed

1. The confirmation of unidimensionality.

E.g. RUMM2030 provides the pair sample t-test, the number of the significant tests should less than 5% of the overall tests

2. Over-fit statistic examined by the Chi-square test; a non-significant p-value was acceptable.
3. To remedy the disordered thresholds, strategies such as collapsing the adjacent response options should be complied with during the calculation of total scores for the corresponding version of NDI.
4. There was no differential item functioning (DIF) existing including both uniform and non-uniform DIF in the revised version.
5. Item correlations ranged from -0.3 to 0.3 or within a narrow range for the test of local dependency.
6. We adopted an appropriate level of the person separation index (PSI) (PSI > 0.7)

Data Extraction

The score transformation algorithm was obtained if the revised version achieved an acceptable level of model fit identified by the eligibility criteria

Study Selection

An independent reviewer (ZL) performed the systematic electronic searches in all the databases. ZL also identified and removed the duplicate studies. The independent reviewer then carried out the screening of the titles/abstracts and identifying the full text articles. The final stage included an independent full text review by ZL to determine final article eligibility.

Statistical procedure

The demographic statistics of the sample including age, sex, total score of the original NDI 10-item version, NDI 8-item version, and NDI 5-item version were described by mean, standard deviation (SD), minimum and maximum value.

Agreement of Rasch solutions

The normal distribution of mean differences of all three comparisons were inspected by the histogram. Using the B&A plots, we summarized the individual agreement between each of the identified NDI versions by the mean difference and the 95% limits of agreement (LoA; ±1.96 times the standard deviation).

To test the average agreement and differences between each NDI score, we examined the mean differences by one-sample t-test. We reported the sample size for each comparison, the degree of freedom, mean differences, standard error of differences (SE), p-values, and 95% confidence interval (CI).

Transformation including logarithmic and linear transformation would be applied to normalize the non-uniform pattern of the bias on the plot. For instance, when the B&A plot shows a linear relationship between differences and means as the differences or measurement bias starts with negative value and then become positive while the magnitude of the mean increases, we can regress differences between the methods (D) on the average of the two methods (A) by $D = b1 \times A + b0$. The 95% LoA for the regression should build on the SD of the residual (SDres) from the established model (±1.96 times SDres) (Bland & Altman, 1999).

All analysis was performed by IBM SPSS statistics, Version 25.0 (IBM Corporation, Armonk, NY). We considered a significance level of $p \leq 0.05$ as statistically significant.

Results

Study Selection and NDI version identification

Initially, our search yielded 303 publications. After removing the duplications, 296 articles were left. Six studies were then selected for full text review after title and abstract review. Of these, two Rasch solutions that met the study criteria were identified from 2 individual studies (a 8 item and a 5 item version). This allowed 3 B&A comparisons (NDI-10 vs. NDI-8, NDI-10 vs.NDI-5, and NDI-8 vs. NDI-5). The flowchart of studies through the selection process is displayed in Figure 1.

Ordinal score transformation

Three NDI scores were calculated for each participant. The first NDI score was derived from the original ordinal scale (maximum of 50)(Vernon, 2008). We calculated second set of NDI scores according to the 8 item solution provided by Van Der Velde and his collogues (Van Der Velde et al., 2009). Two items (headache and lifting) were deleted before summarizing the ordinal score of 8-item version. Afterwards, the ordinal scores were then transferred to linear score with the maximum value of 50. For third score transformation, 5 items regarding person care, concentration, working, driving, and recreation were kept into the total score calculation. A rescoring strategy, as suggested by previous research, was used to remedy the disordered threshold of driving related item (Walton & MacDermid, 2013). The original score of responses (012345) was re-coded by collapsing the fourth and fifth options (012334), while the original structure (012345) was retained for other 4 items. Therefore, the maximum total score of NDI 5-item version was 24 on the ordinal scale. Finally, this ordinal total score was transferred on the linear scale based on the transformation list (Walton & MacDermid, 2013). Please see Appendix 2 for a summary of transformations.

Sample

Table 1 describes the demographic information including age, pain intensity, total scores of NDI 10-item, NDI 8-item, and NDI 5-item version and stratified by sex. Thirty-one subjects experienced injury or trauma related neck-pain including car accident, sports injury, and fall. Other reasons are arthritis, pinched nerve, and disc problem. The normal distribution of the mean differences of comparisons were confirmed by inspecting the histogram. See Figure 2,3,4.

Agreement of Rasch solutions

Table 2 demonstrated both average and individual agreement results of all three comparisons.

Through pairwise comparisons, we identified the mean difference was approximately 10% of the total score between the 10-item and 5-item (-4.6 points), whereas the 10-item versus 8-item and 8-item versus 5-item had similar mean differences that were about half (-2.3 points). We considered the traditional 10-

item as the reference method during comparisons, negative mean differences indicating that both 8-item and 5-item were systematically scored higher than standard 10-item version. The B&A plots displayed wider 95% LoA for the agreement between 10-item and 8-item (-12.0, 7.4) and 5-item (-14.9, 5.8) compared with the agreement of 8-item and 5-item (-7.8, 3.3).

Through visual inspection of the Bland-Altman plot, the bias between 10-item and 8-item tended to be in opposite directions as negative value of differences were found in the lower end (before scores of 20) and positive values in the high end of the scale (between 20 and 40). Similar tendency was identified in the comparison between 10-item and 5-item version. However, such patterns disappeared in the plot comparing 8-item with 5-item version, where a fair number of bias located below zero line except for fewer outliers in the low end of the range. Meanwhile, plots representing bias between two versions clustered in the high end of the range. Please see figure 5, 6, 7.

The linear relationship showing on the B&A plot when compared 8- and 5-item version was confirmed by the simple linear regression giving $D = -0.2 \times A + 2.2$ with a significant p value for the over model and regression coefficient ($p < 0.001$) (Bland & Altman, 1999). We then plotted 95% LoA based on the SDres which was equal to 2.4 from the regression model. The new upper and lower limited was constructed as $D = -0.2 \times A + 2.189 \pm 1.96 \times 2.4$. Please see Figure 8.

Discussion

Summary of evidence

We identified two Rasch approved versions of NDI (10- and 8-item) through a comprehensive literature review and revealed disagreements within comparable versions (10-item vs. 8-, and 5- item) using B&A plot analysis. (Giavarina, 2015; Van Der Velde et al., 2009; Walton & MacDermid, 2013).

The wide range of the 95% LoA established surrounding the point estimate of the agreement would threaten the interchangeable application of different versions. When compared 10- with 8-item version, a difference of plus 7.4 or less 12.0 units accounting for nearly 15% to 25% of the total score was important for a measurement of 50 units, since 9 units of change would significantly influence the classification of the disability level(Vernon, 2008). For example, a participant who obtained a raw score of 20 as moderate level of neck disability would be inaccurately categorized into mild or severe level given a bias ranged from -12.0 to 7.4 units. The bias within agreement increased to 30% (-14.9 for lower limit) of the total score comparing 10- with 5-item versions. Although differences between 8- and 5-item version were uniform after linear transformation, a variation of 4.7 units (10% of the total score) would mislead the judgement of the level of neck disability using NDI.

Another evidence that the unstable variance identified in error patterns on B&A plot also suggest that Rasch approved NDI including both 8-item and 5-item versions might not be appropriate to be used interchangeably with the traditional 10-item version even with small errors (-2.3 and -4.6). Overestimation might occur in the lower range of scores and underestimation might occur in the upper range of the scale

when apply modified versions. Through visual inspection, the direction of bias reverted when the scores approaching 20 points which were categorized as the moderate level of disability. Attempts including both logarithmic and linear transformation failed to normalize the bias pattern. Therefore, the systematic bias within the agreement was not predictable and would vary across different levels of neck disability.

The bias between 10- and 8-item version were nearly half of that comparing 10- with 5-item version. One potential explanation could be the various constructs captured by different short NDI versions. Specifically, the modification of 8-item version was driven by the Rasch modelling. Items (headache and lifting) were deleted to achieve the optimal model fit (Van Der Velde et al., 2009). Nevertheless, for the 5-item version, the authors conducted two steps of the modification including both conceptual evaluation and Rasch analysis. The International Classification of Functioning, Disability and Health (ICF) was used to determine the conceptual relevance of individual items. According to the ICF framework, the authors kept function-related aspects during the initial item reduction (Walton & MacDermid, 2013; WHO, 2002). Symptom-based item such as pain intensity was removed at this stage (Walton & MacDermid, 2013). Such conceptual differences are retained in 8-item version but not in the 5-item version and result in more systematic errors between the scores. Researchers might determine the choice of 8- and 5-item NDI based on specialized conceptual needs. For example, 8-item provides the evaluation of neck disability regarding pain intensity, sleeping, and reading.

Strengths & limitations

The literature review only examined studies published in the English language, which may limit the identification of potential Rasch solutions of NDI. The study sample was recruited from community clinics in a single city in Canada which restricts the generalizability of study findings.

Implications

Future studies should examine our study results on a generalize sample across multiple time points since we only analyzed the cross-sectional data from a community-based population. A consensus on the selection of different versions should be achieved by researchers and clinicians. Measurement of neck-related function seems to be a delightful direction for future research considering the disagreement found within three versions.

Conclusion

The traditional 10-item version should not be used interchangeably with other two Rasch-approved versions. Different coverage of content should be considered during the decision of 8- and 5-item NDI.

Abbreviations

NDI: The neck disability index

B&A: Bland-Altman

PROM: patient-reported outcome measure

CTT: classical test theory

PET: problem elicitation technique

IRT: Item response theory

DIF: differential item functioning

PSI: person separation index

SD: standard deviation

SE: standard error

CI: confidence interval

SDres: SD of the residual

ICF: International Classification of Functioning, Disability and Health

Declarations

Ethics approval and consent to participate: McMaster REB #03-145 and HiREB #13-300. All participants provided written, signed consent, the consent form will be submitted upon request.

Consent publication: All authors consent publication.

Availability of data and materials: The datasets during and/or analysed during the current study available from the corresponding author on reasonable request.

Competing interests: The authors report no conflicts of interest.

Funding: The authors report no funding.

Author's contribution: ZL and JM conducted the systematic electronic searches independently in each database. The same investigators then proceeded to identify and remove the duplicate studies. In the next stage, we independently screened the titles and abstracts and obtained all full-text articles marked as "include" or "uncertain". In the final stage, the same two reviewers independently performed the full

text reviews to assess final article eligibility. In case of disagreement, a third reviewer (GN), facilitated a consensus through discussion. The first author (ZL) performed the data analysis. ZL wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgements: JM was supported by a CIHR Chair in Gender, Work and Health and the Dr. James Roth Research Chair in Musculoskeletal Measurement and Knowledge Translation.

References

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research 1999;*, *8*, 135–160.

Bland, J. M., & Altman, D. G. (2010). Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies*, *47*(8), 931–936. https://doi.org/10.1016/j.ijnurstu.2009.10.001

Bobos, P., Macdermid, J. C., Walton, D. M., Gross, A., & Santaguida, P. L. (2018). Patient-reported outcome measures used for neck disorders: An overview of systematic reviews. *Journal of Orthopaedic and Sports Physical Therapy*, *48*(10), 775–788. https://doi.org/10.2519/jospt.2018.8131

Bovim, G., Schrader, H., & Sand, T. (1994). Neck pain in the general population. *Spine*. https://doi.org/10.1097/00007632-199406000-00001

Cano, S. J., Barrett, L. E., Zajicek, J. P., & Hobart, J. C. (2011). Beyond the reach of traditional analyses: Using Rasch to evaluate the DASH in people with multiple sclerosis. *Multiple Sclerosis Journal*, *17*(2), 214–222. https://doi.org/10.1177/1352458510385269

Covic, T., Pallant, J. F., Conaghan, P. G., & Tennant, A. (2007). A longitudinal evaluation of the Center for Epidemiologic Studies-Depression scale (CES-D) in a Rheumatoid Arthritis Population using Rasch Analysis. *Health and Quality of Life Outcomes*, *5*, 1–8. https://doi.org/10.1186/1477-7525-5-41

Croft, P. R., Lewis, M., Papageorgiou, A. C., Thomas, E., Jayson, M. I. V., Macfarlane, G. J., & Silman, A. J. (2001). Risk factors for neck pain: A longitudinal study in the general population. *Pain*, *93*(3), 317–325. https://doi.org/10.1016/S0304-3959(01)00334-7

Evans, R., Bronfort, G., Schulz, C., Maiers, M., Bracha, Y., Svendsen, K., … Transfeldt, E. (2012). Supervised exercise with and without spinal manipulation performs similarly and better than home exercise for chronic neck pain: A randomized controlled trial. *Spine*, *37*(11), 903–914. https://doi.org/10.1097/BRS.0b013e31823b3bdf

Gabel, C. P., Cuesta-Vargas, A. I., Osborne, J. W., Burkett, B., & Melloh, M. (2014). Confirmatory factory analysis of the Neck Disability Index in a general problematic neck population indicates a one-factor

model. *Spine Journal*, *14*(8), 1410–1416. https://doi.org/10.1016/j.spinee.2013.08.026

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia Medica*, *25*(2), 141–151. https://doi.org/10.1201/b16720-35

Hogg-Johnson, S., Van Der Velde, G., Carroll, L. J., Holm, L. W., Cassidy, J. D., Guzman, J., … Peloso, P. (2008). The Burden and Determinants of Neck Pain in the General Population Results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders Task Force on Neck Pain and Its Associated Disorders per-formed a systematic search and . *Eur Spine J*, *33*(1), 17. https://doi.org/10.1007/s00586-008-624-y

Hoving, J. L., O' Leary, E., Niere, K., Sally, G., & Buchbinder, R. (2003). Validity of the neck disability index, Northwick Park neck pain questionnaire, and problem elicitation technique for measuring disability associated with whiplash-associated disorders. *International Association for the Study of Pain*, *102*, 273–281. https://doi.org/doi:10.1016/S0304-3959(02)00406-2

Hung, M., Cheng, C., Hon, S. D., Franklin, J. D., Lawrence, B. D., Neese, A., … Brodke, D. S. (2015). Challenging the norm: Further psychometric investigation of the neck disability index. *Spine Journal*, *15*(11), 2440–2445. https://doi.org/10.1016/j.spinee.2014.03.027

Iyer, S., Koltsov, J. C. B., Steinhaus, M., Ross, T., Stein, D., Yang, J., … Kim, H. J. (2019). A Prospective, Psychometric Validation of National Institutes of Health Patient-Reported Outcomes Measurement Information System Physical Function, Pain Interference, and Upper Extremity Computer Adaptive Testing in Cervical Spine Patients: Successes and . *Spine*, *44*(22), 1539–1549. https://doi.org/10.1097/BRS.0000000000003133

Jerosch-Herold, C., Chester, R., Shepstone, L., Vincent, J. I., & MacDermid, J. C. (2018). An evaluation of the structural validity of the shoulder pain and disability index (SPADI) using the Rasch model. *Quality of Life Research*, *27*(2), 389–400. https://doi.org/10.1007/s11136-017-1746-7

MacDermid, J. C., Walton, D. M., Avery, S., Blanchard, A., Etruw, E., McAlpine, C., & Goldsmith, C. H. (2009). Measurement properties of the neck disability index: a systematic review. *The Journal of Orthopaedic and Sports Physical Therapy*, *39*(5), 400–417. https://doi.org/10.2519/jospt.2009.2930

Nazari, G., MacDermid, J. C., Sinden, K. E., Richardson, J., & Tang, A. (2019). Inter-Instrument Reliability and Agreement of Fitbit Charge Measurements of Heart Rate and Activity at Rest, during the Modified Canadian Aerobic Fitness Test, and in Recovery. *Physiotherapy Canada*, *71*(3), 197–206. https://doi.org/10.3138/ptc.2018-25

Packham, T., & Macdermid, J. C. (2013). Measurement properties of the Patient-Rated Wrist and Hand Evaluation: Rasch analysis of responses from a traumatic hand injury population. *Journal of Hand Therapy*, *26*(3), 216–224. https://doi.org/10.1016/j.jht.2012.12.006

Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, *46*(1), 1–18. https://doi.org/10.1348/014466506X96931

Smith, A. B., Rush, R., Fallowfield, L. J., Velikova, G., & Sharpe, M. (2008). Rasch fit statistics and sample size considerations for polytomous data. *BMC Medical Research Methodology*, *8*, 1–11. https://doi.org/10.1186/1471-2288-8-33

Van Der Velde, G., Beaton, D., Hogg-Johnston, S., Hurwitz, E., & Tennant, A. (2009). Rasch analysis provides new insights into the measurement properties of the neck disability index. *Arthritis Care and Research*, *61*(4), 544–551. https://doi.org/10.1002/art.24399

Vernon, H. (2008). The Neck Disability Index: State-of-the-Art, 1991-2008. *Journal of Manipulative and Physiological Therapeutics*, *31*(7), 491–502. https://doi.org/10.1016/j.jmpt.2008.08.006

Walton, D. M., & MacDermid, J. C. (2013). A brief 5-item version of the Neck Disability Index shows good psychometric properties. *Health and Quality of Life Outcomes*, *11*(1), 5–13. https://doi.org/10.1186/1477-7525-11-108

WHO. (2002). Towards a Common Language for Functioning , Disability and Health ICF. *Who*, *1149*, 1–22. https://doi.org/WHO/EIP/GPE/CAS/01.3

Wlodyka-Demaille, S., Poiraudeau, S., Catanzariti, J. F., Rannou, F., Fermanian, J., & Revel, M. (2004). The ability to change of three questionnaires for neck pain. *Joint Bone Spine*, *71*(4), 317–326. https://doi.org/10.1016/j.jbspin.2003.04.004

Tables

Table 1 Demographic characteristic of the sample

	Male (N=43)		Female (N=158)	
	Mean (SD)	Range, min-max	Mean (SD)	Range, min-max
Age, y	49.2 (12.2)	19 - 74	45.7 (12.8)	19 - 74
Pain intensity	2.0 (1.4)	0 -5	2.1 (1.2)	0 - 5
NDI 10-item	14.6 (10.7)	2 – 44	17.0 (9.8)	0 – 41
NDI 8 item	17.8 (7.0)	0 – 33	19.0 (6.0)	0 – 31.5
NDI 5 item	20.1 (8.3)	0 - 35	21.2 (7.6)	0 - 33

SD: standard deviation
Min: minimum
Max: maximum

Table 2 Individual and average agreement of three comparisons

Comparison	Sample size	Degree of freedom	Individual agreement				Average agreement
			Mean of difference (d)* (95% CI)	SD of difference	Upper d+1.96SD	Lower d-1.96SD	SE
10-item vs. 8-item	201	200	-2.3 (-4.6 - 0)	5.0	7.4	-12.0	0.4
10-item vs. 5-item	201	200	-4.6 (-9.1 - 0)	5.3	5.8	-14.9	0.5
8-item vs. 5-item	201	200	-2.3 (-4.6 - 0)	2.8	3.3	-7.8	0.2

*p< 0.05
SD: standard deviation
SE: standard error base on the mean of difference
10-item: The total score of NDI 10-item (original) version on ordinal scale with maximum of 50 points.
8-item: The total score of NDI 8-item version on linear scale with maximum of 50 points.
5-item: The total score of NDI 5-item version on linear scale with maximum of 50 points.
SD: standard deviation
SE: standard error

Figures

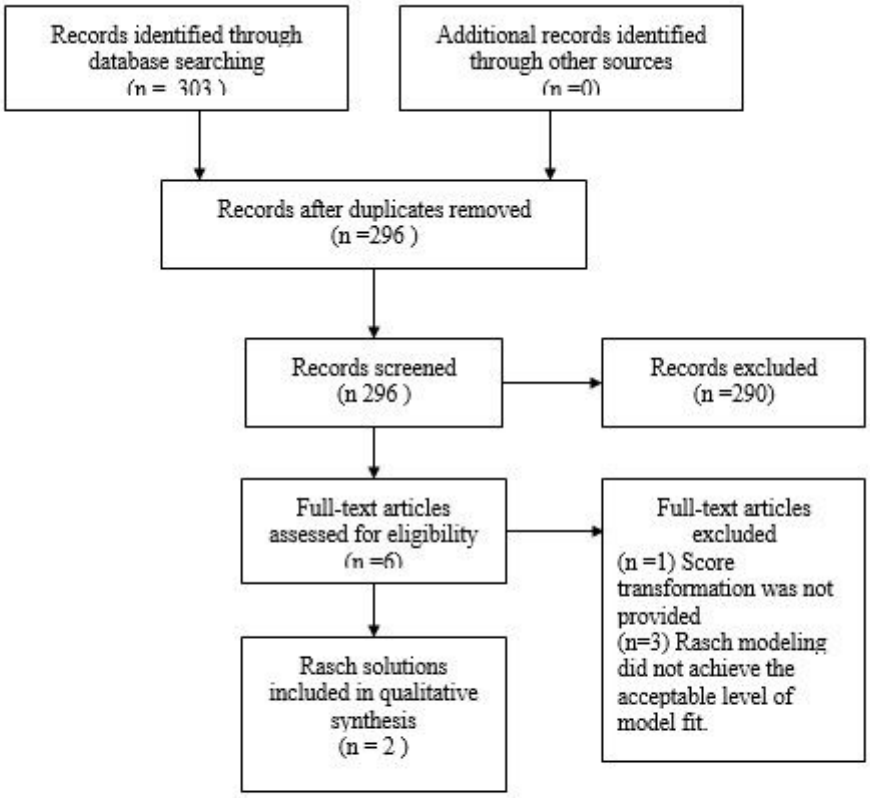

Figure 1

Flow Diagram of study selection results based on PRISMA guideline.

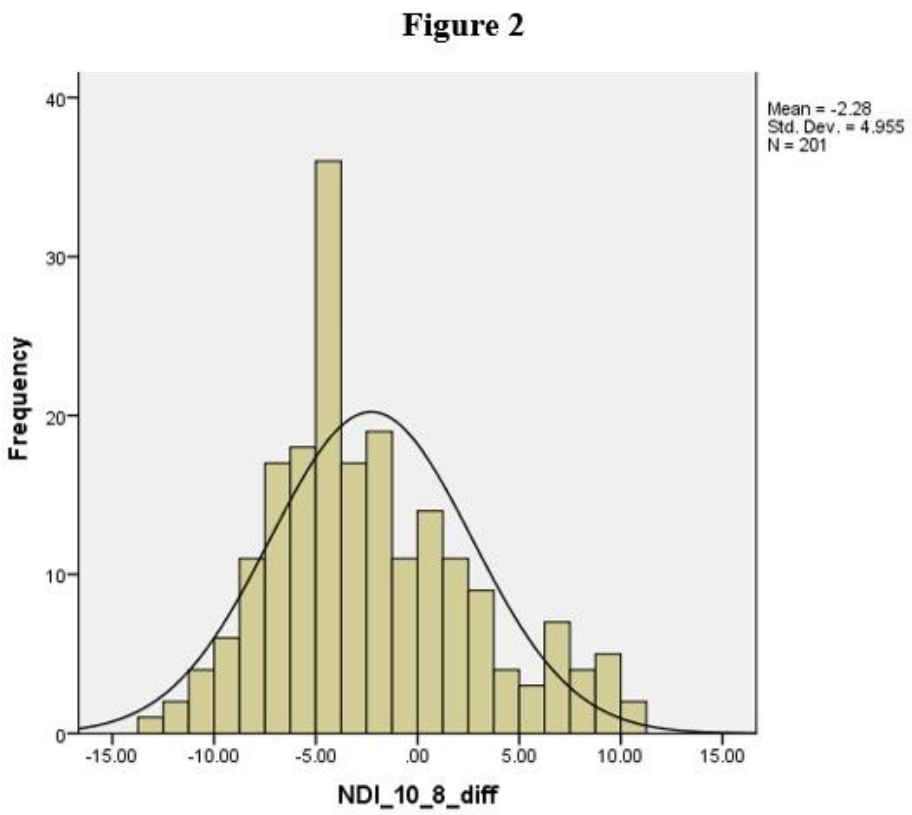

Figure 2

Histogram of the difference comparing NDI 10-item total score with NDI 8-item total score. NDI: neck disability index.

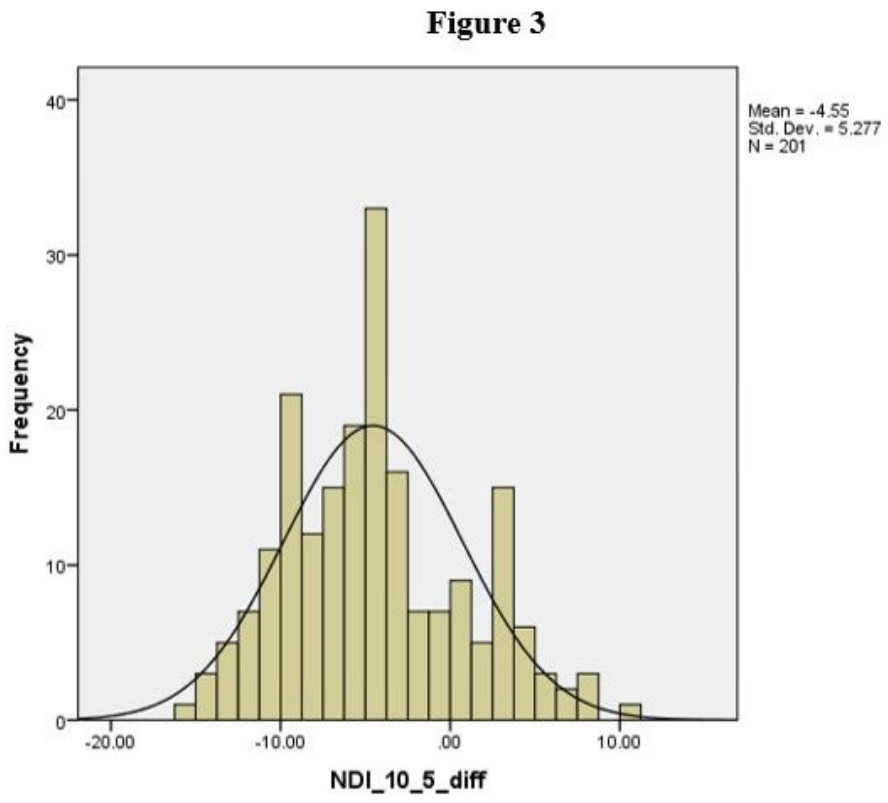

Figure 3

Histogram of the difference comparing NDI 10-item total score with NDI 5-item total score. NDI: neck disability index.

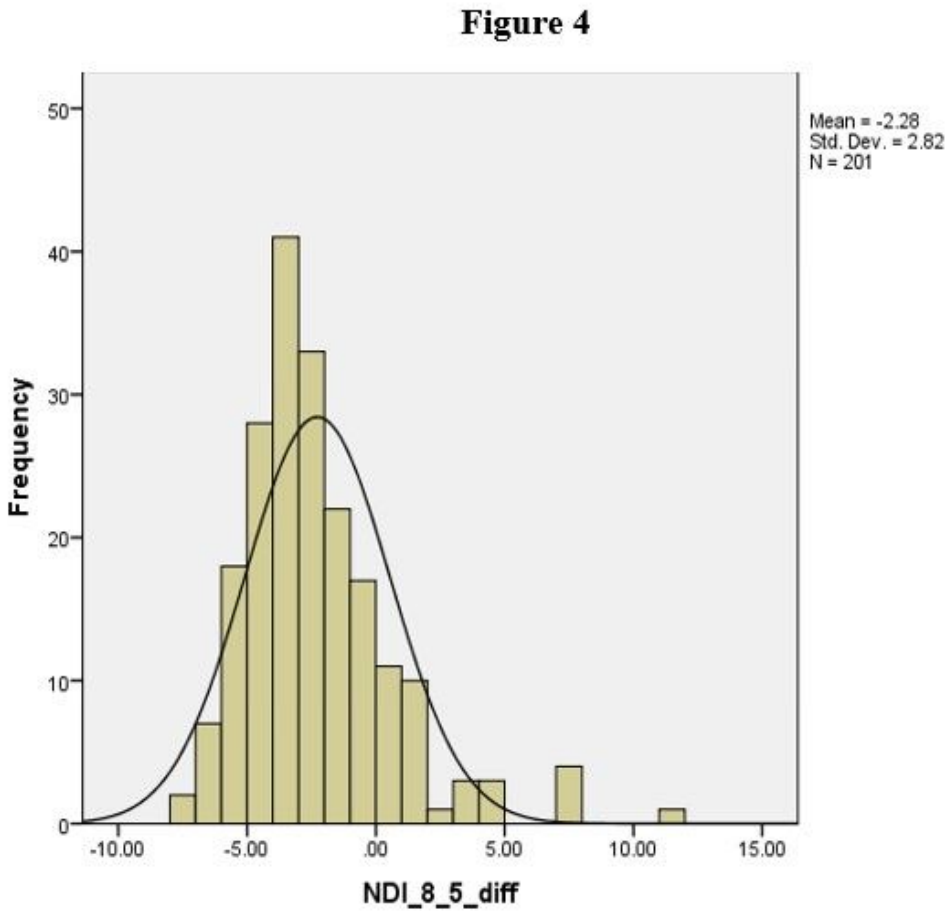

Figure 4

Histogram of the difference comparing NDI 8-item total score with NDI 5-item total score. NDI: neck disability index.

Figure 5

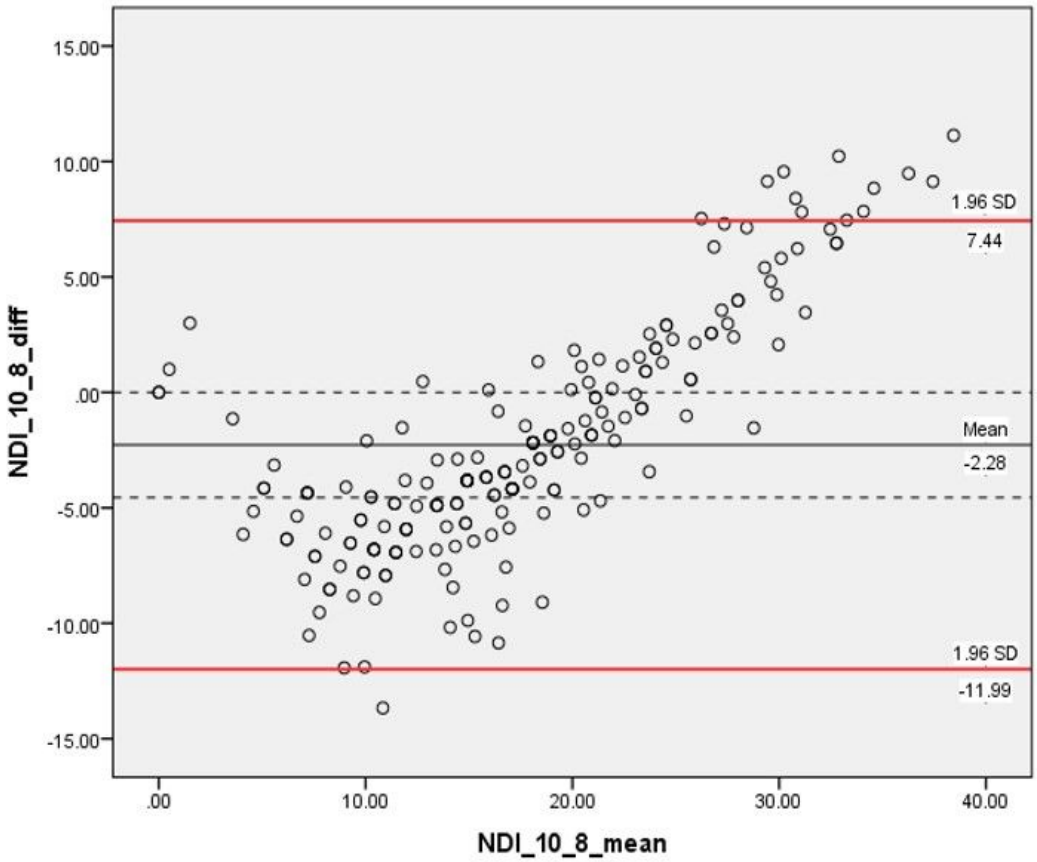

Figure 5

Bland–Altman plots displaying 95% LoA in pair-wise comparison between NDI 10-item with NDI 8-item version. LoA: limits of agreement. NDI: neck disability index.

Figure 6

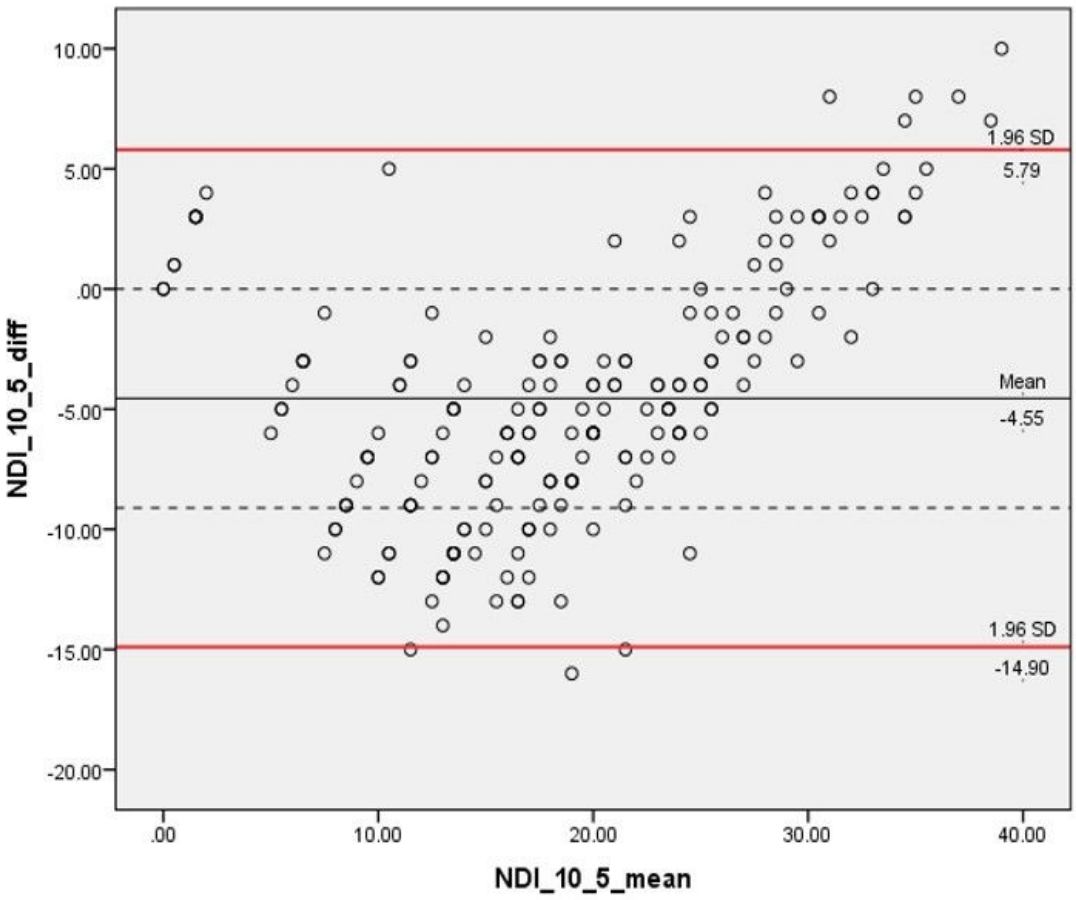

Figure 6

Bland–Altman plots displaying 95% LoA in pair-wise comparison between NDI 10-item with NDI 5-item version. LoA: limits of agreement. NDI: neck disability index.

Figure 7

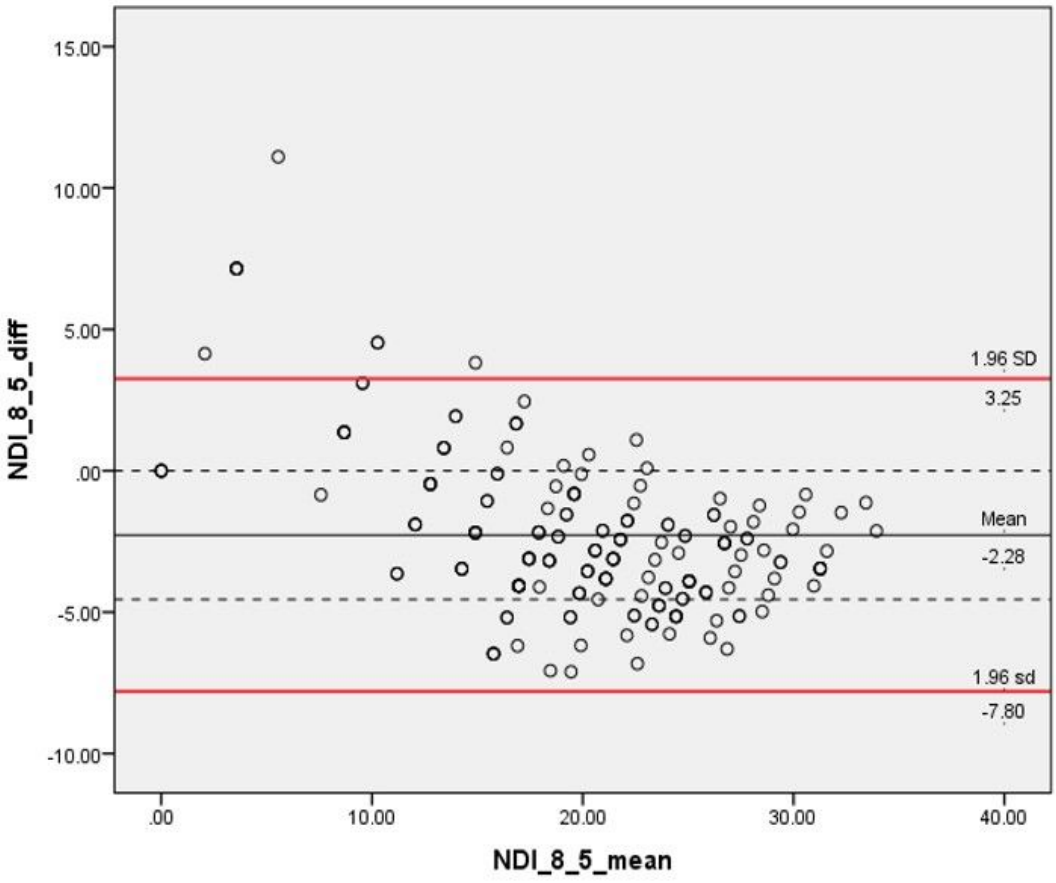

Figure 7

Bland–Altman plots displaying 95% LoA in pair-wise comparison between NDI 8-item with NDI 5-item version. LoA: limits of agreement. NDI: neck disability index.

Figure 8

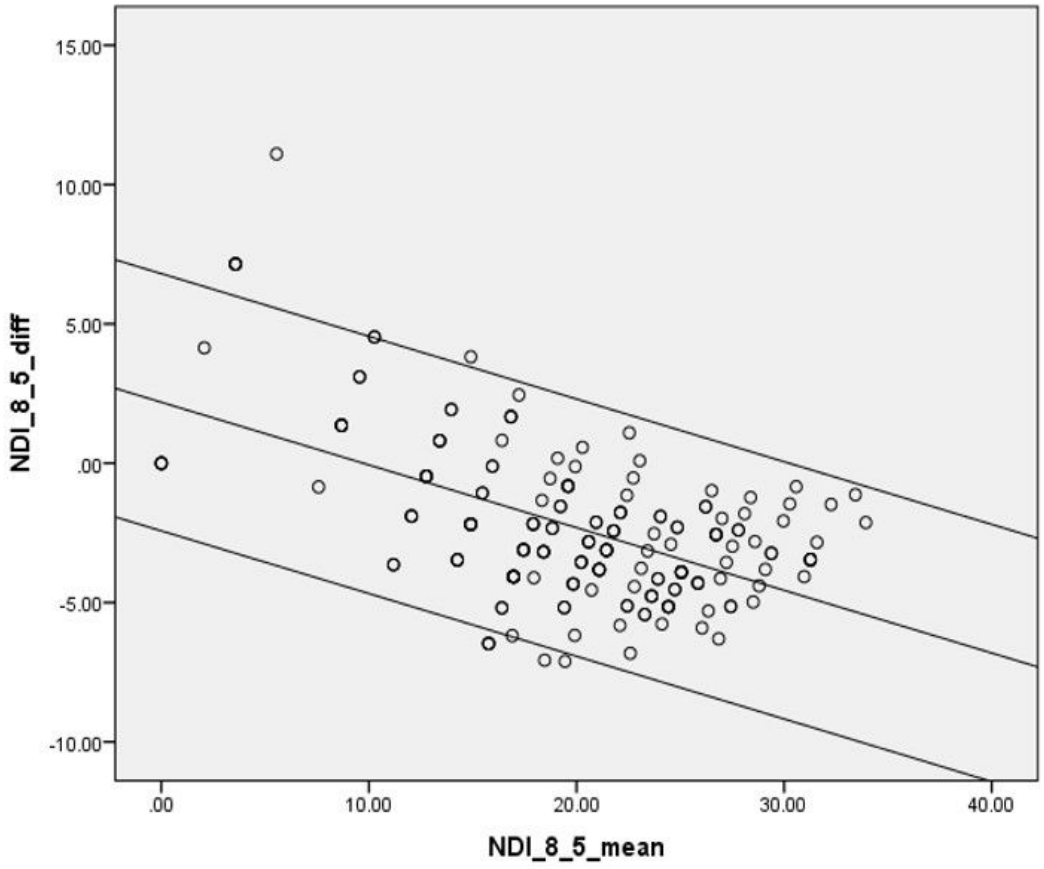

Figure 8

Bland–Altman plots displaying 95% LoA in regression between NDI 8-item with NDI 5-item version as this varies across the range of the scores. LoA: limits of agreement. NDI: neck disability index.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PRISMAchecklist.doc
- Appendix1ZeLu.docx
- Appendix2ZeLu.docx