**NIH Quality Assessment Tool for Before-After (Pre-Post) Studies With No Control Group**

*Note: the tool has not been standardised and independently published. Researchers have to determine their own parameters for making judgements.*

*CD, cannot determine; NA, not applicable; NR, not reported*

**Di Consiglio, 2021**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Criteria** | **Yes** | **No** | **CD, NA, NR** | **Argumentation/evidence** |
| 1.Was the study question or objective clearly stated? | X |  |  | “The effectiveness of NoiBene on students’ well-being, emotional awareness, emotion regulation, assertiveness, and transdiagnostic factors (perfectionism, repetitive thinking and experiential avoidance) was investigated”. |
| 2.Were eligibility/selection criteria for the study population prespecified and clearly described? |  | X |  | There was no pre-registration available, but criteria were well-described:  “Then, they had to complete a ‘Preselection’ module. In this module, the student had to answer the MCMI-III and SCL-90-R questionnaire. This step allowed the early detection of any relevant conditions. If clinical relevance was suggested, the student was contacted for a second follow-up interview and directed towards a treatment more suited to his or her needs. In the latter case, NoiBene was used as a support of their therapy, but they were excluded from the study”. |
| 3.Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest? | X |  |  |  |
| 4.Were all eligible participants that met the prespecified entry criteria enrolled? | X |  |  | “If the screening did not suggest any psychopathology, students carried out the NoiBene programme (see Table S3 for further details about the programme). At the beginning and the end of the programme, students filled out the questionnaires indicated above.” |
| 5.Was the sample size sufficiently large to provide confidence in the findings? |  |  | NR | No formal power calculation and target was pre-specified to determine what sample size would be sufficiently large:  “A total of 178 students completed the first screening, but only 154 students started the programme; 59 students completed it, 40 students dropped out, and 55 students are still working on the programme”. |
| 6.Was the test/service/intervention clearly described and delivered consistently across the study population? | X |  |  | See table S3 – fully automated digital delivery across all participants.  Only difference is that some participants appear to have taken longer completing the program. |
| 7.Were the outcome measures prespecified, clearly defined, valid, reliable, and assessed consistently across all study participants? |  | X | CD | No pre-registration is available but validated standardised questionnaire used across all participants:  “At the beginning and end of the intervention, students completed a series of standardised questionnaires to assess NoiBene efficacy. And in study 1, we used the PWB to assess psychological well-being … (see the Appendix S2 for further details about questionnaires).” |
| 8.Were the people assessing the outcomes blinded to the participants' exposures/interventions? |  | X |  | No participants assessed the outcomes through self-report and were aware they received the Noibene intervention. |
| 9.Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis? |  | X |  | 59 out of 154 students completed the intervention at post-measures.  Analysis did not account for this:  “In the group of students who completed the programme (n = 59), a series of paired-samples t-test was conducted to compare the selected measures at baseline and post-intervention. As shown in Table 4, the analysis revealed significant differences in different dimensions of psychological well-being”. |
| 10.Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes? | X |  |  | “In the group of students who completed the programme (n = 59), a series of paired-samples t-test was conducted to compare the selected measures at baseline and post-intervention. As shown in Table 4, the analysis revealed significant differences in different dimensions of psychological well-being”.  Table 4 provides p-values. |
| 11.Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention (i.e., did they use an interrupted time-series design)? |  | X | NA |  |
| 12.If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis take into account the use of individual-level data to determine effects at the group level? |  |  | NA |  |

|  |  |
| --- | --- |
| Overall quality rating: | High risk of bias / low quality |
| Reasoning: | High attrition, no pre-registration, no blinding of participants/outcome assessors, no sample size calculation. |