Peer Review Report

Review Report on Predicting Low Cognitive Ability at Age 5 – Feature Selection Using Machine Learning Methods and Birth Cohort Data
Original Article, Int J Public Health

EVALUATION

Q 1   **Please summarize the main findings of the study.**

This is certainly an interesting study. Applying machine learning methods for predicting intelligence in age 5.

Q 2   **Please highlight the limitations and strengths.**

This is certainly an interesting study. Applying machine learning methods for predicting intelligence in age 5.

Q 3   **Please provide your detailed review report to the authors. The editors prefer to receive your review structured in major and minor comments. Please consider in your review the methods (statistical methods valid and correctly applied (e.g. sample size, choice of test), is the study replicable based on the method description?), results, data interpretation and references. If there are any objective errors, or if the conclusions are not supported, you should detail your concerns.**

This is certainly an interesting study. Applying machine learning methods for predicting intelligence in age 5.

General:
The referencing style applied by the International Journal of Public Health is very impractical (if an author adds or deletes one reference the entire numbering has to be changed). This leads to errors, wastes a lot of valuable time and makes it more difficult for readers to understand where the ideas come from. It would be better if International Journal of Public Health switches to APA style.

Abstract:
A multiple correlation R should be added – how good is the final model in predicting (low vs. high) intelligence in age 5. Of course, a correlation for a continuous dependent variable can be contrasted to the correlation for a dichotomous dependent variable.

Introduction:
The statistical approach of machine learning is certainly innovative, but it says nothing about the underlying causal factors that produce these associations. Specifically: Genetic factors must be mentioned. E.g.:
Plomin, R. (2018). Blueprint: How DNA makes us who we are. London: Penguin/Allen Lane.

Method:
Add average IQs of the groups.
Is there any information on the sample regarding migration background or race/ethnicity?

Results:
Add a correlation matrix between all relevant variables.
For the importance of parental education in predicting children's cognitive performance levels, see the multi-country study:

Rindermann, H. & Ceci, S. J. (2018). Parents' education is more important than their wealth in shaping their children's intelligence: Results of 19 samples in seven countries at different developmental levels. Journal for the Education of the Gifted, 41(4), 298-326.
or
Lemos, G. C., Almeida, L. S. & Colom, R. (2011). Intelligence of adolescents is related to their parents' educational level but not to family income. Personality and Individual Differences, 50(7), 1062-1067.
Table 1: It is not possible to evaluate the size of the effects. Add standardized effect sizes, best a bivariate correlation and a standardized beta coefficient, maybe a standardized d effect size.

PLEASE COMMENT

Q 4 — **Is the title appropriate, concise, attractive?**

yes

Q 5 — **Are the keywords appropriate?**

ok

Q 6 — **Is the English language of sufficient quality?**

yes

Q 7 — **Is the quality of the figures and tables satisfactory?**

Yes.

Q 8 — **Does the reference list cover the relevant literature adequately and in an unbiased manner?)**

mainly yes

QUALITY ASSESSMENT

Question		Rating (of 5)
Q 9	Originality	4
Q 10	Rigor	3
Q 11	Significance to the field	3
Q 12	Interest to a general audience	4
Q 13	Quality of the writing	4
Q 14	Overall scientific quality of the study	4

REVISION LEVEL

Q 15 — **Please make a recommendation based on your comments:**

Minor revisions.