Reviewer's report

Title: Early Identification of Young Children at Risk for Poor Academic Achievement: Preliminary Development of a Parent-Report Prediction Tool

Version: 1 Date: 18 March 2011

Reviewer: Mary Eamon

Reviewer's report:

1. Discretionary Revisions

Perhaps some clarification could be made in both the abstract and the first sentence of the “Background section” (and citations appear to be needed for the latter) regarding the relationship between early school success and later health. That is, to what “health” outcomes are the authors referring?

Perhaps all of the citations in the “Initial Item Pool Creation” section are unnecessary, as they are presented in Table 1.

It’s very confusing to be referred to a Table 4 and then provided a Table 5 (in an Appendix A), which occur before the introduction of the previously numbered tables.

In the text, I think that the authors also should introduce Table 1 before Table 2, and also briefly explain what the table contains.

The authors combined family structure and conflict about child rearing in a way that I have not seen before. Perhaps some justification beyond stating that conflict does not apply to single-parent families could be given, or a citation could be provided where such variables were used in previous research.

2. Minor Essential Revisions

In the first paragraph, fifth line and sixth lines, the term “medical” in the phrase “the medical home” seems to need some clarification.

There are spacing, punctuation, and citation problems. For example, a space should be placed before the first parenthesis for the citation; the punctuation mark (comma, period) should be placed after the final parentheses for the citation, not before the first one; and the author’s name in the citation should not be repeated if it was placed in the sentence (e.g., “subsequently redefined in the work of Zaslow et al. [Zaslow, et al., 2000]).

The last line under “Initial Item Pool Creation” states that “we then categorized predictors as a child characteristic, family characteristic or home environment variable.” But, it appears that “neighborhood environment” was also used as a category (at least on Table 2).
In the “Data Source” section (8th) line, the authors referred to the mothers surveyed in the NLSY as “baby boomer mothers.” These females were sampled in 1978 when they were between the ages of 14 and 21. That would mean they were born between 1957 and 1964, which to my knowledge are not referred to as “baby boomers.”

In the second paragraph under “Predictor Variables” (first sentence), the authors state that all but one of the 32 NLSY variables “relied solely on parental report” (Table 2). But according to Table 2, four of the variables under the HOME Emotional Subscale relied on interviewer observation.

The authors state in Appendix B that they “created four domain-specific multivariable regression models” composed of “socio-demographic and neighborhood characteristics, family and child care environment, child factors, and health care receipt.” But, the labels for the groups of variables are inconsistent with these four domains in Table 2 and Table 3, and the labels are inconsistent between the two tables. This is very confusing. In addition, in Table 3, the first label is “SOCIODEMOGRAPHICS & NEIGHBORHOOD CHARACTERISTICS,” but no neighborhood characteristics are presented.

As one of their limitations, I think the authors should acknowledge that their analysis cannot establish causation between the independent variables and the three PIAT scores.

3. Major Compulsory Revisions

Under the “Methods Section” the reader is referred to Figure 1. However, I did not download such a figure. Does it exist?

In Appendix A, the rationale for the procedures used to rank these predictors are not clear; for example, is this a standard way of comparing results across studies? If so, the authors need to cite an appropriate reference. I also don’t understand how coefficients for specific variables can be compared across studies if effect sizes are not calculated. Did the authors calculate effect sizes from the coefficients for all of these independent variables across these studies?

In the second paragraph under “Predictor Variables,” the authors listed the HOME emotional subscale as one of the instruments from which items were taken; but weren’t items also taken from the HOME cognitive stimulation subscale (e.g., number of books child has of his/her own, frequency of reading to child; hours watched TV on average school day)? And, if these didn’t come from that scale, why were not items from this scale used? I would think that they would be more relevant than items from the HOME emotional subscale when predicting achievement test scores. Finally, the authors might note that the HOME-Short Form is used in the NLSY.

The authors need to explain why n = 2919 in Table 2 when the descriptives are provided, but range from n=2200 to n=1384 in Table 3. Why the discrepancy? In addition, why aren’t the descriptive results weighted? Given my knowledge of the
NLSY, I believe that they should be.

In the Results section, on what basis do the authors conclude that the child/family sociodemographic variables remained strongly associated with the PIAT scores? These coefficients do not appear to be standardized nor are effect sizes presented; therefore, the authors need to clarify on what basis they determined the strength of these relationships.

Finally, although I can see how this “index” might identify children at risk of lower academic achievement, it is unclear how many of the variables in this index could assist in “customizing the delivery of preventive care services.” This is because very little can be done about many of the factors found to be related to PIAT scores (i.e., child’s gender, number of children in the household, maternal age at child’s birth, maternal race/ethnicity, family structure, maternal educational attainment). The authors should expand upon how they believe this index can do more than assist in identifying children at risk for lower academic achievement.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.