Author's response to reviews

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

Authors:

Susmita Pati (susmita.pati@stonybrook.edu)
Kyleen Hashim (kyleen.hashim@gmail.com)
Brett Brown (bbrown@wrma.com)
Alexander G Fiks (fiks@email.chop.edu)
Christopher B Forrest (forrestc@email.chop.edu)

Version: 2 Date: 20 June 2011

Author's response to reviews: see over
June 20, 2011

BMC Health Services Research
BioMed Central Ltd.
Floor 6, 236 Gray’s Inn Road
London, WC1X 8HB
United Kingdom

Dear Drs. Norton and Galbraith,

We are pleased to submit the revised manuscript entitled “Early Identification of Young Children at Risk for Poor Academic Achievement: Preliminary Development of a Parent-Report Prediction Tool” to BMC Health Services Research.

We have paid careful attention to the reviewer’s and associate editor’s comments and revised the manuscript substantially as detailed in the ensuing pages. As per journal guidelines, all changes in the manuscript have been highlighted in yellow. However, for ease of reading, most changes in tables have been marked with the “comments” feature of Microsoft Word.

The primary goal of this project was to create a prediction tool using information from parents that is predictive of early school success (or failure, operationalized as academic achievement) as the first step in a research program whose overarching goal is to tailor pediatric health supervision to individual needs. In this study, we describe the preliminary development and validation of an early school success prediction tool ultimately designed for use in the pediatric office setting. The tool uses only parental report about 12 child, family, and home environment characteristics of two year-old children and we assess its predictive association with school success at ages 6-7 years.

Given the current movement to strengthen developmental assessments within pediatric health supervision, we believe this work is important, relevant, and timely. Findings from a structured literature review that informed this project were published as a project report by Child Trends. However, the results of our data analyses have not been published nor are being considered for publication elsewhere. This manuscript is being submitted only to BMC Health Services Research; it will not be submitted elsewhere while under consideration, it has not been published elsewhere, and, should it be published in BMC Health Services Research, and it will not be published elsewhere--either in similar form or verbatim--without permission of the editors.

Dr. Pati will handle all correspondence regarding this manuscript.
Susmita Pati, MD, MPH
Associate Professor and Chief, Division of Primary Care Pediatrics
State University of New York at Stony Brook & Stony Brook Long Island Children’s Hospital
Health Sciences Center T11 020
Stony Brook, NY 11794-8111
Phone: 631-444-2709, Fax: 631-444-7865
Email: susmita.pati@stonybrook.edu

We are solely responsible for the conception and design of this work, analysis of the data, and preparation of the manuscript to take public responsibility for it. We believe the manuscript represents original, important, and valid work. We certify that there are no affiliations or involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the manuscript. This project was reviewed and approved by the
Institutional Review Board at The Children’s Hospital of Philadelphia and the State University of New York at Stony Brook.

Thank you for your consideration,
Susmita Pati, MD, MPH
Kyleen Hashim
Brett Brown, PhD
Alexander G. Fiks, MD, MSCE
Christopher B. Forrest, MD, PhD
Response to Reviewer and Editor Comments

Associate Editor comments

In particular, please address the reviewer comments about what this study adds the existing literature on risk factors for poor academic achievement, and address reviewer questions about how clinicians could elicit information about these risk factors and how this information could be used to do more than just identify high-risk children.

Response: We appreciate the editor’s comments and have substantially revised the discussion section of the manuscript to address these points. Please see also responses to all of Reviewer 1’s comments and number 19 of Reviewer 2’s comments.

Version: 1 Date: 7 March 2011 Reviewer: Pamela Davis-Kean Reviewer's report:

Major Compulsory Revisions

1. This manuscript attempts to put together a risk inventory that physicians can use to determine achievement as early as 2 years of age. The authors did a good job reviewing the literature except for referencing work by Arnold Sameroff who has written extensively on the additive risk factor of children and families.

2. Thus, these risk have already been defined in the psychological literature. It is well known that demographic variables are important for identifying risk. Indeed, poverty, is a major risk factor for both achievement and health. I would suggest that many pediatricians already have a sense of these risks when meeting with parents....Thus, the question for the authors is what is the contribution of this article and risk portfolio about that which already known about risk.

3. Also, from multiple experiences having asked these questions, income is not a variable that most parents are comfortable discussing with anyone. How would pediatricians or other clinicians approach these questions?

4. An analysis that was more sensitive to issues of parenting and home environment may be more informative.

Response: We appreciate the reviewer's thoughtful comments and have enumerated our responses to each item in the ensuing list.

1. In our literature review, we now reference Sameroff's work.

2. Though we agree that many pediatricians do have a sense of these risks when interacting with families, there currently is no empiric way for pediatricians to sort children into different groups of risk (e.g., high, medium, low) for early academic problem. In addition, the aggregate effect of risk and resilience factors on early academic achievement has not been fully explored. Though this work does not identify any new factors associated with early academic achievement, the contribution of this work is the application of these known relationships to clinical practice; that is, this work begins the process of creating a preliminary version of an
evidence-based instrument that aggregates multiple factors that can specifically be obtained from parental report into an index that could be used to sort children into different groups of risk for early academic achievement.

3. As the reviewer points out, income is a difficult question to approach with parents. In our field-testing, we are comparing the predictive value of insurance type (i.e. public, private, self-pay) to granular income information to determine the relative need for detailed income information in order to reasonably sort children into different groups of risk for early academic problems. We have revised the manuscript (discussion) to clarify these issues.

4. Though we were unable to fully explore issues of parenting and home environment in this work due to the limitations of the NLSY dataset, we have now field-tested a revised and expanded version of this instrument in a cohort of 2100 urban children and families that includes multiple items from these domains. We have added a statement about this limitation in the penultimate paragraph of discussion.

Minor Essential Revisions
5. Using data dummies for missing data is not the state of the art of missing data. The authors should use multiple imputations or FIML adjustments found in structural equation modeling programs to deal more effectively with the issue of missing data.

Response: Though we agree that multiple imputations or FIML adjustments are effective state-of-the-art methods for handling missing data, we chose to use data dummies in this analysis because these methods cannot be easily implemented in the office setting for scoring purposes whereas data dummies (with mean/median substitution) can be. We now include a statement about this rationale in Appendix B.

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:
Discretionary Revisions
1. 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?

Response: We appreciate the reviewer’s comments and have clarified and referenced the relationship between early school success and later health outcomes that have been documented in the literature.

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

Response: We appreciate the reviewer’s suggestion and have carefully reviewed
and removed some of the citations in the Initial Item Pool Creation section and retained all of these in Table 1.

3. 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.

**Response:** We appreciate the reviewer’s concern and have renumbered the tables for clarity.

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

**Response:** We appreciate the reviewer’s comments. Table 1 is now introduced in the “Initial Item Pool Creation” section of the methods followed by Table 2 introduced in the “Predictor Variable” section of the methods.

5. 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.

**Response:** We appreciate the reviewers’ comment. In the NLSY, the conflict item was not asked of single parents. However, both the conflict and marital status appeared important in bivariate and multivariate regressions by topic area. In creating the final model, these two variables were combined in this way to ensure that the effects of both could be determined. We have clarified this in the methods (section entitled “predictor variables”) and Appendix B.

**Minor Essential Revisions**

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

**Response:** We appreciate this concern and, due to space concerns, have referenced the Agency for Healthcare and Research Quality’s website that defines the “medical home” as a model of the organization of primary care that delivers the core functions of primary health care: accessible, continuous, comprehensive and coordinated, patient-centered care that are managed centrally by a primary care physician with the active involvement of non-physician staff.

7. 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]).

**Response:** We appreciate these concerns and have corrected these spacing, punctuation, and citation issues. However, for ease of reading, we have not highlighted these changes in the text.
8. 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).

Response: We appreciate the reviewer’s comment and have revised the “Initial Item Pool Creation” section to include the neighborhood environment domain.

9. 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.”

Response: We appreciate the reviewer’s comment. Because the NLSY documents and press releases (see below) refer to these respondents as “baby boomers,” we have retained this descriptor in the manuscript.

“You are a NLSY79 respondent. You were born in the years 1957 to 1964, the latter years of the “baby boom” that occurred in the United States from 1946 to 1964.” [https://nlsy79.norc.org/interestingFacts.html]

The NLSY press releases also refer to them as the youngest baby boomers: [http://www.bls.gov/nls/nlsy79.htm](http://www.bls.gov/nls/nlsy79.htm) (see bottom of page)

10. 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.

Response: We appreciate the reviewer’s comment. We have corrected the number of NLSY variables (31 items/multi-item subscales) and clarified that the multi-item HOME emotional subscale relies on interviewer observation (methods section entitled “predictor variables”). Results reported in Table 2 show the distribution of the responses for these items for descriptive purposes. However, the final model shown in Table 3 does not include the HOME subscale items or score.

11. 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 and Neighborhood Characteristics,” but no neighborhood characteristics are presented.

Response: We appreciate the reviewer’s comment. Because there were no significant neighborhood or health care receipt predictors in our final predictive models, the remaining predictors were re-categorized from the original 4 domains (sociodemographic and neighborhood characteristics, family and child care environment, child factors, and health care receipt) into 3 domains (child characteristics, family characteristics, and home environment) for ease of translation into practice; this domain classification is utilized in Tables 2 and 3. We
have revised Tables 2 and 3 to use consistent terminology in referring to the
domains and included a statement about the domain reclassification in Appendix B.
For ease of reading, we have not highlighted these changes in the text.

12. 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.
Response: We agree with the reviewer’s comments and have revised our
limitations section accordingly.

Major Compulsory Revisions

14. Under the “Methods Section” the reader is referred to Figure 1. However, I did
not download such a figure. Does it exist?
Response: To the best of our knowledge, Figure 1 was uploaded with the original
submission and has been uploaded again with this revised version.

15. 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?
Response: There is not standard approach for comparing effect sizes across studies
that use different data-sets and different types of statistical models. Nonetheless, we
developed an approximate method that categorized predictors into a 3 level ordinal
scale based on the size of their beta coefficient as simple heuristic.

16. 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.
Response: We chose variables from the NLSY based on our findings from the
literature. As much as possible, we tried to use single parent-response items that
related to the literature, and only resorted to using multiple-item scales when no
other proxy was available. The literature consistently pointed to maternal warmth
and sensitivity as a potentially influential factor in early child development. Because
no single item seemed to capture this variable, we used the HOME emotional
subscale in this case. In contrast, the literature specifically identified “number of
books” and “reading to child” as potentially important stand-alone items. Since these
specific items were available in the NLSY, we pulled them from the HOME subscale.
We have revised the methods section entitled “predictor variables” to note that the
HOME-Short Form is used in the NLSY and that specific items were selected because the literature has specifically identified these factors as important stand-alone items.

17. 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.

**Response:** We appreciate the reviewer's comments and have clarified (in Methods section entitled “outcome variables” and with annotations to Tables 2 and 3) that the results in Table 3 were weighted results restricted by response rates for each of the three PIAT scores (Math, Reading Recognition, and Reading Comprehension) whereas Table 2 provides unweighted sample sizes and weighted results.

18. 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.

**Response:** One approach to determine the strength of these relationships is to calculate effect sizes. In this study, this can be done by dividing the OLS regression estimates by the standard deviation of the PIAT score. Because the standard deviation of the PIAT (15 as noted in the Methods section entitled “outcome variables”) does not vary within subscales (i.e. for math, reading, etc.), this division by a constant does not alter the relative strength of the relationships observed in the regression. In addition, because there are several child/family sociodemographic variables in this study where the absolute value of the regression estimate exceeds 2, it is reasonable to describe these relationships as strong.

19. 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.

**Response:** Though we agree that many of the factors in this preliminary index are immutable, there is strong evidence that, for children at risk for lower academic achievement, early intervention programs (e.g., Head Start) can optimize performance. Thus, this index can assist in further refinement of targeting and outreach efforts for these programs. In addition, for mutable factors (e.g., maternal depression, etc.), preventive care provided by the pediatrician can be augmented by referral to evidence-based programs known to be effective. We have revised our discussion (fourth paragraph) to discuss the importance of providing evidence-based customized pediatric health supervision and, using children with asthma as
an example, describe how clinicians routinely judge patients’ disease severity and other health risk factors to determine the mix and intensity of services that should be applied on an individual basis.