PLOS ONE

RESEARCH ARTICLE

A new model of wheezing severity in young children using the validated ISAAC wheezing module: A latent variable approach with validation in independent cohorts

Steven M. Brunwasser[1,2]*, Tebeb Gebretsadik[3], Diane R. Gold[4,5], Kedir N. Turi[1], Cosby A. Stone, Jr.[1], Soma Datta[5], James E. Gern[6], Tina V. Hartert[1]

1 Vanderbilt University Medical Center, Department of Medicine, Division of Allergy, Pulmonary, and Critical Care Medicine, T-1218 Medical Center North, Nashville, TN, United States of America, **2** Vanderbilt University Medical Center, Department of Obstetrics and Gynecology, B-1118 MCN, Nashville, TN, United States of America, **3** Vanderbilt University Medical Center, Department of Biostatistics, West End, Nashville, TN, United States of America, **4** Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Department of Environmental Health, Boston, MA, United States of America, **5** Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, United States of America, **6** University of Wisconsin-Madison, School of Medicine and Public Health, Department of Pediatrics, K4/918 CSC, Madison, WI, United States of America

* steven.brunwasser@vanderbilt.edu

Check for updates

OPEN ACCESS

Citation: Brunwasser SM, Gebretsadik T, Gold DR, Turi KN, Stone CA, Jr., Datta S, et al. (2018) A new model of wheezing severity in young children using the validated ISAAC wheezing module: A latent variable approach with validation in independent cohorts. PLoS ONE 13(4): e0194739. https://doi.org/10.1371/journal.pone.0194739

Editor: Yungling Leo Lee, National Taiwan University College of Public Health, TAIWAN

Received: January 6, 2018

Accepted: March 8, 2018

Published: April 17, 2018

Copyright: © 2018 Brunwasser et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data were obtained from the third party International Study of Asthma and Allergies in Childhood and are available at: http://isaac.auckland.ac.nz/phases/phaseone/results/resultsIndv.php. The authors confirm that they did not have any special access privileges to these data. All other relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by U19 AI 095227, National Institutes of Health/National

Abstract

Background

The International Study of Asthma and Allergies in Children (ISAAC) Wheezing Module is commonly used to characterize pediatric asthma in epidemiological studies, including nearly all airway cohorts participating in the Environmental Influences on Child Health Outcomes (ECHO) consortium. However, there is no consensus model for operationalizing wheezing severity with this instrument in explanatory research studies. Severity is typically measured using coarsely-defined categorical variables, reducing power and potentially underestimating etiological associations. More precise measurement approaches could improve testing of etiological theories of wheezing illness.

Methods

We evaluated a continuous latent variable model of pediatric wheezing severity based on four ISAAC Wheezing Module items. Analyses included subgroups of children from three independent cohorts whose parents reported past wheezing: infants ages 0–2 in the INSPIRE birth cohort study (Cohort 1; $n = 657$), 6-7-year-old North American children from Phase One of the ISAAC study (Cohort 2; $n = 2{,}765$), and 5-6-year-old children in the EHAAS birth cohort study (Cohort 3; $n = 102$). Models were estimated using structural equation modeling.

Results

In all cohorts, covariance patterns implied by the latent variable model were consistent with the observed data, as indicated by non-significant χ^2 goodness of fit tests (no evidence of

Institute of Allergy and Infectious Diseases (https://www.niaid.nih.gov/) to TVH; UG3 OD023282, National Institutes of Health/Office of the Director (https://www.nih.gov/institutes-nih/nih-office-director) to JEG; K24 AI 077930, National Institutes of Health Midcareer Investigator Award (https://www.nih.gov/) to TVH; K12 HS022990, Agency for Healthcare Research and Quality (https://www.ahrq.gov/) to SMB; R01 AI035786, National Institute of Allergy and Infectious Diseases to DRG. Research reported in this publication was supported by the CREW consortium, Office of The Director, National Institutes of Health under award number UG3 OD023282. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

model misspecification). Cohort 1 analyses showed that the latent factor structure was stable across time points and child sexes. In both cohorts 1 and 3, the latent wheezing severity variable was prospectively associated with wheeze-related clinical outcomes, including physician asthma diagnosis, acute corticosteroid use, and wheeze-related outpatient medical visits when adjusting for confounders

Conclusion

We developed an easily applicable continuous latent variable model of pediatric wheezing severity based on items from the well-validated ISAAC Wheezing Module. This model prospectively associates with asthma morbidity, as demonstrated in two ECHO birth cohort studies, and provides a more statistically powerful method of testing etiologic hypotheses of childhood wheezing illness and asthma.

Introduction

Background

Asthma is among the most common and costly diseases affecting childhood [1,2]. Large epidemiological studies aimed at advancing understanding of the development and sequelae of asthma often rely on brief questionnaires that can be broadly implemented to characterize symptom severity [3,4]. Investigators for the International Study of Asthma and Allergies in Children (ISAAC) [3] developed and validated a research questionnaire assessing wheezing symptoms [3,5] that has since been used worldwide to assess the prevalence and etiology of asthma [4,6–10]. The ISAAC Wheezing Module (ISAAC-WM) includes eight items, four specifically focused on wheezing, either parent-reported for younger children or self-reported by older children [3]. The questionnaire and manual are available at http://isaac.auckland.ac.nz/story/methods/methods.php.

There is no consensus method for operationalizing wheezing severity using the ISAAC-WM. Respondents endorsing past-year wheezing complete three items assessing the frequency of wheezing episodes and negative sequelae in the past 12 months. Additionally, all respondents report on whether or not their child experienced exercise-induced wheezing [3]. Researchers typically stratify children into discrete severity groups based on frequency of wheezing episodes (item 3), wheeze-related sleep disturbance (item 4), wheeze-related speech disturbance (item 5) [7], exercise-induced wheezing (item 7), or a combination of these items [11].

Discrete severity variables are suboptimal in explanatory research as they artificially reduce variability, grouping individuals with truly differing severity levels into the same data bins [12]. This reduces power, likely resulting in underestimates of etiological associations and suboptimal theory testing [12,13]. A number of research groups have used exploratory techniques, like principal components [14,15] or latent factor approaches [16,17], to develop continuous asthma severity measures [18]. To our knowledge, however, no studies have tested the plausibility of these models using confirmatory analytic techniques. Additionally, we know of no studies, either exploratory or confirmatory, evaluating the ability of the ISAAC-WM to capture pediatric wheezing severity on a continuous scale for research purposes.

Purpose

There are a number of large epidemiological studies evaluating pediatric wheezing illness using the ISAAC-WM [4,8,19,20], including nearly all airway cohorts participating in the

Environmental Influences on Child Health Outcomes (ECHO) initiative [21]. ECHO is an NIH-funded consortium of previously established birth cohorts aimed at understanding the modifiable environmental etiologies of four major diseases, including asthma [21]. These studies contain rich data (e.g., early life exposures and wheeze-related clinical outcomes), providing opportunities to test etiological models of wheezing illness development and outcomes. To conduct strong tests of theory, however, it is critical to measure wheezing severity with optimal precision. The purpose of this study is to demonstrate how studies using the ISAAC-WM can use structural equation modeling (SEM) to estimate wheezing severity as a continuous latent variable. Using a continuous latent variable approach, rather than discrete severity variables, may lead to more precise measurement and stronger tests of etiological theory [22] in studies using the ISAAC-WM. This, in turn, could lead to improved understanding of pediatric wheezing illness and strengthen prevention and early intervention efforts.

SEM is a flexible multivariate statistical approach that supports precise specification of theoretical models and estimation of latent variables representing constructs that we cannot measure directly [22–25]. Latent *factors* capture shared variance among multiple measured variables (*factor indicators*) believed to have a common underlying cause [25,26]. In this study, we conceptualize wheezing illness as a latent (unobserved) variable that has measurable consequences (e.g., wheezing attacks, disturbed sleep). The four wheezing-focused ISAAC-WM items capture and quantify these observable manifestations of wheezing illness. If we assume that wheezing illness (the latent construct) is driving observed correlations among the ISAAC-WM wheezing items, then their shared variance can be used to estimate wheezing illness as a continuous latent variable. Modeling associations with well-specified latent variables generally results in increased statistical power [22,27].

We evaluated our proposed model of wheezing illness severity in three independent cohorts. Two of these cohorts are part of the Children's Respiratory Research and Environment Workgroup (CREW) within the ECHO initiative [21], and the third is a publicly available data set from the ISAAC Phase I study [6]. Nearly all ECHO airway cohorts have used the ISAAC-WM, thus a more powerful approach to estimating severity in explanatory research studies with this instrument may be of broad interest. We address whether the proposed latent factor model of pediatric wheezing severity is:

1. consistent with observed data in all three cohorts
2. positively and concurrently associated with established markers of wheezing illness (convergent validity)
3. stable over repeated time points (longitudinal invariance) and across child sexes
4. positively and prospectively associated with wheeze-related clinical outcomes (predictive validity)

A valid latent variable approach to measuring wheezing severity with the ISAAC-WM may facilitate stronger tests of etiological theory and advancements in our understanding of pediatric wheezing illness.

Materials and methods

Data sources and measures

Table 1 provides demographic characteristics for each cohort.

Cohort 1. INSPIRE birth cohort study. The Infant Susceptibility to Pulmonary Infections and Asthma following RSV Exposure (INSPIRE) study is an ongoing population-based birth cohort that is part of the ECHO/CREW consortium, evaluating the role of early life

Table 1. Participant characteristics for all cohorts.

	n	Proportion
Cohort 1: INSPIRE (*n* = 657)		
Maternal History of Asthma	154	0.23
Child RSV+ Molecular Diagnostic Test	66	0.10
Maternal Prenatal Cigarette Smoking (Any)	139	0.21
Medicaid Insurance at Enrollment	377	0.57
Child Cesarean Birth Delivery	224	0.34
Child Sex: Female	250	0.38
Mother Married at Enrollment	368	0.56
Child Race		
African American	167	0.25
Asian	10	0.02
Native American/Native Alaskan	6	0.01
Native Hawaiian/Pacific Islander	2	0.003
White	510	0.78
Other Race	37	0.06
Child Ethnicity: Hispanic/Latino	58	0.09
Cohort 2: ISAAC Phase 1 (*n* = 2,765)		
Child Sex: Female	1,193	0.43
Cohort 3: EHAAS (*n* = 102)		
Child Sex: Female	34	0.33
Child Race/Ethnicity		
African American	10	0.10
Asian	7	0.07
Hispanic/Latino	4	0.04
White	80	0.78
Other Race	1	0.01

https://doi.org/10.1371/journal.pone.0194739.t001

respiratory infection in pediatric asthma [19,28]. The study enrolled term and otherwise healthy infants (N = 1,951) in Middle Tennessee. The Vanderbilt University Institutional Review Board (IRB) approved all study procedures. Wheezing symptoms are assessed annually using the parent-reported ISAAC-WM [19]. Although all participants complete the ISAAC-WM, given our focus on wheezing severity, we included only children whose parents reported wheezing in the first two years of life (n = 657).

At both the one- and two-year assessments, parents reported whether or not their children used asthma medications and/or had been hospitalized for respiratory illnesses in the prior year. With regard to asthma medications, parents were asked explicitly if their children had been treated with any of the following: Budesonide, Fluticasone, Beclomethasone, Montelukast, Albuterol, Budesonide-Salmeterol, Fluticasone-Salmeterol, Prednisone/Prednisolone, Dexamethasone, or any other asthma medications. At the year-three assessment, parents reported several wheeze-related clinical outcomes: physician diagnosis of asthma at any point in life (*Absent* or *Present*), past-year treatment of asthma symptoms with any form of corticosteroid (*Never*, *1–3 times*, *4+ times*), and past-year frequency of wheeze-related healthcare visits (*None*, *1–3*, *4+*). When asking parents about corticosteroid use, no distinction was made between parenteral steroids or oral corticosteroids in the questioning.

Cohort 2. ISAAC Phase 1 study. Cohort 2 analyses used publicly available data from the ISAAC Study Phase I (http://isaac.auckland.ac.nz/phases/phaseone/results/resultsIndv.php) to

evaluate whether the covariance patterns implied by the latent wheezing severity model established in Cohort 1 were consistent with observed covariance patterns in an independent dataset. Detailed study procedures are available in prior publications [3,5]. ISAAC Phase I included two cohorts of children (ages 6–7 and ages 13–14) from 38 countries. We limited analyses to 6-7-year-old children from the three North American countries: Canada, U.S.A., and Barbados. This was felt to be most representative of the ECHO/CREW consortium sites. Analyses included the subset of children who, by parent report, had ever had wheezing episodes (n = 2,765).

Cohort 3. EHAAS birth cohort. Cohort 3 analyses evaluated whether the latent wheezing severity model established with Cohort 1 was consistent with observed data patterns in the Epidemiology of Home Allergens and Asthma Study (EHAAS) study, and whether the latent factor was prospectively associated with wheezing morbidity. EHAAS is a prospective birth cohort study that is also part of the ECHO/CREW consortium (N = 505) [20,29–31]. Families from a Boston hospital were eligible if the biological mother was 18+ years old at enrollment, at least one biological parent had a history of allergies or asthma, infants were born term without major congenital anomalies and were not admitted to the neonatal intensive care unit. Study procedures were approved by the Brigham and Women's Hospital IRB. We evaluated ISAAC-WM data collected when children were 5 and 6 years old. Analyses were limited to the subsample of children whose parents reported wheezing episodes at some point during the two-year follow-up window (n = 102). Clinical outcomes at age 7 included whether or not the child had: (1) a physician diagnosis of asthma, (2) any prior-year wheezing-related medical visits, and (3) any prior-year wheezing-related urgent visits to a healthcare provider's office or emergency department.

Statistical analyses

Structural equation modeling (SEM) was the primary analytic strategy with parameters estimated via robust weighted least squares [32]. Fig 1A shows the hypothesized latent wheezing severity model. Associations between the latent wheezing severity variable and the ordinal ISAAC-WM severity items were estimated using a latent response variable framework [22,32,33], in which ordered categorical outcomes (y_1, y_2, y_3, y_4) are conceptualized as coarse approximations of underlying continuous variables ($y_1^*, y_2^*, y_3^*, y_4^*$). The models used probit regressions to estimate c-1 thresholds (τ) for ordinal outcomes, where c is the number of ordered levels, dividing the underlying continuous distributions into discrete categories [32]. Analyses were conducted using the lavaan package version 0.5–22 [34] in R [35] and Mplus version 7.4 [36]. Missing data in the response variables were presumed to be missing at random conditional on the model covariates [37].

The ISAAC Phase 1 data (Cohort 2) had a complex structure with children nested within schools (j = 282) within study sites (k = 3). We used multilevel SEM [38,39] to evaluate the latent wheezing severity factor model while accounting for clustering (dependence) within schools (Fig 1B). On the within-school level, the wheezing severity model was identical to the model specified in Cohort 1. On the between-schools level, we specified a saturated model where all indicator variances and covariances were freely estimated and the indicators were regressed on dummy variables coding study site. This approach accounts for between-school and between-site variability, but presumes that wheezing severity is a meaningful construct only on the within-schools level of analysis; that is, wheezing severity is only intended to be measured as a child-level characteristic, not a school-level characteristic [40].

In the Cohort 3 analyses, to avoid sparse cells for the ISAAC-WM indicators, data for each of the ISAAC-WM items were aggregated across the age-5 and age-6 assessments (Table 2). These aggregated items were used in all models.

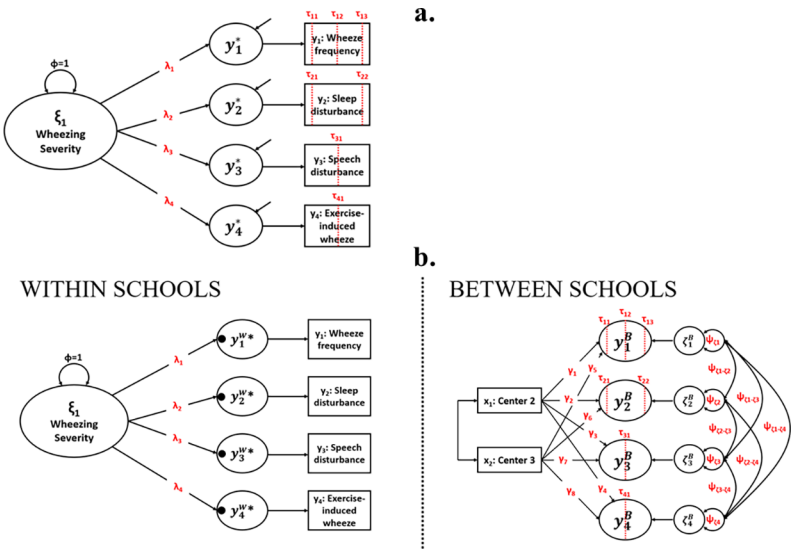


Fig 1. Latent variable model of wheezing illness severity. Panel *a* shows the latent wheezing severity model used in cohorts 1 and 3. The severity of wheezing illness is estimated as a unidimensional latent variable (η_1) with four reflective ordinal indicators: wheezing episode frequency (y_1), frequency of wheeze-related sleep disturbance (y_2), wheeze-related speech disturbance (y_3), and exercise-induced wheeze (y_4). The ordinal indicators are presumed to be coarse measurements of underlying continuous variables (y_1^*-y_4^*). Panel *b* shows the multilevel wheezing severity model used in the Cohort 2 analyses. The within-schools level of the model is identical to panel *a*. The between-schools level of the model accounts for non-independence due to clustering within schools and study sites. Estimated parameters are depicted in red.

https://doi.org/10.1371/journal.pone.0194739.g001

Assessing model adequacy. In all three cohorts, the degree of discrepancy between proposed models and observed data was assessed using a mean- and variance-adjusted χ^2 Goodness of Fit (χ^2_{GOF}) test [32,37,41,42]. Significant χ^2_{GOF} values lead to a rejection of the null hypothesis of a perfect correspondence between the model and the observed data [22]. We reported two additional indices of model adequacy: the root mean error of approximation (RMSEA) [43,44] and comparative fit index (CFI) [45]. The RMSEA provides an estimate of the amount of misspecification per model degree of freedom with a 90% confidence interval (CI), whereas the CFI compares the hypothesized model to a simpler baseline (null) model, penalizing the hypothesized model for each estimated parameter [46]. Although there are no perfect cutoff values for these indices [47], RMSEA values $< .05$ and CFI values $\geq .95$ are generally considered to be indicative of a desirable model [48].

It is noteworthy that, owing to the relatively small sample size, Study 3 models were underpowered to detect significant model misspecification with the χ^2_{GOF} test. Consequently, we relied on data from studies 1 and 2 to assess whether our proposed latent wheezing severity model was consistent with the observed data, whereas Study 3 analyses were focused on evaluating whether the latent wheezing severity variable was prospectively associated with established markers of wheezing illness severity (predictive validity).

Convergent validity. Using Cohort 1 data at both the year-one and year-two assessments, we specified models evaluating concurrent associations between the latent wheezing severity factor and parent-reported markers of wheezing illness: wheezing medication use (*Any* vs. *None*) and respiratory hospitalizations (*Any* vs. *None*). These associations were adjusted for seven potential confounders representing characteristics that have been linked to asthma risk in prior studies [49]: maternal asthma history [50] (*Absent* vs. *Present*), maternal prenatal cigarette use [51] (*None* vs. *Any*), maternal marital status [52] (*Married* vs. *Single*), insurance at

Table 2. Cohort 3 EHAAS study: Derived wheezing severity items characterizing the severity of wheeze across the 60- and 72-month assessments combined.

ISAAC-WM WHEEZING ITEM	DERIVED SEVERITY LEVELS COMBINING ACROSS THE 60- AND 72-MONTH ASSESSSMENTS
Item 3. Number of wheezing attacks in past 12 months	Very infrequent (value = 0) • *None* (60-Month) and *1–3 attacks* (72-Month) • *1–3 attacks* (60-Month) and *None* (72-Month) Infrequent (value = 1) • *1–3 attacks* (60-Month) and *1–3 attacks* (72-Month) • *None* (60-Month) and *4–12 attacks* (72-Month) • *4–12 attacks* (60-Month) and *None* (72-Month) Frequent (value = 2) • *1–3 attacks* (60-Month) and *4–12 attacks* (72-Month) • *4–12 attacks* (60-Month) and *1–3 attacks* (72-Month) • *None* (60-Month) and *13+ attacks* (72-Month) • *13+ attacks* (60-Month) and *None* (72-Month) Very frequent (value = 3) • *4–12 attacks* (60-Month) and *4–12 attacks* (72-Month) • *4–12 attacks* (60-Month) and *13+ attacks* (72-Month) • *13+ attacks* (60-Month) and *4–12 attacks* (72-Month) • *13+ attacks* (60-Month) and *13+ attacks* (72-Month)
Item 4. Child's sleep disturbed due to wheezing	Never (value = 0) • *Never* (60-Month) and *Never* (72-Month) Infrequent (value = 1) • *Never* (60-Month) and *< 1 night/week* (72-Month) • *< 1 night/week* (60-Month) and *Never* (72-Month) Frequent (value = 2) • *Never* (60-Month) and *1+ nights/week* (72-Month) • *1+ nights/week* (60-Month) and *Never* (72-Month) • *<1 night/week* (60-Month) and *<1 night/week* (72-Month) • *<1 night/week* (60-Month) and *1+ nights/week* (72-Month) • *1+ nights/week* (60-Month) and *<1 night/week* (72-Month) • *1+ nights/week* (60-Month) and *1+ nights/week* (72-Month)
Item 5. Wheezing ever severe enough to child's speech to only one or two words at a time between breaths	Absent at both assessments (value = 0) Present at either or both assessments (value = 1)
Item 7. Wheezing or whistling in chest during or after exercise	Absent at both assessments (value = 0) Present at either or both assessments (value = 1)

https://doi.org/10.1371/journal.pone.0194739.t002

enrollment as a proxy for socioeconomic status [53] (*Private* vs. *Medicaid*), birth method [54] (*Vaginal* vs. *Cesarean*), child sex [55], and child race [55] (African American, all other race groups, and white race [reference group]). Both the markers of wheezing illness severity (validity outcomes) and the latent wheezing factor were regressed on these seven asthma risk factors.

Latent factor stability. Using Cohort 1 data, we evaluated whether our latent wheezing severity model was stable across the one- and two-year follow-ups using procedures described in Liu et al. (2016) [56]. An unstable factor structure indicates that the latent construct cannot be measured reliably in the same way across repeated assessments, making it challenging to interpret change over time [56–58]. Four competing models, representing differing degrees of instability, were specified and compared using scaled χ^2 likelihood ratio tests [59]. Significant χ^2 tests are *not* desirable in this context because they indicate that models allowing for instability in the factor structure fit better than more restrictive models that assume stability. The most desirable model allows mean levels of wheezing severity (the latent factor) to change over

time, but holds the measurement of wheezing severity (factor loadings, thresholds, and residual variances) constant over time. In all models, the latent wheezing factor was regressed on the seven asthma risk factors described above, with associations between risk factors and the latent factor permitted to vary freely across time points. We used similar procedures [60] to test whether the latent factor measurement was consistent across sexes (female vs. male children) [61].

Predictive validity. Using Cohort 1 data, we specified longitudinal models evaluating whether the latent wheezing illness factor at the two-year assessment was prospectively associated with year-three clinical outcomes, including physician asthma diagnosis, acute corticosteroid use, and wheeze-related outpatient visits. These prospective associations were adjusted for all seven asthma risk factors described above. Using Cohort 3 data, we evaluated whether the latent wheezing severity variable measured at ages 5–6 was prospectively associated with wheeze-related outcomes at age 7, adjusting for child sex and race.

Comparison to discrete severity models. Using data from Cohort 1, we gauged whether the latent factor approach to estimating wheezing severity as a continuous construct had greater predictive utility relative to a more traditional approach with a discrete severity variable. To accomplish this, we first derived a five-level ordinal wheezing severity variable indicating the number of ISAAC-WM wheezing items endorsed (i.e., present to some degree) at the year-two assessment (*None*, *1*, *2*, *3*, *4*). We then reran the latent variable predictive validity models described in the prior section, using the same covariates and estimation procedures, but replacing the latent wheezing severity factor with the ordinal severity variable. This allowed us to evaluate whether the latent variable modeling approach yielded stronger associations with the future clinical outcomes (physician asthma diagnosis, acute corticosteroid use, and wheeze-related outpatient visits) compared to the more traditional discrete severity approach.

Results

Table 3 provides descriptive statistics for the ISAAC-WM severity items for all three cohorts. Separate online appendixes for Cohort 1 (S1 Appendix), Cohort 2 (S2 Appendix), and Cohort 3 (S3 Appendix) are available providing model code. For cohorts 1 and 3, we provide sufficient summary statistics (covariance matrixes, weighted least squares weight vector, asymptotic variance matrix, and sample size) to allow others to reproduce our models (S1 File).

Model fit

Table 4 provides information about model adequacy for all models. Model numbers are provided in the table and are referenced in the text below.

Cohort 1. INSPIRE birth cohort. The unidimensional latent factor model of wheezing illness severity (Fig 1A) provided a close fit to the observed data at both the one- and two-year assessments, as indicated by non-significant χ^2_{GOF} tests (Table 4, models 1.1 and 1.2). At both assessments, there were strong positive associations between the latent wheezing factor and the ISAAC-WM wheezing indicators (S1 Fig). Consistent with prior research using a different instrument [16], information curves showed that levels of wheezing illness severity factor could be estimated with greatest precision at moderate to high levels of illness severity: The latent factor was less informative for children with mild wheezing symptoms (S2 Fig).

Cohort 2. ISAAC Phase 1. Consistent with Cohort 1 findings, the multilevel SEM model evaluating the wheezing severity as a unidimensional construct yielded a non-significant χ^2_{GOF} test (Table 4, Model 2.1), indicating a close fit to the data. Associations between the latent factor and all four indicators were positive and significant on the within-school level of analysis.

Table 3. Descriptive statistics for ISAAC-wheezing module severity items.

	LEVELS	COUNT	PROP	COUNT	PROP
Cohort 1: INSPIRE		**1st Year of Life**		**2nd Year of Life**	
# of Wheeze Episodes	Never	205	0.31	156	0.24
	1 to 3	328	0.50	274	0.42
	4 to 12	73	0.11	76	0.12
	13+	28	0.04	17	0.03
	Missing	23	0.04	134	0.20
Wheeze Disturbs Sleep	Never	455	0.69	336	0.51
	< 1 night/week	83	0.13	93	0.14
	1+ nights/week	96	0.15	94	0.14
	Missing	23	0.04	134	0.20
Wheeze Disturbs Speech	Present	116	0.18	100	0.15
	Missing	24	0.04	135	0.21
Exercise-Induced Wheeze	Present	67	0.10	76	0.12
	Missing	14	0.02	33	0.05
Cohort 2: ISAAC Phase 1					
# of Wheeze Episodes	Never	1,173	0.42		
	1 to 3	1,074	0.39		
	4 to 12	330	0.12		
	13+	107	0.04		
	Missing	81	0.03		
Wheeze Disturbs Sleep	Never	1,793	0.65		
	< 1 night/week	641	0.23		
	1+ nights/week	228	0.08		
	Missing	103	0.04		
Wheeze Disturbs Speech	Present	259	0.09		
	Missing	91	0.03		
Exercise-Induced Wheeze	Present	714	0.26		
	Missing	119	0.04		
Cohort 3: EHAAS					
# of Wheeze Episodes	Very Infrequent	31	0.30		
	Infrequent	39	0.38		
	Frequent	12	0.12		
	Very frequent	16	0.16		
	Missing	4	0.04		
Wheeze Disturbs Sleep	Never	55	0.54		
	Infrequent	23	0.23		
	Frequent	21	0.21		
	Missing	3	0.03		
Wheeze Disturbs Speech	Present	11	0.11		
	Missing	3	0.03		
Exercise-Induced Wheeze	Present	28	0.27		
	Missing	1	0.01		

PROP = proportion of the total sample.

https://doi.org/10.1371/journal.pone.0194739.t003

Cohort 3. EHAAS birth cohort. Consistent with findings from cohorts 1 and 2, the hypothesized unidimensional factor model yielded a non-significant χ^2_{GOF} test (Table 4, Model 3.1). Thus, there was no evidence of model misspecification. This finding should be interpreted cautiously,

Table 4. Indices of model adequacy.

	χ^2_{GOF}			RMSEA		CFI	Model Description
Cohort 1: INSPIRE Birth Cohort							
	Est	*df*	*p*	Est	90% CI	Est	
Model 1.1	2.35	2	.31	.02	.00, .08	1.0	Y1: Wheeze severity model
Model 1.2	3.76	2	.15	.04	.00, .10	1.0	Y2: Wheeze severity model
Model 1.3	27.19	29	.56	.00	.00, .03	1.0	Y1: Concurrent association with respiratory hospital visit
Model 1.4	31.08	29	.36	.01	.00, .03	.97	Y2: Concurrent association with respiratory hospital visit
Model 1.5	26.75	29	.59	.00	.00, .03	1.0	Y1: Concurrent association with asthma medication
Model 1.6	29.47	29	.44	.01	.00, .03	1.0	Y2: Concurrent association with asthma medication
Model 1.7	29.48	29	.44	.01	.00, .03	.99	Prospective associations with Y3 asthma diagnosis
Model 1.8	29.28	29	.45	.00	.00, .03	1.0	Prospective associations with Y3 corticosteroids
Model 1.9	29.01	29	.46	.00	.00, .03	1.0	Prospective associations with Y3 wheeze medical visits
Cohort 2: ISAAC Phase 1							
Model 2.1	0.48	2	.79	.02	NA	1.0	Wheeze severity model
Model 2.2	7.20	5	.21	.00	NA	1.0	Regression of wheeze severity factor on child sex
Cohort 3: EHAAS Birth Cohort							
Model 3.1	0.83	2	.66	.00	.00, .15	1.0	Wheeze severity model
Model 3.2	4.00	8	.86	.00	.00, .06	1.0	Regression of wheeze severity on child sex and race
Model 3.3	7.01	11	.80	.00	.00, .07	1.0	Predictive validity: 84-month asthma dx.
Model 3.4	7.22	11	.78	.00	.00, .07	1.0	Predictive validity: 84-month wheeze visits
Model 3.5	6.05	11	.87	.00	.00, .05	1.0	Predictive validity: 84-month ER visits

χ^2_{GOF} = model chi-square goodness of fit test; RMSEA = root mean error of approximation; CFI = Comparative Fit Index; Y1-Y3 = Year 1 –Year 3; NA = RMSEA confidence intervals were not available for multilevel SEM models.

https://doi.org/10.1371/journal.pone.0194739.t004

however, as the model was underpowered to detect a significant discrepancy between the model and data because of the relatively small sample size in the Cohort 3 analyses.

Convergent validity

Using Cohort 1 data at both the one- and two-year assessments, the wheezing illness factor was positively associated with hospitalization for respiratory illness (Table 5, models 1.3 and

Table 5. Cohort 1: Associations between the latent wheezing illness severity factors at both year-1 and year-2 and established markers of wheezing illness severity.

LATENT FACTOR	MODEL	EST	95% CI	
Marker of Wheezing Illness Severity	**#**	**β[a]**	**Lower**	**Upper**
Year 1 Wheezing Illness Severity (Ages 0–1)				
Ages 0–1 Wheeze -Related Hospitalizations	1.3	0.68	0.51	0.86
Ages 0–1 Asthma Medication Use	1.5	0.87	0.72	1.01
Year 2 Wheezing Illness Severity (Ages 1–2)				
Ages 1–2 Wheeze -Related Hospitalizations	1.4	0.54	0.35	0.72
Ages 1–2 Asthma Medication Use	1.6	0.90	0.77	1.03
Ages 2–3 Asthma Diagnosis	1.7	0.71	0.53	0.88
Ages 2–3 Acute Corticosteroid Use	1.8	0.63	0.49	0.76
Ages 2–3 Wheeze-Related Medical Visits	1.9	0.64	0.48	0.79

[a]Linear regression coefficient estimating the association between the latent wheezing illness severity factor and the continuous latent variable (y^*) underlying the observed categorical outcome variable (y) that is an established marker of wheezing severity.

https://doi.org/10.1371/journal.pone.0194739.t005

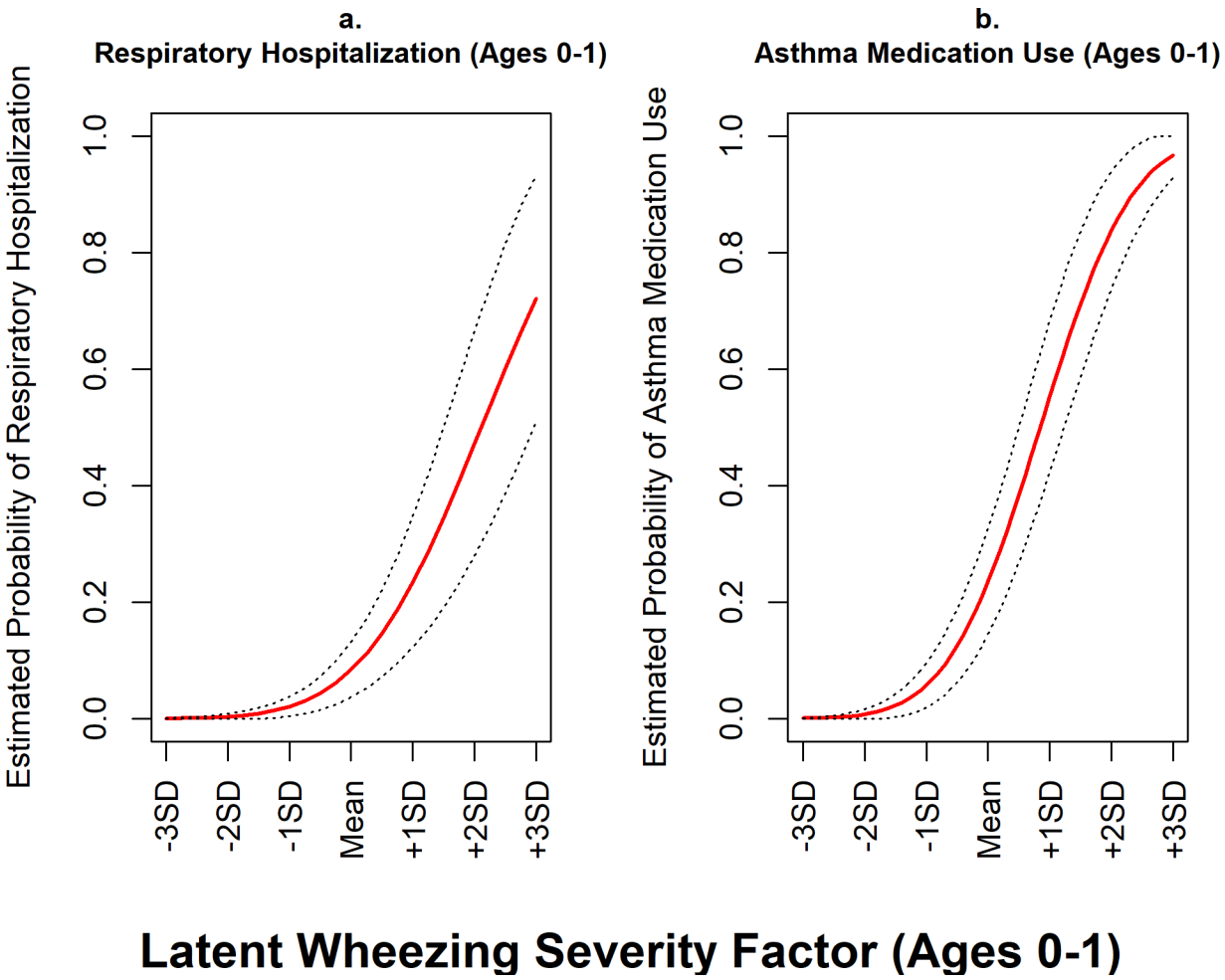


Fig 2. Cohort 1: Concurrent associations between the latent wheezing severity factor and markers of wheezing illness in year-one. This figure shows estimated probabilities of having at least one respiratory hospitalization (panel *a;* model 1.3) and using asthma medication (panel *b*; model 1.5) against levels of the latent wheezing severity variable in the first year of life. As wheezing severity increases so does the estimated probability of respiratory hospitalization and medication use. These models held all covariates constant at their median values. Dotted lines represent 95% confidence intervals for estimated probability estimates.

https://doi.org/10.1371/journal.pone.0194739.g002

1.4) and asthma medication use (Table 5, models 1.5 and 1.6) adjusting for covariates. Fig 2 shows the magnitude of these concurrent associations at year-one.

Latent factor stability

Using Cohort 1 data, there was no evidence from likelihood ratio χ^2 tests (p values $\geq$ .20) that imposing invariance restrictions on the loadings, thresholds, or unique variances across time points worsened model fit (Table 6). Thus, there was no evidence of temporal instability, allowing for meaningful evaluation of change in the latent factor. Additionally, there was no evidence of latent factor instability across child sex at either year-one or year-two (p values $\geq$ .25; Table 6), indicating that the same latent wheezing severity model was reasonable for girls and boys

Consistent with prior research [62], male children scored higher on average than female children on the wheezing illness severity factor in all three cohorts. In Cohort 1 analyses, male children scored 0.20 (95% CI [0.02, 0.45]) standard deviations (*SD*s) higher (more severe

Table 6. Cohort 1: Comparison of models testing stability in the latent wheezing severity factor across time and across child sex.

	Measurement Parameters			χ^2_{GOF}			Model Comparison		
	Factor Loadings	Indicator Thresholds	Residual Variances	Est.	*df*	*p*	Models	χ^2_{DIFF}	*p*
Models Evaluating Measurement Stability Across One- and Two-Year Assessments									
1	Free	Free	Free	61.97	62	.48	-	-	-
2	Constant	Free	Free	64.52	65	.49	2 vs 1	0.22	.97
3	Constant	Constant	Free	68.19	68	.47	3 vs 2	4.10	.20
4	Constant	Constant	Constant	69.58	72	.56	4 vs 3	0.92	.88
Models Evaluating Measurement Stability Across Child Sex (Female vs. Male) at Year-One									
5	Free	Free	Free	46.31	46	.46	-	-	-
6	Constant	Free	Free	49.64	49	.45	6 vs 5	2.97	.25
7	Constant	Constant	Free	51.45	51	.46	7 vs 6	0.63	.60
8	Constant	Constant	Constant	52.69	55	.56	8 vs 7	0.68	.85
Models Evaluating Measurement Stability Across Child Sex (Female vs. Male) at Year-Two									
9	Free	Free	Free	38.95	46	.76	-	-	-
10	Constant	Free	Free	41.01	49	.78	10 vs 9	0.72	.72
11	Constant	Constant	Free	42.95	51	.78	11 vs 10	1.24	.37
12	Constant	Constant	Constant	47.83	55	.74	12 vs 11	2.70	.37

χ^2_{GOF} = chi-square goodness of fit statistic evaluating whether the model is consistent with the observed data; χ^2_{DIFF} = chi-square test comparing the fit of competing nested models; Free = parameter was allowed to vary freely across groups (i.e., time points or child sex); Constant = parameter was constrained to be equal across groups (i.e., time points or child sex).

https://doi.org/10.1371/journal.pone.0194739.t006

symptoms) than females on the latent factor in the second year of life. Similarly, in cohorts 2 (model 2.2) and 3 (model 3.2), male children scored 0.09 *SD*s (95% CI [0.003, 0.18]) and 0.68 *SD*s (95% CI [0.01, 1.35]) higher on wheezing severity than females, respectively.

Predictive validity

In Cohort 1 longitudinal models, the severity of wheezing illness in the second year of life was uniquely and prospectively associated with year-three physician asthma diagnosis (model 1.7), acute corticosteroid use (model 1.8), and wheeze-related outpatient visits (model 1.9), adjusting for covariates (Table 5). Using Cohort 3 data, adjusting for covariates, the latent wheezing illness factor representing wheezing severity measured at ages 5–6 was prospectively associated with all clinical outcomes measured at age 7 (Fig 3): physician asthma diagnosis (model 3.3; β = 0.52, 95% CI [0.22, 0.82]), wheeze -related medical visits (model 3.4; β = 0.58, 95% CI [0.20, 0.96]), and urgent wheeze-related medical visits (model 3.5; β = 0.82, 95% CI [0.41, 1.24]).

Comparison with discrete wheezing severity model

Using Cohort 1 data, the discrete wheezing severity variable coding the number of ISAAC-WM severity items endorsed at the year-two assessment was significantly and positively associated with year-three asthma diagnosis, acute corticosteroid use, and wheeze/asthma-related outpatient visits. Thus, the discrete severity exposure variable had utility in predicting future wheezing morbidity. However, Fig 4 illustrates the potential benefits of modeling wheezing severity as a latent continuous variable in explanatory models. Fig 4A plots the estimated probability of a child having at least one wheeze-related medical visit in the third year of life against the year-2 discrete wheezing severity variable, adjusting for covariates. As the number of ISAAC-WM severity items endorsed increased in the second year of life (x-

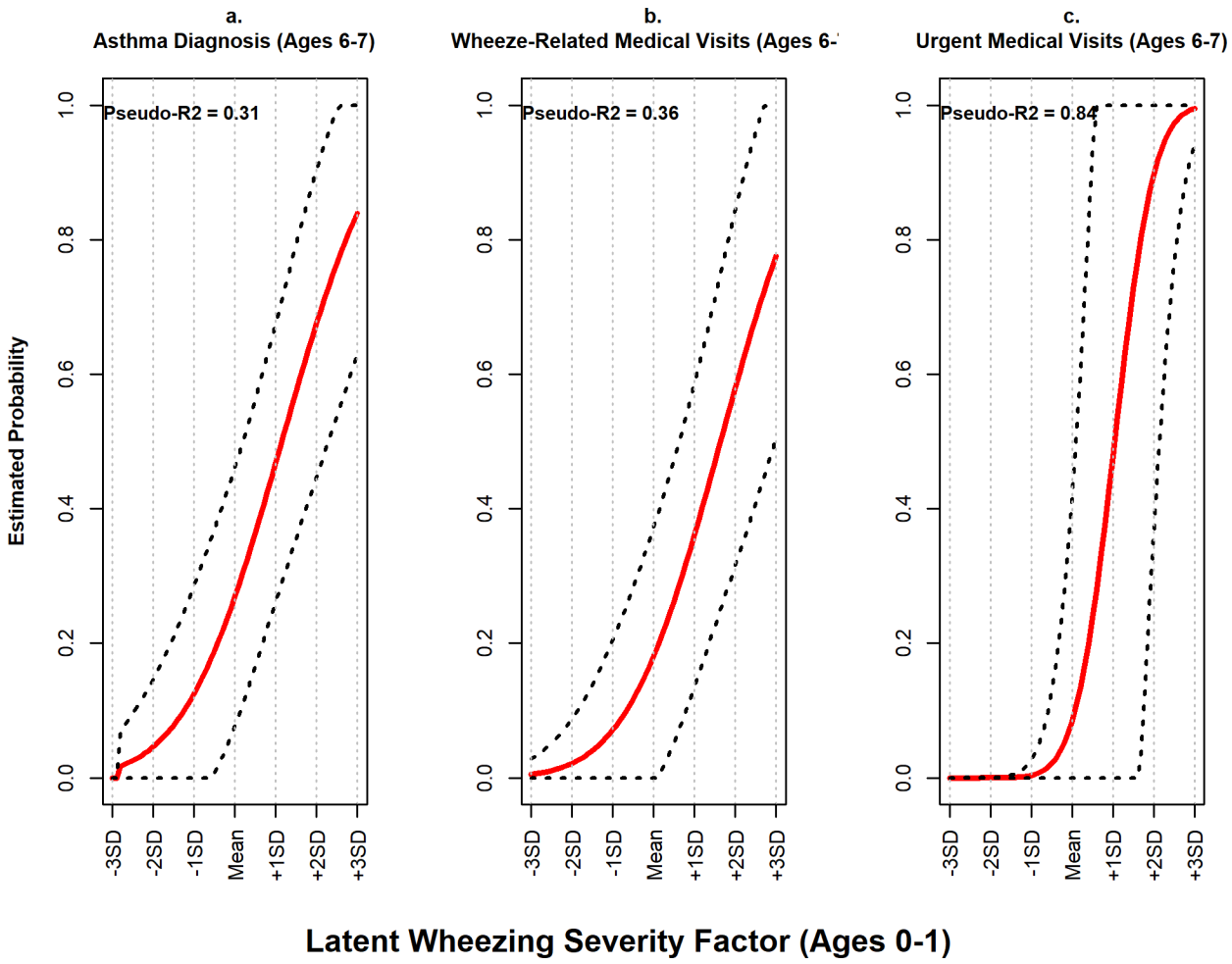


Fig 3. Cohort 3: Prospective associations between wheezing illness severity and wheeze-related morbidity outcomes. In the EHAAS birth cohort study, there were strong prospective associations between the latent wheezing illness severity factor representing wheezing illness severity from 60–72 months and three asthma morbidity outcomes at the 84-month follow-up: physician diagnosis of asthma (panel *a*; model 3.3), any asthma medical visits (panel *b*; model 3.4), and urgent asthma/wheeze-related medical visits to either a doctor's office or the emergency department (panel *c*; model 3.5). These models held all covariates constant at their median values. Dotted lines represent 95% confidence intervals. Pseudo-R^2 values represent the approximate proportion of variance in the outcome accounted for by the predictors.

https://doi.org/10.1371/journal.pone.0194739.g003

axis) the estimated probability of a child having a wheeze-related medical visit in the third year of life (y-axis) increased. In the discrete severity model, however, there are only five possible values of the estimated probability of year-three wheeze-related medical visits, ranging from 0.05 to 0.63, corresponding to the five levels of the discrete severity variable.

In contrast, Fig 4B shows the estimated probabilities of year-3 wheeze-related medical visits as a function of the latent wheezing severity factor, with estimated probabilities following a smooth curve ranging from 0.01 (at -3 *SD*s) to 0.82 (at +3 *SD*s). Going from the lowest level of the discrete wheezing severity variable to the highest results in a 0.59 (95% CI [0.34, 0.83]) increase in the estimated probability of year-3 wheeze-related medical visits; whereas going from 3 *SD*s below the mean to 3 *SD*s above the mean on the latent severity variable results in an increase in the estimated probability of 0.81, 95% CI [0.67, 0.94]. Similar patterns were observed for year-3 physician asthma diagnosis (Fig 5) and acute corticosteroid treatment (Fig 6). Also, the degree of uncertainty around the estimated probabilities was smaller when using the latent wheezing severity model. This is consistent with our central premise that a latent

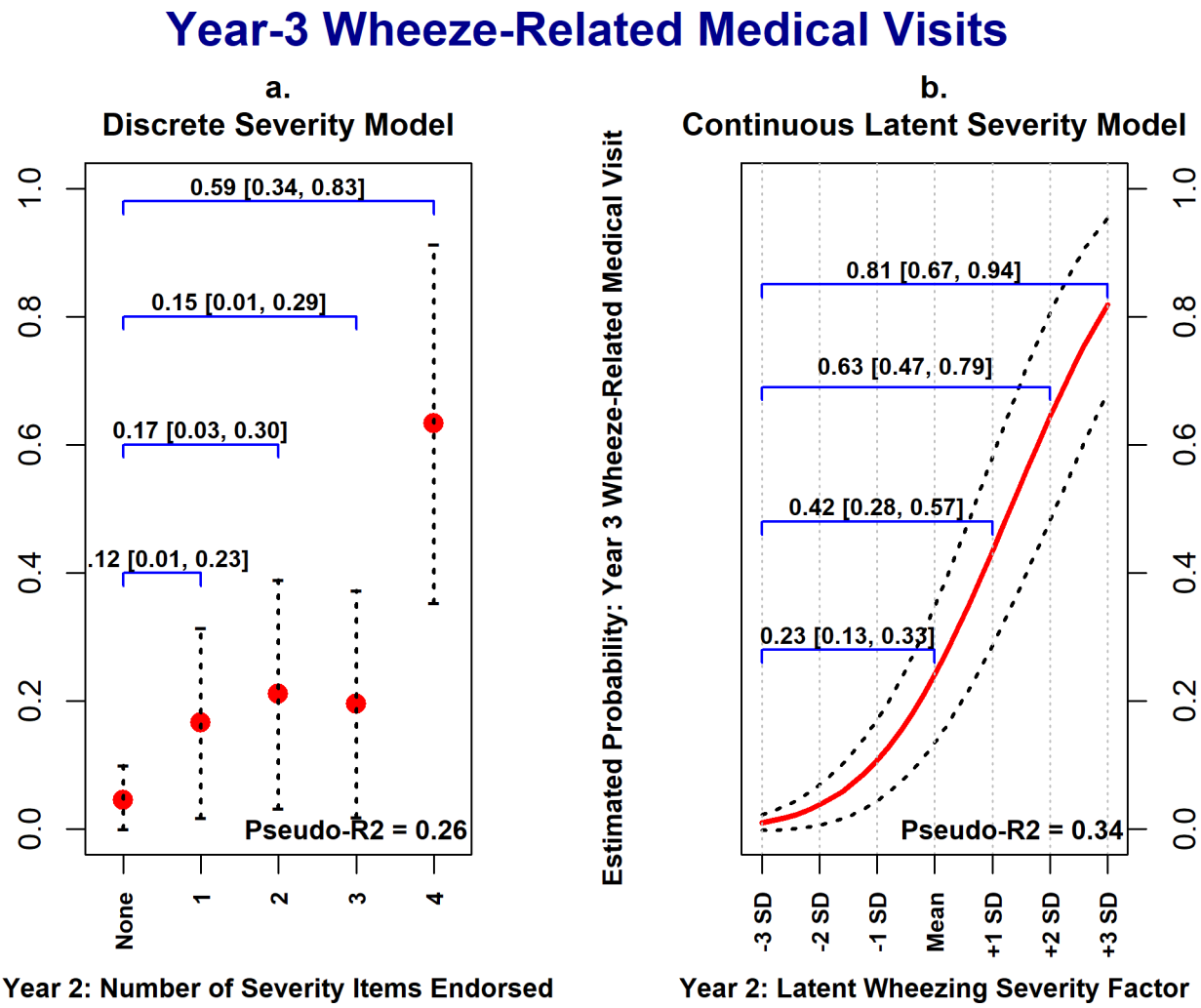


Fig 4. Cohort 1: Estimated probability of year-three wheeze-related medical visits as a function of wheezing severity in year-2. These plots show the strength of associations between year-2 wheezing severity (x-axis) and year-3 wheeze-related medical visits (y-axis), with all covariates held constant at their median values. Panels *a* shows estimated probabilities of corticosteroid treatment being present in the third year of life as a function of the discrete severity exposure variable; whereas panel *b* shows estimated probabilities vs. the latent continuous severity factor. In both models, as year-2 wheezing severity increases, so does the estimated probability of acute corticosteroid treatment, though the range of estimated probabilities is larger in the latent severity model. Values and 95% confidence intervals above the blue brackets show the increase in the estimated probabilities for a given increase in wheezing severity. Dotted lines represent 95% confidence intervals for estimated probability estimates.

https://doi.org/10.1371/journal.pone.0194739.g004

severity model can provide improved estimation of severity, resulting in stronger associations with clinical outcomes.

Discussion

Summary of key findings

The findings from three independent studies support the utility of a unidimensional latent variable approach to measuring pediatric wheezing illness severity based on items from the ISAAC-WM in explanatory research studies. We found:

1. a close correspondence between the latent wheezing severity model and the observed data (cohorts 1 and 2);

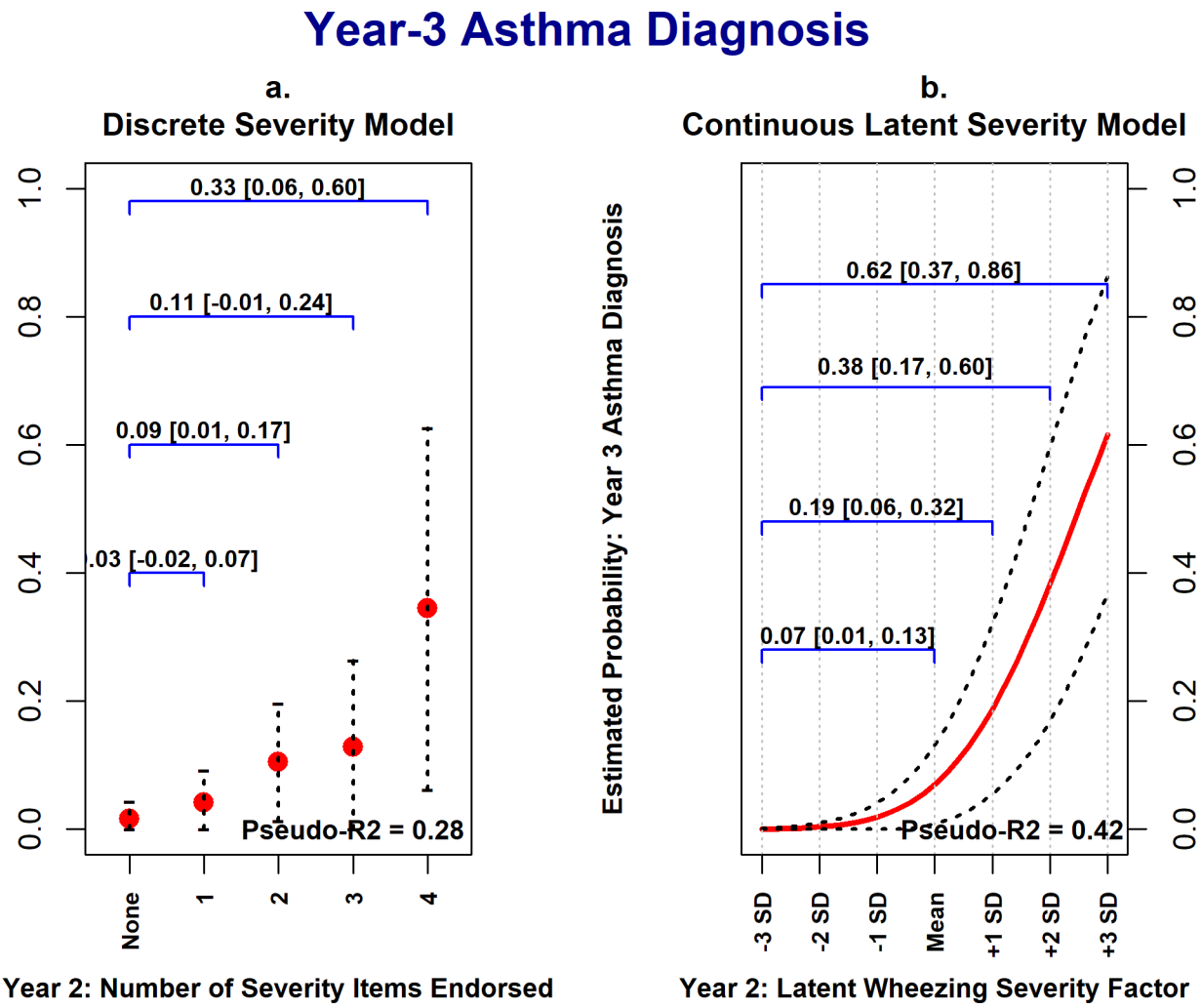


Fig 5. Cohort 1: Year-3 physician asthma diagnosis as a function of wheezing severity in year-2. These plots show the strength of associations between year-2 wheezing severity (x-axis) and year-3 physician asthma diagnosis (y-axis), with all covariates held constant at their median values. Panels *a* shows estimated probabilities of a physician asthma diagnosis being present in the third year of life as a function of the discrete severity exposure variable; whereas panel *b* shows estimated probabilities vs. the latent continuous severity factor. In both models, as year-2 wheezing severity increases, so does the estimated probability of a physician asthma diagnosis being present, though the range of estimated probabilities is larger in the latent severity model. Values and 95% confidence intervals above the blue brackets show the expected increase in the estimated probabilities for a given increase in wheezing severity. Dotted lines represent 95% confidence intervals for estimated probability estimates.

https://doi.org/10.1371/journal.pone.0194739.g005

2. concurrent (Cohort 1) and prospective (cohorts 1 and 3) associations with wheeze-related clinical outcomes;
3. stability in the latent wheezing severity factor structure across time points and child sex (Cohort 1);
4. stronger prospective associations with wheezing outcomes compared to a discrete severity variable (Cohort 1).

Interpretation and implications

Wheezing is the hallmark of asthma and an indicator of asthma control [63,64]. The ISAAC questionnaire is a validated instrument used in epidemiologic studies worldwide [3,4]. The use

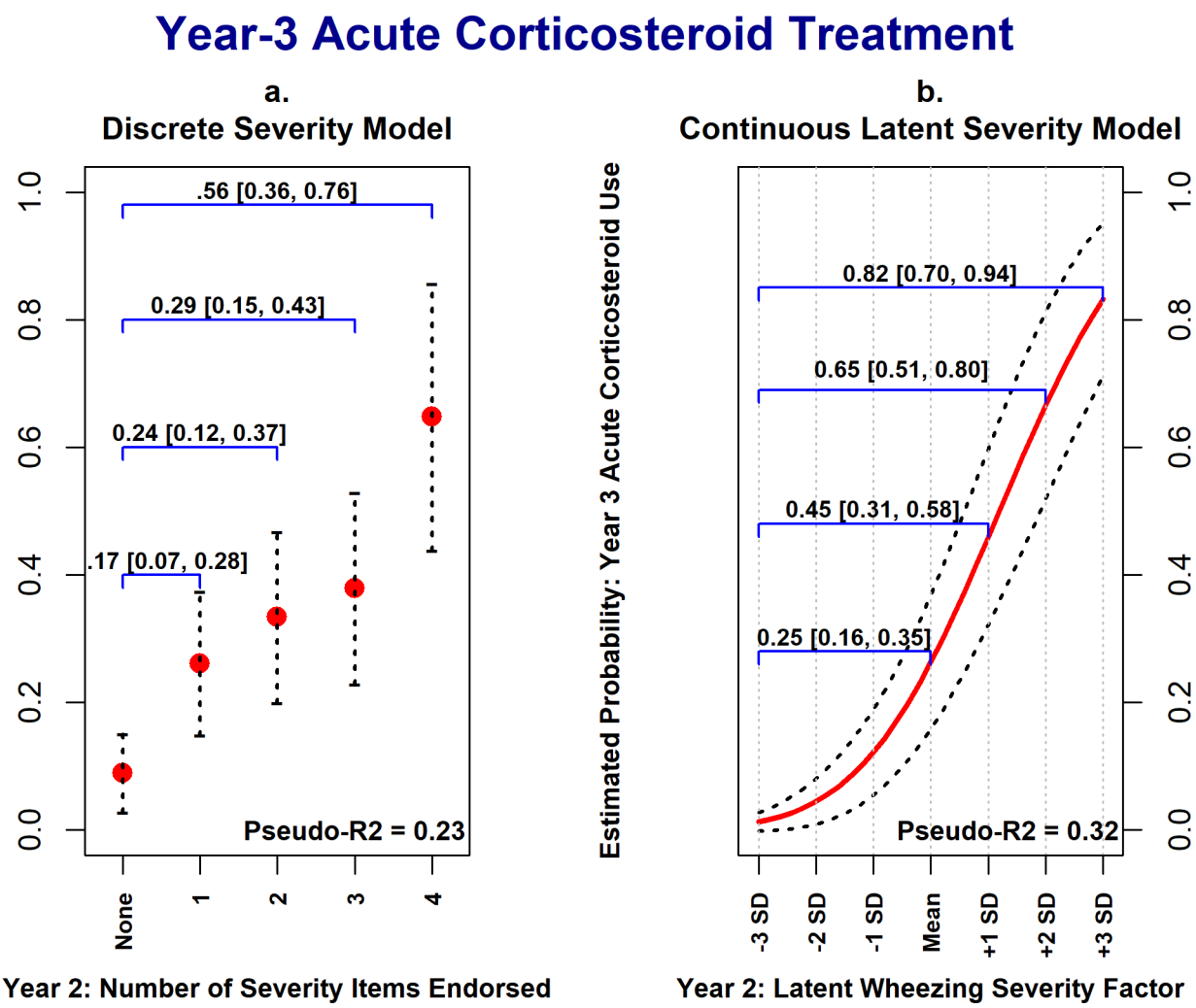


Fig 6. Cohort 1: Year-3 estimated probability of acute corticosteroid treatment as a function of wheezing severity in year-2. These plots show the strength of associations between year-2 wheezing severity (x-axis) and year-3 acute corticosteroid treatment (y-axis), with all covariates held constant at their median values. Panels *a* shows estimated probabilities of corticosteroid treatment being present in the third year of life as a function of the discrete severity exposure variable; whereas panel *b* shows estimated probabilities vs. the latent continuous severity factor. In both models, as year-2 wheezing severity increases, so does the estimated probability of acute corticosteroid treatment, though the range of estimated probabilities is larger in the latent severity model. Values and 95% confidence intervals above the blue brackets show the increase in the estimated probabilities for a given increase in wheezing severity. Dotted lines represent 95% confidence intervals for estimated probability estimates.

https://doi.org/10.1371/journal.pone.0194739.g006

of wheezing questions to assess disease severity in etiological research is highly desirable, particularly for the new ECHO consortium in which nearly all of the asthma birth cohorts utilized the ISAAC-WM. Our latent variable approach yields a continuous measure of wheezing illness severity that capitalizes on shared variance from the four wheezing-focused ISAAC-WM items. Our findings suggest that this latent variable approach results in stronger associations with established measures of wheezing illness severity than models using discrete severity variables based on the ISAAC-WM. Thus, for studies using the ISAAC-WM, the proposed latent variable approach may result in stronger tests of etiological theory than models using discrete severity variables.

It is important to note that the proposed SEM-based latent variable approach is intended for use in explanatory research models, where the goal is to test specific etiological hypotheses and theories. Latent variables do not yield unique scores for individual respondents [65].

Consequently, our proposed SEM-based latent variable model of wheezing severity could not be used in clinical practice to aid in diagnostic decisions or to make patient-specific predictions about risk for future adverse outcomes [65]. There are, however, a number of existing questionnaires measuring wheezing severity that were designed and validated for these purposes [18]. However, for researchers testing specific etiological models of wheezing illness development and/or outcomes with the ISAAC-WM, this approach could lead to stronger tests of theory. We provided programming code for both lavaan (open-source) and Mplus (proprietary) in online appendices to facilitate implementation in datasets utilizing the ISAAC-WM.

Despite its widespread use, this was the first study, to our knowledge, to test a model of wheezing illness severity using the ISAAC-WM. Our model was founded in the hypothesis that wheezing severity can be conceptualized as a continuous unidimensional latent variable and that the categorical ISAAC-WM items are surrogates, capturing measurable consequences of this unobserved variable. In each cohort, the χ^2_{GOF} test, which is highly sensitive to model misspecifications with large sample sizes [66,67], found no significant evidence of discrepancies between the covariance matrixes implied by our proposed model of wheezing severity and the observed covariance matrixes.

This was also the first study to explicitly evaluate whether a model of pediatric wheezing severity was consistent over time (longitudinal invariance) and across child sexes. Our Cohort 1 analyses suggested that the factor structure (i.e., the relationship between the latent factor and its measured indicators) was stable over the first two years of life, allowing for meaningful evaluation of change in pediatric wheezing severity. Although males and females differed in their mean levels of wheezing severity, there was no evidence that the latent factor structure was inconsistent across child sexes. This means that a single model of wheezing severity in early childhood for girls and boys may be sufficient as long as analyses account for mean differences across sexes.

Finally, we demonstrated the benefits of using a continuous latent factor model within a structural equation modeling framework compared to a discrete wheezing severity variable in terms of its power to detect associations with future clinical outcomes. Going from the lowest to highest levels of the latent factor was associated with a greater increase in the estimated probability of future wheezing morbidity outcomes compared to going from the lowest to highest levels of a discrete severity variable. Thus, the latent factor approach may yield more powerful tests of etiological associations than categorical severity variables.

While there are a number of strengths of this study, including evaluation of the latent variable approach in three independent datasets, there are several limitations which should be considered. First, we relied exclusively on parent report for the measurement of both wheezing symptoms and validity outcomes; though the ISAAC-WM was designed and validated for this purpose. Second, we tested our factor model using just four ordinal indicators. More precise measurements of the indicators (e.g., exact counts of wheezing episodes) should result in better severity estimation. Greater precision might also be achieved by combining the ISAAC-WM items with other indicators of wheezing severity, assuming that they are theoretically and empirically commensurate with the proposed model. Third, our models were unidimensional and only captured variability in wheezing symptoms. Prior studies suggest that asthma is a multidimensional construct [16–18], with wheezing severity represented in one of the dimensions. Future studies should explore whether the proposed latent wheezing severity model can be combined with other symptom dimensions to provide a more comprehensive representation of asthma severity [68,69]. Fourth, the latent factor model does not discriminate as well among children who are low on the severity spectrum compared to those with moderate to severe symptoms. Adding indicators that better distinguish between children with mild

symptoms would be desirable in future studies. Lastly, it is important to note that SEM generally requires large samples and may not be feasible in smaller studies [70].

Conclusions

In conclusion, the proposed latent variable model of pediatric wheezing illness estimates wheezing severity as a continuous construct, is consistent with data from three independent cohorts, and is prospectively associated with asthma morbidity. This modeling approach can be applied in cohorts with ISAAC-WM data with a sufficient sample size by adapting the provided code. Using a latent severity approach provides more powerful testing of etiological hypotheses.

Supporting information

S1 Fig. Cohort 1: Associations between the latent wheezing severity factor and the ISAAC-WM indicators. The probability of being in each severity category on the ISAAC-WM severity items in year-one as a function of the latent wheezing illness factor. The probability of being in more severe categories increases with increasing levels of the latent factor for each ISAAC-WM item.
(TIFF)

S2 Fig. Cohort 1 INSPIRE item information curves. This plot shows item information curves for each ISAAC-WM observed wheezing severity indicator at year one (panel a) and year-two (panel b). The reciprocal of the variance in the estimates (i.e., information) is plotted against levels of the latent factor, with higher levels of information indicating greater measurement precision. The latent factors are measured with greater precision at moderate to high levels of wheezing severity and are less good at distinguishing between individuals with low severity wheezing illness.
(TIFF)

S1 Appendix. Model code and covariance matrixes for Cohort 1 analyses.
(DOCX)

S2 Appendix. Model code and covariance matrixes for Cohort 2 analyses.
(DOCX)

S3 Appendix. Model code and covariance matrixes for Cohort 3 analyses.
(DOCX)

S1 File. Data to reproduce Cohort 1 and Cohort 3 models. This zip file contains R data files with following information for all Cohort 1 and Cohort 3 models: observed covariance matrixes, thresholds, weighted least squares weight matrixes (the "wls.v" matrix in lavaan), sample asymptotic variance-covariance matrixes (the "gamma" matrix in lavaan), and sample sizes (n) associated with each model. As all models were covariance-structure models, these data should be sufficient to allow others to reproduce our results. Matrixes for each model are stored as lists within the R environment. The accompanying S1Code.R file provides R code for extracting summary statistics for each model.
(ZIP)

Author Contributions

Conceptualization: Steven M. Brunwasser, Kedir N. Turi, Cosby A. Stone, Jr., Tina V. Hartert.

Data curation: Tebeb Gebretsadik, Soma Datta.

Formal analysis: Steven M. Brunwasser.

Funding acquisition: Diane R. Gold, James E. Gern, Tina V. Hartert.

Methodology: Steven M. Brunwasser, Kedir N. Turi.

Project administration: Diane R. Gold, Tina V. Hartert.

Resources: Diane R. Gold, James E. Gern, Tina V. Hartert.

Supervision: Tina V. Hartert.

Validation: Tebeb Gebretsadik.

Writing – original draft: Steven M. Brunwasser.

Writing – review & editing: Tebeb Gebretsadik, Diane R. Gold, Kedir N. Turi, Cosby A. Stone, Jr., Soma Datta, James E. Gern, Tina V. Hartert.

References

1. Barnett S, Nurmagambetov T. Costs of asthma in the United States: 2002–2007. J Allergy Clin Immunol. 2011; 127: 145–152. https://doi.org/10.1016/j.jaci.2010.10.020 PMID: 21211649
2. Pearson W, Goates S, Harrykissoon S, Miller S. State-based medicaid costs for pediatric asthma emergency department visits. Prev Chronic Dis. 2014; 11: E108. https://doi.org/10.5888/pcd11.140139 PMID: 24967830
3. Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, et al. International Study of Asthma and Allergies in Childhood (ISAAC): Rationale and methods. Eur Respir J. 1995; 8: 483–491. PMID: 7789502
4. Asher I. ISAAC International Study of Asthma and Allergies in Childhood. Pediatr Pulmonol. 2007; 42: 100. https://doi.org/10.1002/ppul.20525 PMID: 17123321
5. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Worldwide variations in the prevalence of asthma symptoms: the International Study of Asthma and Allergies in Childhood (ISAAC). Eur Respir J. 1998; 12: 315–335. https://doi.org/10.1183/09031936.98.1202031 PMID: 9727780
6. Asher M, Montefort S, Björkstén B, Lai C, Strachan D, Weiland S, et al. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. The Lancet. 2006; 368: 733–743. https://doi.org/10.1016/S0140-6736(06)69283-0
7. Correia Junior M, Sarinho E, Rizzo J, Sarinho S. Lower prevalence and greater severity of asthma in hot and dry climate. J Pediatr (Rio J). 2017; 93: 148–155. https://doi.org/10.1016/j.jped.2016.05.006 PMID: 27500595
8. Ellwood P, Asher M, Billo N, Bissell K, Chiang C-Y, Ellwood E, et al. The Global Asthma Network rationale and methods for Phase I global surveillance: Prevalence, severity, management and risk factors. Eur Respir J. 2017; 49: 1601605. https://doi.org/10.1183/13993003.01605-2016 PMID: 28077477
9. Magzoub A, Musa O, Elsony A, Alawad A, Dawod OY. Validation of the modified International Study of Asthma and Allergies in Childhood questionnaire: Is wheeze alone enough for determination of asthma symptoms prevalence? Int J Med Sci Public Health. 2017; 6. Available: http://www.ejmanager.com/mnstemps/67/67-1478141266.pdf
10. Toledo MF, Saraiva-Romanholo BM, Oliveira RC, Saldiva PHN, Silva LFF, Nascimento LFC, et al. Changes over time in the prevalence of asthma, rhinitis and atopic eczema in adolescents from Taubaté, São Paulo, Brazil (2005–2012): Relationship with living near a heavily travelled highway. Allergol Immunopathol (Madr). 2016; 44: 439–444. https://doi.org/10.1016/j.aller.2016.02.006 PMID: 27395326
11. Oluwole O, Arinola G, Huo D, Olopade C. Household biomass fuel use, asthma symptoms severity, and asthma underdiagnosis in rural schoolchildren in Nigeria: a cross-sectional observational study. BMC Pulm Med. 2017; 17: 3. https://doi.org/10.1186/s12890-016-0352-8 PMID: 28056916
12. Taylor A, West S, Aiken L. Loss of power in logistic, ordinal logistic, and probit regression when an outcome variable Is coarsely categorized. Educ Psychol Meas. 2006; 66: 228–239. https://doi.org/10.1177/0013164405278580

13. O'Byrne PM, Reddel HK, Eriksson G, Östlund O, Peterson S, Sears MR, et al. Measuring asthma control: a comparison of three classification systems. Eur Respir J. 2010; 36: 269–276. https://doi.org/10.1183/09031936.00124009 PMID: 20110397

14. Horner S, Kieckhefer G, Fouladi R. Measuring asthma severity: instrument refinement. J Asthma. 2006; 43: 533–538. https://doi.org/10.1080/02770900600857192 PMID: 16939994

15. Bailey W, Higgins D, Richards B, Richards J. Asthma severity: a factor analytic investigation. Am J Med. 1992; 93: 263–269. https://doi.org/10.1016/0002-9343(92)90231-Y PMID: 1524077

16. Rosier MJ, Bishop J, Nolan T, Robertson CF, Carlin JB, Phelan PD. Measurement of functional severity of asthma in children. Am J Respir Crit Care Med. 1994; 149: 1434–1441. https://doi.org/10.1164/ajrccm.149.6.8004295 PMID: 8004295

17. Wildfire J, Gergen P, Sorkness C, Mitchell H, Calatroni A, Kattan M, et al. Development and validation of the Composite Asthma Severity Index—an outcome measure for use in children and adolescents. J Allergy Clin Immunol. 2012; 129: 694–701. https://doi.org/10.1016/j.jaci.2011.12.962 PMID: 22244599

18. Cloutier M, Schatz M, Castro M, Clark N, Kelly H, Mangione-Smith R, et al. Asthma outcomes: Composite scores of asthma control. J Allergy Clin Immunol. 2012; 129: S24–S33. https://doi.org/10.1016/j.jaci.2011.12.980 PMID: 22386507

19. Larkin EK, Gebretsadik T, Moore ML, Anderson LJ, Dupont WD, Chappell JD, et al. Objectives, design and enrollment results from the Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure Study (INSPIRE). BMC Pulm Med. 2015; 15: 45. https://doi.org/10.1186/s12890-015-0040-0 PMID: 26021723

20. Gold D, Burge H, Carey V, Milton D, Platts-Mills T, Weiss S. Predictors of repeated wheeze in the first year of life: the relative roles of cockroach, birth weight, acute lower respiratory illness, and maternal smoking. Am J Respir Crit Care Med. 1999; 160: 227–236. https://doi.org/10.1164/ajrccm.160.1.9807104 PMID: 10390405

21. Schmidt C. Growing a new study: environmental influences on child health outcomes. Environ Health Perspect. 2015; 123: A260–A263. https://doi.org/10.1289/ehp.123-A260 PMID: 26421459

22. Bollen KA. Structural equations with latent variables. Oxford, England: John Wiley & Sons; 1989.

23. Bagozzi R, Yi Y. Specification, evaluation, and interpretation of structural equation models. J Acad Mark Sci. 2012; 40: 8–34. https://doi.org/10.1007/s11747-011-0278-x

24. Bentler PM, Stein JA. Structural equation models in medical research. Stat Methods Med Res. 1992; 1: 159–181. https://doi.org/10.1177/096228029200100203 PMID: 1341656

25. Bollen KA. Latent variables in psychology and the social sciences. Annu Rev Psychol. 2002; 53: 605–634. https://doi.org/10.1146/annurev.psych.53.100901.135239 PMID: 11752498

26. Bollen K, Bauldry S. Three Cs in measurement models: Causal indicators, composite indicators, and covariates. Psychol Methods. 2011; 16: 265–284. http://dx.doi.org.proxy.library.vanderbilt.edu/10.1037/a0024448 PMID: 21767021

27. Cole D, Preacher K. Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. Psychol Methods. 2014; 19: 300–315. http://dx.doi.org/10.1037/a0033805 PMID: 24079927

28. Rosas-Salazar C, Shilts M, Tovchigrechko A, Schobel S, Chappell J, Larkin E, et al. Differences in the nasopharyngeal microbiome during acute respiratory tract infection with human rhinovirus and respiratory syncytial virus in infancy. J Infect Dis. 2016; 214: 1924–1928. https://doi.org/10.1093/infdis/jiw456 PMID: 27923952

29. Tse S, Coull B, Sordillo J, Datta S, Gold D. Gender- and age-specific risk factors for wheeze from birth through adolescence. Pediatr Pulmonol. 2015; 50: 955–962. https://doi.org/10.1002/ppul.23113 PMID: 25348842

30. Tse S, Rifas-Shiman S, Coull B, Litonjua A, Oken E, Gold D. Sex-specific risk factors for childhood wheeze and longitudinal phenotypes of wheeze. J Allergy Clin Immunol. 2016; 138: 1561–1568.e6. https://doi.org/10.1016/j.jaci.2016.04.005 PMID: 27246527

31. Ly N, Gold D, Weiss S, Celedón J. Recurrent wheeze in early childhood and asthma among children at risk for atopy. Pediatrics. 2006; 117: e1132–e1138. https://doi.org/10.1542/peds.2005-2271 PMID: 16740815

32. Muthén BO. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika. 1984; 49: 115–132.

33. Muthén BO. Mplus Technical Appendices. [Internet]. Los Angeles, CA; 1998 2004. Available: http://statmodel.com/download/techappen.pdf

34. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. J Stat Softw. 2012; 48: 1–36.

35. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available: http://www.R-project.org/

36. Muthén L, Muthén B. Mplus user's guide. Seventh edition. Los Angeles, CA: Muthén & Muthén; 1998.

37. Asparouhov T, Muthén BO. Weighted least squares estimation with missing data. Mplus Tech Append. 2010; 1–10.

38. Bauer DJ. Estimating multilevel linear models as structural equation models. J Educ Behav Stat. 2003; 28: 135–167.

39. Hox J, Maas C. Multilevel structural equation models: The limited information approach and the multivariate multilevel approach. Recent developments on structural equation models. Springer Science & Business Media; 2004. pp. 135–149. Available: http://link.springer.com/chapter/10.1007/978-1-4020-1958-6_8

40. Stapleton L, Yang J, Hancock G. Construct meaning in multilevel settings. J Educ Behav Stat. 2016; 41: 481–520. https://doi.org/10.3102/1076998616646200

41. Muthén BO, Asparouhov T. Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. Mplus Web Notes. 2002; 4: 1–22.

42. Muthén B, Du Toit S, Spisic D. Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Psychometrika. 1997; 75: 1–45.

43. Steiger J. Point estimation, hypothesis testing, and interval estimation using the RMSEA: Some comments and a reply to Hayduk and Glaser. Struct Equ Model. 2000; 7: 149–162. http://dx.doi.org/10.1207/S15328007SEM0702_1

44. Steiger J, Lind J. Statistically-based tests for the number of common factors. Paper presented at: Meeting of the Psychometric Society; 1980 May; Iowa City, IA, USA.

45. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull. 1980; 88: 588–606.

46. Rigdon E. CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. Struct Equ Model Multidiscip J. 1996; 3: 369–379. https://doi.org/10.1080/10705519609540052

47. Chen F, Curran PJ, Bollen KA, Kirby J, Paxton P. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. Sociol Methods Res. 2008; 36: 462. https://doi.org/10.1177/0049124108314720 PMID: 19756246

48. Hu L-T, Bentler P. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychol Methods. 1998; 3: 424–453. https://doi.org/10.1037/1082-989X.3.4.424

49. Castro-Rodriguez J, Forno E, Rodriguez-Martinez C, Celedón J. Risk and protective factors for childhood asthma: what is the evidence? J Allergy Clin Immunol Pract. 2016; 4: 1111–1122. https://doi.org/10.1016/j.jaip.2016.05.003 PMID: 27286779

50. Lim R, Kobzik L, Dahl M. Risk for asthma in offspring of asthmatic mothers versus fathers: a meta-analysis. PloS One. 2010; 5: e10134. https://doi.org/10.1371/journal.pone.0010134 PMID: 20405032

51. Burke H, Leonardi-Bee J, Hashim A, Pine-Abata H, Chen Y, Cook DG, et al. Prenatal and passive smoke exposure and incidence of asthma and wheeze: systematic review and meta-analysis. Pediatrics. 2012; 129: 735–744. https://doi.org/10.1542/peds.2011-2196 PMID: 22430451

52. Moncrief T, Beck A, Simmons J, Huang B, Kahn R. Single parent households and increased child asthma morbidity. J Asthma Off J Assoc Care Asthma. 2014; 51: 260–266. https://doi.org/10.3109/02770903.2013.873806 PMID: 24320709

53. Thakur N, Oh S, Nguyen E, Martin M, Roth L, Galanter J, et al. Socioeconomic status and childhood asthma in urban minority uouths. the GALA II and SAGE II studies. Am J Respir Crit Care Med. 2013; 188: 1202–1209. https://doi.org/10.1164/rccm.201306-1016OC PMID: 24050698

54. Thavagnanam S, Fleming J, Bromley A, Shields MD, Cardwell CR. A meta-analysis of the association between Caesarean section and childhood asthma. Clin Exp Allergy. 2008; 38: 629–633. https://doi.org/10.1111/j.1365-2222.2007.02780.x PMID: 18352976

55. Geier D, Kern J, Geier M. Demographic and neonatal risk factors for childhood asthma in the USA. J Matern Fetal Neonatal Med. 2017;Published online ahead of print. https://doi.org/10.1080/14767058.2017.1393068 PMID: 29034748

56. Liu Y, Millsap RE, West SG, Tein J-Y, Tanaka R, Grimm KJ. Testing measurement invariance in longitudinal data with ordered-categorical measures. Psychol Methods. 2016;Advance online publication: 1–21. http://dx.doi.org/10.1037/met0000075

57. Vandenberg R, Lance C. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organ Res Methods. 2000; 3: 4–70. https://doi.org/10.1177/109442810031002

58. Little T, Preacher K, Selig J, Card N. New developments in latent variable panel analyses of longitudinal data. Int J Behav Dev. 2007; 31: 357–365. https://doi.org/10.1177/0165025407077757

59. Satorra A. Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In: Heijmans RDH, Pollock DSG, Satorra A, editors. Innovations in multivariate statistical analysis: A festschrift for Heinz Neudecker. Dordrecht, The Netherlands: Springer Science+Business Media; 2000. pp. 233–247. https://doi.org/10.1007/978-1-4615-4603-0_17

60. Hirschfeld G, Von Brachel R. Multiple-Group confirmatory factor analysis in R: a tutorial in measurement invariance with continuous and ordinal indicators. Pract Assess Res Eval. 2014; 19. Available: http://pareonline.net/getvn.asp?v=19&n=7

61. Millsap R, Tein J-Y. Assessing factorial invariance in ordered-categorical measures. Multivar Behav Res. 2004; 39: 479–515. https://doi.org/10.1207/S15327906MBR3903_4

62. Zein J, Udeh B, Teague W, Koroukian S, Schlitz N, Bleecker E, et al. Impact of age and sex on outcomes and hospital cost of acute asthma in the United States, 2011–2012. PLOS ONE. 2016; 11: e0157301. https://doi.org/10.1371/journal.pone.0157301 PMID: 27294365

63. Castro-Rodríguez J, Holberg C, Wright A, Martinez F. A clinical index to define risk of asthma in young children with recurrent wheezing. Am J Respir Crit Care Med. 2000; 162: 1403–1406. https://doi.org/10.1164/ajrccm.162.4.9912111 PMID: 11029352

64. Liu A, Zeiger R, Sorkness C, Mahr T, Ostrom N, Burgess S, et al. Development and cross-sectional validation of the Childhood Asthma Control Test. J Allergy Clin Immunol. 2007; 119: 817–825. https://doi.org/10.1016/j.jaci.2006.12.662 PMID: 17353040

65. Rigdon E. Rethinking partial least squares path modeling: In praise of simple methods. Long Range Plann. 2012; 45: 341–358. https://doi.org/10.1016/j.lrp.2012.09.010

66. Bentler P. On tests and indices for evaluating structural models. Personal Individ Differ. 2007; 42: 825–829. https://doi.org/10.1016/j.paid.2006.09.024

67. Steiger J. Understanding the limitations of global fit assessment in structural equation modeling. Personal Individ Differ. 2007; 42: 893–898. https://doi.org/10.1016/j.paid.2006.09.017

68. Nathan R, Sorkness C, Kosinski M, Schatz M, Li J, Marcus P, et al. Development of the asthma control test: A survey for assessing asthma control. J Allergy Clin Immunol. 2004; 113: 59–65. https://doi.org/10.1016/j.jaci.2003.09.008 PMID: 14713908

69. Cockcroft D, Swystun V. Asthma control versus asthma severity. J Allergy Clin Immunol. 1996; 98: 1016–1018. PMID: 8977499

70. Westland J. Lower bounds on sample size in structural equation modeling. Electron Commer Res Appl. 2010; 9: 476–487. https://doi.org/10.1016/j.elerap.2010.07.003

S1 Appendix: Cohort 1 Analyses

Table of Contents

Contents 2
Variable key 2
Model 1.1 3
m1.1: lavaan code 3
m1.1: Sample statistics (observed covariance matrix and thresholds) 3
m1.1: Model implied covariance matrix 3
m1.1: Residual correlation matrix for model 1.1 4
m1.1: Weight vector used in weighted least squares estimation for model 1.1 4
Model 1.2 5
m1.2: lavaan code 5
m1.2: Sample statistics (observed covariance matrix and thresholds) 5
m1.2: Model implied covariance matrix 5
m1.2: Residual correlation matrix for model 1.2 5
m1.2: Weight vector used in weighted least squares estimation for model 1.2 6
Model 1.3 6
m1.3: lavaan code 6
m1.3: Sample statistics (observed covariance matrix and thresholds) 7
m1.3: Model implied covariance matrix 7
m1.3: Residual correlation matrix for model 1.3 8
Model 1.4 8
m1.4: lavaan code 8
m1.4: Sample statistics (observed covariance matrix and thresholds) 8
m1.4: Model implied covariance matrix 9
m1.4: Residual correlation matrix for model 1.4 9
Model 1.5 10
m1.5: lavaan code 10
m1.5: Sample statistics (observed covariance matrix and thresholds) 10
m1.5: Model implied covariance matrix 11
m1.5: Residual correlation matrix for model 1.5 11
Model 1.6 11

m1.6: lavaan code 11
m1.6: Sample statistics (observed covariance matrix and thresholds) 11
m1.6: Model implied covariance matrix 12
m1.6: Residual correlation matrix for model 1.6 12
Model 1.7 13
m1.7: lavaan code 13
m1.7: Sample statistics (observed covariance matrix and thresholds) 13
m1.7: Model implied covariance matrix 14
m1.7: Residual correlation matrix for model 1.7 14
Model 1.8 14
m1.8: lavaan code 14
m1.8: Sample statistics (observed covariance matrix and thresholds) 15
m1.8: Model implied covariance matrix 15
m1.8: Residual correlation matrix for model 1.8 16
Model 1.9 16
m1.9: lavaan code 16
m1.9: Sample statistics (observed covariance matrix and thresholds) 16
m1.9: Model implied covariance matrix 17
m1.9: Residual correlation matrix for model 1.9 17

Contents

This file provides supplemental modeling information for the Cohort 1 analyses. We provide *lavaan* code for the primary Cohort 1 models (see Table 4 in the primary text). Additionally, we provide observed and model-implied covariance marixes for each model, as well as the residual correlation matrix. For models 1.1 and 1.2, we provide the weight matrix used in Weighted Least Squares estimation.

Variable key

Below we define the observed variables used in the analyses. For detailed information about *lavaan* modeling code, see the lavaan website.

- dnwhez12 = ISAAC-WM item from the one-year follow-up assessment asking about the frequency of wheezing attacks in the past year (4-level ordinal variable)

- dawake12 = ISAAC-WM item from the one-year follow-up assessment asking about the frequency of wheeze-induced sleep disturbance in the past year (3-level ordinal variable)
- dspeech12 = ISAAC-WM item from the one-year follow-up assessment asking about the frequency of wheeze-induced speech disturbance in the past year (2-level ordinal variable)
- dexwhez12 = ISAAC-WM item from the one-year follow-up assessment asking whether the child has experienced exercise-induced wheeze

Model 1.1

m1.1: lavaan code

```
m1.1syn <- '
whzsev12 =~ dnwhez12 + dawake12 + dspeech12 + dexwhez12
'

m1.1 <- cfa( m1.1syn, data = insp.isaac3, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise' )
```

m1.1: Sample statistics (observed covariance matrix and thresholds)

```
## $cov
##            dnwh12 dawk12 dspc12 dxwh12
## dnwhez12   1.000
## dawake12   0.590  1.000
## dspeech12  0.539  0.592  1.000
## dexwhez12  0.455  0.465  0.559  1.000
##
## $mean
##  dnwhez12  dawake12 dspeech12 dexwhez12
##         0         0         0         0
##
## $th
##  dnwhez12|t1  dnwhez12|t2  dnwhez12|t3  dawake12|t1  dawake12|t2
##       -0.458        0.997        1.704        0.576        1.030
## dspeech12|t1 dexwhez12|t1
##        0.903        1.258
```

m1.1: Model implied covariance matrix

```
##            dnwh12 dawk12 dspc12 dxwh12
## dnwhez12   1.000
## dawake12   0.563  1.000
```

```
## dspeech12 0.576  0.602  1.000
## dexwhez12 0.475  0.497  0.509  1.000
```

m1.1: Residual correlation matrix for model 1.1

```
##            dnwh12 dawk12 dspc12 dxwh12
## dnwhez12    0.000
## dawake12    0.027  0.000
## dspeech12 -0.036 -0.009  0.000
## dexwhez12 -0.021 -0.032  0.051  0.000
```

m1.1: Weight vector used in weighted least squares estimation for model 1.1

```
inspect( m1.1, 'wls.v' )

##       [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11]
##  [1,] 0.58 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00
##  [2,] 0.00 0.432 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00
##  [3,] 0.00 0.000 0.203 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00
##  [4,] 0.00 0.000 0.000 0.554 0.000 0.000 0.000 0.000 0.000 0.000  0.00
##  [5,] 0.00 0.000 0.000 0.000 0.421 0.000 0.000 0.000 0.000 0.000  0.00
##  [6,] 0.00 0.000 0.000 0.000 0.000 0.462 0.000 0.000 0.000 0.000  0.00
##  [7,] 0.00 0.000 0.000 0.000 0.000 0.000 0.349 0.000 0.000 0.000  0.00
##  [8,] 0.00 0.000 0.000 0.000 0.000 0.000 0.000 0.554 0.000 0.000  0.00
##  [9,] 0.00 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.247 0.000  0.00
## [10,] 0.00 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.336  0.00
## [11,] 0.00 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.58
## [12,] 0.00 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00
## [13,] 0.00 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  0.00
##       [,12] [,13]
##  [1,] 0.000  0.00
##  [2,] 0.000  0.00
##  [3,] 0.000  0.00
##  [4,] 0.000  0.00
##  [5,] 0.000  0.00
##  [6,] 0.000  0.00
##  [7,] 0.000  0.00
##  [8,] 0.000  0.00
##  [9,] 0.000  0.00
## [10,] 0.000  0.00
## [11,] 0.000  0.00
## [12,] 0.331  0.00
## [13,] 0.000  0.33
```

Model 1.2

m1.2: lavaan code

```
m1.2syn <- '
whzsev24 =~ dnwhez24 + dawake24 + dspeech24 + dexwhez24
'

m1.2 <- cfa( m1.2syn, data = insp.isaac3, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise' )
```

m1.2: Sample statistics (observed covariance matrix and thresholds)

```
## $cov
##            dnwh24 dawk24 dspc24 dxwh24
## dnwhez24   1.000
## dawake24   0.672  1.000
## dspeech24  0.599  0.624  1.000
## dexwhez24  0.594  0.471  0.387  1.000
##
## $mean
##  dnwhez24  dawake24 dspeech24 dexwhez24
##         0         0         0         0
##
## $th
##  dnwhez24|t1  dnwhez24|t2  dnwhez24|t3  dawake24|t1  dawake24|t2
##       -0.529        0.924        1.845        0.365        0.916
## dspeech24|t1 dexwhez24|t1
##        0.872        1.166
```

m1.2: Model implied covariance matrix

```
##            dnwh24 dawk24 dspc24 dxwh24
## dnwhez24   1.000
## dawake24   0.690  1.000
## dspeech24  0.630  0.585  1.000
## dexwhez24  0.537  0.499  0.456  1.000
```

m1.2: Residual correlation matrix for model 1.2

```
##            dnwh24 dawk24 dspc24 dxwh24
## dnwhez24    0.000
## dawake24   -0.018  0.000
## dspeech24  -0.031  0.039  0.000
## dexwhez24   0.057 -0.028 -0.069  0.000
```

m1.2: Weight vector used in weighted least squares estimation for model 1.2

```
inspect( m1.2, 'wls.v' )

##       [,1]  [,2] [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11]
##  [1,] 0.48 0.000 0.00 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00
##  [2,] 0.00 0.387 0.00 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00
##  [3,] 0.00 0.000 0.14 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00
##  [4,] 0.00 0.000 0.00 0.507 0.000 0.000 0.000 0.000 0.000  0.00  0.00
##  [5,] 0.00 0.000 0.00 0.000 0.389 0.000 0.000 0.000 0.000  0.00  0.00
##  [6,] 0.00 0.000 0.00 0.000 0.000 0.401 0.000 0.000 0.000  0.00  0.00
##  [7,] 0.00 0.000 0.00 0.000 0.000 0.000 0.381 0.000 0.000  0.00  0.00
##  [8,] 0.00 0.000 0.00 0.000 0.000 0.000 0.000 0.614 0.000  0.00  0.00
##  [9,] 0.00 0.000 0.00 0.000 0.000 0.000 0.000 0.000 0.253  0.00  0.00
## [10,] 0.00 0.000 0.00 0.000 0.000 0.000 0.000 0.000 0.000  0.41  0.00
## [11,] 0.00 0.000 0.00 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.55
## [12,] 0.00 0.000 0.00 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00
## [13,] 0.00 0.000 0.00 0.000 0.000 0.000 0.000 0.000 0.000  0.00  0.00
##       [,12] [,13]
##  [1,] 0.000 0.000
##  [2,] 0.000 0.000
##  [3,] 0.000 0.000
##  [4,] 0.000 0.000
##  [5,] 0.000 0.000
##  [6,] 0.000 0.000
##  [7,] 0.000 0.000
##  [8,] 0.000 0.000
##  [9,] 0.000 0.000
## [10,] 0.000 0.000
## [11,] 0.000 0.000
## [12,] 0.351 0.000
## [13,] 0.000 0.243
```

Model 1.3

m1.3: lavaan code

```
m1.3syn <- '
whzsev12 =~ dnwhez12 + dawake12 + dspeech12 + dexwhez12
whzsev12 ~ masthma + msmokepreg + medicaid + cesarean + male + married + cbla
ck + cother
resphosp12 ~ whzsev12 + masthma + msmokepreg + medicaid + cesarean + male +
             married + cblack + cother
'
```

```
m1.3 <- cfa( m1.3syn, data = insp.isaac3, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise' )
```

m1.3: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##            dnwh12 dawk12 dspc12 dxwh12 rsph12
## dnwhez12   1.000
## dawake12   0.590  1.000
## dspeech12  0.541  0.595  1.000
## dexwhez12  0.460  0.473  0.566  1.000
## resphosp12 0.300  0.299  0.481  0.422  1.000
##
## $res.int
##   dnwhez12   dawake12  dspeech12  dexwhez12 resphosp12
##          0          0          0          0          0
##
## $res.th
##   dnwhez12|t1   dnwhez12|t2   dnwhez12|t3   dawake12|t1   dawake12|t2
##        -0.317         1.162         1.882         0.705         1.164
##  dspeech12|t1  dexwhez12|t1 resphosp12|t1
##         0.551         1.243         1.147
##
## $res.slopes
##            masthm msmkpr medicd cesarn   male marrid cblack cother
## dnwhez12    0.225  0.110  0.118  0.139  0.113 -0.117 -0.128 -0.163
## dawake12    0.017  0.202  0.156  0.015 -0.003  0.035 -0.070 -0.174
## dspeech12   0.062  0.174 -0.177 -0.092 -0.067 -0.306 -0.155 -0.439
## dexwhez12   0.030  0.028 -0.063  0.037 -0.006 -0.123  0.285 -0.302
## resphosp12 -0.081  0.299  0.093 -0.102 -0.098 -0.016  0.377 -0.157
##
## $cov.x
##            masthm msmkpr medicd cesarn male   marrid cblack cother
## masthma     0.180
## msmokepreg  0.007  0.167
## medicaid    0.001  0.052  0.245
## cesarean    0.007 -0.001 -0.024  0.225
## male        0.005 -0.006 -0.005  0.000  0.236
## married    -0.008 -0.058 -0.142  0.011  0.002  0.247
## cblack      0.012 -0.001  0.067 -0.004 -0.001 -0.086  0.190
## cother      0.004 -0.005  0.010 -0.006 -0.003 -0.012 -0.002  0.074
```

m1.3: Model implied covariance matrix

```
##            dnwh12 dawk12 dspc12 dxwh12 rsph12
## dnwhez12   1.000
## dawake12   0.497  1.000
```

```
## dspeech12  0.696  0.584  1.000
## dexwhez12  0.507  0.425  0.595  1.000
## resphosp12 0.391  0.328  0.459  0.334  1.000
```

m1.3: Residual correlation matrix for model 1.3

```
##            dnwh12 dawk12 dspc12 dxwh12 rsph12
## dnwhez12    0.000
## dawake12    0.093  0.000
## dspeech12  -0.155  0.011  0.000
## dexwhez12  -0.047  0.047 -0.030  0.000
## resphosp12 -0.090 -0.029  0.022  0.088  0.000
```

Model 1.4

m1.4: lavaan code

```
m1.4syn <- '
whzsev24 =~ dnwhez24 + dawake24 + dspeech24 + dexwhez24
whzsev24 ~ masthma + msmokepreg + medicaid + cesarean + male + married + cbla
ck + cother
resphosp24 ~ whzsev24 + masthma + msmokepreg + medicaid + cesarean + male + m
arried + cblack + cother
'

m1.4 <- cfa( m1.4syn, data = insp.isaac3, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise' )
```

m1.4: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##            dnwh24 dawk24 dspc24 dxwh24 rsph24
## dnwhez24   1.000
## dawake24   0.664  1.000
## dspeech24  0.610  0.631  1.000
## dexwhez24  0.576  0.457  0.437  1.000
## resphosp24 0.271  0.254  0.539  0.344  1.000
##
## $res.int
##   dnwhez24   dawake24  dspeech24  dexwhez24 resphosp24
##          0          0          0          0          0
##
## $res.th
##   dnwhez24|t1   dnwhez24|t2   dnwhez24|t3   dawake24|t1   dawake24|t2
##        -0.357         1.128         2.083         0.577         1.139
```

```
##  dspeech24|t1  dexwhez24|t1 resphosp24|t1
##         1.006         1.470         1.753
##
## $res.slopes
##            masthm msmkpr medicd cesarn   male marrid cblack cother
## dnwhez24    0.402  0.058 -0.004  0.102  0.145 -0.094  0.063 -0.155
## dawake24    0.276 -0.175  0.114 -0.015  0.169 -0.052  0.147 -0.001
## dspeech24   0.373 -0.241 -0.044  0.000  0.374 -0.134 -0.252 -0.185
## dexwhez24   0.230  0.016  0.102  0.149  0.063 -0.187  0.506  0.185
## resphosp24 -0.106 -0.020  0.390  0.141 -0.084  0.081  0.150 -0.600
##
## $cov.x
##            masthm msmkpr medicd cesarn male   marrid cblack cother
## masthma     0.180
## msmokepreg  0.007  0.167
## medicaid    0.001  0.052  0.245
## cesarean    0.007 -0.001 -0.024  0.225
## male        0.005 -0.006 -0.005  0.000  0.236
## married    -0.008 -0.058 -0.142  0.011  0.002  0.247
## cblack      0.012 -0.001  0.067 -0.004 -0.001 -0.086  0.190
## cother      0.004 -0.005  0.010 -0.006 -0.003 -0.012 -0.002  0.074
```

m1.4: Model implied covariance matrix

```
##            dnwh24 dawk24 dspc24 dxwh24 rsph24
## dnwhez24   1.000
## dawake24   0.541  1.000
## dspeech24  0.738  0.621  1.000
## dexwhez24  0.550  0.463  0.632  1.000
## resphosp24 0.363  0.306  0.418  0.311  1.000
```

m1.4: Residual correlation matrix for model 1.4

```
##            dnwh24 dawk24 dspc24 dxwh24 rsph24
## dnwhez24    0.000
## dawake24    0.123  0.000
## dspeech24  -0.128  0.010  0.000
## dexwhez24   0.026 -0.007 -0.195  0.000
## resphosp24 -0.093 -0.052  0.122  0.033  0.000
```

Model 1.5

m1.5: lavaan code

```
m1.5syn <- '
whzsev12 =~ dnwhez12 + dawake12 + dspeech12  + dexwhez12
whzsev12 ~ masthma + msmokepreg + medicaid + cesarean + male + married + cbla
ck + cother
whzmed12 ~ whzsev12 + masthma + msmokepreg + medicaid + cesarean + male + mar
ried + cblack + cother
'

m1.5 <- cfa( m1.5syn, data = insp.isaac3, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise' )
```

m1.5: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##           dnwh12 dawk12 dspc12 dxwh12 whzm12
## dnwhez12  1.000
## dawake12  0.590  1.000
## dspeech12 0.541  0.595  1.000
## dexwhez12 0.460  0.473  0.566  1.000
## whzmed12  0.584  0.452  0.506  0.447  1.000
##
## $res.int
##  dnwhez12  dawake12 dspeech12 dexwhez12  whzmed12
##         0         0         0         0         0
##
## $res.th
##  dnwhez12|t1  dnwhez12|t2  dnwhez12|t3  dawake12|t1  dawake12|t2
##       -0.317        1.162        1.882        0.705        1.164
## dspeech12|t1 dexwhez12|t1  whzmed12|t1
##        0.551        1.243        0.616
##
## $res.slopes
##           masthm msmkpr medicd cesarn   male marrid cblack cother
## dnwhez12   0.225  0.110  0.118  0.139  0.113 -0.117 -0.128 -0.163
## dawake12   0.017  0.202  0.156  0.015 -0.003  0.035 -0.070 -0.174
## dspeech12  0.062  0.174 -0.177 -0.092 -0.067 -0.306 -0.155 -0.439
## dexwhez12  0.030  0.028 -0.063  0.037 -0.006 -0.123  0.285 -0.302
## whzmed12   0.212  0.104  0.048  0.100  0.167 -0.142 -0.261 -0.025
##
## $cov.x
##            masthm msmkpr medicd cesarn male   marrid cblack cother
## masthma     0.180
## msmokepreg  0.007  0.167
## medicaid    0.001  0.052  0.245
## cesarean    0.007 -0.001 -0.024  0.225
```

```
## male        0.005 -0.006 -0.005  0.000  0.236
## married    -0.008 -0.058 -0.142  0.011  0.002  0.247
## cblack      0.012 -0.001  0.067 -0.004 -0.001 -0.086  0.190
## cother      0.004 -0.005  0.010 -0.006 -0.003 -0.012 -0.002  0.074
```

m1.5: Model implied covariance matrix

```
##           dnwh12 dawk12 dspc12 dxwh12 whzm12
## dnwhez12  1.000
## dawake12  0.600  1.000
## dspeech12 0.727  0.575  1.000
## dexwhez12 0.542  0.428  0.519  1.000
## whzmed12  0.576  0.455  0.551  0.411  1.000
```

m1.5: Residual correlation matrix for model 1.5

```
##           dnwh12 dawk12 dspc12 dxwh12 whzm12
## dnwhez12   0.000
## dawake12  -0.010  0.000
## dspeech12 -0.186  0.020  0.000
## dexwhez12 -0.081  0.045  0.047  0.000
## whzmed12   0.008 -0.003 -0.046  0.036  0.000
```

Model 1.6

m1.6: lavaan code

```
m1.6syn <- '
whzsev24 =~ dnwhez24 + dawake24 + dspeech24 + dexwhez24
whzsev24 ~ masthma + msmokepreg + medicaid + cesarean + male + married + cbla
ck + cother
whzmed24 ~ whzsev24 + masthma + msmokepreg + medicaid + cesarean + male + mar
ried + cblack + cother
'

m1.6 <- cfa( m1.6syn, data = insp.isaac3, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise' )
```

m1.6: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##           dnwh24 dawk24 dspc24 dxwh24 whzm24
## dnwhez24  1.000
## dawake24  0.664  1.000
```

```
## dspeech24 0.610  0.631  1.000
## dexwhez24 0.576  0.457  0.437  1.000
## whzmed24  0.672  0.604  0.602  0.426  1.000
##
## $res.int
##  dnwhez24  dawake24 dspeech24 dexwhez24  whzmed24
##         0         0         0         0         0
##
## $res.th
##  dnwhez24|t1  dnwhez24|t2  dnwhez24|t3  dawake24|t1  dawake24|t2
##       -0.357        1.128        2.083        0.577        1.139
## dspeech24|t1 dexwhez24|t1  whzmed24|t1
##        1.006        1.470        0.087
##
## $res.slopes
##           masthm msmkpr medicd cesarn  male marrid cblack cother
## dnwhez24   0.402  0.058 -0.004  0.102 0.145 -0.094  0.063 -0.155
## dawake24   0.276 -0.175  0.114 -0.015 0.169 -0.052  0.147 -0.001
## dspeech24  0.373 -0.241 -0.044  0.000 0.374 -0.134 -0.252 -0.185
## dexwhez24  0.230  0.016  0.102  0.149 0.063 -0.187  0.506  0.185
## whzmed24   0.107  0.109 -0.157  0.141 0.251 -0.124 -0.102 -0.167
##
## $cov.x
##            masthm msmkpr medicd cesarn male   marrid cblack cother
## masthma     0.180
## msmokepreg  0.007  0.167
## medicaid    0.001  0.052  0.245
## cesarean    0.007 -0.001 -0.024  0.225
## male        0.005 -0.006 -0.005  0.000  0.236
## married    -0.008 -0.058 -0.142  0.011  0.002  0.247
## cblack      0.012 -0.001  0.067 -0.004 -0.001 -0.086  0.190
## cother      0.004 -0.005  0.010 -0.006 -0.003 -0.012 -0.002  0.074
```

m1.6: Model implied covariance matrix

```
##           dnwh24 dawk24 dspc24 dxwh24 whzm24
## dnwhez24  1.000
## dawake24  0.671  1.000
## dspeech24 0.708  0.622  1.000
## dexwhez24 0.556  0.488  0.515  1.000
## whzmed24  0.671  0.589  0.621  0.488  1.000
```

m1.6: Residual correlation matrix for model 1.6

```
##           dnwh24 dawk24 dspc24 dxwh24 whzm24
## dnwhez24   0.000
## dawake24  -0.007  0.000
```

```
## dspeech24 -0.098  0.009  0.000
## dexwhez24  0.020 -0.032 -0.078  0.000
## whzmed24   0.001  0.015 -0.020 -0.062  0.000
```

Model 1.7

m1.7: lavaan code

```
m1.7syn <- '
whzsev24 =~ dnwhez24 + dawake24 + dspeech24 + dexwhez24
whzsev24 ~ masthma + msmokepreg + medicaid + cesarean + male + married + cbla
ck + cother
y3asthdx ~ whzsev24 + masthma + msmokepreg + medicaid + cesarean + male + mar
ried + cblack + cother
'

m1.7 <- cfa(m1.7syn, data = insp.isaac3, estimator = 'wlsmv', ordered = 'y3as
thdx', missing = 'pairwise' )
```

m1.7: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##           dnwh24 dawk24 dspc24 dxwh24 y3sthd
## dnwhez24  1.000
## dawake24  0.664  1.000
## dspeech24 0.610  0.631  1.000
## dexwhez24 0.576  0.457  0.437  1.000
## y3asthdx  0.432  0.503  0.420  0.327  1.000
##
## $res.int
##  dnwhez24  dawake24 dspeech24 dexwhez24  y3asthdx
##         0         0         0         0         0
##
## $res.th
##  dnwhez24|t1  dnwhez24|t2  dnwhez24|t3  dawake24|t1  dawake24|t2
##       -0.357        1.128        2.083        0.577        1.139
## dspeech24|t1 dexwhez24|t1  y3asthdx|t1
##        1.006        1.470        1.595
##
## $res.slopes
##           masthm msmkpr medicd cesarn  male marrid cblack cother
## dnwhez24   0.402  0.058 -0.004  0.102 0.145 -0.094  0.063 -0.155
## dawake24   0.276 -0.175  0.114 -0.015 0.169 -0.052  0.147 -0.001
## dspeech24  0.373 -0.241 -0.044  0.000 0.374 -0.134 -0.252 -0.185
## dexwhez24  0.230  0.016  0.102  0.149 0.063 -0.187  0.506  0.185
## y3asthdx   0.425  0.007  0.121 -0.079 0.011 -0.054  0.615  0.118
```

```
##
## $cov.x
##             masthm msmkpr medicd cesarn male   marrid cblack cother
## masthma      0.180
## msmokepreg   0.007  0.167
## medicaid     0.001  0.052  0.245
## cesarean     0.007 -0.001 -0.024  0.225
## male         0.005 -0.006 -0.005  0.000  0.236
## married     -0.008 -0.058 -0.142  0.011  0.002  0.247
## cblack       0.012 -0.001  0.067 -0.004 -0.001 -0.086  0.190
## cother       0.004 -0.005  0.010 -0.006 -0.003 -0.012 -0.002  0.074
```

m1.7: Model implied covariance matrix

```
##            dnwh24 dawk24 dspc24 dxwh24 y3sthd
## dnwhez24   1.000
## dawake24   0.650  1.000
## dspeech24  0.712  0.620  1.000
## dexwhez24  0.557  0.485  0.531  1.000
## y3asthdx   0.458  0.399  0.437  0.342  1.000
```

m1.7: Residual correlation matrix for model 1.7

```
##            dnwh24 dawk24 dspc24 dxwh24 y3sthd
## dnwhez24    0.000
## dawake24    0.014  0.000
## dspeech24  -0.102  0.011  0.000
## dexwhez24   0.019 -0.029 -0.094  0.000
## y3asthdx   -0.026  0.105 -0.017 -0.015  0.000
```

Model 1.8

m1.8: lavaan code

```
m1.8syn <- '
whzsev24 =~ dnwhez24 + dawake24 + dspeech24 + dexwhez24
whzsev24 ~ masthma + msmokepreg + medicaid + cesarean + male + married + cbla
ck + cother
y3steroidso ~ whzsev24 + masthma + msmokepreg + medicaid + cesarean + male +
married + cblack + cother
'

m1.8 <- cfa(m1.8syn, data = insp.isaac3, estimator = 'wlsmv', ordered = 'y3st
eroidso', missing = 'pairwise' )
```

m1.8: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##             dnwh24 dawk24 dspc24 dxwh24 y3strd
## dnwhez24    1.000
## dawake24    0.664  1.000
## dspeech24   0.610  0.631  1.000
## dexwhez24   0.576  0.457  0.437  1.000
## y3steroidso 0.466  0.417  0.351  0.363  1.000
##
## $res.int
##    dnwhez24    dawake24   dspeech24   dexwhez24 y3steroidso
##           0           0           0           0           0
##
## $res.th
##    dnwhez24|t1    dnwhez24|t2    dnwhez24|t3    dawake24|t1    dawake24|t2
##         -0.357          1.128          2.083          0.577          1.139
##   dspeech24|t1   dexwhez24|t1 y3steroidso|t1 y3steroidso|t2
##          1.006          1.470          0.534          1.654
##
## $res.slopes
##             masthm msmkpr medicd cesarn   male marrid cblack cother
## dnwhez24     0.402  0.058 -0.004  0.102  0.145 -0.094  0.063 -0.155
## dawake24     0.276 -0.175  0.114 -0.015  0.169 -0.052  0.147 -0.001
## dspeech24    0.373 -0.241 -0.044  0.000  0.374 -0.134 -0.252 -0.185
## dexwhez24    0.230  0.016  0.102  0.149  0.063 -0.187  0.506  0.185
## y3steroidso  0.454  0.083 -0.215 -0.094 -0.017 -0.153  0.073 -0.130
##
## $cov.x
##            masthm msmkpr medicd cesarn male   marrid cblack cother
## masthma     0.180
## msmokepreg  0.007  0.167
## medicaid    0.001  0.052  0.245
## cesarean    0.007 -0.001 -0.024  0.225
## male        0.005 -0.006 -0.005  0.000  0.236
## married    -0.008 -0.058 -0.142  0.011  0.002  0.247
## cblack      0.012 -0.001  0.067 -0.004 -0.001 -0.086  0.190
## cother      0.004 -0.005  0.010 -0.006 -0.003 -0.012 -0.002  0.074
```

m1.8: Model implied covariance matrix

```
##             dnwh24 dawk24 dspc24 dxwh24 y3strd
## dnwhez24    1.000
## dawake24    0.636  1.000
## dspeech24   0.702  0.605  1.000
## dexwhez24   0.564  0.486  0.537  1.000
## y3steroidso 0.454  0.392  0.433  0.347  1.000
```

m1.8: Residual correlation matrix for model 1.8

```
##              dnwh24 dawk24 dspc24 dxwh24 y3strd
## dnwhez24      0.000
## dawake24      0.028  0.000
## dspeech24    -0.092  0.026  0.000
## dexwhez24     0.012 -0.029 -0.099  0.000
## y3steroidso  0.011  0.025 -0.081  0.015  0.000
```

Model 1.9

m1.9: lavaan code

```
m1.9syn <- '
whzsev24 =~ dnwhez24 + dawake24 + dspeech24 + dexwhez24
whzsev24 ~ masthma + msmokepreg + medicaid + cesarean + male + married + cbla
ck + cother
y3whzvisito ~ whzsev24 + masthma + msmokepreg + medicaid + cesarean + male +
married + cblack + cother
'

m1.9 <- cfa(m1.9syn, data = insp.isaac3, estimator = 'wlsmv', ordered = 'y3wh
zvisito', missing = 'pairwise' )
```

m1.9: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##              dnwh24 dawk24 dspc24 dxwh24 y3whzv
## dnwhez24     1.000
## dawake24     0.664  1.000
## dspeech24    0.610  0.631  1.000
## dexwhez24    0.576  0.457  0.437  1.000
## y3whzvisito 0.461  0.423  0.397  0.331  1.000
##
## $res.int
##    dnwhez24    dawake24   dspeech24   dexwhez24 y3whzvisito
##           0           0           0           0           0
##
## $res.th
##    dnwhez24|t1    dnwhez24|t2    dnwhez24|t3    dawake24|t1    dawake24|t2
##         -0.357          1.128          2.083          0.577          1.139
##   dspeech24|t1   dexwhez24|t1 y3whzvisito|t1 y3whzvisito|t2
##          1.006          1.470          0.447          1.610
##
## $res.slopes
```

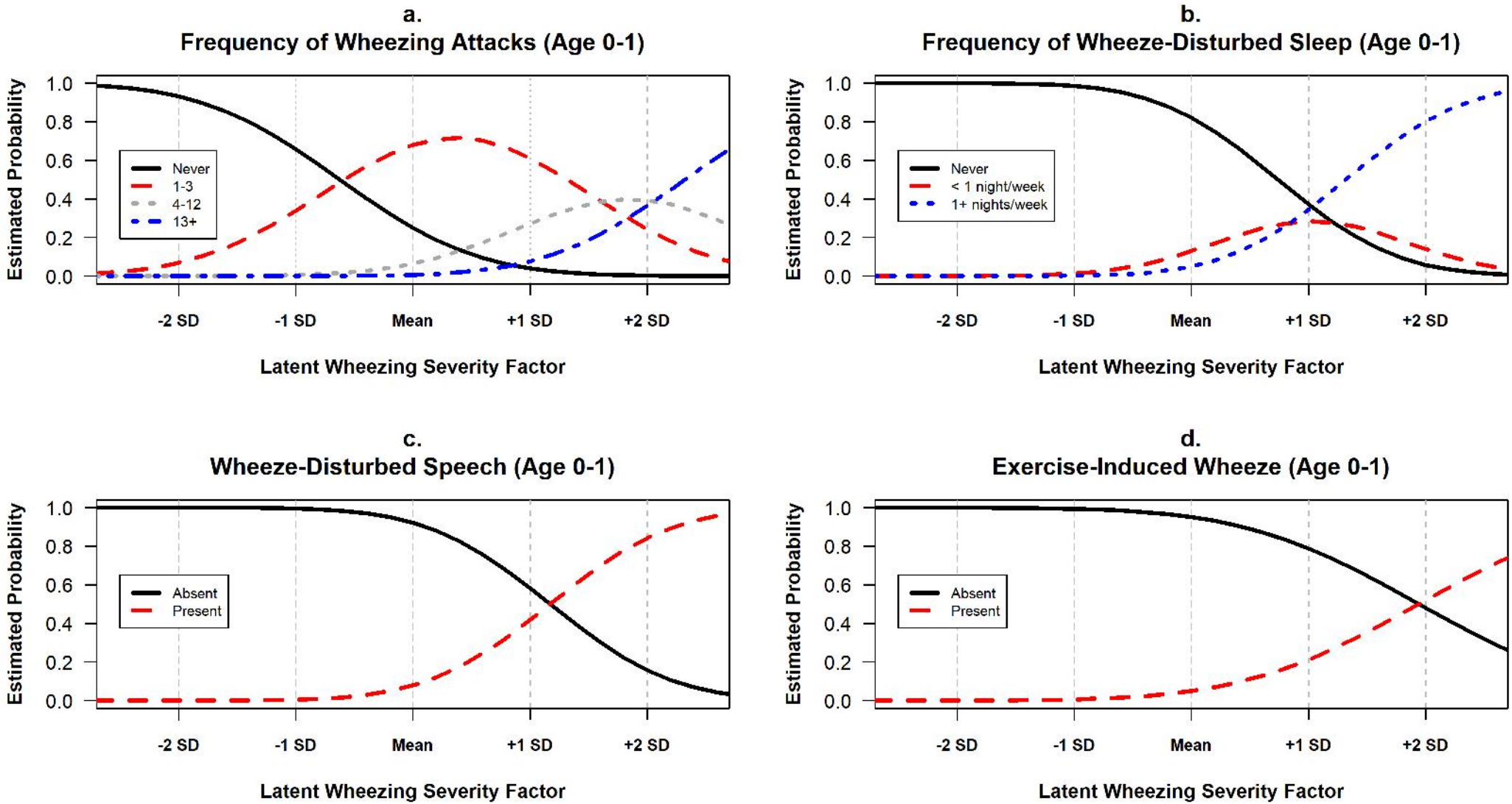
```
##              masthm msmkpr medicd cesarn  male marrid cblack cother
## dnwhez24      0.402  0.058 -0.004  0.102 0.145 -0.094  0.063 -0.155
## dawake24      0.276 -0.175  0.114 -0.015 0.169 -0.052  0.147 -0.001
## dspeech24     0.373 -0.241 -0.044  0.000 0.374 -0.134 -0.252 -0.185
## dexwhez24     0.230  0.016  0.102  0.149 0.063 -0.187  0.506  0.185
## y3whzvisito   0.375 -0.010 -0.223 -0.195 0.115 -0.345  0.231 -0.165
##
## $cov.x
##             masthm msmkpr medicd cesarn male   marrid cblack cother
## masthma      0.180
## msmokepreg   0.007  0.167
## medicaid     0.001  0.052  0.245
## cesarean     0.007 -0.001 -0.024  0.225
## male         0.005 -0.006 -0.005  0.000  0.236
## married     -0.008 -0.058 -0.142  0.011  0.002  0.247
## cblack       0.012 -0.001  0.067 -0.004 -0.001 -0.086  0.190
## cother       0.004 -0.005  0.010 -0.006 -0.003 -0.012 -0.002  0.074
```

m1.9: Model implied covariance matrix

```
##              dnwh24 dawk24 dspc24 dxwh24 y3whzv
## dnwhez24     1.000
## dawake24     0.650  1.000
## dspeech24    0.703  0.608  1.000
## dexwhez24    0.562  0.487  0.526  1.000
## y3whzvisito 0.459  0.397  0.430  0.344  1.000
```

m1.9: Residual correlation matrix for model 1.9

```
##              dnwh24 dawk24 dspc24 dxwh24 y3whzv
## dnwhez24      0.000
## dawake24      0.014  0.000
## dspeech24    -0.093  0.023  0.000
## dexwhez24     0.014 -0.030 -0.089  0.000
## y3whzvisito   0.003  0.026 -0.032 -0.013  0.000
```


a.
Frequency of Wheezing Attacks (Age 0-1)
Estimated Probability
1.0
0.8
0.6
0.4
0.2
0.0
Never
1-3
4-12
13+
-2 SD
-1 SD
Mean
+1 SD
+2 SD
Latent Wheezing Severity Factor
b.
Frequency of Wheeze-Disturbed Sleep (Age 0-1)
Estimated Probability
1.0
0.8
0.6
0.4
0.2
0.0
Never
< 1 night/week
1+ nights/week
-2 SD
-1 SD
Mean
+1 SD
+2 SD
Latent Wheezing Severity Factor
c.
Wheeze-Disturbed Speech (Age 0-1)
Estimated Probability
1.0
0.8
0.6
0.4
0.2
0.0
Absent
Present
-2 SD
-1 SD
Mean
+1 SD
+2 SD
Latent Wheezing Severity Factor
d.
Exercise-Induced Wheeze (Age 0-1)
Estimated Probability
1.0
0.8
0.6
0.4
0.2
0.0
Absent
Present
-2 SD
-1 SD
Mean
+1 SD
+2 SD
Latent Wheezing Severity Factor

S2 Appendix: Cohort 2 Analyses

Contents

ISAAC Phase I (Cohort 2) data preparation in R…… 2-3

Mplus code for Model 2.1……... 4

Mplus code for Model 2.2……... 5

ISAAC Phase I (Cohort 2) data preparation in R

ISAAC Phase I data were publicly available and downloaded on November 7, 2016 from: http://isaac.auckland.ac.nz/phases/phaseone/results/resultsIndv.php. The Coding and Data Transfer Manual is available at: http://isaac.auckland.ac.nz/phases/phaseone/phaseone.html.

R Code for Data Preparation

```
##############################
# Read in and merge data sets
##############################

## Load ISAAC Phase 1 data sets for 6-7 year-old N. American children into R
barb67 <- read.csv( 'barb67.csv', header = T )
ham67 <- read.csv( 'ham67.csv', header = T )
sask67 <- read.csv( 'sask67.csv', header = T )

## Merge the data sets together (by rows) to create a single data frame
isaac <- rbind( barb67, ham67, sask67 )

##############################
# Recode relevant variables
##############################

## Make sure there is no overlap in school numbers across countries
## by adding 1000 to all number within school 173
isaac$dschool <- isaac$school
isaac$dschool[ isaac$country == 173 ] <- isaac$school[ isaac$country == 173 ] + 1000

## Child sex: make female = 0 and male = 1
isaac$dsex <- ifelse( isaac$sex == 9, NA, 2 - isaac$sex )
with(isaac, table( sex, dsex ) )

## ISAAC-WM item 1: Has your child ever wheezed?
# Recode: missing = NA; No = 0; 1 = Yes
table( isaac$dwhezev <- ifelse( isaac$whezev == 9, NA, 2 - isaac$whezev ) )

## ISAAC-WM item 2: Has your child wheezed in past 12 months?
# Recode: missing = NA; No = 0; 1 = Yes
table( isaac$dwhez12 <- ifelse( isaac$whez12 == 9, NA, 2 - isaac$whez12 ) )
# If parent denied ever wheezing (dwhezev==0) then code 0 for this item
table( isaac$dwhez12 <- ifelse( isaac$dwhezev == 0, 0, isaac$dwhez12 ) )

## ISAAC-WM item 3: How many attacks of wheezing in past 12 months?
# Recode: missing = NA; None = 0; 1-3 = 1; 4-12 = 2; 13+ = 3
table( isaac$dnwhez12 <- ifelse( isaac$nwhez12 == 9, NA, isaac$nwhez12 - 1 ) )
# If denied ever wheezing (dwhezev==0) or wheezing in past 12 months (dwhez12==0) then code 0
table( isaac$dnwhez12 <- ifelse( isaac$dwhezev == 0 | isaac$dwhez12 == 0, 0, isaac$dnwhez12 ) )
```

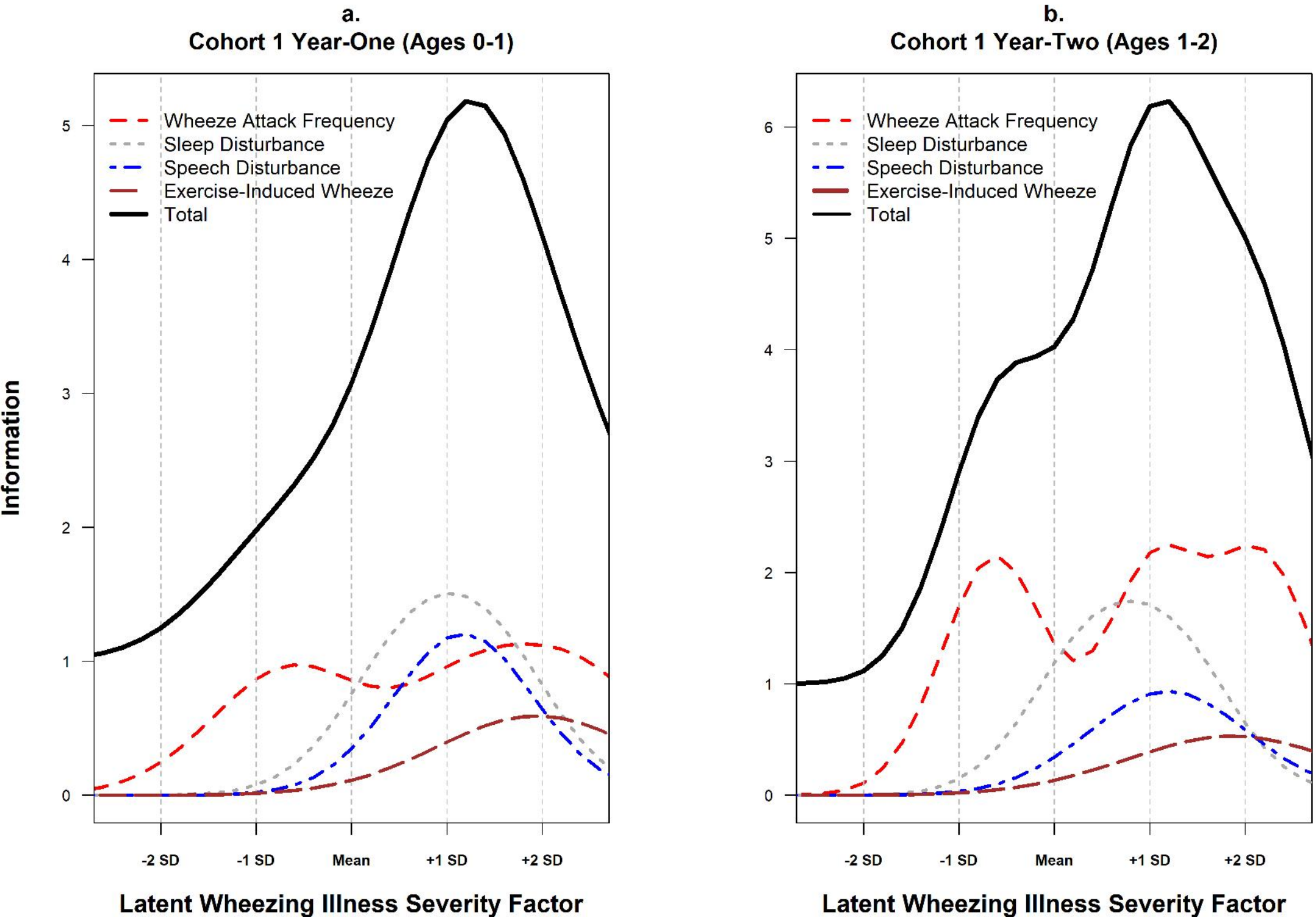
```
## ISAAC-WM item 4: Sleep disturbed due to wheezing in past 12 months?
# Recode: missing = NA; Never = 0; < 1 night/week = 1; 1+ nights/week = 2
table( isaac$dawake12 <- ifelse( isaac$awake12 == 9, NA, isaac$awake12 - 1 ) )
# If denied ever wheezing (dwhezev==0) or wheezing in past 12 months (dwhez12==0) then code 0
table( isaac$dawake12 <- ifelse( isaac$dwhezev == 0 | isaac$dwhez12 == 0, 0, isaac$dawake12 ) )

## ISAAC-WM item 5: Speech ever limited due to wheezing in past 12 months?
# Recode: missing = NA; No = 0; Yes = 1
table( isaac$dspeech12 <- ifelse( isaac$speech12 == 9, NA, 2 - isaac$speech12 ) )
# If denied ever wheezing (dwhezev==0) or wheezing in past 12 months (dwhez12==0) then code 0
table( isaac$dspeech12 <- ifelse( isaac$dwhezev == 0 | isaac$dwhez12 == 0, 0, isaac$dspeech12 ) )

## ISAAC-WM item 7: Exercise induced wheeze?
# Recode: missing = NA; No = 0; Yes = 1
table( isaac$dexwhez12 <- ifelse( isaac$exwhez12 == 9, NA, 2 - isaac$exwhez12 ) )
with(isaac, table( exwhez12, dexwhez12 ) )
summary( isaac$dexwhez12 )

## Recode study centre variable as three dummy variables with values 0 and 1
table( isaac$c1 <- ifelse( isaac$centre == 1, 1, 0 ) )
table( isaac$c2 <- ifelse( isaac$centre == 2, 1, 0 ) )
table( isaac$c99 <- ifelse( isaac$centre == 99, 1, 0 ) )

##########################################################################
# Create new data frame (isaac1) with only children who have ever wheezed
##########################################################################
isaac1 <- subset( isaac, dwhezev == 1 )

#################################################################################
# Create Mplus data and input files using the MplusAutomation R package (Hallquist & Wiley, 2017)
#################################################################################
Install.packages( 'MplusAutomation' )
library( MplusAutomation )
prepareMplusData( isaac1, file = 'isaac1.dat' )
```

Mplus code for Model 2.1: Cohort 2 (ISAAC Phase 1) multilevel SEM model of wheezing severity (see www.statmodel.com for coding details; see Code E1 for variable key)

```
TITLE: m2.1 -- saturated between-schools model;
DATA: FILE = "isaac1.dat";
VARIABLE:
NAMES = form version country centre agegrp school serial dint age agemonths sex
    whezev whez12 nwhez12 awake12 speech12 asthmaev exwhez12 cough12 pnoseev
    pnose12 ieyes12 pnosejan pnosefeb pnosemar pnoseapr pnosemay pnosejun pnosejul
    pnoseaug pnosesep pnoseoct pnosenov pnosedec iactiv12 hfeverev rashev rash12
    sitesev rashage rclear12 rawake12 eczemaev brthev brth12 brth1m exbrthev
    exbrth12 exbrth1m wwokenev wwoken12 wwoken1m cwokenev cwoken12 cwoken1m
    sabrthev sabrth12 sabrth1m dschool dsex dwhezev dwhez12 dnwhez12 dawake12
    dspeech12 dexwhez12 c1 c2 c99;
MISSING=.;
CLUSTER = dschool;
USEVAR=dnwhez12 dawake12 dspeech12 dexwhez12 c2 c99;
BETWEEN = c2 c99;
CATEGORICAL = dnwhez12 dawake12 dspeech12 dexwhez12;
ANALYSIS: type = twolevel; estimator = wlsmv;
MODEL:
 %WITHIN%
whzw BY dnwhez12* dawake12* dspeech12* dexwhez12*;
whzw @1;
%BETWEEN%
dnwhez12 WITH dawake12 dspeech12 dexwhez12;
dawake12 WITH dspeech12 dexwhez12;
dspeech12 WITH dexwhez12;
dnwhez12 dawake12 dspeech12 dexwhez12 ON c2 c99;
OUTPUT: sampstat cinterval residual stand(stdy);
PLOT: type = plot1 plot2 plot3;
```

Mplus code for Model 2.2: Cohort 2 (ISAAC Phase 1) multilevel SEM model regression wheezing severity on child sex on the within-schools level of analysis.

```
TITLE: m2.1 -- saturated between-schools model;
DATA: FILE = "isaac1.dat";
VARIABLE:
NAMES = form version country centre agegrp school serial dint age agemonths sex
    whezev whez12 nwhez12 awake12 speech12 asthmaev exwhez12 cough12 pnoseev
    pnose12 ieyes12 pnosejan pnosefeb pnosemar pnoseapr pnosemay pnosejun pnosejul
    pnoseaug pnosesep pnoseoct pnosenov pnosedec iactiv12 hfeverev rashev rash12
    sitesev rashage rclear12 rawake12 eczemaev brthev brth12 brth1m exbrthev
    exbrth12 exbrth1m wwokenev wwoken12 wwoken1m cwokenev cwoken12 cwoken1m
    sabrthev sabrth12 sabrth1m dschool dsex dwhezev dwhez12 dnwhez12 dawake12
    dspeech12 dexwhez12 c1 c2 c99;
MISSING=.;
CLUSTER = dschool;
USEVAR=dnwhez12 dawake12 dspeech12 dexwhez12 c2 c99 dsex;
WITHIN = dsex;
BETWEEN = c2 c99;
CATEGORICAL = dnwhez12 dawake12 dspeech12 dexwhez12;
ANALYSIS: type = twolevel; estimator = wlsmv;
MODEL:
 %WITHIN%
whzw BY dnwhez12* dawake12* dspeech12* dexwhez12*;
whzw @1;
whzw ON dsex;
%BETWEEN%
dnwhez12 WITH dawake12 dspeech12 dexwhez12;
dawake12 WITH dspeech12 dexwhez12;
dspeech12 WITH dexwhez12;
dnwhez12 dawake12 dspeech12 dexwhez12 ON c2 c99;
OUTPUT: sampstat cinterval residual stand(stdy);
PLOT: type = plot1 plot2 plot3;
```


a.
Cohort 1 Year-One (Ages 0-1)
b.
Cohort 1 Year-Two (Ages 1-2)
Wheeze Attack Frequency
Sleep Disturbance
Speech Disturbance
Exercise-Induced Wheeze
Total
Information
0
1
2
3
4
5
6
-2 SD
-1 SD
Mean
+1 SD
+2 SD
Latent Wheezing Illness Severity Factor

S3 Appendix: Cohort 3 Analyses

Table of Contents

Contents 2
Variable key 2
Model 3.1 2
m3.1: lavaan code 2
m3.1: Sample statistics (observed covariance matrix and thresholds) 2
m3.1: Model implied covariance matrix 3
m3.1: Residual correlation matrix for Model 3.1 3
Model 3.2 3
m3.2: lavaan code 3
m3.2: Sample statistics (observed covariance matrix and thresholds) 4
m3.2: Model implied covariance matrix 4
m3.2: Residual correlation matrix for Model 3.2 4
Model 3.3 5
m3.3: lavaan code 5
m3.3: Sample statistics (observed covariance matrix and thresholds) 5
m3.3: Model implied covariance matrix 6
m3.3: Residual correlation matrix for Model 3.3 6
Model 3.4 6
m3.4: lavaan code 6
m3.4: Sample statistics (observed covariance matrix and thresholds) 6
m3.4: Model implied covariance matrix 7
m3.4: Residual correlation matrix for Model 3.4 7
Model 3.5 7
m3.5: lavaan code 7
m3.5: Sample statistics (observed covariance matrix and thresholds) 8
m3.5: Model implied covariance matrix 8
m3.5: Residual correlation matrix for model 1.5 9

Contents

This file provides supplemental modeling information for the Cohort 3 analyses. We provide *lavaan* code for the primary Cohort 3 models (see Table 4 in the primary text). Additionally, we provide observed and model-implied covariance marixes for each model, as well as the residual correlation matrix.

Variable key

Below we define the observed variables used in the analyses. For detailed information about *lavaan* modeling code, see the lavaan website.

- dnwhez12 = ISAAC-WM item from the one-year follow-up assessment asking about the frequency of wheezing attacks in the past year (4-level ordinal variable)
- dawake12 = ISAAC-WM item from the one-year follow-up assessment asking about the frequency of wheeze-induced sleep disturbance in the past year (3-level ordinal variable)
- dspeech12 = ISAAC-WM item from the one-year follow-up assessment asking about the frequency of wheeze-induced speech disturbance in the past year (2-level ordinal variable)
- dexwhez12 = ISAAC-WM item from the one-year follow-up assessment asking whether the child has experienced exercise-induced wheeze

Model 3.1

m3.1: lavaan code

```
m3.1.syn <- '
whzsev =~ dnwhez + dawake + dspeech + dexwhez
'

m3.1 <- cfa( m3.1.syn, data = ehaas.wide1, estimator = 'wlsmv', std.lv = T, m
issing = 'pairwise',
                          ordered = c( 'dnwhez','dawake','dspeech','dexwhez' )
)
```

m3.1: Sample statistics (observed covariance matrix and thresholds)

```
## $cov
##         dnwhez dawake dspech dexwhz
## dnwhez  1.000
```

```
## dawake  0.337  1.000
## dspeech 0.307  0.304  1.000
## dexwhez 0.326  0.114  0.147  1.000
##
## $mean
##  dnwhez  dawake dspeech dexwhez
##       0       0       0       0
##
## $th
##  dnwhez|t1  dnwhez|t2  dnwhez|t3  dawake|t1  dawake|t2 dspeech|t1
##     -0.478      0.566      0.981      0.140      0.799      1.221
## dexwhez|t1
##      0.591
```

m3.1: Model implied covariance matrix

```
##         dnwhez dawake dspech dexwhz
## dnwhez  1.000
## dawake  0.341  1.000
## dspeech 0.339  0.199  1.000
## dexwhez 0.296  0.174  0.173  1.000
```

m3.1: Residual correlation matrix for Model 3.1

```
##         dnwhez dawake dspech dexwhz
## dnwhez   0.000
## dawake  -0.003  0.000
## dspeech -0.032  0.106  0.000
## dexwhez  0.030 -0.059 -0.025  0.000
```

Model 3.2

m3.2: lavaan code

```
m3.2syn <- '
whzsev =~ dnwhez + dawake + dspeech + dexwhez
whzsev ~ white + male
'

m3.2 <- sem( m3.2syn, data = ehaas.wide1, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise',
                        ordered = c( 'dnwhez','dawake','dspeech','dexwhez')
)
```

m3.2: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##         dnwhez dawake dspech dexwhz
## dnwhez  1.000
## dawake  0.310  1.000
## dspeech 0.281  0.263  1.000
## dexwhez 0.294  0.069  0.099  1.000
##
## $res.int
##  dnwhez  dawake dspeech dexwhez
##       0       0       0       0
##
## $res.th
##  dnwhez|t1  dnwhez|t2  dnwhez|t3  dawake|t1  dawake|t2 dspeech|t1
##     -0.258      0.799      1.223      0.585      1.262      2.004
## dexwhez|t1
##      0.812
##
## $res.slopes
##          white  male
## dnwhez   0.018 0.317
## dawake   0.118 0.511
## dspeech  0.534 0.452
## dexwhez -0.153 0.485
##
## $cov.x
##       white  male
## white  0.171
## male  -0.003  0.224
```

m3.2: Model implied covariance matrix

```
##         dnwhez dawake dspech dexwhz
## dnwhez  1.000
## dawake  0.294  1.000
## dspeech 0.272  0.254  1.000
## dexwhez 0.209  0.196  0.181  1.000
```

m3.2: Residual correlation matrix for Model 3.2

```
##         dnwhez dawake dspech dexwhz
## dnwhez   0.000
## dawake   0.016  0.000
## dspeech  0.009  0.009  0.000
## dexwhez  0.084 -0.127 -0.082  0.000
```

Model 3.3

m3.3: lavaan code

```
m3.3syn <- '
whzsev =~ dnwhez + dawake + dspeech + dexwhez
whzsev ~ white + male
asthmadx.84 ~ whzsev + white + male
'

m3.3 <- cfa( m3.3syn, data = ehaas.wide1, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise',
                          ordered = c( 'dnwhez','dawake','dspeech','dexwhez','
asthmadx.84' ) )
```

m3.3: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##             dnwhez dawake dspech dexwhz ast.84
## dnwhez      1.000
## dawake      0.310  1.000
## dspeech     0.281  0.263  1.000
## dexwhez     0.294  0.069  0.099  1.000
## asthmadx.84 0.480  0.098  0.054  0.145  1.000
##
## $res.int
##      dnwhez      dawake     dspeech     dexwhez asthmadx.84
##           0           0           0           0           0
##
## $res.th
##      dnwhez|t1      dnwhez|t2      dnwhez|t3      dawake|t1      dawake|t2
##         -0.258          0.799          1.223          0.585          1.262
##     dspeech|t1     dexwhez|t1 asthmadx.84|t1
##          2.004          0.812          0.217
##
## $res.slopes
##              white  male
## dnwhez       0.018 0.317
## dawake       0.118 0.511
## dspeech      0.534 0.452
## dexwhez     -0.153 0.485
## asthmadx.84 -0.240 0.333
##
## $cov.x
##       white  male
## white  0.171
## male  -0.003  0.224
```

m3.3: Model implied covariance matrix

```
##              dnwhez dawake dspech dexwhz ast.84
## dnwhez       1.000
## dawake       0.309  1.000
## dspeech      0.286  0.142  1.000
## dexwhez      0.284  0.141  0.130  1.000
## asthmadx.84 0.411  0.204  0.188  0.187  1.000
```

m3.3: Residual correlation matrix for Model 3.3

```
##              dnwhez dawake dspech dexwhz ast.84
## dnwhez        0.000
## dawake        0.000  0.000
## dspeech      -0.005  0.121  0.000
## dexwhez       0.010 -0.072 -0.031  0.000
## asthmadx.84  0.069 -0.106 -0.134 -0.042  0.000
```

Model 3.4

m3.4: lavaan code

```
m3.4syn <- '
whzsev =~ dnwhez + dawake + dspeech + dexwhez
whzsev ~ white + male
whzdr12.84 ~ whzsev + white + male
'

m3.4 <- cfa( m3.4syn, data = ehaas.wide1, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise',
                          ordered = c( 'dnwhez','dawake','dspeech','dexwhez','
whzdr12.84' ) )
```

m3.4: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##             dnwhez dawake dspech dexwhez w12.84
## dnwhez       1.000
## dawake       0.310  1.000
## dspeech      0.281  0.263  1.000
## dexwhez      0.294  0.069  0.099  1.000
## whzdr12.84  0.483  0.248  0.130 -0.033  1.000
##
## $res.int
##     dnwhez     dawake    dspeech    dexwhez whzdr12.84
```

```
##          0          0          0          0          0
##
## $res.th
##     dnwhez|t1     dnwhez|t2     dnwhez|t3     dawake|t1     dawake|t2
##        -0.258         0.799         1.223         0.585         1.262
##    dspeech|t1    dexwhez|t1 whzdr12.84|t1
##         2.004         0.812         1.223
##
## $res.slopes
##             white  male
## dnwhez      0.018 0.317
## dawake      0.118 0.511
## dspeech     0.534 0.452
## dexwhez    -0.153 0.485
## whzdr12.84  0.250 0.549
##
## $cov.x
##       white  male
## white  0.171
## male  -0.003  0.224
```

m3.4: Model implied covariance matrix

```
##            dnwhez dawake dspech dexwhz w12.84
## dnwhez     1.000
## dawake     0.332  1.000
## dspeech    0.290  0.189  1.000
## dexwhez    0.205  0.134  0.117  1.000
## whzdr12.84 0.412  0.269  0.235  0.166  1.000
```

m3.4: Residual correlation matrix for Model 3.4

```
##            dnwhez dawake dspech dexwhz w12.84
## dnwhez      0.000
## dawake     -0.023  0.000
## dspeech    -0.009  0.074  0.000
## dexwhez     0.089 -0.065 -0.018  0.000
## whzdr12.84  0.071 -0.021 -0.104 -0.199  0.000
```

Model 3.5

m3.5: lavaan code

```
m3.5syn <- '
whzsev =~ dnwhez + dawake + dspeech + dexwhez
```

```
whzsev ~ white + male
whzemerg12.84 ~ whzsev + white + male
'

m3.5 <- cfa( m3.5syn, data = ehaas.wide1, estimator = 'wlsmv', std.lv = T, mi
ssing = 'pairwise',
                          ordered = c( 'dnwhez','dawake','dspeech','dexwhez','
whzemerg12.84' ) )
```

m3.5: Sample statistics (observed covariance matrix and thresholds)

```
## $res.cov
##                dnwhez dawake dspech dexwhz w12.84
## dnwhez         1.000
## dawake         0.310  1.000
## dspeech        0.281  0.263  1.000
## dexwhez        0.294  0.069  0.099  1.000
## whzemerg12.84 0.628  0.369  0.304  0.104  1.000
##
## $res.int
##        dnwhez        dawake       dspeech       dexwhez whzemerg12.84
##             0             0             0             0             0
##
## $res.th
##        dnwhez|t1        dnwhez|t2        dnwhez|t3        dawake|t1
##           -0.258            0.799            1.223            0.585
##        dawake|t2       dspeech|t1       dexwhez|t1 whzemerg12.84|t1
##            1.262            2.004            0.812            1.169
##
## $res.slopes
##                white  male
## dnwhez         0.018 0.317
## dawake         0.118 0.511
## dspeech        0.534 0.452
## dexwhez       -0.153 0.485
## whzemerg12.84  0.182 0.611
##
## $cov.x
##       white  male
## white  0.171
## male  -0.003  0.224
```

m3.5: Model implied covariance matrix

```
##                dnwhez dawake dspech dexwhz w12.84
## dnwhez         1.000
## dawake         0.324  1.000
```

```
## dspeech       0.297  0.206  1.000
## dexwhez       0.206  0.143  0.131  1.000
## whzemerg12.84 0.562  0.390  0.357  0.248  1.000
```

m3.5: Residual correlation matrix for model 1.5

```
##               dnwhez dawake dspech dexwhz w12.84
## dnwhez         0.000
## dawake        -0.014  0.000
## dspeech       -0.016  0.057  0.000
## dexwhez        0.088 -0.074 -0.032  0.000
## whzemerg12.84  0.066 -0.021 -0.053 -0.144  0.000
```