Check for updates

CLINICAL RESEARCH ARTICLE Wisconsin Upper Respiratory Symptom Survey for Kids: Validation of an Illness-specific Quality of Life Instrument

Kathryn M. Schmit¹, Roger Brown², Supriya Hayer³, Mary M. Checovich⁴, James E. Gern⁵, Ellen R. Wald¹ and Bruce Barrett⁴

BACKGROUND: Acute respiratory infections (ARIs) are the most common illness seen in the pediatric ambulatory setting. Research in this area is hampered by the lack of validated ARI measures. The aim of this study was to assess the reliability and validity of the Wisconsin Upper Respiratory Symptom Survey for Kids (WURSS-K), a 15-item instrument, which measures illness-specific symptoms and impact on quality of life during an ARI.

METHODS: WURSS-K was administered to two populations: (1) children aged 4–10 years recruited from the local community and (2) 9- and 10-year-old children from an ongoing study, the Urban Environment and Childhood Asthma.

RESULTS: Overall, 163 children with 249 ARI episodes completed WURSS-K. WURSS-K was analyzed using multiple models to evaluate reliability and validity for a two-factor structure (symptom and functionality) and a single global structure. These models provided evidence of reliability and validity with omega of 0.72 and 0.91 for symptoms and functionality along with the single structure with omega of 0.90.

CONCLUSIONS: WURSS-K shows strong psychometric properties for validity and reliability as either a single global factor or a twofactor structure. This instrument will be useful in both therapeutic trials and observational studies among children with ARI in ambulatory settings.

Pediatric Research (2021) 90:1207-1214; https://doi.org/10.1038/s41390-021-01395-9

IMPACT:

- WURSS-K is a valid and reliable illness-specific quality of life instrument that evaluates the impacts of ARIs on children.
- WURSS-K is designed for children 4–10 years of age, for whom there is a lack of validated assessment tools.
- This now validated instrument will be useful for future observational studies and therapeutic trials among children with ARIs in ambulatory settings.

INTRODUCTION

Acute respiratory infections (ARIs) in children are a major reason for healthcare visits, especially in ambulatory settings. In a national survey, ARIs were the third most common primary diagnosis in office visits after hypertension and routine well-child checks in children and adults.¹ Children have between 1–8 ARIs per year depending on age, exposures, and individual susceptibility.^{2–5} In the United States, these non-influenza viral respiratory tract infections cost over 40 billion dollars annually, including both direct healthcare costs and indirect costs related to missed days of school and work.⁶ Three billion dollars are spent annually on overthe-counter medications for symptomatic relief of ARIs despite little evidence to support their efficacy in children.⁶

Accurate measurement of ARI symptoms is critical for both epidemiologic and clinical trials to further understand illness severity, duration, and response to treatment. Various validated assessment instruments exist for children with bronchiolitis,^{7–17} but few are available for children older than 2 years of age with ARIs. Randomized controlled trials in these children,^{18,19} predominantly focusing on symptomatic therapies, usually use their

own assessment measures that lack previous validation. This lack of consistent and validated measurement tools for assessing children in the ambulatory setting creates a challenge for the scientific community when comparing or conducting observational or therapeutic studies.

Existing assessment instruments for children with ARIs concentrated solely on symptom severity. The Canadian Acute Respiratory Illness and Flu Scale was the first ARI severity measure in children to assess functional and parental impact as well as symptom severity.²⁰ The addition of these components is essential to understanding the full scope of illness in children who are not always able to communicate their symptoms effectively.

Recognizing the importance of including both symptom severity and functional impact, the Wisconsin Upper Respiratory Symptom Survey (WURSS) was created.²¹ WURSS is an illness-specific quality of life instrument that was developed and validated to assess the impact of ARIs in adults.²² It is available in different length versions, more than 20 languages, and has been used in dozens of clinical trials in several countries.^{23–26}

Received: 6 August 2020 Accepted: 14 January 2021 Published online: 24 February 2021

¹Department of Pediatrics, Division of Infectious Diseases, University of Wisconsin-Madison, Madison, WI, USA; ²Research Design & Statistics Unit, University of Wisconsin-Madison School of Nursing, Madison, WI, USA; ³Medical nonprofit, The India Project Inc., Santa Rosa, CA, USA; ⁴Department of Family Medicine and Community Health, University of Wisconsin-Madison, Madison, WI, USA and ⁵Department of Pediatrics, Division of Allergy and Immunology, University of Wisconsin-Madison, MI, USA Correspondence: Kathryn M. Schmit (kschmit@uwhealth.org)

Wisconsin Upper Respiratory Symptom Survey for kids-daily symptom report

Day:	Date:		Time:	ID:					
Please fill in one circle for each question:									
		Not sick	A little sic	k Sick	Very sick				
		\bigcirc			\bigcirc				
How sick do you feel toda	y?	0	0	0	0				
How bad are your cold	symptoms	? (O	verall, since yesterday)						
	Do	not have this	A little bad	Bad	Very bad				
		(((
Runny nose		0	0	0	0				
Stuffy nose		0	0	0	0				
Sneezing		0	0	0	0				
Sore throat (hurts to swall	ow)	0	0	0	0				
Cough		0	0	0	0				
Feeling tired		0	0	0	0				
Since yesterday, how ha	ard has it b	een to:							
		Not at all	A little hard	d Hard	Very hard				
		($(\cdot \cdot)$	(
Think		0	0	0	0				
Sleep		0	0	0	0				
Breathe		0	0	0	0				
Talk		0	0	0	0				
Walk, climb stairs, exercis	e	0	0	0	0				
Go to school		0	0	0	0				
Play with friends		0	0	0	0				
Compared to yesterday	, I feel my o	cold is							
A lot better	lot better A little better The		e same	A little worse	A lot worse				
			••• 						
0	0		0	0	0				
I completed th	is page:	all by myse	lf 🗌 with s	some help	a lot of help				

Fig. 1 Wisconsin Upper Respiratory Symptom Survey for Kids (WURSS-K)—daily symptom report. WURSS-K questionnaire items include global illness severity (item 1), severity of symptoms (items 2-7), functional impacts (items 8-14), and evaluation of change over time (item 15).

The aim of this study was to assess the validity and reliability of a pediatric version of the WURSS for Kids (WURSS-K). WURSS-K is a 15-item instrument that focuses on both illness-specific symptoms and impact on quality of life, and is designed for use in children 4–10 years of age.

METHODS

Phase 1: Development of WURSS-K

Phase 1 involved the development of the initial WURSS-K instrument aimed at assessing symptoms and quality of life using visual representation of happy and sad faces to assist Likert scale ratings, similar to Wong–Baker faces.²⁷ The WURSS-21 provided the initial questionnaire format.²³ Iterative review and refinement of the evolving questionnaire instrument was carried out by a multidisciplinary team, including two pediatricians (Wald, Gern), a family physician (Barrett), a psychometrics expert (Brown), a physician in postdoctoral research fellowship (Hayer), and an expert in questionnaire design from the University of Wisconsin Survey Center (Dykema). Research nurses trained in cognitive debriefing strategies then interviewed children regarding content,

format, and ability to understand each questionnaire item. Audio recordings of these interviews were reviewed individually by team members and then discussed in a group format. Following these discussions and revisions, the WURSS-K daily symptom report was finalized in 2014.

Participants in this development process included 14 children, 4–6 years of age, who were enrolled in an ongoing research study.³ These children had upper respiratory symptoms starting 1–3 days before the interviews. The child's parent was present during the interview. For children requiring assistance, the parent was allowed to explain questions or response options.

WURSS-K includes 15 items; 14 answered on a 4-point ordinal Likert-type scale (0 = not sick/do not have this/not at all, 3 = very sick/very bad/very hard) (Fig. 1). Questionnaire items included those related to global illness severity (item 1), severity of symptoms (items 2–7), functional impacts (items 8–14), and comparison in order to evaluate change over days of illness (item 15). Happy and sad face representations are included along with the ordinal scales to facilitate survey completion by children. There was a question at the end of the form for

children to indicate whether the questionnaire was completed: (1) all by myself, (2) with some help, or (3) with a lot of help.

Scoring of the daily survey in Fig. 1 uses the sum of unweighted scores for each item. The sum of the 14 items on a 4-point ordinal scale results in a global total score ranging from 0 (low symptoms and low functional impairment) to 42 (high symptoms and high functional impairment). Item 15, which evaluates for change between days of illness, is not included in the overall score. For the two factors of symptoms and functionality, the total symptoms span from 0 (low) to 18 (high) and the total functionality spans from 0 (low impairment) to 21 (high impairment).

Phase 2: Administration of WURSS-K

Following development of the WURSS-K instrument, psychometric testing was done using two separate populations: (1) locally recruited sample in Madison, Wisconsin (Madison sample) and (2) part of an ongoing multicity research project (Urban Environment and Childhood Asthma (URECA) sample).²⁸

For the Madison sample, volunteers were recruited by advertisements posted in the local community from October 2014 to December 2016. Inclusion criteria were an age of 4–10 years and willingness to answer daily symptom surveys during an ARI episode. The majority of participants were enrolled before exhibiting ARI symptoms, and were asked to call study personnel at the first sign of cold symptoms.

Study team members met with parents and children to obtain parental informed consent and child assent, review study design, explain the WURSS-K instrument, and collect baseline demographic information. When the child developed ARI symptoms, a parent called the study phone and was asked "Do you believe that your child has a cold?". After the parent answered "Yes," the parent was asked questions about symptoms using modified Jackson criteria²⁹ (Supplementary Table 1). The child was classified as having an ARI if the parent answered "Yes" to the initial question and to one of the questions regarding nasal symptoms or sore throat. Once the child was classified as having an ARI, the child (or parent as proxy) would fill out WURSS-K booklet daily during an ARI episode. An ARI episode was considered complete when the child answered "Not sick" 2 days in a row on item 1 of WURSS-K. Booklets were mailed back and data entered into a REDCap database.³⁰ Participants received minimal compensation for joining the study and for each booklet returned during an ARI.

In the URECA sample, the WURSS-K was administered to 9- and 10-year-old children at their routine study follow-up visits. URECA is an observational prospective study that enrolled pregnant women living in urban areas throughout the United States (Baltimore, Boston, New York City, and St. Louis) and has followed their children from birth.²⁸ This cohort consists mainly of ethnic minorities; enrollment criteria included having at least one parent with allergic disease or asthma, and residence in an area in which at least 20% of the families have a household income below the poverty line. URECA study team members asked participants the first item of WURSS-K ("How sick do you feel today?") and noted their responses. If the response was other than "Not sick," the remainder of WURSS-K was completed by the participant (or a proxy). In order to avoid potential confounders, children with a diagnosis of asthma were removed from our analysis.

This study was approved by the University of Wisconsin School of Medicine and Public Health Institutional Review Board.

Phase 3: Psychometric analysis of WURSS-K

The WURSS-K questionnaire items are grouped into symptoms (six items) and functionality (seven items). Descriptive statistics of these 13 items included measures of central tendency, dispersion, and a polychoric correlation matrix, which optimizes the analysis of ordinal data. Missing data were assessed as either missing completely at random (MCAR) or missing at random.³¹ Multiple imputation by chained equations of missing values were used if

the assumption of MCAR was accepted.³² Since some of the participants experienced multiple ARIs during the study period, psychometric structural analysis was based on ARI episodes and not on subjects. A cluster factor analysis³³ was then used to model the ARIs clustered within subjects to adjust the standard errors for nested dependency.

Analyses were conducted in five different stages using classical test theory for the first four and invariance across the two-sample populations for the last one. Each of these stages is important in the evaluation of an instrument to determine (1) construct validity in the first stage, (2) internal reliability in the second stage, (3) whether to use as a global or a two-factor assessment (construct validity) in the third stage, (4) difficulty of items in the fourth stage, and (5) how it performs in different populations in the fifth stage.

In the first stage, confirmatory factor analysis was used to model the factor (congeneric) structure of WURSS-K. The congeneric model is one type of measurement in which a factor is measured by several observed items.³⁴ WURSS-K was defined as a two-factor correlated model, with factors of (1) symptoms and (2) functionality.

In the second stage, reliability was assessed by a series of hierarchical measurement models, ranging from least to most restrictive: factor structure, tau equivalent, partial tau equivalent, and parallel.³⁵ Tau equivalent assumes that each of the individual items measures the same factor on the same scale with the same degree of precision, but with different amounts of error. The corresponding reliability coefficient of tau equivalent is Cronbach's alpha. Byrne, Shavelson, and Muthen³⁶ have proposed relaxing some of the tau-equivalency restrictions to partial tauequivalent model. Partial tau equivalent occurs when equal constraints are imposed on some but not all loadings. This less restrictive model then allows the model fit to be reassessed after some loadings are freely estimated. The parallel model assumes that each individual item measures the same domain, on the same scale with the same degree of precision and the same amount of error. Each of these models was assessed for goodness of fit to determine the best representation of the data, and to assess which measure of reliability would be considered the most appropriate.

In the third stage, bifactor analysis³⁷ was used to evaluate the unidimensional or global nature of WURSS-K. Along with parameter and fit measures, explained common variance (ECV) and factor determinacy (FD) were estimated. ECV provides the extent to which an item's responses are accounted for by variation on the single global factor. ECV > 0.70 demonstrates that the common variance can be regarded as unidimensional with a slight relative bias.³⁸ FD assesses the correlation between the estimated score and the underlying factor; values range from 0 to 1 with FD > 0.90 being ideal.³⁹ Finally, relative parameter bias was estimated, which is the difference between an item's loading in the unidimensional solution and its general factor loading in the bifactor, divided by the general factor loading in the bifactor. According to Muthén et al., parameter bias <10–15% is acceptable.^{40,41}

The fourth stage addresses item importance based on item response theory (IRT), using a two-faceted Rasch partial credit model. This model uses infit mean square (INFIT MNSQ) and outfit mean square (OUTFIT MNSQ) statistics to compare the fit of the observed data to values expected by the Rasch model.^{42–44} While item fit indices of 1.0 are ideal,⁴⁵ Wright and Linacre have suggested that items are "fit" if their MNSQ falls within the range 0.6–1.4, with the range 0.5–1.5 considered acceptable.⁴⁴ Higher scores indicate a wider variation in the response to an item, which suggests more difficulty, whereas lower (or negative) scores indicate less difficulty.

In the last stage, measurement invariance analysis was used to compare the URECA and Madison samples to see if the structure remained the same despite any potential differences in the two study population. If measurement invariance fails, it is evidence



Fig. 2 Flow chart of participants. Children included in the validation of WURSS-K with the Madison sample (on the left) and URECA sample (on the right).

Gender Male

Race

Female

that the items do not measure the same factor in the same way in different situations. The most basic requirement is configural or pattern invariance, which addresses whether the configuration of the model is similar across groups. If this is not accepted, the usefulness of the scale in alternate populations should be questioned.

RESULTS

There were 249 ARI episodes with a complete WURSS-K instrument, comprising 130 and 119 ARI episodes among 58 and 105 children, respectively, from the Madison and URECA samples (Fig. 2). Children in the Madison group completed a WURSS-K symptom report once daily throughout each illness episode, up to a maximum of 21 days; there was a mean of 2.2 (SD 1.5) ARI illness episodes per participant during the study period.

Participant demographics are shown in Table 1. The mean age at time of WURSS-K completion was 6.7 (SD 1.8) years in the Madison sample and 9.5 (SD 0.5) years in URECA sample. In addition to being older, the URECA sample had more participants who were female, African American, of Hispanic ethnicity, and with a lower annual household income.

WURSS-K total mean scores from the Madison sample were 9.6 (SD 5.0) on day 1, 6.8 (SD 5.4) on day 3, and 4.6 (SD 4.9) on day 10 (Fig. 3). The URECA sample had a mean global score of 8.2 (SD 6.5), with symptom and functionality scores of 4.8 and 3.4. Psychometric structural analysis in the Madison sample focused on day 3 of symptoms as this was the day during an ARI episode when the symptom complex stabilized. On day 3, the Madison sample had a mean global score of 6.8 (SD 5.4), with the symptom and functionality scores of 4.7 and 2.0. Missing data analysis showed completeness of 96% (p = 0.02) for day 3 in the Madison sample and 95% (p = 0.04) for the URECA sample. Based on Little's MCAR test,³¹ the missing values in both samples indicated that the assumption of MCAR was accepted, allowing imputation using multiple imputation by chained equations.³²

There were 1029 and 119 total days of ARI illness with a completed WURSS-K from the Madison and URECA samples, respectively. The last item on this report ("I completed this page: all by myself, with some help, or with a lot of help") was completed by 97% of children with 39% (433/1114) "all by myself," 44% (490/1114) "with some help," and 17% (191/1114) "with a lot of help." Mean ages of children were 8.0 (SD 1.5), 6.7 (SD 1.9), and 6.1 (SD 1.6) years for responses of "all by myself," "with some help," and "with a lot of help," respectively.

African American 2 (3) 77 (73) Asian 6 (10) 0 (0) White 49 (84) 0 (0) Multiple 1(2)5 (5) Other 0 (0) 23 (22) Ethnicity Hispanic 0 (0) 21 (20) Non-Hispanic 58 (100) 84 (80) Annual household income (\$) <15.000 0 (0) 51 (49) 15-25.000 3 (5) 24 (23) 25-50,000 25 (24) 4 (7) 50-75,000 8 (14) 5 (5) >75,000 43 (72) 0 (0) Parent education^a Some high school 0 (0) 22 (21) High school graduate 1 (2) 36 (34) College or tech school credits 7 (12) 45 (43) Bachelor's degree 23 (40) 0 (0) Post-graduate degree 27 (46) 0 (0) Age at WURSS-K completion Number of ARI^b (n) Number of ARI (n) Age (years) 4 14 (10) 5 28 (16) 6 18 (11) 7 21 (14) 8 24 (14) 9 19 (14) 63 (63) 10 6 (3) 56 (56) ^aIn the Madison sample, parent education refers to the highest education

Table 1. Participant demographics for the children included in the

Madison, n (%)

36 (62)

22 (38)

URECA, n (%)

41 (39)

64 (61)

Madison and URECA sample populations.

^aIn the Madison sample, parent education refers to the highest education level achieved by the parent who completed the baseline information. In the URECA sample, parent education refers to maternal education. ^bARI is acute respiratory infection episodes.

050....

1210

Wisconsin Upper Respiratory Symptom Survey for Kids: Validation of an... KM Schmit et al.



Fig. 3 Change of WURSS-K scores over time during ARI episodes in the Madison sample using notched boxplots. a WURSS-K global score; b WURSS-K total symptom score; c WURSS-K total functionality score. n number of ARI episodes included on each day. Closed circles symbols represent outliers.

Psychometric results

Table 2 includes the goodness of fit for the models from the first four stages. The two-factor congeneric model (Model 1) was supported with high FDs of 0.98 for symptom factor and 0.92 for the functionality factor.

The tau-equivalent model (Model 2) was not supported based on rejection of χ^2 difference test (χ^2 diff 40.05, p < 0.001). Cronbach's alpha coefficients were 0.67 and 0.82 for symptom and functionality factors. Since the two-factor tau-equivalency model was rejected, the use of Cronbach's alpha may be biased, underestimating the reliability. Therefore, a partial tau-equivalent model (Model 3) was subsequently pursued, in which two items were freed, which was accepted by the χ^2 difference test (χ^2 diff 16.85, p = 0.052). The corresponding omega reliability assessment coefficients are 0.72 and 0.91 for symptoms and functionality.

Since evidence supported two separate factors (symptom and functionality), global assessment with a bifactor model was assessed (Model 5). The bifactor model fit indices yield an FD of 0.98, with reliability coefficients of Cronbach 0.85 and omega 0.90. ECV was 0.72, greater than the acceptable value of 0.70. The relative parameter bias resulted in average absolute relative bias across items of 10.4%, which is within the accepted range^{40,41} and allows the use of WURSS-K as either a single global factor or a two-factor assessment.

Item response theory analysis of the WURSS-K indicated adequate item fit with all INFIT and OUTFIT MNSQ having values of 0.5–1.5 (Table 3). The Rasch analysis indicated strong support for unidimensional integrity explaining 59% of the variance; Reckase suggests that this level is acceptable for unidimensional assumption.^{46,47} The easiest items were "How bad are your cold symptoms"—"stuffy nose" (item 3), "runny nose" (item 2), and "cough" (item 6). The items that appeared to be most difficult for participants were "How hard has it been to think" (item 8), "How hard has it been to go to school" (item 13), "How hard has it been to play with friends" (item 14), and "How hard has it been to walk, climb stairs, exercise" (item 12).

The two-sample populations (Madison vs URECA) were compared using a two-factor configural invariance model, which was accepted (Supplementary Table 4). The weak invariance based on partial tau-equivalent model estimated that the two groups fit the data fairly well (χ^2 difference 13.97, p = 0.23) according to suggested fit criteria.⁴⁸ Further analyses of invariance across groups, both strong and strict were not supported with χ^2 difference of 45.68 (p = 0.004) and 54.84 (p < 0.001), respectively. These findings suggest that the measurement of the two-factor WURSS-K model did not differ substantially across the two-sample populations, thereby demonstrating that the instrument was robust across groups with different demographics.

1212

Table 2. I sychometric drarysis of works it dsing goodless of it, with the models.										
Model	χ ² (d.f.)	P value	χ^2 /d.f. ratio	CFI	TLI	RMSEA	SRMR	χ^2 diff (d.f.)	P value	Decision
Model 1: Factor congeneric Two factors	113.14 (64)	<0.001	1.76	0.931	0.916	0.077	0.093	_	—	Accept
FD 1: symptom	0.983									
FD 2: functionality	0.921									
Model 2: Factor tau equivalent	150.39 (75)	<0.001	2.00	0.894	0.890	0.088	0.123	40.05 (11)	<0.001	Reject
Model 3: Factor partial tau equivalent	121.51 (73)	0.0003	1.66	0.932	0.927	0.071	0.107	16.81 (8)	0.052	Accept
Model 4: Factor parallel based on partial tau equivalent	204.62 (84)	<0.001	2.43	0.831	0.843	0.105	0.133	148.98 (11)	<0.001	Reject
Model 5: Bifactor single-factor congeneric	77.64 (52)	0.012	1.49	0.964	0.946	0.062	0.070			Accept
FD	0.983									

The first 4 measurement models (Models 1–4) were used to assess the two-domain structure of symptom and functionality. The bifactor model (Model 5) assessed the global structure of WURSS-K.

FD factor determinacy (>0.90),³⁸ χ^2 chi square, *d.f.* degrees of freedom, $\chi^2/d.f.$ *ratio* (<3 good fit), *CFI* comparative fit index (>0.95 good fit, ≥0.90 acceptable fit), *TLI* Tucker–Lewis index (>0.95 good fit, ≥0.90 acceptable fit), *RMSEA* root mean square error of approximation (<0.06 good fit, >0.10 unacceptable fit), *SRMR* standardized root mean square residual (<0.08 good fit). ⁴⁶

 Table 3.
 WURSS-K items sorted by item difficulty from easiest to hardest using infit mean square (INFIT MNSQ) and outfit mean square (OUTFIT MNSQ) statistics.

ltem	S		Item difficulty	INFIT MNSQ	INFIT Z value	OUTFIT MNSQ	OUTFIT Z value
3	Stuffy nose	Easy	-1.38	0.90	-0.80	0.87	-0.90
2	Runny nose		-1.30	1.10	0.80	1.26	1.80
6	Cough	Т	-1.27	1.14	1.10	1.13	1.00
7	Feeling tired		-0.43	1.03	0.30	0.89	-0.60
5	Sore throat		-0.27	1.25	1.60	1.06	0.40
9	Sleep		-0.10	1.26	1.60	1.08	0.50
4	Sneezing		0.04	1.09	0.60	1.07	0.40
10	Breathe		0.37	0.77	-1.40	0.61	-2.00
11	Talk		0.79	1.19	1.00	0.92	-0.20
12	Walk, climb stairs, exercise		0.82	1.04	0.30	0.69	-1.20
14	Play with friends	_	0.88	1.49	2.20	0.89	-0.30
13	Go to school		0.92	1.19	0.90	0.68	-1.20
8	Think	Difficult	0.93	1.15	0.80	0.88	-0.40

DISCUSSION

These analyses demonstrate that WURSS-K is a valid and reliable instrument for assessing ARI in children 4–10 years old. The congeneric (Model 1) and bifactor (Model 5) model results support the validity of WURSS-K as either a two-factor structure or a single global structure. The acceptance of partial tau equivalent (Model 3), with strong corresponding Cronbach alpha and omega coefficients, demonstrates internal consistency across the two factors of symptoms and functionality.

Comparing the WURSS-K global structure across the Madison and URECA samples, structural invariance was accepted for both configural and weak invariance, providing evidence for the robustness of this instrument across two pediatric populations with different demographic characteristics.

Item response analysis suggested that the functionality items were more difficult for children to answer. Rating the impact of the illness on ability to "think" was the most difficult, followed by impact on ability to "go to school." This is understandable, as these assessments require a certain level of comprehension that may be difficult for younger children to evaluate or for parents to assess. In this study, 38% completed WURSS-K by themselves without any help (mean age 8.0 years) and 17% required a lot of help (mean age 6.1 years). To our knowledge, this is the first standardized and validated assessment tool for ARI that children can complete by themselves if able.

This study has several limitations, including sample size and potential selection bias. Another limitation is that the URECA sample population collected information at a single time point, thus limiting information about the illness over time. Further research is needed to address instrument responsiveness and the clinical significance of cumulative score values.

CONCLUSION

WURSS-K shows strong psychometric properties for validity and reliability in children aged 4–10 years during ARI episodes. This instrument will be useful in clinical trials or observational studies when assessing children with ARI in the ambulatory setting.

WURSS-K and other WURSS instruments are available free of charge for educational and public health purposes with

registration at the following website: https://www.fammed.wisc. edu/wurss/. Commercial for-profit use will need to be licensed by the Wisconsin Alumni Research Foundation.

ACKNOWLEDGEMENTS

This work was supported by the following grants NIAID RO1 Al097172, NIH/NIAID UM1 Al114271, NCCIH K24AT006543, NCCIH T32AT006956, and HRSA T32HP10010. During the study, B.B. was supported by a mid-career research and mentoring grant from NCCIH (K24AT006543); S.H. received support from a research training grant from NCCIH (K24AT006543); S.H. received support from a grant from University of Wisconsin Department of Family Medicine and Community Health was awarded to S.H. K.M.S. was supported by the University of Wisconsin Department of Family Medicine and Community Health Primary Care Research Fellowship, funded by grant T32HP10010 from the Health Resources and Services Administration. E.R.W. was supported by NIAID RO1 Al097172 and J.E.G. was supported by NIH/NIAID UM1 Al114271.

We are grateful to the Department of Pediatrics study nurses—Stacey Moyer, Andrea Blom, and Patricia Filas-Mortenson—and the URECA study coordinators for their assistance with the initial administration of WURSS-K. Also, we thank our undergraduate students Lily Comp, Daniel Fehrenbach, and Ina Kansariwala for assisting with subject recruitment and data collection and Jennifer Dykema from the University of Wisconsin Survey Center for helping to create the daily symptom form.

AUTHOR CONTRIBUTIONS

K.M.S.: analysis and interpretation of data, drafting and revising the article, and final approval of the version to be published. R.L.B.: conceptualization, design, analysis and interpretation of data, drafting and revising the article, final approval of the version to be published. S.H.: conceptualization, design, acquisition of data, revising the article, and final approval of the version to be published. M.C.: conceptualization, design, acquisition of data, revising the article, and final approval of the version to be published. J.E.G.: conceptualization, design, acquisition of data, revising the article, and final approval of the version to be published. J.E.G.: conceptualization, design, acquisition of data, revising the article, and final approval of the version to be published. E.R.W.: conceptualization, design, revising the article, and final approval of the version to be published. B.B.: conceptualization, design, acquisition of data, ravising the article, and final approval of the version to be published.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41390-021-01395-9.

Competing interests: The authors declare no competing interests.

Statement of consent: Parental consent and child assent were obtained for all participants in this study.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Hsiao, C. J., Cherry, D. K., Beatty, P. C. & Rechtsteiner, E. A. National Ambulatory Medical Care Survey: 2007 summary. *Natl Health Stat. Rep.* 27, 1–32 (2010).
- Dingle, J., Badjer, G. & Jordan, W. Illness in the home. A study of 25,000 illnesses in a Group of Cleveland Families (Western Reserve University, 1964).
- Demuri, G. P. et al. Clinical features, virus identification, and sinusitis as a complication of upper respiratory tract illness in children ages 4-7 years. J. Pediatr. 171, 133–139.e1 (2016).
- Monto, A. S., Malosh, R. E., Petrie, J. G., Thompson, M. G. & Ohmit, S. E. Frequency of acute respiratory illnesses and circulation of respiratory viruses in households with children over 3 surveillance seasons. J. Infect. Dis. 210, 1792–1799 (2014).
- Revai, K. et al. Incidence of acute otitis media and sinusitis complicating upper respiratory tract infection: the effect of age. *Pediatrics* 119, e1408–e1412 (2007).
- Fendrick, A. M., Monto, A. S., Nightengale, B. & Sarnes, M. The economic burden of non-influenza-related viral respiratory tract infection in the United States. *Arch. Intern. Med.* **163**, 487–494 (2003).
- Chong, S. L. et al. The modified respiratory index score (RIS) guides resource allocation in acute bronchiolitis. *Pediatr. Pulmonol.* 52, 954–961 (2017).
- Davies, C. J., Waters, D. & Marshall, A. A systematic review of the psychometric properties of bronchiolitis assessment tools. J. Adv. Nurs. 73, 286–301 (2017).
- 9. Destino, L. et al. Validity of respiratory scores in bronchiolitis. *Hosp. Pediatr.* 2, 202–209 (2012).

- Fernandes, R. M. et al. Validity of bronchiolitis outcome measures. *Pediatrics* 135, e1399–e1408 (2015).
- Golan-Tripto, I. et al. Modified Tal score: validated score for prediction of bronchiolitis severity. *Pediatr. Pulmonol.* 53, 796–801 (2018).
- Justicia-Grande, A. J. & Martinón-Torres, F. The ReSVinet score for bronchiolitis: a scale for all seasons. Am. J. Perinatol. 36, S48–S53 (2019).
- Justicia-Grande, A. J. et al. Development and validation of a new clinical scale for infants with acute respiratory infection: the resvinet scale. *PLoS ONE* **11**, 1–15 (2016).
- Mussman, G. M. et al. Respiratory scores as a tool to reduce bronchodilator use in children hospitalized with acute viral bronchiolitis. *Hosp. Pediatr.* 7, 279–286 (2017).
- Rödl, S. et al. Prospective evaluation of clinical scoring systems in infants with bronchiolitis admitted to the intensive care unit. *Eur. J. Clin. Microbiol. Infect. Dis.* 31, 2667–2672 (2012).
- Santanello, N. C. et al. Validation of a pediatric caregiver diary to measure symptoms of postacute respiratory syncytial virus bronchiolitis. *Pediatr. Pulmonol.* 40, 31–38 (2005).
- 17. van Miert, C. et al. Development and validation of the liverpool infant bronchiolitis severity score: a research protocol. J. Adv. Nurs. **70**, 2353–2362 (2014).
- Anheyer, D., Cramer, H., Lauche, R., Saha, F. J. & Dobos, G. Herbal medicine in children with respiratory tract infection: systematic review and meta-analysis. *Acad. Pediatr.* 18, 8–19 (2018).
- Fahey, T., Stocks, N. & Thomas, T. Systematic review of the treatment of upper respiratory tract infection. Arch. Dis. Child. 79, 225–230 (1998).
- Jacobs, B. et al. Canadian Acute Respiratory Illness and Flu Scale (CARIFS): development of a valid measure for childhood respiratory infections. J. Clin. Epidemiol. 53, 793–799 (2000).
- Barrett, B. et al. The Wisconsin Upper Respiratory Symptom Survey (WURSS): a new research instrument for assessing the common cold. J. Fam. Pract. 51, 265 (2002).
- Barrett, B. et al. The Wisconsin Upper Respiratory Symptom Survey is responsive, reliable, and valid. J. Clin. Epidemiol. 58, 609–617 (2005).
- 23. Barrett, B. et al. Validation of a short form Wisconsin upper respiratory symptom survey (WURSS-21). *Health Qual. Life Outcomes* **7**, 76 (2009).
- 24. Barrett, B. et al. Echinacea for treating the common cold a randomized trial. Ann. Intern. Med. **153**, 769–777 (2010).
- 25. Barrett, B. et al. Meditation or exercise for preventing acute respiratory infection: a randomized controlled trial. *Ann. Fam. Med.* **10**, 339–346 (2012).
- Barrett, B. et al. Placebo effects and the common cold: a randomized controlled trial. Ann. Fam. Med. 9, 312–322 (2011).
- 27. Wong, L. D. & Baker, C. M. Pain in children: comparison of assessment scales. *Pediatr. Nurs.* **14**, 9–17 (1988).
- Gern, J. E. et al. The Urban Environment and Childhood Asthma (URECA) birth cohort study: design, methods, and study population. *BMC Pulm. Med.* 9, 1–15 (2009).
- 29. Jackson, G., Dowling, H. & Muldoon, R. Present concepts of the common cold. *Am. J. Public Health* **52**, 940–945 (1962).
- Harris, P. A. et al. Research electronic data capture (REDCap)- a metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inform. 42, 377–381 (2009).
- Little, R. J. A. A test of missing completely at random for multivariate data with missing values. J. Am. Stat. Assoc. 83, 1198–1202 (1988).
- Royston, P. & White, I. Multiple Imputation by Chained Equations (MICE): implementation in Stata. J. Stat. Software 45, 1–20 (2011).
- Asparouhov, T. Sampling weights in latent variable modeling. *Struct. Equ. Model.* 12, 411–434 (2005).
- Polit, D. F. & Beck, C. T. Nursing Research: Generating and Assessing Evidence for Nursing Practice 9th edn. (Wolters Kluwer, 2012).
- Graham, J. M. Congeneric and (essentially) tau-equivalent estimates of score reliability: what they are and how to use them. *Educ. Psychol. Meas.* 66, 930–944 (2006).
- Byrne, B. M., Shavelson, R. J. & Muthen, B. Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychol. Bull.* **105**, 456–466 (1989).
- Yung, Y. F., Thissen, D. & McLeod, L. D. On the relationship between the higherorder factor model and the hierarchical factor model. *Psychometrika* 64, 113–128 (1999).
- Stucky, B. D., Thissen, D. & Edelen, M. O. Using logistic approximations of marginal trace lines to develop short assessments. *Appl. Psychol. Meas.* 37, 23–39 (2013).
- 39. Gorsuch, R. L. Factor Analysis 2nd edn. (Erlbaum, 1983).
- Muthén, B., Kaplan, D. & Hollis, M. On structural equation modeling with data that are not missing completely at random. *Psychometrika* 52, 431–462 (1987).
- Rodriguez, A., Reise, S. P. & Haviland, M. G. Evaluating bifactor models: calculating and interpreting statistical indices. *Psychol Methods.* 21, 137–150 (2016).

- Linacre, J. M. Winsteps[®] Rasch Measurement Computer Program (Beaverton, 2013).
 Linacre, J. M. What do infit and outfit, mean-square and standardized mean? Rasch Meas. Trans. 16, 878 (2002).
- Wright, B. D. & Linacre, J. M. Reasonable mean-square fit values. *Rasch Meas. Trans.* 8, 370 (1994).
- Smith, R. M., Schumacker, R. E. & Bush, M. J. Using item mean squares to evaluate fit to the Rasch model. J Outcome Meas. 2, 66–78 (1998).
- Brentari, E. & Golia, S. Unidimensionality in Rasch model: how to detect and interpret. *Statistica* 67, 253–261 (2007).
- Reckase, M. D. Unifactor latent trait models applied to multi-factor tests: results and implications. J. Educ. Stat. 4, 207–230 (1979).
- Hu, L. & Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55 (1999).