# Rasch Analysis of The WURSS-21 Dimensional Validation and Assessment of Invariance

### Abstract

**Background:** The purpose of this study is to use Rasch analysis to explore the validity of considering self-report scores from Wisconsin Upper Respiratory Symptom Survey (WURSS-21) as a single global illness severity domain. The WURSS-21 is a widely used questionnaire instrument that assesses symptom severity and functional impact of common cold and flu-like illness.

**Methods:** This study applies item response theory, specifically Rasch modeling, to investigate dimensional and measurement properties of the WURSS-21, and looks at invariance over time. The data assessed represents 1167 people, each scoring the WURSS-21 once daily for up to seven consecutive days of acute upper respiratory infection (URI) illness.

**Results:** Rasch analysis supports a single domain WURSS-21 global symptom score. Assessment of differential item functioning across seven days of illness provides evidence for measurement invariance. While individual items rating physical symptoms were somewhat variable, items rating functional impairment and quality of life impact appeared quite consistent across a single domain over seven days of illness.

**Conclusion:** Rasch analysis of WURSS-21 items provides evidential support for a single invariant domain. These findings support the practice of using a simply summed daily global illness severity score to represent the overall symptomatic and functional impairments arising from URI.

**Keywords:** Common cold; Patient reported outcomes; Quality of life; URI; Validation

**Roger L. Brown PhD**
*University of Wisconsin School of Nursing, Research Design & Statistics Unit*
*Department of Family Medicine and Community Health, University of Wisconsin School of Medicine and Public Health*
Email: Roger.brown@wisc.edu

**Chidi N. Obasi MD PhD**
*Department of Family Medicine and Community Health, University of Wisconsin School of Medicine and Public Health*
Email: cobasi@wisc.edu

**Bruce Barrett MD PhD**
*Department of Family Medicine and Community Health, University of Wisconsin School of Medicine and Public Health*
Email: bruce.barrett@fammed.wisc.edu

**\*Corresponding author:** Bruce Barrett, MD, PhD, Professor, Department of Family Medicine and Community Health, University of Wisconsin – Madison, 1100 Delaplaine, Ct. Madison, WI 53715, Phone: (608) 263 – 2220, Fax: (608) 263– 5813, Email: bruce.barrett@fammed.wisc.edu

**Abbreviations:** WURSS: Wisconsin Upper Respiratory Symptom Survey; URI: Acute Upper Respiratory Infection; CTT: Classical Test Theory; IRT: Item Response Theory; RSM: Rating Scale Model; MNSQ: Mean Square; PCA: Principal Components Analysis; DIF: Differential Item Functioning

## Background

Acute upper respiratory infection (URI) illness is a clinical syndrome produced from viral infection of the upper respiratory tract. A wide variety of etiological agents are involved, including rhinovirus, coronavirus, adenovirus, influenza, parainfluenza and respiratory syncytial virus [1]. Influenza is often classified separately, but usually causes an illness syndrome very similar to other URIs [2]. In the U.S., non-influenza URI has an estimated annual health cost of $40 billion, including 40 million missed work and school days [3]. In a single year, influenza can lead to 31 million outpatient visits, 3 million hospitalizations, 610,000 life-years lost, and an estimated economic impact as high as $87 billion [4].

The WURSS-21 is a valid and reliable self-report research tool incorporating specific symptoms and functional impairments common to URI illness [5-7]. Ten specific symptoms assessed include runny nose, plugged nose, sneezing, sore throat, scratchy throat, cough, hoarseness, head congestion, chest congestion, and feeling tired. The instrument also includes nine functional items rating ability to think clearly, sleep well, breathe easily, exercise, work inside and outside the home, accomplish daily activities, interact with others and live one's personal life. An introductory item rates overall illness severity, and a concluding item rates change-since-yesterday. All items are scored on an 8-point Likert scales from 0 (absent or no impairment) through 1(very mild), 3(mild), 5(moderate) and 7(severe) (Table 1).

The original WURSS was developed using individual face-to-face interviews and focus groups among people recruited from the community [8] with Jackson-defined colds [9]. Semi-structured interviews included open-ended questions aimed at eliciting terminology for assessing symptoms and quality of life values related to experienced cold illness. Of more than 150 terms used to define symptomatic or functional impairment, 42 were chosen for the original WURSS instrument. Adding an introductory global severity item (How sick do you feel today?) and a concluding daily change item (Compared to yesterday, I feel that my cold is.....) led to the WURSS-44 [8]. Subsequently, assessment of item-level assessment of responsiveness and importance to patients produced the WURSS-21 [5]. The WURSS-44 and WURSS-21 have been independently validated [5,7]. Desire to reduce the

time and burden associated with WURSS-21 completion led to the development of the WURSS-11 with similar dimensional construct [6].

While the initial validation of the WURSS-44 using factor analysis based on classical test theory suggested a 10-dimensional structure [7], and the WURSS-21 to WURSS-11 derivation work indicated a 3-dimensional structure [6], we have generally recommended using a simple sum of the 19 items as the most appropriate daily global severity score. Numerous studies have used this simple sum global severity WURSS-21 score, including several NIH-sponsored randomized clinical trials [10-15]. To

date, investigators at more than 250 institutions in more than 50 countries have registered to use one or more versions of WURSS. For non-profit and educational purposes, use of WURSS is free. License fees for commercial use go through the Wisconsin Alumni Research Foundation.

This current study uses Rasch analysis to explore the validity of the common practice of using a global score, and assesses invariance of the measure across the 7-day timeline of a typical URI illness. Our purpose is therefore to use the Rasch item response theory method to assess the validity of treating the WURSS-21 as a single global measurement domain [16].

**Table 1:** WURSS-21 symptom items.

| Symptom: Please rate the average severity of your cold symptoms over the last 24 hours for each symptom: | | Function: Over the last 24 hours, how much has your cold interfered with your ability to: | |
|---|---|---|---|
| **Item Number** | **Symptom** | **Item Number** | **Symptom** |
| I1 | Runny Nose | I11 | Think Clearly |
| I2 | Plugged Nose | I12 | Sleep Well |
| I3 | Sneezing | I13 | Breathe Easily |
| I4 | Sore Throat | I14 | Walk, Climb Stairs, Exercise |
| I5 | Scratchy Throat | I15 | Accomplish Daily Activities |
| I6 | Cough | I16 | Work Outside the Home |
| I7 | Hoarseness | I17 | Work Inside the Home |
| I8 | Head Congestion | I18 | Interact with Others |
| I9 | Chest Congestion | I19 | Live your Personal Life |
| I10 | Feeling Tired | | |

## Methods

### Data sources

Data for this paper come from four studies using the WURSS-21 instrument. These include the WURSS-44 validation (WURSS-21 development) study [7], the WURSS-21 validation study [5], and two clinical trials that used the WURSS-21 [11,12]. Together these total a cohort of n=1167 people with URI illnesses self-reported daily on the WURSS-21, with sample size decreasing over time as people recover from illness: day1 = 1167, day2 = 1157, day3 = 1153, day4 = 1144, day5 = 1112, day6 = 1048, and day7 = 945. Inclusion of a viable URI illness was defined as having self-identified common cold, at least one nasal or throat symptom, and a score ≥2 points on the Jackson scale [9]. Histories of allergy and asthma were reasons for exclusion if active symptoms were observed at enrollment. Use of antibiotics or immune related medication was also reason for exclusion. The end of the illness was confirmed by at least two subsequent days of not having symptoms. All study protocols were approved and monitored by the University of Wisconsin-Madison Institutional Review Board.

### Statistical analysis approach

While Classical test Theory (CTT) was originally used in assessing psychometric properties of the WURSS-21, this method is limited by its focus on whole test assessments of

multi-dimensional structure, and by the assumption that items are equally difficult for participants to understand and respond to. The Item Response Theory (IRT) approach may be more appropriate, as it assesses properties of the individual items across populations and within individuals over time, allowing for differential item difficulties [16-19]. IRT offers important advantages over CTT, especially when employing the Rasch Model [16]. Rasch analysis, in contrast to the CTT approach, allows for the assumption that the set of symptom items is intended to measure a single domain. This fits with assessment of URI illness episodes, usually considered discrete events, and is consistent with the operational basis of the WURSS-21, which is used to provide a single score for each day of illness. It is not uncommon to first use a CTT factor analysis model to explore item-domain structure, followed by a Rasch-based IRT analysis to assess the quality of the items in the larger unitary domain [20].

### Rasch model

A Rasch Model [16,21,22] was used to assess person and item reliability, item statistics and ordering of response categories, using the WINSTEP software Version 3.80 [23]. A special case of the Rasch Model for use with polytomously scored items, known as the Rating Scale Model (RSM), was employed. The RSM assumes all items are equally discriminating and have the same number of response categories and estimates a person's probability of

responding to a certain item category [19]. The model may be written as:

$$Prob\{X_{vi}=x|\theta_v,\beta_i,\omega_h\}=\frac{\exp\left[\sum_{h=0}^{x}\left(\theta_v-(\beta_i-\omega_h)\right)\right]}{\sum_{x=0}^{m}\exp\left[\sum_{h=0}^{x}\left(\theta_v-(\beta_i-\omega_h)\right)\right]} \quad (1)$$

Where $v$ = 1, …, $n$ are the respondents, i=1, …., k are the items, h= 0,…,m -1 are the number of thresholds, x = 0,…,m are the response categories, and $X_{vi}$ is the response vector for each respondent. The parameters of $\theta$ and $\beta$ represent the respondent and item parameters respectively, with the parameter $\omega_h$ the common set of thresholds applied to all items.

Two fit indices were used to investigate item concurrence with the overall symptom domain, which were the mean square (MNSQ) outfit and in fit. The outfit MNSQ measures the average mismatch between the Rasch model and the data, and is sensitive to extreme values. The infit MNSQ is more sensitive to patterns of responses to items targeted for the subject matter. The expected value for both outfit and infit MNSQ is 1, with a range of values from 0 to infinity. Values near 1 indicate little distortion of the measurement system, and values greater than 2 indicates that the item fails to define the same construct as the other items do in a domain, and degrades the measurement. MNSQ values lower than 0.5 may be an indication of item redundancy with values 0.5 to 1.5 considered satisfactory [24,25]. In addition to assessing item-domain integrity, these measures may help detect problematic symptom items.

An underlying assumption of the RSM approach is uni-dimensionality [16], in that the symptom items measure only a single domain. To test this assumption, we used a post-hoc approach based on a principal component analysis (PCA) conducted on the standardized residuals produced from the RSM [26]. When there is the presence of a dominant factor with over 20% of variance explained, Reckase [27] suggests that the unidimensional assumption may be considered acceptable. In conjunction with RSM, confirmatory factor analysis was also employed to assess the assumption of uni-dimensionality.

## Invariance over time

A major challenge in conducting longitudinal assessments is the possibility that measures developed for a given domain at one particular time may not be assessing the same domain at other points in time or differential item functioning (DIF). DIF refers to the condition in which an item displays different properties at different time periods after controlling for the abilities of the groups [28]. Investigating invariance over time assesses whether the WURSS-21 measures the same underlying symptom severity domain across the duration of URI illness. This allows both comparability and a meaningful interpretation of respondents' symptom severity scores in longitudinal studies. Assessment of measurement invariance in the RSM context can be conceptualized as asking whether item parameters are applicable to the multiple assessments over time, and whether individual items have stable relationships to the domain of interest across longitudinal time measures [29]. When self-reports occur over multiple time periods, response dependency may impact underlying RSM assumptions [30]. In order to construct a "repetition-bias-free" RSM for multi-item longitudinal instruments, the first time measurement can be considered the benchmark, with subsequent time points randomized. Since the measurement framework is anchored, this controls for within-person over-time dependency, allowing all time-points to be analyzed together.

## Results and Discussion

The WURSS-21 data were obtained from four studies (n=1167 total) spanning the years 2002-2010, as outlined above. The characteristics of the participants are presented in (Table 2).

**Table 2:** Characteristics of the study participants.

|  | Combined Study Data | Study Data-1 [7] | Study Data-2[5] | Study Data-3[11] | Study Data-4 [12]* |
|---|---|---|---|---|---|
| Participants | 1167 | 149 | 232 | 718 | 66 |
| (% Female) | 66% | 70% | 66% | 64% | 80% |
| Mean Age In Years (SD†) | 35 (15) | 35.5 (15) | 34.1 (14) | 33.7 (14) | 59.3 (7) |
| Race | 1024 | 129 | 199 | 631 | 65 |
| (% Caucasian) | 88% | 85% | 87% | 88% | 98% |
| Smoking Status | 756 | 88 | 143 | 468 | 57 |
| (% Non Smoking) | 65% | 59% | 62% | 65% | 86% |
| Education (%≥College Graduate)‡ | 549 (47%) | 88 (62%) | 105 (46%) | 314 (44%) | 42 (64%) |
| Income | 465 | 54 | 76 | 301 | 34 |
| (% ≥$50,000 )# | 40% | 36% | 33% | 42% | 52% |

*66 participants had been ill from the total of 149 participants monitored during study-4;

† SD=Standard deviation;

‡ (4% combined missing data on education; 6% missing data on education during study-3)

# (3% combined missing data on income; 5% missing data on income during study-3; 2% missing data on income during stud

## Rasch model results

Rasch model principal components analysis using standardized residuals for data across all seven days showed that a single dominant factor explained 57% of total variance. When a dominant factor explains over 20% of variance, Reckase [27] suggests the use of uni dimensional RSM model. The single domain was also supported by confirmatory factor analysis, providing fit indices of 0.924 for the confirmatory factor index, and 0.915 for the Tucker Lewis Index.

IRT analysis was performed to estimate the goodness-of-fit (infit and outfit) of the observed data to the model-expected data and the item symptom rarity of the 19 items from WURSS-21 (Table 3). The fit indices determine how well each item contributes to a single common construct. The infit MNSQ index is more sensitive to unexpected responses to an item near a person's ability level, and the outfit MNSQ index is sensitive to unexpected responses to more distant items [31]. According to Wright and Linacre [25], item fit indices of 1.0 are ideal, and values between 0.5 and 1.5 considered satisfactory indications of model-data fit [24,32]. While typical Rasch modeling uses the terminology of difficult/easy items in the assessment of an instrument, Linacre [33] suggests adjusting the terminology for symptom measures as rarely observed (=difficult item) and often observed (=easy item), which is appropriate here, as the WURSS-21 is a symptom instrument.

Strongly supporting unidimensional integrity, all of the infit MNSQ over time (19 items over seven days) were in the productive range of 0.5 to 1.5 (Table 3). Outfit statistics were also strongly supportive of this model. The exception was the symptom item of "Runny nose," item I1, indicated rarity later on in the progression of the cold, as indicated by higher outfit MNSQ values. Although rare (difficult) in that specific time frame, it was still considered a productive (useful) item in the assessment of the overall domain.

The log odds of the probability (item rarity) are shown for each of the seven days, reflecting the rarity for an individual to assess the symptom (Figure 1). Higher scores indicate rarer observations with a symptom, and lower scores indicate a more prevalent occurrence of the symptom. The symptom that respondents indicated most problematic (rare) in the early stages of the cold (days 1-3) was chest congestion (I9), with the most highly observed (prevalent, easy, useful) symptom being feeling tired (I10). The pattern of symptom observation (rare $v_s$ prevalent) across each of the seven days of the illness was strikingly similar (Figure 1).

## Time invariance

Rasch modeling was used to assess item equivalence or DIF over the span of seven days of the URI [28]. The requirement that an instrument works invariantly across time ensures that changes in symptom measures reflect real improvements in experienced symptoms and not just differences in the measurement of the symptoms. (Figure 2) show the DIF size (difference between the individual day log odds and an overall log odds), indicating that more item difference across the seven days of the URI was noticed for the physical symptoms of sneezing (I3), cough (I6), chest congestion (I9), etc. and least different across time for the more social-type of symptoms, (e.g., feeling tired (I10), interacting with others (I18). While more symptom differences were encountered with the reporting of physical symptoms (I1-I9) rather than functional impairments (I10-I19), all differences were considered reasonable for an assessment of item invariance across the seven day time period (Figure 2). It was noteworthy that participants were variable relative to the symptom of cough (I6), with the symptom reported as rare initially in the progression of the illness (days 1, 2, and 3), but then becoming more prevalent latter on (days 4-7).
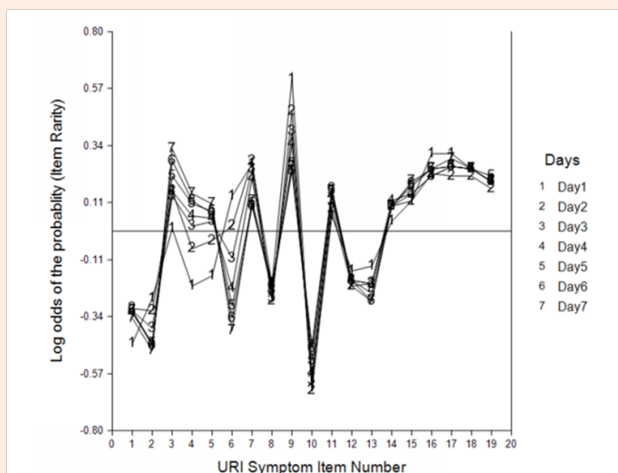


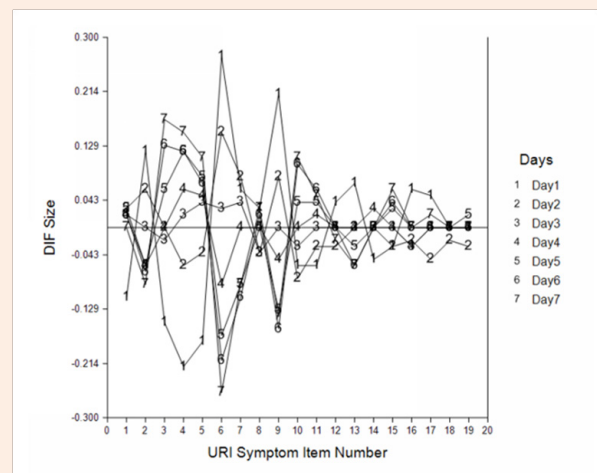**Figure 1:** Log odds of the probability (Item rarity) for each day.



**Figure 2:** URI Item symptom difference across days of the cold.

**Table 3:** Item fit statistics from the WURSS-21 across the 7 days.

| Item | Day 1 INFIT MNSQ | Day 1 OUTFIT MNSQ | Day 2 INFIT MNSQ | Day 2 OUTFIT MNSQ | Day 3 INFIT MNSQ | Day 3 OUTFIT MNSQ | Day 4 INFIT MNSQ | Day 4 OUTFIT MNSQ | Day 5 INFIT MNSQ | Day 5 OUTFIT MNSQ | Day 6 INFIT MNSQ | Day 6 OUTFIT MNSQ | Day 7 INFIT MNSQ | Day 7 OUTFIT MNSQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I1 | 1.28 | 1.4 | 1.28 | 1.46 | 1.29 | 1.51 | 1.27 | 1.44 | 1.42 | 1.68 | 1.39 | 1.74 | 1.45 | 1.88 |
| I2 | 1.19 | 1.2 | 1.18 | 1.24 | 1.19 | 1.27 | 1.13 | 1.27 | 1.09 | 1.19 | 1.04 | 1.21 | 1.1 | 1.26 |
| I3 | 1.32 | 1.44 | 1.25 | 1.33 | 1.28 | 1.36 | 1.27 | 1.31 | 1.28 | 1.27 | 1.37 | 1.43 | 1.34 | 1.35 |
| I4 | 1.47 | 1.54 | 1.45 | 1.47 | 1.44 | 1.48 | 1.44 | 1.4 | 1.42 | 1.4 | 1.45 | 1.31 | 1.37 | 1.18 |
| I5 | 1.25 | 1.34 | 1.18 | 1.2 | 1.27 | 1.28 | 1.26 | 1.2 | 1.23 | 1.17 | 1.26 | 1.07 | 1.19 | 0.99 |
| I6 | 1.08 | 1.11 | 1.28 | 1.34 | 1.24 | 1.26 | 1.25 | 1.35 | 1.25 | 1.29 | 1.32 | 1.38 | 1.35 | 1.39 |
| I7 | 1.18 | 1.16 | 1.26 | 1.21 | 1.29 | 1.2 | 1.31 | 1.31 | 1.3 | 1.19 | 1.36 | 1.19 | 1.42 | 1.25 |
| I8 | 0.94 | 0.92 | 1.01 | 0.99 | 1.01 | 0.97 | 1.05 | 0.99 | 1.12 | 1.1 | 1.03 | 0.93 | 1.1 | 0.94 |
| I9 | 1.21 | 1.07 | 1.25 | 1.16 | 1.26 | 1.12 | 1.27 | 1.08 | 1.3 | 1.09 | 1.21 | 0.91 | 1.27 | 0.94 |
| I10 | 0.85 | 0.86 | 0.91 | 0.91 | 0.88 | 0.89 | 0.83 | 0.89 | 0.85 | 0.9 | 0.87 | 0.91 | 0.96 | 1.02 |
| I11 | 0.8 | 0.77 | 0.76 | 0.75 | 0.8 | 0.8 | 0.83 | 0.76 | 0.84 | 0.73 | 0.88 | 0.89 | 0.9 | 0.88 |
| I12 | 1.09 | 1.07 | 1.1 | 1.09 | 1.06 | 1.03 | 1.06 | 0.97 | 1.02 | 0.89 | 0.96 | 0.85 | 0.96 | 0.81 |
| I13 | 0.89 | 0.85 | 0.81 | 0.8 | 0.8 | 0.79 | 0.85 | 0.81 | 0.78 | 0.72 | 0.8 | 0.76 | 0.82 | 0.73 |
| I14 | 0.83 | 0.78 | 0.74 | 0.73 | 0.76 | 0.71 | 0.8 | 0.73 | 0.8 | 0.7 | 0.84 | 0.69 | 0.79 | 0.6 |
| I15 | 0.59 | 0.56 | 0.6 | 0.57 | 0.6 | 0.55 | 0.66 | 0.57 | 0.61 | 0.53 | 0.68 | 0.51 | 0.66 | 0.48 |
| I16 | 1 | 0.87 | 1.02 | 0.9 | 0.97 | 0.84 | 0.96 | 0.79 | 0.95 | 0.71 | 1 | 0.7 | 0.97 | 0.67 |
| I17 | 0.85 | 0.75 | 0.75 | 0.68 | 0.71 | 0.63 | 0.75 | 0.63 | 0.76 | 0.58 | 0.81 | 0.65 | 0.8 | 0.56 |
| I18 | 0.78 | 0.71 | 0.76 | 0.75 | 0.72 | 0.64 | 0.77 | 0.67 | 0.8 | 0.63 | 0.87 | 0.62 | 0.85 | 0.61 |
| I19 | 0.78 | 0.71 | 0.74 | 0.69 | 0.82 | 0.72 | 0.79 | 0.67 | 0.89 | 0.72 | 0.88 | 0.63 | 0.89 | 0.62 |

Mean Square (MNSQ); Estimates of fit (outfit-outlier-sensitive fit and infit- inlier-sensitive fit)

Infit and outfit MNSQ over time (19 items over seven days) were in the productive range of 0.5 to 1.5

However, the symptom item of "'Runny nose," item I1, indicated rarity later on in the progression of the cold, with higher outfit MNSQ values above 1.50.

## Participant and symptom map

The map of items and individual responders (participants) is illustrated in (Figures 3a-3c) (maps per day). Ideally, location values of each symptom item in the WURSS-21 should cover the continuum, with their distribution sufficiently wide to collect the variability of the URI symptoms. These figures show that both participant and item distributions do provide sufficient variability. The participant-symptom map here shows the distribution of participant scores (left side), and the symptom item level of observation (right side) for the single domain WURSS-21.

Participants with higher WURSS-21 score and items with "rarely observed symptoms" are located on the positive side of the graphic, at the top of the map. One may observe the decline in symptoms as the cold progresses from day 1 to day 7, as shown in changes of WURSS-21 score distributions (left side of the figure).

From an item standpoint, one may see that symptom item I9 (chest congestion) was rarely observed early on in the cold, especially the first three days. Whereas, item I10 (feeling tired), was reported more frequently, and was consistent throughout the seven day period.
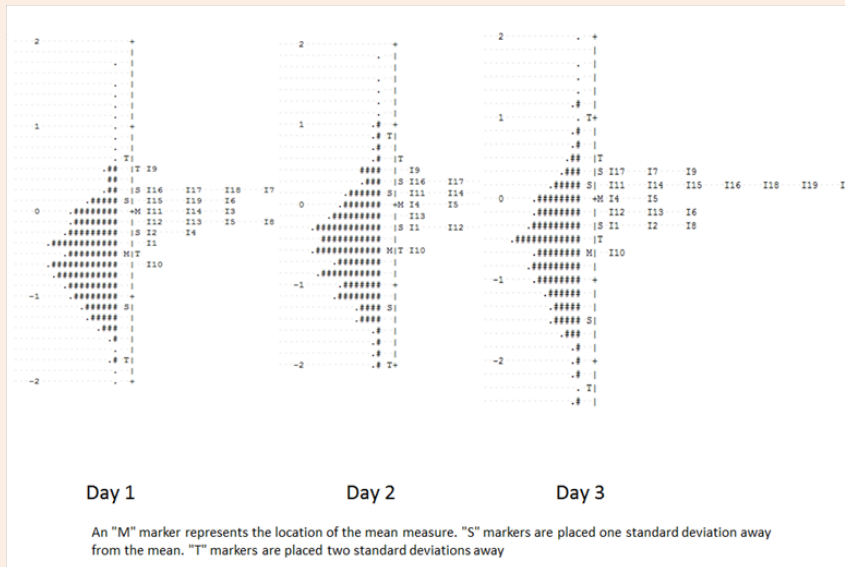


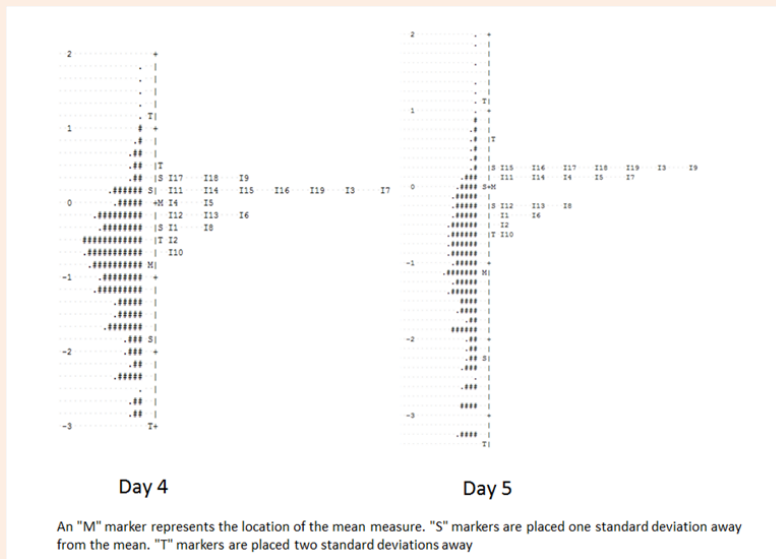**Figure 3a:** Patient and Item Symptom Maps.



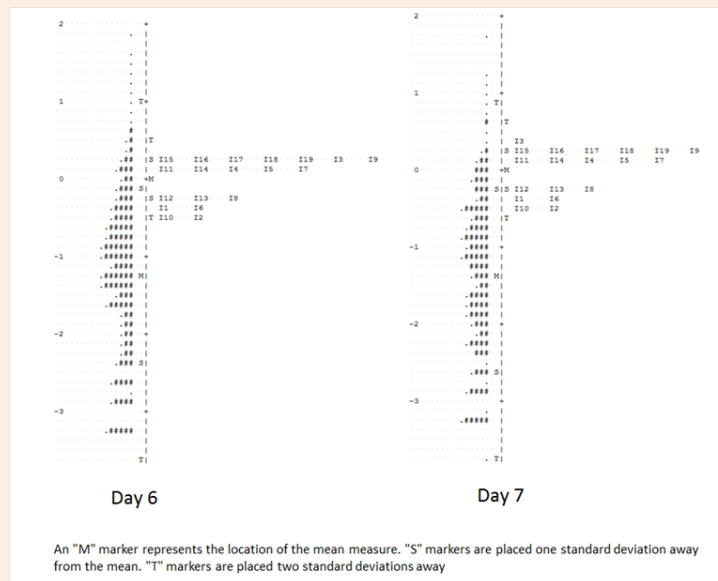**Figure 3b:** Patient and Item Symptom Maps.

**Figure 3c:** Patient and Item Symptom Maps.

## Conclusion

The use of the item response theory Rasch model provided useful detailed insight as well as support for a single domain WURSS-21 symptom score. Assessment of DIF scores provided evidence for measurement invariance across the first seven days of the illness. While physical symptoms (items I1-I9) were slightly more variable relative to occurrence in the unitary domain structure, the functional and quality of life responses appeared very consistent in the domain over the seven days of illness. In general, analysis of individual items provided strong evidential support for an invariant domain measure. This supports the use of a simply-summed global severity score for the WURSS-21, consistent with its common use.

## Acknowledgment

## Authorship Contribution

All authors contributed substantially to the development of this manuscript and approved the final version of this work.

## Conflict of Interest

Drs. Barrett and Brown have transferred copyright of the WURSS-21 licensed to Wisconsin Alumni Research Foundation (WARF). The WURSS-21 is free of charge for all educational and non-profit work, but is occasionally licensed by WARF to companies for research. A small portion of those license fees comes back to Drs. Barrett and Brown as royalties. The authors declare no other conflicts of interests pertaining to this report.

## References

1. Monto AS (2002) Epidemiology of viral respiratory infections. Am J Med 22(112 Suppl 6A): 4S-12S.

2. Eccles R (2005) Understanding the symptoms of the common cold and influenza. Lancet Infect Dis 5(11): 718-725.

3. Fendrick A, Monto A, Nightengale B, Sarnes M (2003) The economic burden of non-influenza-related viral respiratory tract infection in the United States. Arch Intern Med 163(4): 487-494.

4. Molinari NA, Ortega SIR, Messonnier ML,Thompson WW, Wortley PM, et al. (2007) The annual impact of seasonal influenza in the US: measuring disease burden and costs. Vaccine 25(27): 5086-5096.

5. Barrett B, Brown RL, Mundt MP, Thomas GR, Barlow SK, et al. (2009) Validation of a short form Wisconsin Upper Respiratory Symptom Survey (WURSS-21). Health Qual Life Outcomes 12(7): 76.

6. Obasi CN, Brown RL, Barrett BP (2013) Item reduction of the Wisconsin Upper Respiratory Symptom Survey (WURSS-21) leads to the WURSS-11. Qual Life Res 23(4): 1293-1298.

7. Barrett B, Brown R, Mundt M, Safdar N, Dye L, et al. (2005) The Wisconsin Upper Respiratory Symptom Survey is responsive, reliable, and valid. J Clin Epidemiol 58(6): 609-617.

8. Barrett B, Locken K, Maberry R, Schwamman J, Brown R, et al. (2002) The Wisconsin Upper Respiratory Symptom Survey (WURSS): a new research instrument for assessing the common cold. J Fam Pract 51(3): 265.

9. Jackson GG, Dowling HF, Muldoon RL (1962) Acute respiratory diseases of viral etiology. VII. Present concepts of the common cold. Am J Public Health Nations Health 52: 940-945.

10. Heinz SA, Henson DA, Austin MD, Jin F, Nieman DC (2010) Quercetin supplementation and upper respiratory tract infection: A randomized community clinical trial. Pharmacol Res 62(3): 237-242.

11. Barrett B, Brown R, Rakel D, Rabago D, Marchand L, et al. (2011) Placebo effects and the common cold: a randomized controlled trial. Ann Fam Med 9(4): 312-322.

12. Barrett B, Hayney MS, Muller D, Rakel D, Ward A, et al. (2012) Meditation or Exercise for Preventing Acute Respiratory Infection: A Randomized Controlled Trial. Ann Fam Med 10(4): 337-346.

13. Murdoch DR, Slow S, Chambers ST, Jennings LC, Stewart AW, et al. (2012) Effect of vitamin D3 supplementation on upper respiratory tract infections in healthy adults: the VIDARIS randomized controlled trial. JAMA 308(13): 1333-1339.

14. Byun JS, Yang SY, Jeong IC, Hong ke, Kang W, et al. (2011) Effects of So-cheong-ryong-tang and Yeon-gyo-pae-dok-san on the common cold: randomized, double blind, placebo controlled trial. J Ethnopharmacol 133(2): 642-646.

15. Tiralongo E, Lea RA, Wee SS, Hanna MM, Griffiths LR (2012) Randomised, double blind, placebo-controlled trial of echinacea supplementation in air travellers. Evid Based Complement Alternat Med 2012: 1-9.

16. Georg Rasch (1992) Probabilistic Models for Some Intelligence and Attainment Tests.

17. De Boeck P, Wilson M (2005) Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach. Journal of Educational Measurement 42(3): 303-307.

18. Lord FM (1981) Applications of Item Response Theory to Practical Testing Problems. Journal of Educational Measurement 18(3): 178-180.

19. Embretson SE, Reise SP (2000) Item Response Theory for Psychologists. Lawrence Erlbaum Associates 384.

20. Randal J, Engelhard G (2010) Using confirmatory factor analysis and the Rasch model to assess measurement invariance in a high stakes reading assessment. Applied Measurement in Education 23(3): 286-306.

21. Andrich D (1978) Application of a psychometric rating model to ordered categories which are scored with successive integers. Applied Psychological Measurement 2(4): 581-594.

22. Andrich D (1978) Rating formulation for ordered response categories. Psychometrika 43(4): 561-573.

23. Linacre JM (2013a) Winsteps® Rasch measurement computer program. Beaverton.

24. Linacre JM (2002) What do Infit and Outfit, Mean-square and Standardized mean? Rasch Measurement Transactions 16(2): 878.

25. Wright BD, Linacre JM (1994) Reasonable mean-square fit values. Rasch Measurement Transactions 8(3): 370.

26. Brentari E, Golia S (2007) Unidimensionality in the Rasch Model: How to detect and interpret. Statistica 67(3): 253-261.

27. Reckase MD (1979) Unifactor Latent Trait Models Applied to Multi-Factor Tests: Results and Implications. Journal of Educational Statistics 4(3): 207-230.

28. Angoff WH (1992) Perspectives on differential item functioning methodology.

29. Reise S, Widman K, Pugh R (1993) Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. Psychol Bull 114(3): 552-566.

30. Chien TW (2008) Repeated Measure Designs (Time Series) and Rasch. Rasch Measurement Transactions 22(3): 1171.

31. Piquero NL, MacIntosh R, Hickman M (2001) Applying Rasch modeling to the validity of a control balance scale. Journal of Criminal Justice 29(6): 493-505.

32. Smith RM, Schumacker, Bush MJ (1998) Using item mean squares to evaluate fit to the Rasch model. J Outcome Meas 2(1): 66-78.

33. Linacre JM (2013b) Rasch Measurment Forum.