WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

Sleepless Nights, Better Measures? A Head-to-Head Comparison of Bayesian and Frequentist Approaches to Validate Quality-of-Life Instruments in Sleep Disorder Epidemiology.

Billingsley Kaambwa

Flinders University, Adelaide, Australia

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

Sleepless Nights, Better Measures? A Head-to-Head Comparison of Bayesian and Frequentist Approaches to Validate Quality-of-Life Instruments in Sleep Disorder Epidemiology.

Billingsley Kaambwa

Flinders University, Adelaide, Australia

Background

- Item response theory (IRT) and classical test theory (CTT) are common approaches for validating QoL instruments.
- In Bayesian validation, uncertainty expressed by assigning a prior distribution that reflects prior knowledge considered natural approach.
- But little is known about performance of Bayesian performance of IRT/CTT validation vs Frequentist IRT/CTT validation in sleep epidemiology.

Why Sleep?

- Societal Consequences:
  - Productivity, safety, health, and well-being
  - 4-23% for insomnia, 9-38% for Obstructive Sleep Apnoea (OSA)
  - Economic Impact: $1.8 billion (Portugal) - $207.5 billion (USA) in 2017).

Objectives

- Compare the performance of Bayesian IRT/CTT validation vs Frequentist IRT/CTT validation approaches.

- Does this performance differ by type of instrument (generic versus condition(sleep)-specific)?

Methods

Data (1,510 Australians with self-reported sleep disorders)

- Survey, including QoL instruments:
  - Assessment of Quality of Life 4 Dimensions (AQoL-4D)
  - EQ5D 5 Dimensions 5 Levels (EQ5D-5L)
  - ICEpop CAPability measure for Adults (ICECAP-A)
  - Epworth Sleepiness Scale (ESS)
  - Functional Outcomes of Sleep Questionnaire - 10 Dimensions (FOSQ-10)
- Mean age: 46 years.

Flinders University

WCE WORLD CONGRESS OF EPIDEMIOLOGY 2024

Methods

Data (1,510 Australians with self-reported sleep disorders)

- Survey, including QoL instruments:
  - Assessment of Quality of Life 4 Dimensions (AQoL-4D)
  - EQ5D 5 Dimensions 5 Levels (EQ5D-5L)
  - ICEpop CAPability measure for Adults (ICECAP-A)
  - Epworth Sleepiness Scale (ESS)
  - Functional Outcomes of Sleep Questionnaire 10 Dimensions (FOSQ-10)

Generic

- Mean age: 46 years.

Flinders University

WCE WORLD CONGRESS OF EPIDEMIOLOGY 2024

Methods

Data (1,510 Australians with self-reported sleep disorders)

- Survey, including QoL instruments:
  - Assessment of Quality of Life 4 Dimensions (AQoL-4D)
  - EQ5D 5 Dimensions 5 Levels (EQ5D-5L)
  - ICEpop CAPability measure for Adults (ICECAP-A)

  Generic
  - Epworth Sleepiness Scale (ESS)
  - Functional Outcomes of Sleep Questionnaire 10 Dimensions (FOSQ-10)

  Sleep-specific
- Mean age: 46 years.

Flinders University

WCE WORLD CONGRESS OF EPIDEMIOLOGY 2024

Overview of instruments

	AQoL-4D Health-related QoL **4 Dimensions**	**EQ-5D 5L** Health-related QoL **5 Dimensions**	**ICECAP-A** Capability **5 attributes**	**ESS** Sleepiness scale **8 subscales**	**FOSQ-10** Functional status **5 subscales**
Dimensions	1.Independent Living 2.Mental Health 3.Relationships 4.Senses	1.Mobility 2.Self-care; 3.Usual activities 4.Anx/Depression 5.Pain/discomfort	1.Attachment 2.Stability 3.Achievement 4.Enjoyment 5.Autonomy	1.Sitting & Reading 2.Watching TV 3.Sitting inactive 4.Passenger in car 5.Lying down to rest 6.Sitting and talking 7.Sitting quietly 8.Sitting in stationery car	1.General productivity 2.Activity level 3.Vigilance 4.Social outcomes 5.Intimacy and sexual relationships
# Questions & Scoring	**12 Questions** **Utilities (-0.04 – 1)**	**5 Questions** **Utilities (-0.59 – 1)**	**5 Questions** **Utilities (0 – 1)**	**8 Questions** **Summative score (0-24)**	**10 Questions** **Summative score (5-20)**

Overview of instruments

	AQoL-4D Health-related QoL **4 Dimensions**	**EQ-5D 5L** Health-related QoL **5 Dimensions**	**ICECAP-A** Capability **5 attributes**	**ESS** Sleepiness scale **8 subscales**	**FOSQ-10** Functional status **5 subscales**
Dimensions	1.Independent Living 2.Mental Health 3.Relationships 4.Senses	1.Mobility 2.Self-care; 3.Usual activities 4.Anx/Depression 5.Pain/discomfort	1.Attachment 2.Stability 3.Achievement 4.Enjoyment 5.Autonomy	1.Sitting & Reading 2.Watching TV 3.Sitting inactive 4.Passenger in car 5.Lying down to rest 6.Sitting and talking 7.Sitting quietly 8.Sitting in stationery car	1.General productivity 2.Activity level 3.Vigilance 4.Social outcomes 5.Intimacy and sexual relationships
# Questions & Scoring	**12 Questions** **Utilities (-0.04 – 1)**	**5 Questions** **Utilities (-0.59 – 1)**	**5 Questions** **Utilities (0 – 1)**	**8 Questions** **Summative score (0-24)**	**10 Questions** **Summative score (5-20)**

Overview of instruments

	AQoL-4D Health-related QoL **4 Dimensions**	**EQ-5D 5L** Health-related QoL **5 Dimensions**	**ICECAP-A** Capability **5 attributes**	**ESS** Sleepiness scale **8 subscales**	**FOSQ-10** Functional status **5 subscales**
Dimensions	1.Independent Living 2.Mental Health 3.Relationships 4.Senses	1.Mobility 2.Self-care; 3.Usual activities 4.Anx/Depression 5.Pain/discomfort	1.Attachment 2.Stability 3.Achievement 4.Enjoyment 5.Autonomy	1.Sitting & Reading 2.Watching TV 3.Sitting inactive 4.Passenger in car 5.Lying down to rest 6.Sitting and talking 7.Sitting quietly 8.Sitting in stationery car	1.General productivity 2.Activity level 3.Vigilance 4.Social outcomes 5.Intimacy and sexual relationships
# Questions & Scoring	**12 Questions** **Utilities (-0.04 – 1)**	**5 Questions** **Utilities (-0.59 – 1)**	**5 Questions** **Utilities (0 – 1)**	**8 Questions** **Summative score (0-24)**	**10 Questions** **Summative score (5-20)**

Overview of instruments

	AQoL-4D Health-related QoL **4 Dimensions**	**EQ-5D 5L** Health-related QoL **5 Dimensions**	**ICECAP-A** Capability **5 attributes**	**ESS** Sleepiness scale **8 subscales**	**FOSQ-10** Functional status **5 subscales**
Dimensions	1.Independent Living 2.Mental Health 3.Relationships 4.Senses	1.Mobility 2.Self-care; 3.Usual activities 4.Anx/Depression 5.Pain/discomfort	1.Attachment 2.Stability 3.Achievement 4.Enjoyment 5.Autonomy	1.Sitting & Reading 2.Watching TV 3.Sitting inactive 4.Passenger in car 5.Lying down to rest 6.Sitting and talking 7.Sitting quietly 8.Sitting in stationery car	1.General productivity 2.Activity level 3.Vigilance 4.Social outcomes 5.Intimacy and sexual relationships
# Questions & Scoring	**12 Questions** **Utilities (-0.04 – 1)**	**5 Questions** **Utilities (-0.59 – 1)**	**5 Questions** **Utilities (0 – 1)**	**8 Questions** **Summative score (0-24)**	**10 Questions** **Summative score (5-20)**

Methods

Classical Test Theory (CTT) Approach

- Construct (convergent and discriminant) validity
- Confirmatory Factor Analysis.
- Based on International Classification of Functioning, Disability, and Health (ICF) framework.

Flinders University

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

Methods

Classical Test Theory (CTT) Approach

- Convergent Validity
  - AVE > 0.5
  - AVE (average variance extracted) = amount of variation in a latent variable that can be explained by its own indicators.

- Discriminant validity
  - AVE > squared correlation ($Corr^2$)
  - $Corr^2$ = amount of variation in latent variable that can be explained by indicators of another/other latent variables.

Methods

Item Response Theory (IRT) Approach

- Estimate calculated: Discrimination parameters

- Instrument dimensions ranked according to their discrimination parameter values by Factor and Type of instrument.

Methods

Bayesian vs Frequentist Methods

Bayesian

- Bayesian Priors
  - Dirichlet distribution
  - Weakly informative priors on the hierarchical SDs

Frequentist

- Maximum likelihood estimation

Methods

Bayesian vs Frequentist Methods

- Agreement
  - Concordance correlation Coefficients
  - Kendall's Tau-B correlations
  - Krippendorf's alpha

Results

CTT - Convergent and Discriminant Validity Results

			Type of approach	
Instrument	**Factor**	**Validity type**	**Bayesian**	**Frequentist**
EQ5D-5L	1	Convergent	No Problem	No Problem
		Discriminant	No Problem	Problem
	2	Convergent	Problem	Problem
		Discriminant	No Problem	Problem
AQoL-4D	1	Convergent	Problem	Problem
		Discriminant	Problem	Problem
	2	Convergent	Problem	Problem
		Discriminant	Problem	Problem
ICECAP-A	1	Convergent	Problem	No Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
FOSQ-10	1	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
ESS	1	Convergent	Problem	Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem

CTT - Convergent and Discriminant Validity Results

			Type of approach	
Instrument	**Factor**	**Validity type**	**Bayesian**	**Frequentist**
EQ5D-5L	1	Convergent	No Problem	No Problem
		Discriminant	No Problem	Problem
	2	Convergent	Problem	Problem
		Discriminant	No Problem	Problem
AQoL-4D	1	Convergent	Problem	Problem
		Discriminant	Problem	Problem
	2	Convergent	Problem	Problem
		Discriminant	Problem	Problem
ICECAP-A	1	Convergent	Problem	No Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
FOSQ-10	1	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
ESS	1	Convergent	Problem	Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem

AGREEMENT

CC coefficients

- Generic: 0.40
- Sleep: 1.00
- Both: 0.66

Kendall's Tau B

- Generic: 0.41
- Sleep: 1.00
- Both: 0.66

Krippendorf's alpha

- Generic: 0.42
- Sleep: 1.00
- Both: 0.67

CTT - Convergent and Discriminant Validity Results

			Type of approach	
Instrument	**Factor**	**Validity type**	**Bayesian**	**Frequentist**
EQ5D-5L	1	Convergent	No Problem	No Problem
		Discriminant	No Problem	Problem
	2	Convergent	Problem	Problem
		Discriminant	No Problem	Problem
AQoL-4D	1	Convergent	Problem	Problem
		Discriminant	Problem	Problem
	2	Convergent	Problem	Problem
		Discriminant	Problem	Problem
ICECAP-A	1	Convergent	Problem	No Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
FOSQ-10	1	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem
ESS	1	Convergent	Problem	Problem
		Discriminant	Problem	Problem
	2	Convergent	No Problem	No Problem
		Discriminant	Problem	Problem

AGREEMENT

CC coefficients

- Generic: 0.40
- Sleep: 1.00
- Both: 0.66

Kendall's Tau B

- Generic: 0.41
- Sleep: 1.00
- Both: 0.66

Krippendorf's alpha

- Generic: 0.42
- Sleep: 1.00
- Both: 0.67

Item Response Theory (IRT): Discrimination parameters ranks

Instrument	Factor	domain	Bayesian	Frequentist
EQ5D-5L	1	Mobility	1	1
	1	Activity	2	2
	2	Selfcare	3	2
	2	Pain	1	1
	2	Anxiety	2	3
AQoL-4D	1	Self	4	2
	1	House	6	3
	1	Mobility	3	1
	1	Rel	7	6
	1	Friend	8	7
	1	Fam	2	4
	1	Sleep	9	9
	1	General	5	8
	1	Pain	1	5
	2	Eye	1	3
	2	Hear	3	1
	2	Comm	2	2
ICECAP-A	1	Secure	2	2
	1	Independent	3	3
	1	Achieve	1	1
	2	Love	2	2
	2	Joy	1	1

Item Response Theory (IRT): Discrimination parameters ranks

Instrument	Factor	domain	Bayesian	Frequentist
ESS	1	Read	2	1
	1	TV	1	2
	1	Rest	3	3
	2	Sit	1	3
	2	Car	5	5
	2	Talk	3	1
	2	Lunch	4	4
	2	Traffic	2	2
FOSQ-10	1	Concentration	1	6
	1	Memory	4	7
	1	Hobby	2	3
	1	House	3	5
	1	Done	5	1
	1	Finance	7	4
	1	Work	6	2
	2	Meal	3	3
	2	Short	1	1
	2	Long	2	2

Item Response Theory (IRT): Discrimination parameters ranks

Bayesian values lower by about 15%

Item Response Theory (IRT): Discrimination parameters ranks

Bayesian values lower by about 15%

AGREEMENT

CC coefficients

- Sleep: 0.76

Kendall's Tau B

- Sleep: 0.45

Krippendorf's alpha

- Sleep: 0.34

Flinders University

WCE WORLD CONGRESS OF EPIDEMIOLOGY 2024

Item Response Theory (IRT): Discrimination parameters ranks

Bayesian values lower by about 15%

AGREEMENT

CC coefficients

- Generic: 0.18
- Sleep: 0.76

Kendall's Tau B

- Generic: 0.19
- Sleep: 0.45

Krippendorf's alpha

- Generic: 0.28
- Sleep: 0.34

Flinders University

WCE WORLD CONGRESS OF EPIDEMIOLOGY 2024

Summary & Conclusions

- Bayesian validity estimates were more conservative: Bayesian Discrimination Parameters were smaller than frequentist values.

- Agreement between Bayesian and frequentist results was higher for sleep-specific than generic instruments in both the CTT and IRT approaches.

- Selecting between Bayesian and frequentist CTT/IRT techniques may be crucial when validating generic instruments, especially in condition-specific populations

Flinders University

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024

Thank You!

Billingsley Kaambwa, PhD
Associate Professor in Health Economics

Billingsley.Kaambwa@flinders.edu.au

Flinders University

WCE
WORLD CONGRESS OF EPIDEMIOLOGY 2024