www.ispor.org

ISPOR
Improving healthcare decisions

Health State Utility (HSU) Good Practices Task Force

Recommendations for Identification and Use of HSU Data in Cost-effectiveness in Decision Modelling

ISPOR Europe 2018
12 November Barcelona, Spain

Task Force Members:

John Brazier, MSc, PhD, FMedSci, Dean, School of Health and Related Research (ScHARR) University of Sheffield, Sheffield, England, UK

- **Roberta Ara, MSc**, Senior Research Fellow, University of Sheffield, Sheffield, England, UK
- **Ismail Azzabi, MSc, MAS**, Senior Manager Global Health Economics, Takeda Pharmaceutical International AG, Zurich, Switzerland
- **Hélène Chevrou-Séverac, PhD**, Director HEOR Medical Affairs, Celgene International, Sàrl, Switzerland
- **Bruce Crawford, MA, MPH**, Vice President, Real World Evidence, APAC, Syneos Health, Tokyo, Japan
- **Luciane Cruz, ScD, PhD**, Health Technology Assessment Institute, Federal University of Rio Grande du Sol, Porto Alegre, Brazil
- **John Kamon, MSc, PhD**, Professor of Health Economics, The University of Adelaide, Adelaide, Australia

Task Force Members continued…

- **Luciane Cruz, ScD, PhD**, Health Technology Assessment Institute, Federal University of Rio Grande du Sol, Porto Alegre, Brazil
- **John Karnon, MSc, PhD**, Professor of Health Economics, The University of Adelaide, Adelaide, Australia
- **Andrew Lloyd, DPhil**, Director, Acaster Lloyd Consulting Ltd, Oxford, England, UK
- **Suzy Paisley, MA, PhD**, Senior Research Fellow, Director of Information Resources, University of Sheffield, Sheffield, England, UK
- **A. Simon Pickard, PhD**, Professor & Director, Graduate Studies, Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago, Chicago, IL, USA
- **Jan Busschbach, PhD**, Professor & Chair of the section Medical Psychology & Psychotherapy, Erasmus University Medical Center; Senior Investigator, Viersprong Institute for Studies of Personality Disorders (VISPD), Netherlands

ISPOR www.ispor.org

Current Reporting Standards: recent review

Example: review of reporting standards associated with HSUs used in recently published (2015-2018) cost-effectiveness analyses using case-study in cardiovascular disease

See summary in:

Ara R, **Brazier J**, Lloyd A, Chevrou-Severac H. How health state utilities used in cost-effectiveness models are currently identified, reviewed and reported. *Value and Outcomes Spotlight*. September/October 2018, 4(5): 31-33.

Current approaches to reporting HSUs used in models

- No literature review (too many hits, too much effort, no-one else does)
- Use outdated evidence (or just cite a value used in previous NICE submission)
- Don't cite original source (requires a detective to identify actual source: 3/4 iterations)
- Don't report actual values used (just cite any study & let reader find them if they really want to)
- Never, ever provide details of source study (e.g. patient characteristics, measure used, sample size etc.)
- 'Tweak' the values (several times)

ISPOR www.ispor.org

Current approaches to reporting HSUs used in models

- Assume comparator has no benefit (equivalent to baseline of full health)
- Ignore adverse events (or make something up, or use evidence from a different treatment)
- Add clinical effects together when using two interventions, or ignore one (equivalent to using additive or minimum method to combine HSUs for concurrent events)
- Combine effects measured using different units (e.g. mmol/L & mg/dL) (equivalent to using SF-6D, HUI & EQ-5D within same model)
- Don't bother using the values in the report in the cost-effectiveness model

www.ispor.org

ISPOR
Improving healthcare decisions

Overview of session

Presentations:

1. *Searching, identifying & reviewing HSU for use in CE models*
2. *Synthesis of health state utilities*
3. *Using health state utilities in cost-effectiveness models*
4. *Minimum reporting standards - The SpRUCE checklist*

With some future considerations

SECTION

1

www.ispor.org

Searching, identifying & reviewing HSU for use in CE models

Suzy Paisley, PhD.
School of Health and Related
University of Sheffield

A working example – C/E analysis in cardiovascular disease

- Examined HSUs used in recently published (2015-2018) CE models
- Also examined the reporting standards from studies included in the review
- How thorough, transparent and reliable was reporting of values from literature?

9

ISPOR www.ispor.org

Current Reporting Standards: Cardiovascular disease example

- 1/24 reported undertook literature review (limited details), 1 other referred to this review for their evidence
- 6/24 referenced original sources for all HSUs
  - 18/24 referenced previous CE studies (as opposed to original sources)
  - 7/24 required at least 3 iterations to track down original sources
  - 13/24 used at least some HSU collected in 1990s (studies all published after 2014)
- 4/24 correctly reported all HSUs when checked against original sources
- 20/24 at least some HSUs could not be matched in references or original sources, or original sources could not be identified due to incorrect referencing
- 0/24 provided all basic details of sources (e.g. study type, sample size, age, details of health condition, time of data collection etc.)

Ara R et al., Are current reporting standards used to describe health state utilities used in cost-effectiveness models satisfactory? Does this matter? Unpublished manuscript

Searching

- Wide range of sources may be appropriate
  - Clinical trials
  - Observational studies
  - Registries
  - Surveys
  - Previous economic evaluations
- Methods must accommodate
  - Needs of the model
  - Needs of the decision maker

11

Model development

- Model development often proceeds in an iterative fashion
- Searching & identifying utilities is similar, the scope may change as model develops
  - Modeller and information specialist must work together closely
- Initial focused searches may be broadened in later iterations

12

Scope of searches

- Evidence relevant to a model
- All health states and aspects of treatment and management that might impact on health-related quality of life (HRQL)
  - E.g. prophylactic treatments must include possible events in the future
- Different to standard SR approach
  - Standard SR approach Population – Intervention – Comparison – Outcomes (PICO).
  - Scope of search typically Population + Intervention
  - Scope of search for models defined by all aspects of condition and treatment pathway that might impact of HRQL

13

Search scope should account for

- Health state descriptions within the model
- Treatment effects of interventions and comparators of interest (treatment benefits and adverse effects)
- Treatment effects and management at all stages of the clinical pathway included in the model
- Carer utilities
- Comorbidities
- Concurrent clinical events/sequelae
- General population norms
- Moderators that might affect quality of life e.g., method of administration, treatment setting

14

Considerations when searching

- Exhaustive searching vs searching the full scope of HRQL impact
- Inconsistent search vocabulary and indexing
- Which states drive model results?
  - Focus resources here
  - *Value of information* could inform when to stop searching
- Transparency
  - Not cherry picking

15

Selection and Reviewing

- Standard SR processes
- Selection (relevance)
  - Study population matches model?
  - Appropriateness of the measure (e.g. EQ-5D etc)
- Quality assessment
  - Sample size/ response rates/ loss to follow up/ missing data
  - Proximity of data collection to event?
- Data extraction

16

Fitting SR process to modelling process

- Build in iterations
- Identify key characteristics of options available
- Ideal vs best available evidence
- May require a trade off between different issues
- Evidence to inform scenarios and sensitivity analysis
- Selection rationale must be documented

17

Future developments

- How much evidence are we missing?
- Artificial intelligence
  - Inform scope of HRQL impact
  - Ranking evidence to support selection
  - Automated extraction to identify key study characteristics

18

Selection

- Standardisation
  - Same measure used for all states in a model
  - May be better to prefer a source study which describes more of the required health states
  - But maintain minimum standards
- May require a trade off between different issues
- Selection rationale must be documented

SECTION

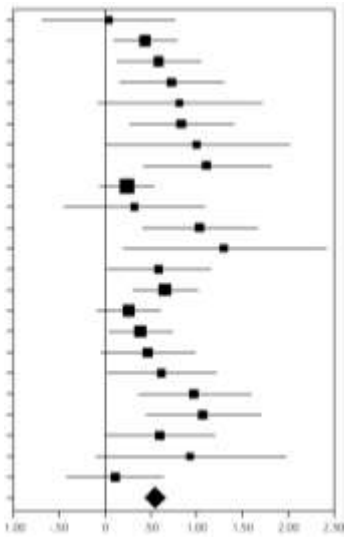
2

Synthesis of health state utilities

Bruce Crawford, MA, MPH
Syneos Health

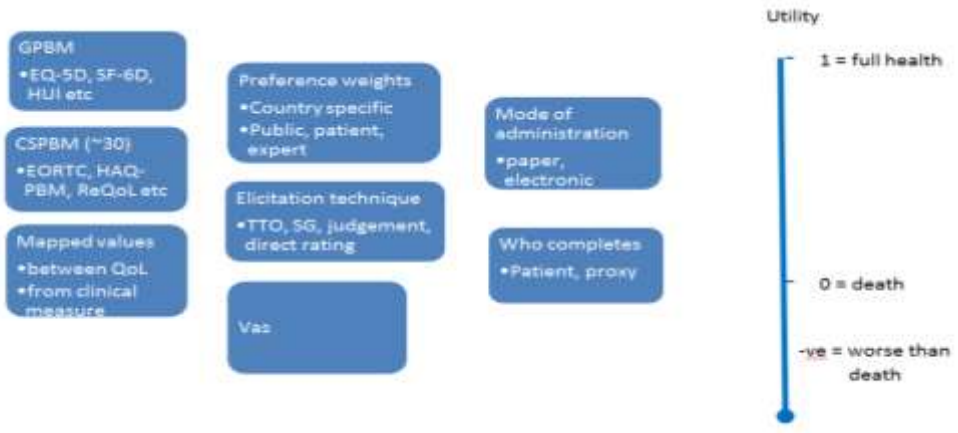
Why undertake synthesis?

- Often multiple published values for a given health state
- Want to use all available evidence
- Synthesis should generate a more accurate estimate of the mean value and uncertainty and improve generalizability



21

Measures and methods used to quantify utility



Should we undertake HSU synthesis?

- Difference in HSUs instruments / methods
- General preference based measures / condition-specific preference based measures
- Techniques used to elicit weights differ
- Mode of collection may differ
- Who completes the questionnaires – proxy, patient, general public, clinician
- Preference weights
- Statistical techniques

EQ-5D-5L differences by country value set

- Different value sets on the same data produce significantly different results
- Authors conclude that different country value sets are not interchangeable

Table 3 EQ-5D-5L utility scores derived from Chinese, Japanese, Korean, and UK tariffs

Tariffs	n	Mean	Standard deviation	Median	Minimum	Maximum
Chinese	608	0.828	0.184	0.879	−0.297	1.000
Japanese	608	0.802	0.164	0.823	0.062	1.000
Korean	608	0.831	0.137	0.829	0.010	1.000
UK	608	0.838	0.154	0.866	−0.213	1.000

Note: The difference among the four national tariffs was statistically significant (χ^2=438.952, P<0.001).
Abbreviation: EQ-5D-5L, five-level EuroQol-5 dimensions.

 Liu et al. Patient Preference and Adherence 2017; 11: 1049-1056

Different measures can produce vastly different results

Table 3. Comparison of Health State Utility Scores by Medical and Socioeconomic Factors

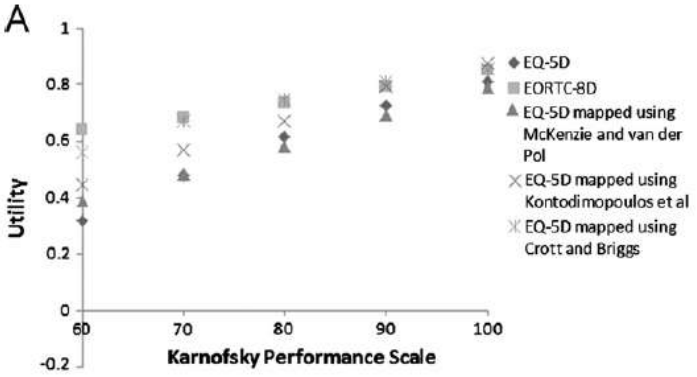
Variable (No. of Cases)	Health State Utility Score by Measurement Method				
	SG	TTO	VAS	EQ-5D	HUI3
Surgery type					
Primary (n = 54)	0.93 (0.17)	0.95 (0.13)	0.76 (0.20)[a]	0.83 (0.19)[a]	0.78 (0.22)[a]
Salvage (n = 5)	0.98 (0.04)	0.98 (0.04)	0.48 (0.13)[a]	0.62 (0.17)[a]	0.37 (0.29)[a]
Chemotherapy					
Yes (n = 13)	0.92 (0.10)	0.99 (0.03)	0.66 (0.19)[b]	0.76 (0.17)[b,c]	0.57 (0.38)
No (n = 87)	0.91 (0.18)	0.94 (0.14)	0.77 (0.18)[b]	0.83 (0.18)[b,c]	0.78 (0.21)
Oral cavity tumor stage					
T1 or T2 (n = 47)	0.95 (0.13)[a]	0.96 (0.09)	0.77 (0.19)	0.83 (0.18)	0.8 (0.21)
T3 or T4 (n = 20)	0.87 (0.22)[a]	0.88 (0.21)	0.7 (0.20)	0.83 (0.09)	0.74 (0.21)
Tracheotomy and/or feeding tube					
Yes (n = 6)	0.99 (0.02)	0.91 (0.12)	0.69 (0.23)	0.78 (0.14)	0.73 (0.25)
No (n = 94)	0.91 (0.17)	0.95 (0.14)	0.76 (0.19)	0.82 (0.18)	0.75 (0.29)

0.07 vs 0.21 diff in utility

25 *Noel et al. JAMA Otolaryngol Head Neck Surg 2015; 141(8): 696-703*

Different mapping methods produce varied utilities for the same performance level

- Models may utilize functional levels such as HAQ DI levels of KPS levels, but the utility value assigned to that level can vary widely depending on the method used for utility



26 *Rowen et al. ViH 2012; 15: 1059-1068*

When undertake synthesis?

- Are there enough HSU estimates?
- Are studies sufficiently homogeneous?
- Are they using the same HSUV measurement system?

Methods of synthesis of HSUVs

- Apply strict eligibility criteria to reduce heterogeneity
  - Same measurement instrument used
  - Same patient group (e.g., mild, moderate, severe depression; same KPS or HAQ-DI range)
- Strict eligibility is useful but considerable heterogeneity often remains
- Often there are not enough studies to restrict to a single measurement instrument

- Use meta-regression
  - Limited to parameters reported in each manuscript
  - It has been suggested that 10 studies per covariate should be used, but this is often not feasible

Conclusions

- Currently, formal synthesis is very limited for HSUVs due to heterogeneity of valuation methods, country differences, population differences, etc.
- Research still needed for methods of meta-regression in HSUV synthesis and when appropriate
- When conducting synthesis:
  - Limit to a single measure if possible; converting / mapping to a single measure may be useful but be aware of the additional variance introduced
  - Carefully evaluate heterogeneity and model the impact if using meta-regression

SECTION

3

Using health state utilities in cost-effectiveness models

Helene Chevrou-Severac, PhD,
HEOR Director, Medical Affairs, Celgene International, Boudry, Switzerland

Identifying the most appropriate health state utility (HSU) is just the beginning .....

Issue	Recommendation
Sensitivity analysis (SA)	Always begin with SA
Discrete health states or discrete event simulation	
Individual mean HSUs or function-based HSUs	
Comorbidities & age	
Concurrent clinical events or conditions	
Treatment related adverse effects	
Acute clinical events	
Sensitivity analyses (again)	Always end with SA

31

Begin: sensitivity analysis & model health states

- Decide the format of the model in terms of numbers and types of health states that differ according to HSU
  - Use clinical expertise & availability of evidence to inform decision
  - If simple model structure implemented that may not represent all important HSUs
  - Examine & discuss expected effects of such omissions => sensitivity analysis
- Look at the effect of each health state in model individually
  - Which HSUs influence the ICER?
  - Important to inform level of literature searches

=> Communicate with search/review team which HSUs really matter

32

Individual HSUs or function based:
Clinical events – acute? chronic? fluctuating? progressive?

Simple 'acute' events:

- Bone fractures (e.g. hip, wrist, vertebrae)
- CVD (e.g. MI, angina, stroke)

Less obvious in chronic progress conditions:

- Arthritic conditions 'flares'
- Asthma exacerbations (hospitalised?)
- Crohn's disease
- Age related macular degeneration

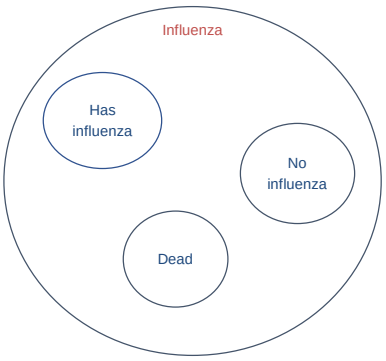
Consider:

- Frequency of events
- Severity of effect on QoL
- Duration of time effect lasts
- 'Rebound' HSU after recovery
- Timing of data collection matches acute period of event?
- Explore likely effect if collection time does not match event

33

Individual HSUs, or function based?

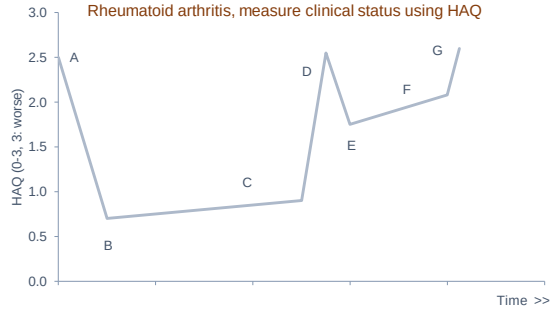
Discrete HSUs:



34

Discrete mean HSUs, or function based HSUs

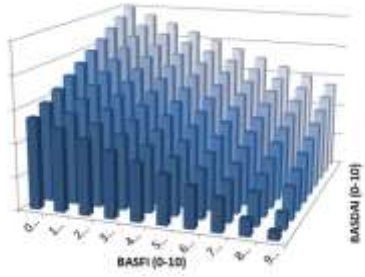
Rheumatoid arthritis, measure clinical status using HAQ



A, Baseline HAQ=2.5, QoL=0.3, Start Tx1, HAQ & QoL improve
B, HAQ & QoL improve
C, HAQ progression on Tx1
D, Stop Tx1, HAQ rebounds
E, Start Tx2, HAQ & QoL improve
F, HAQ progression on Tx2
G, Stop Tx2, HAQ rebounds

RA: HAQ strong relationship with utility.
Could use 3-30 discrete HS, depending on sample size for subgroups.

Ankylosing Spondylitis, measure clinical status using both BASFI & BASDAI

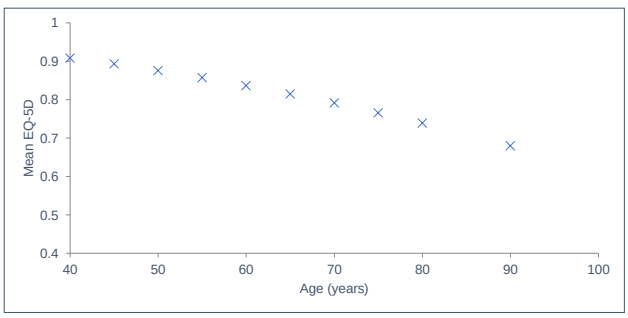


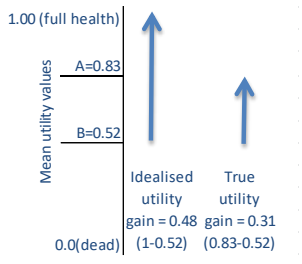
AS: Both BASDAI & BASFI have strong relationship with utility.
3-D matrix, substantial sample size reqd. Use function to predict utility conditional on BASDAI & BASFI (plus age, gender ……. .

35

Baseline HSUs vary over model lifetime horizon

Mean EQ-5D scores by age from HSE



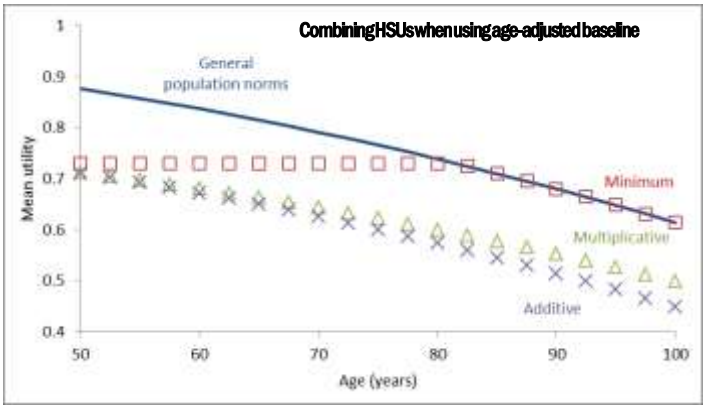


A: mean utility, age 69, without condition X
B: mean utility, age 69, with condition X

Mean HSU is never equal to full health irrespective of measure, preference weights, age, condition or sample. Baseline of full health overestimates effects of avoiding events or conditions.

36

Comorbidities, age & role of general population norms



NB. For multiplicative DO NOT multiply HSUs together!
Estimate 'multiplier' = observed mean HSU / baseline mean HSU (at age of sample).

37

Treatment related adverse events (AE)

CONSIDER.

- Are AEs prevalent?
- Do treatment related AEs have substantial effect on QoL?
  - Check potential effect on ICER using SA in model.
- Is effect captured in evidence used for HSUs in model?
  - Those experiencing AE less likely to complete QoL measure.
  - Take care not to double count effect if use trial data
- What is likely duration of effect ?
  - Chronic long term/ discrete short term & treatment withdrawal.

TEST extreme values in model

39

End: sensitivity analysis (again)

- ALL HSU represented by **parametric probability distributions**
  - Sample random values for PsA using Beta: rescale if –ve values possible
  - Consistency: for ordered values, sample using 'difference method' [Ren 2017]
- **Univariate: test which HSUs influence ICER using CIs**
- **Multivariate: test LCI & UCI that move ICER in same direction**

40

Recommendations when using health state utilities in models

Issue	Recommendation
Sensitivity analysis	Use to determine which HSUs influence the ICER & report these analysis if applicable.
Discrete health states	**Expected effects of excluding potentially relevant HSUs** from a cost-effectiveness model structure should be examined & discussed.
Individual mean HSUs or function-based HSUs	Relevance of available data (study population, utility measures, alignment with model's health states) Reliability of analyses (e.g. precision of mean HSUs, validity of estimated utility functions) *see Wailoo et al ISPOR TF*
Comorbidities & age (baseline)	Mean HSU values represent comorbidity utility effects at the mean age of the utility study population. **Age-specific comorbidity effects** should be estimated using age-specific population norms.
Concurrent clinical events or conditions	**The multiplicative method should be used** to handle the utility effects of concurrent clinical events.
Treatment related adverse effects	Assess whether utility effects of AEs are captured by the utility data used into a model. If not, and AE HSUs are required, the range of HSUs required should be informed by expected effect on ICER.
Acute clinical events	In the absence of data collected around the event, plausible HSUs for the direct effects of acute events should be multiplied by the expected duration to assess the sensitivity of ICER to HSUs-Acute
Sensitivity analyses (again)	One way & multi-way SA of HSU, Beta (scaled) distributions. Difference method to account for ordering in PsA.

41

SECTION

4

www.ispor.org

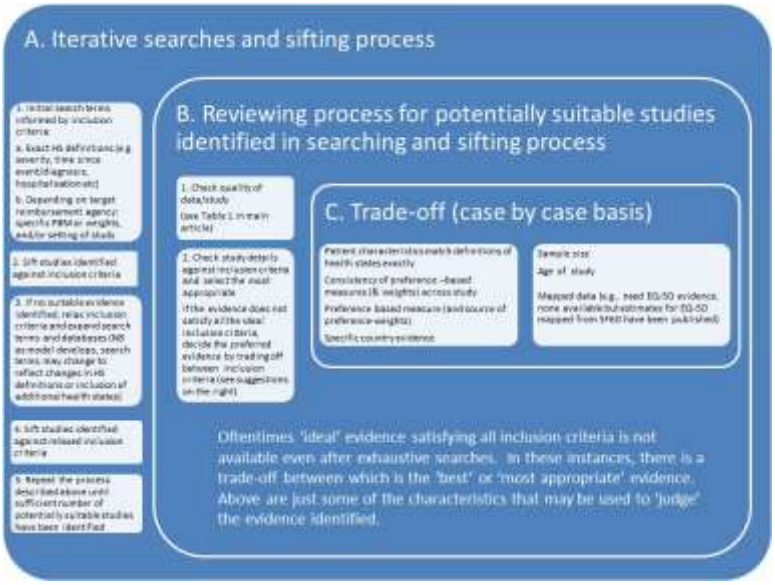
Minimum reporting standards
The SpRUCE checklist

John Brazier, PhD.
School of Health and Related Research
University of Sheffield

Current approaches to reporting HSUs used in models

- No literature review (too many hits, too much effort, no-one else does)
- Use outdated evidence (or just cite a value used in previous NICE submission)
- Don't cite original source (requires a detective to identify actual source: 3/4 iterations)
- Don't report actual values used (just cite any study & let reader find them if they really want to)
- Never, ever provide details of source study (e.g. patient characteristics, measure used, sample size etc.)
- 'Tweak' the values (several times)
- Assume comparator has no benefit (equivalent to baseline of full health)
- Ignore adverse events (or make something up, or use evidence from a different treatment)
- Add clinical effects together when using two interventions, or ignore one (equivalent to using additive or minimum method to combine HSUs for concurrent events)
- Combine effects measured using different units (e.g. mmol/L & mg/dL) (equivalent to using SF-6D, HUI & EQ-5D within same model)
- Don't bother using the values in the report in the cost-effectiveness model

- HSU identification and selection is often not a straightforward process
  - Challenges to being explicit
  - Iterative search and sifting process



44

ISPOR www.ispor.org

SpRUCE Checklist

- Provides minimum reporting standards for the **S**ystematic **R**eview of **U**tilities for **C**ost-**E**ffectiveness (SpRUCE checklist).
- Criteria intended to help model reviewers identify if HSU selection for the model was transparent and appropriate.
- A greater level of detail is likely needed to proceed to peer-reviewed publication of a systematic review
- Five sections:
  1. search strategy
  2. review process
  3. data extracted from each study
  4. basis for obtaining the final HSU
  5. use in cost-effectiveness models

45

SpRUCE Checklist: Search Strategy

Search Strategy	
Search terms and scope	The final search strategy should be adequately defined and appropriate databases included in the search.
Study selection criteria	Explicit criteria for study identification/inclusion should be described and applied, such as patient group of interest, relevant age range and stage of disease/severity etc.

46

ISPOR www.ispor.org

SpRUCE Checklist – Review Process

Review Process	
Quality check	Quality criteria for reviewing studies explicitly stated and applied.
Assessment of relevance	Relevance of HSUs to model and target reimbursement agency described.

47

SpRUCE Checklist

Data Extracted (reporting of variables)	
Population/patient characteristics	Relevant patient characteristics (e.g. age, sex, comorbidities, diagnosis, severity of condition).
Measure used to describe the HSUs	The measure used to obtain the HSUs.
Valuation technique for preference weights	The technique used to value the heath state (e.g. TTO, SG), and the country (provide reference).
Descriptive statistics of HSUs	The mean and variance around any HSU used in the model.
Response rates to the measure used*	Report if response rates are likely to be a threat to validity.
Loss to follow-up/ missing data*	Loss to follow-up (e.g. 1 year after fracture) and missing data should be reported, especially if they may threaten the representativeness of the HSUs.
Original reference	The original source for the HSUs should be referenced (not a previous economic study that has used the evidence).

48

ISPOR www.ispor.org

SpRUCE Checklist

Selection/estimation of final health state value	
Basis for selecting HSUs	The rationale for selecting the HSUs used in the model should be justified.
Method used to combine estimates	Where HSUs are combined, the analytic methods should be described e.g. meta-analysis.

49

SpRUCE Checklist

Methods used when applying the health state utilities in model	
Actual HSUs used	Report all actual HSUs used in the model together with associated measure.
Adjustments or assumptions	Clearly describe any adjustments or assumptions relating to the use of HSUs in the model, reporting both the raw and final values used with worked examples if required to clarify the method used to adjust the data.

50

ISPOR www.ispor.org

Highlights

- Current practice is poor
- This report provides recommendations on the iterative nature of searching for HSUs, reviewing and synthesis of the evidence identified, and the application of the HSUs in cost-effectiveness models.
- It provides the minimum acceptable reporting standards for HSUs used in cost-effectiveness models (SpRUCE checklist)
- It helps those undertaking a systematic review of HSUs for a cost-effectiveness model.
- The SpRUCE checklist should be used by reviewers of manuscripts and reports of modelling work to determine the suitability and validity of the HSUs, to ensure the quality of results is sufficient to be used to inform healthcare policy making

This is an evolving area where there is considerable scope for development and innovation

51

COMING IN JANUARY / FEBRUARY 2019!

Identification, Review and Use of Health State Utilities in Cost-Effectiveness Models: An ISPOR Good Practices for Outcomes Research Task Force Report

52

FOR MORE ISPOR HEALTH STATE UTILITY GOOD PRACTICES TASK FORCE

https://www.ispor.org/heor-resources/good-practices-for-outcomes-research

Wailoo AJ, Hernandez-Alava M, Manca A, et al. **Mapping to Estimate Health-State Utility from Non–Preference-Based Outcome Measures**: An ISPOR Good Practices for Outcomes Research Task Force Report. Value Health 2017; 20(1):18-27.

Wolowacz SE, Briggs A, Belozeroff V, et al. **Estimating Health-State Utility for Economic Models in Clinical Studies**: An ISPOR Good Research Practices Task Force Report. Value Health. 2016; 19(6):704-719.

53

Slides are available on the ISPOR Europe 2018 webpage

ISPOR Europe 2018

10-14 November 2018 | Barcelona, Spain

New Perspectives for Improving 21st Century Health Systems

54

Discussion

55

References

Ara R, Brazier J, Bussbach J, Lloyd A, Chevrou-Severac H. How health state utilities used in cost-effectiveness models are currently identified, reviewed and reported. *Value and Outcomes Spotlight* September/October 2018, 4(5): 31-33.

Brazier J, Ara R, Azzabi I, Bussbach J, Chevrou-Severac H, Crawford B, Cruz L, Kamon J, Lloyd A, Paisley S, Pickard S. Identification, review and use of utility data in Cost-Effectiveness Models: Good Practices for Outcomes Research (in review at Value in Health)

Fryback DG, Lawrence WG. Dollars may not buy as many QALYs as we think: A problem with defining quality of life adjustments. MDM 1997;17;276-284

Ara R, Brazier J. Using health state utility values from the general population to approximate baselines in decision analytic models when condition specific data are not available. ViH 2011;14(4):539-545

Ara R, Brazier J. Populating an economic model with health state utility values: moving toward better practice. ViH 2011;13(5):509-518

Ara R, Wailoo AJ NICE DSU Technical Support Document 12: The use of health state utility values in decision models. 2011.

Ara R, Wailoo AJ. Estimating health state utility values for comorbid health conditions: a synopsis of the current evidence base. MDM, 2012.

Ren S, Minton J, Whyte S, Latimer N & Stevenson M (2017) A New Approach for Sampling Ordered Parameters in Probabilistic Sensitivity Analysis. PharmacoEconomics. View this article in WRRO