UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

A Comparison of the Validity and Responsiveness of the EQ-5D-5L and SF-6D for Measuring Health Spillovers: A Study of the Family Impact of Meningitis

Bhadhuri, Arjun; Jowett, Sue; Jolly, Kate; Al-Janabi, Hareth

DOI:
10.1177/0272989X17706355

License:
None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):
Bhadhuri, A, Jowett, S, Jolly, K & Al-Janabi, H 2017, 'A Comparison of the Validity and Responsiveness of the EQ-5D-5L and SF-6D for Measuring Health Spillovers: A Study of the Family Impact of Meningitis', *Medical Decision Making*. https://doi.org/10.1177/0272989X17706355

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 19. Apr. 2024

A comparison of the validity and responsiveness of the EQ-5D-5L and SF-6D for measuring health spillovers: a study of the family impact of meningitis

Running header: Including health spillovers in economic evaluation

Arjun Bhadhuri MSc [a], Sue Jowett PhD [a], Kate Jolly PhD [b], Hareth Al-Janabi PhD [a]*

[a] *Health Economics Unit, Institute of Applied Health Research, University of Birmingham, UK*

[b] *Institute of Applied Health Research, University of Birmingham, UK*

*Corresponding author: Hareth Al-Janabi, Public Health Building, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; Tel: +441214158483; Email: h.aljanabi@bham.ac.uk

Financial support for this study was provided entirely by a University of Birmingham MDS college doctoral studentship. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Word count: 4456 words

A COMPARISON OF THE VALIDITY AND RESPONSIVENESS OF THE EQ-5D-5L AND SF-6D FOR MEASURING HEALTH SPILLOVERS: A STUDY OF THE FAMILY IMPACT OF MENINGITIS

Arjun Bhadhuri [1], Sue Jowett [1], Kate Jolly [2], Hareth Al-Janabi [1]

[1] Health Economics Unit, Institute of Applied Health Research, University of Birmingham, UK

[2] Department of Public Health, Institute of Applied Health Research, University of Birmingham, UK

Background: The 'health spillover' of patient illness on family members is important to capture in economic evaluation. This study compares the construct validity and responsiveness of two widely used health-related quality of life instruments, the EQ-5D-5L and SF-6D in capturing health spillover effects for family members with and without an informal care role (carers and non-carers).

Methods: Construct validity and responsiveness were assessed using data from a 2012 UK survey of the family impact of meningitis-related sequelae. Construct validity was assessed by testing associations between family members' health status and variables anticipated to be associated with spillover effects (patient health status and informal care). Responsiveness was assessed by testing associations between the longitudinal change in family members' health status and longitudinal change in patient health and caring hours.

Results: Among non-carers, both the EQ-5D-5L and the SF-6D exhibited construct validity with 10 out of the 11 associations that were hypothesised being statistically significant on both measures. There was less clear evidence of responsiveness of the measures for non-carers. Among carers, the EQ-5D-5L exhibited greater construct validity as well as responsiveness, with respect to spillovers from patient health. This was evidenced by the EQ-5D-5L detecting 9 significant associations significant compared with 4 on the SF-6D. However, the SF-6D exhibited greater construct validity,

with respect to spillovers generated from informal care provision (5 associations significant compared with 2 on the EQ-5D-5L).

Conclusion: Both the EQ-5D-5L and the SF-6D exhibited a degree of validity that could justify their use as measures of health-related quality of life spillovers on family members in economic evaluation.

Keywords: economic evaluation; informal care; spillovers; validity; responsiveness; EQ-5D; SF-6D

INTRODUCTION

Economic evaluations in health care aim to compare the costs and benefits of health interventions in order to determine whether they offer value-for-money (1). Cost-utility analysis is a common type of economic evaluation in health care that usually limits the assessments of the benefits of the intervention to health of the patients. However a cost-utility analysis should theoretically also include the health effects of an intervention to carers and family members of patients (2). Although this is rarely done at present, there is ongoing discussion in the literature advocating the regular inclusion of the health effects of interventions on family members in cost-utility (and cost-effectiveness) analysis (3, 4). This is because the health of family members is also affected by illnesses and interventions since family members often share the negative emotional distress of an illness, and there is frequently a need for family members to provide emotionally and physically draining informal care for patients with chronic illness and disability (5). These mechanisms result in 'spillovers' on family members' physical and mental health (5, 6). By routinely including health spillovers, economic evaluations in health care can better guide health technology assessment decisions towards judgements that maximise health across patients and their family networks rather than for just the patients themselves (4). Notably, the Second Panel on Cost-Effectiveness in Health and Medicine recommended that the 'QALYs of patients, carers and 'other affected parties' (such as non-caregiving family members) should be included in economic evaluation (7).

The EQ-5D is the most commonly used instrument to measure the health status of patients in cost-utility analysis in order to calculate quality-adjusted life years (QALYs), and is the recommended instrument for National Institute of Health and Care Excellence (NICE) technology appraisals in the UK (8, 9). However the EQ-5D may not be an appropriate instrument for measuring health spillovers of interventions on family members. Previous studies suggest that it is predominantly the mental health of carers and family members that suffers when a loved one is ill (10, 11). This suggests only

one item of the EQ-5D ('anxiety and depression') may be suitable to capture changes in family members' health status arising from the illness of a patient (11).

An alternative health status instrument is the SF-12 (12); this offers a slightly more detailed measure of the individual's health status with many items related to an individual's mental and psychological health, and also their lifestyle. Data collected from the SF-12 questionnaire can be converted into a SF-6D utility score to calculate Quality Adjusted Life Years (QALYs) (12).

Assessing the validity of health status measures

Instrument validation explores whether the variability in the values elicited from an instrument is compatible with existing knowledge about how the instrument should or is likely to vary according to some other observable variables (13). Health status measures that are compatible with existing hypotheses about how health is expected to vary according to predicting factors, may be seen to exhibit a high level of validity, and vice versa. There are different types of validity used in the psychometric assessments of health status measures; these include construct validity and responsiveness. The construct validity of an instrument is assessed by testing mini-theories that are developed to explain the relationship between the instrument and associated factors (13). Responsiveness is the ability of an instrument to respond to a meaningful or clinically important external change over time (13).

The EQ-5D-5L and SF-6D instruments cannot be used interchangeably to measure an individual's health status (14). There is substantial validity literature comparing the EQ-5D (3 level version) with the SF-6D among patient populations, with variable findings reported regarding the validity of the instruments depending on the patient population being assessed (15-18). Although the EQ-5D and the SF-6D have been used in a range of studies to measure family health spillovers in different clinical contexts (11, 19-21), we are aware of only two studies assessing the validity of a generic

health instrument for measuring health spillovers (21, 22).   These studies assessed the convergent validity and known-groups validity of the EQ-5D-3L and SF-6D for measuring health spillovers in carers of sick children (21, 22). Convergent validity assesses how closely one instrument is related to other instruments that measure the same construct) (13). However these studies did not investigate instrument responsiveness, only briefly investigated construct validity, and did not explore the use of the measures amongst non-caring family members. This new study addresses these gaps.

Research objective

The purpose of this research is to compare the construct validity and responsiveness of the EQ-5D-5L and the SF-6D for capturing the health effects of patient illness on carers and 'non-caring' family members. A case study of families, post-meningitis, where a variety of physical and mental health problems create a range of caring contexts, is used as the basis of the study.

METHODS

The survey dataset of family members of meningitis survivors covered different aspects of the family member experience of living with and caring for the patient, including family member health status measured by the two instruments. This enabled an analysis of the EQ-5D-5L and SF-6D in terms of their ability to detect quality of life effects generated from caring about, and providing informal care for, an individual close to them with long-term impairments. A comparison is not made in this study about the relative scale of health effects in these two groups.

Long term family impact of meningitis case study

Meningitis is an illness that usually infects young people, and especially infants. A longitudinal study was carried out with postal surveys administered 12 months apart (in 2012 and 2013) to the family members of meningitis survivors for self-completion (19). Meningitis is a condition that can result in a number of disabling and often life-long sequelae related to behaviour, mental and physical health impairment of the patient, resulting in a range of caring situations.

In the family impact of meningitis study, 3417 potentially eligible family members of meningitis survivors were contacted to participate using a database held by the Meningitis Research Foundation (a large UK charity). This sampling frame does disproportionately focus on families at the more severe end of the illness spectrum. However, this meant that there were a higher number of cases of informal care which increased the power to examine instrument validity in caregiving family members. A specific power calculation was not used for the validity study as the sample size was determined by the requirements of the original family impact study (19). However the resulting sample size is consistent with other studies measuring validity (18, 23). Each potential eligible family member was sent two questionnaires; they were asked to complete the first and to pass on the second questionnaire to an additional person close to the survivor. The survey was completed by the family members themselves and measured family members' EQ-5D-5L and SF-6D scores, carer wellbeing (using the Carer Experience Scale (24)), and elicited information about informal care provided by the family members (such as hours of care provided and whether personal care was provided). Also, information on the impact of meningitis on aspects of family members' lives was assessed via a bespoke question enquiring whether "meningitis had no effect, a negative effect or a positive effect" on the family member's life. Domains of life (finances, social life, family life, work, exercise and personal health) were selected based on a focus group discussion with members of the Meningitis Research Foundation (19) . Additionally, family members were also asked to complete a

section on the patient’s health. This involved family members providing a proxy report of the patient’s EQ-5D-5L.

This dataset generated from the survey allowed us to not only look at family members who provide care, but also family members who were classified as non-carers. Non-carers can be defined as family members who do not provide informal care for the patient, but may nevertheless experience health spillover resulting from anxiety and distress from witnessing the illness of a loved one. This study does not aim to make a comparison of the impact of illness between carers and non-carers, but rather aims to assess validity within the two groups. In the analysis, ‘carers’ were distinguished from ‘non-carers’ if they reported spending any amount of time ‘providing care as the result of meningitis’ in the baseline survey. It is possible that some family members we classified as non-carers for this study were providing care for other individuals or for meningitis survivors with non-meningitis related conditions.

General approach (caring ‘about’ and ‘for’ the patient)

The sample used for the analysis was constrained in two ways. First, we focused on a single close family member for each patient, selected on the basis of the highest degree of social contact (19). This was done in order to eliminate correlation effects between multiple family members of the same patient. Second, families where the patient had made a complete recovery from meningitis were excluded. This was done to ensure that we only included family members where there was some degree of potential spillover from the meningitis sequelae or caring role. The sequelae most commonly reported were behavioural or emotional problems, mild or moderate learning difficulties, and scarring or tissue damage (19).

Following Bobinac et al (2011), the health spillover of illness on family members can be described as the product of two different effects (5). The first effect is the psychological distress from ‘caring

about' a loved one with an illness. The second effect is the physical and mental strain of providing informal care for a patient (or 'caring for' a patient). The tests of construct validity and responsiveness carried out in this study therefore reflect these two different effects; firstly testing associations between participants' health status responses and a range of characteristics that reflect the severity of the patient's condition (and therefore the likely strength of the 'caring about' spillover), and secondly testing associations between participants' health status responses and characteristics reflecting the burden of caring for the patient.

It was anticipated that the SF-6D would be more valid and responsive than the EQ-5D-5L in detecting health spillovers in family members, by detecting larger effect sizes and stronger correlation coefficients for the hypothesised associations. This is because the SF-6D contains more items than the EQ-5D-5L related to mental health and social functioning, and these items are expected to be particularly sensitive in detecting health spillovers in family members generated from the psychological and informal care burden of meningitis.

Construct validity assessment

Assessment of construct validity firstly compared the EQ-5D-5L and the SF-6D for measuring health spillovers generated from 'caring about' the patient, and secondly for spillovers from 'caring for' the patient.

For non-carers and carers, hypotheses were tested predicting that better family member health status would be associated with better patient health and less negative experiences of meningitis illness, as observed in previous empirical studies of 'caring about' effects (Table 1) (5, 10, 25-27).

<u>Table 1 here</u>

For carers only, hypotheses were tested predicting that the family member EQ-5D-5L and SF-6D were negatively associated with larger volumes of care provision, greater work and finance related pressures from caregiving and worse carer experiences, as observed in previous studies of 'caring for' effects (Table 1). The EQ-5D-5L and SF-6D were compared in terms of the effect sizes and the statistical significance of the associations tested for (further details in 'Statistical analysis' section). (27-41).

Responsiveness analysis

We hypothesised that over the course of 12 months, the change in family members' EQ-5D-5L and SF-6D scores would be positively associated with changes in patient EQ-5D-5L scores and negatively associated with changes in the number of hours family members spent providing informal care.

The responsiveness analysis of the EQ-5D-5L and SF-6D used the baseline data (from 2012) and follow-up data (from 2013) for family members of patients. The analysis was again split to cover the carers and the non-carers separately (as in the construct validity analysis), in order to investigate the performance of the measures in carers and non-carers. Anchor based methods were implemented to assess whether the EQ-5D-5L and SF-6D responded in expected directions to changes in the following two factors over the 12 month period(42):

- patient EQ-5D-5L score (carers and non-carers)
- number of hours per week spent on caring activities (assistance with daily living/organisational support/extra household activity) (carers only)

The objective of an anchor-based analysis is to examine whether scores on the measure of interest change in the expected direction when compared with changes in the scores of a related construct or measure (the 'anchor' measure) (42, 43). Patient health status and informal care hours were

selected as anchors based on their conceptual relationship with family members' quality of life. The anchors were sub-divided into 3 levels to indicate whether the 'anchor' had increased, decreased, or not changed in an important way over time(44). It was predicted that an important improvement in patient health or reduction in caring hours would be associated with a statistically significant increase in family members' EQ-5D-5L and SF-6D score from baseline to follow-up assessment, and vice versa. An 'important' increase/decrease in the patient EQ-5D-5L score was determined by the measurement of a minimal clinically important difference (MCID) in scores between the two periods of at least 0.074, derived from a literature estimate of this difference (45). This estimate was obtained from an EQ-5D-3L study and used as a proxy for the EQ-5D-5L in this study as consistent with other studies (46, 47); as there are only limited empirical estimates of the EQ-5D-5L MCID available which vary widely from 0.051 to 0.10 (48, 49). In the absence of an agreed 'important' change in caring hours, we assumed that a change of 5 or more hours / week was important.

Statistical analysis

Spearman's Rank Correlation coefficients were computed to assess the strength, statistical significance and directions of associations between the family health status measures, and ordinal independent variables of more than two groups including patient health status variables (EQ-5D-5L, EQ-VAS, mobility, self-care, usual activity, anxiety, pain), hours of care provided and Carer Experience Scale. The Mann-Whitney test was used to establish any statistically significant differences in health status between two groups within the sample, and the direction of these differences. The Spearman's Rank Correlation test and Mann-Whitney test are non-parametric tests that only take into account the existence of a difference between two data points (i.e. how they rank) rather than the magnitude of the difference. This is an appropriate method for handling highly skewed variables- like the EQ-5D-5L and SF-6D composite scores (50). However in the tests of

responsiveness, t-tests were used (instead of non-parametric tests) because the changes in EQ-5D-5L and SF-6D scores between 2012 and 2013 were approximately normally distributed.

Assessments were also made about the magnitude of associations by calculating effect sizes (Cohen's D) where independent variables consisted of two groups only, and correlation coefficients (Spearman's) where independent variables were ordinal and consisted of more than two groups. Spearman's rank correlation coefficients of between 0.3 and 0.5 are considered weak, between 0.5 and 0.7 moderate and > 0.7 strong. For Cohen's D effect sizes between 0.2 and 0.5 are considered small, between 0.5 and 0.8 moderate and > 0.8 large (51). The same interpretations apply for negative associations and effect sizes.
The instruments were then compared to find out whether the EQ-5D-5L was able to capture larger effect sizes and stronger associations than the SF-6D, or vice versa. Only individuals that had a complete set of item responses for a validity test were included in the analysis.

Participants were excluded from the study if the person they were close to had subsequently died, as the health losses experienced by bereaved family members are different to those experienced by the family members of living patients (52), and not the focus of this study. Participants were *not* excluded on the basis of whether they shared a household, or how they were related to the person with meningitis.

RESULTS

3417 members of the Meningitis Research Foundation (MRF) were invited to participate in the original family impact of meningitis study (19). This resulted in 1587 eligible family members of 1218 survivors (36% of family units) returning the baseline survey in 2012.

For the present study, at baseline, 1546 (97%) family members completed the EQ-5D-5L, and 1485 (94%) family members completed the SF-6D. 1053 family members (66% of the whole sample) reported being exposed to patient sequelae from meningitis at baseline, and 847 of these family members were included in the construct validity analysis (as they were the closest surveyed family member to the patient). 1022 (64%) of family members responded to the follow-up questionnaire in 2013, and 536 of these family members were included in the responsiveness analysis.

Table 2 documents the descriptive statistics in 2012 for the whole family member sample, and the carer and non-carer sub-samples used in the validity analyses. The patients receiving informal care for meningitis had a much worse mean health status (0.50) than the patients who did not receive informal care for meningitis (0.87).

Table 2 here

Construct validity .

Table 3 details the results for the tests of construct validity; split between non-carers and carers; and further split among carers between hypotheses either relating to 'caring about' or 'caring for' the patient, and hypotheses solely related to 'caring for' the patient. .

Table 3 here

In the ‘caring about’ tests for the non-carers in Table 3, both the EQ-5D-5L and SF-6D each detected statistically significant associations with ten out of the eleven constructs, with all of these associations falling in the expected directions that were hypothesised prior to testing. Statistically significant associations were reported for patient health constructs, and these associations were below the threshold for a small effect. Moderate-to-large effect sizes were reported for constructs relating to the negative impact of meningitis on areas of family members’ lives.

In the tests for carers either relating to ‘caring about’ or ‘caring for’ the patient, the EQ-5D-5L generally detected larger effect sizes and stronger associations than the SF-6D, and more statistically significant associations (nine out of eleven) than the SF-6D (4/11) .

In the tests of carers solely related to ‘caring for’ the patient the SF-6D detected a statistically significant effect size or association five out of ten times, and the EQ-5D-5L two out of ten times. These effect sizes were either small or below the conventional threshold of a small effect size. For the variable ‘hours of care provided’, statistically significant associations ($p<0.01$) were detected using both the EQ-5D-5L and SF-6D, and both associations were below the conventional threshold for a small effect.

Responsiveness

This section details the results of the tests of responsiveness of the family member EQ-5D-5L and SF-6D to clinically relevant external changes between 2012 and 2013, tested among the non-carers and carers separately.

In table 4, there are no clearly observed ‘gradients’ of effect in the non-carers’ EQ-5D-5L or SF-6D moving between an improvement through to a decline in patient health status. This is the result of

there being few significant changes in the expected direction in non-carers' health status when the patients' health improved/did not change/worsened.

Table 4 here

.

In Table 5, the EQ-5D-5L detects a gradient of effect in the expected direction in carer health status as patient health status moves from an improvement to a decline, whereas the SF-6D does not detect such an effect. This is evidenced by the carer EQ-5D-5L score improving by 0.04 between 2012 and 2013 as the patients' health improves, not changing when patients' health does not change, and declining by 0.06 as patients' health worsens (i.e. a gradient from positive change through to negative change, in line with patient health status change). Neither the carer EQ-5D-5L or the SF-6D detected a gradient of effect as caring hours moves from an increase to a decrease.

Table 5 here

DISCUSSION

This study systematically explored whether two commonly used health status measures are valid and responsive measures of health effects (spillovers) amongst carers and non-carers in patients' family networks. The findings from the results suggest that the EQ-5D-5L and SF-6D both exhibit some degree of validity in measuring health spillovers of meningitis on family members. This is because in terms of construct validity, the scores of both instruments were frequently statistically associated with variables that were hypothesised to generate spillovers on family members' health (particularly in the tests of construct validity among the larger non-carer sub-sample), and all the statistically significant relationships were found to be in the predicted directions that were hypothesised prior to analysis. These findings complement previous studies which found that the EQ-5D-3L and SF-6D demonstrated convergent validity in measuring the health status of carers of ill children(21, 22). The findings also complement existing validity literature which indicates that the EQ-5D-5L and SF-6D adequately cover relevant symptoms of anxiety and depression that affect the family members of sick patients (18, 21).

Comparing the relative validity of the two instruments was made more complex by contrasting findings. In the carer sub-sample, the EQ-5D-5L exhibited greater construct validity by detecting stronger associations than the SF-6D for spillovers resulting from poor patient health (and also detecting an anticipated gradient in the responsiveness analysis as patients' health declined over time). However the SF-6D detected more statistically significant associations than the EQ-5D-5L for spillovers resulting from caring burden. It was expected that an instrument that is more socially oriented such as the SF-6D would be better at picking up associations relating to aspects of the caring situation. What was unexpected was that the EQ-5D-5L would be better than the SF-6D at detecting spillovers relating to patient health among the carers, particularly in terms of construct validity. One factor that may partially explain this result is that the EQ-5D-5L was used to measure

patient health status. As a result, there may some degree of greater alignment in scores obtained from the same instrument administered to both patients and family members, than if different instruments are administered.

Some small gradients were observed in the responsiveness analysis. For instance for carers, a small and statistically significant health status improvement was observed where patients' health was reported to have improved, and a small decline was observed in carers' health where reported patients' health also declined. Apart from this case, neither the EQ-5D-5L nor the SF-6D exhibited clear responsiveness to changes over the course of a year in patient health or the caring situation with lack of gradient of effect. This suggests a need to use a longer time period (>12 months) for future studies in this area. Furthermore, as the spillover effect (on the average family member) is likely to be a small proportion of the direct effect (53), it may be too small to be detected even when the changes in patient health exceeded the threshold for a clinically important difference. This was also evidenced in this study by the small effect sizes that were reported in the construct validity analysis. It is important to note in the responsiveness analysis that there was a general worsening in the health of family members between 2012 and 2013 that had a sizeable downward effect on all of the mean differences in family member health status between follow-up and baseline assessment: effect sizes and clinically important differences need to be interpreted with this in mind.

One disadvantage of the SF-6D instrument from this study was that it appears to be more prone to missing data than the EQ-5D-5L. This may exacerbate the problem of missing data on family health spillovers within the context of health intervention trials, where the focus is more likely to be on achieving high response rates from the patients themselves.

Even though in this study the SF-6D exhibited greater validity in detecting associations solely related to 'caring for' the patient, the EQ-5D-5L may yet be chosen for measuring family member health status if the EQ-5D-5L is a preferred measure for patient health. This is because it may be considered

inappropriate to use different health status measures to elicit patient QALYs and family member QALYs for subsequent aggregation in an economic evaluation (54). For instance, this may be the case for economic evaluations conducted in England and Wales for NICE which recommend using the EQ-5D-5L for measuring the health of patients (9).

The positive associations between patient health status and family member health status in this study may not be completely attributed to spillover from the patient to the family member. However, the previous study of the family impact of meningitis demonstrated that the positive association between patient health status and family member health status remains when controlling for a wide range of potentially confounding factors related to the characteristics of the two individuals and the shared environment (19).

Aside from the choice of instrument for measuring health spillovers, there are other issues that need to be addressed to enable regular inclusion of health spillovers in economic evaluation, beyond the scope of this study. Unresolved issues are whether to include informal care costs alongside family member health status due to the issue of potential double counting (55, 56), whether to include the health decrements of bereaved family members, and how many family members should be included in the analysis (4). The 2016 US Panel of Cost-effectiveness in Health and Medicine recommended that both carer time costs and QALYs should be included in economic evaluations (7), perhaps indicating they perceived only a small risk of double counting from including both (55).

There are a number of strengths of this study. The study used a large sample of family members, and data completion of the surveys was generally high. This study is the first to assess the responsiveness of generic instruments for measuring health spillovers, and the first study to look at instrument validity specifically in non-carers.

Some limitations of the study are also acknowledged. There was a relatively small sub-sample of informal carers (n=199) compared with non-carers (n=648) used in analysis. The analysis only related

to long-term effects on health of meningitis. Although meningitis is a condition that creates a wide range of symptoms among young individuals, and therefore a range of caring situations, the findings of this study may not be generalisable to other health conditions, especially where patients are older and care is mostly provided by spouses. Another limitation is that validity and responsiveness were not assessed in relation to a healthcare intervention. Further research addressing these limitations could be informative.

In conclusion, both the EQ-5D-5L and SF-6D appear to be satisfactory instruments for measuring family members' health status in an economic evaluation. This is because both instruments exhibit construct validity in capturing family member health spillovers. However further research is required to assess the validity and responsiveness of the instruments in capturing health spillovers generated from other illnesses and from health interventions. Such research will be important for determining whether the two instruments can be used interchangeably for family members in response to a clinical intervention.

Table 1: Hypotheses for associations between constructs and family members' health status

Survey variable	Predicted direction of family member EQ-5D-5L or SF-6D change
Variables relating to 'caring about' or 'caring for' a patient	
Patient EQ-5D-5L index score Patient Visual Analogue Scale (VAS) score	Positive. Better patient health expected to be associated with lower psychological and care burden in family members thus better health status(5, 25, 26)
Patient EQ-5D-5L item responses	Negative. Higher item response indicates worse patient health which is expected to be associated with worse family member health status (5, 26)
Family members' self-perceived impact of meningitis on areas of life*	Negative. Negative experiences of illness on non-carers and carers in these areas expected to translate to worse family member health status (5, 10, 26, 27)
'Caring for' variables	
Hours of care provided Shares house with patient Daily care for the patient Constant daytime supervision for patient Main carer for patient Provides majority of care	Negative. Greater volumes of informal care provision expected to result in worse carer health (27-32)
Provides personal care/toileting for patient	Negative. Providing ADLs (assistance with daily living) is associated with high informal care burden and increased chance of carer distress, resulting in impaired carer health. (39-41)
Carer Experience Scale	Positive. Higher score indicates better carer experience which is expected to result in better carer health (4, 28, 33)
Family members' self-perceived impact of meningitis on a) work, b) finances.	Negative. Informal carers frequently experience loss of household income and increased care costs, which can cause stress and impaired mental health. (27, 33, 37, 38)

* Areas of life measured were (1) family and relationships, (2) social life, (3) exercise, and (4) views on personal health

Table 2: Descriptive statistics for baseline sample, non-carer sample and carer sample

Characteristic	Full sample (n=1587)	Non-carer sample (n=648)	Carer sample (n=199)
Family member			
Female (n, %)	1152 (72)	556 (86)	166 (83.8)
Age (years, mean (SD))	51.1 (12.8)	51.2 (12.1)	45.9 (11.9)
Health in 2012 (EQ-5D-5L, mean (SD))	0.88 (0.16)	0.87 (0.18)	0.83 (0.17)
Health in 2013 (EQ-5D-5L, mean (SD))	0.86 (0.18)	0.85 (0.19)	0.80 (0.20)
Survivor (patient)			
Female (n, %)	732 (46)	292 (45.2)	100 (50.3)
Age (years, mean (SD))	23.3 (16.1)	24.1 (16.2)	24.1 (20.3)
Time since infection (years, mean (SD))	12.0 (7.3)	12.3 (7.3)	10.4 (8.7)
Health in 2012 (EQ-5D-5L, mean (SD))	0.84 (0.26)	0.87 (0.19)	0.50 (0.35)
Health in 2013 (EQ-5D-5L, mean (SD))	0.83 (0.25)	0.85 (0.20)	0.52 (0.36)
Context			
Relationship to patient (parent, n (%))	1193 (75)	510 (79)	147 (74)
Lives with patient (n, %)	964 (60)	390 (60.5)	166 (83)
Provides care for patient (n, %)	246 (15)	0 (0)	199 (100)
Caring hours/week in 2012 (hours, mean (SD))	3.68 (14.5)	n/a	28.8 (31.7)
Caring hours/week in 2013 (hours, mean (SD))	2.81 (11.9)	n/a	21.2 (27.5)

Note: Carer and non-carer sample statistics presented here are only for the family members used in the validity analysis (that is, family members exposed to meningitis sequelae and assessed as the closest family member to the patient).

Table 3. Effect sizes and correlation coefficients for tests of construct validity of the EQ-5D-5L and SF-6D for measuring spillovers

Constructs associated with family member health spillover	**FAMILY MEMBER INDEX SCORES** **EQ-5D-5L (95% CI)**	**SF-6D (95% CI)**
'Caring about' hypotheses for non- carer sub-sample (n=648)		
Patient EQ-5D-5L	0.22*** (0.14 to 0.29)	0.19*** (0.11 to 0.26)
Patient VAS	0.19*** (0.11 to 0.26)	0.24*** (0.17 to 0.32)
Patient Mobility	-0.09* (-0.16 to -0.01)	-0.04 (-0.12 to 0.04)
Patient Self-Care	-0.14***(-0.22 to -0.06)	-0.13** (-0.21 to -0.05)
Patient Usual activity	-0.07 (-0.15 to 0.00)	-0.09* (-0.17 to -0.01)
Patient Anxiety	-0.23***(-0.30 to -0.15)	-0.20*** (-0.28 to -0.12)
Patient Pain	-0.18***(-0.26 to -0.10)	-0.15*** (-0.23 to -0.07)
Family life	-0.28* (-0.48 to -0.09)	-0.45*** (-0.66 to -0.26)
Social life	-0.52***(-0.74 to -0.31)	-0.56*** (-0.79 to -0.34)
Exercise	-0.82** (-1.11 to -0.53)	-0.59*** (-0.89 to -0.30)
Personal health	-0.95***(-1.31 to -0.59)	-0.83*** (-1.29 to -0.46)
Hypotheses for carer sub-sample related to 'caring about' or 'caring for' the patient (n=199)		
Patient EQ-5D-5L	0.26*** (0.12 to 0.39)	0.09 (-0.05 to 0.24)
Patient VAS	0.24*** (0.10 to 0.37)	0.15* (0.01 to 0.29)
Patient mobility	-0.19** (-0.32 to -0.05)	-0.06 (-0.21 to 0.08)
Patient self-care	-0.18** (-0.32 to -0.05)	-0.08 (-0.22 to 0.06)
Patient usual activity	-0.24***(-0.38 to -0.11)	-0.05 (-0.20 to 0.09)
Patient anxiety	-0.14 (-0.27 to 0.01)	-0.17* (-0.31 to -0.03)
Patient pain	-0.07 (-0.21 to 0.07)	-0.03 (-0.17 to 0.11)
Family life	-0.30* (-0.59 to -0.01)	-0.09 (-0.38 to 0.21)
Social life	-0.45** (-0.74 to -0.15)	-0.34* (-0.64 to -0.05)
Exercise	-0.55***(-0.85 to -0.24)	-0.48***(-0.79 to -0.18)
Personal health	-0.88** (-1.33 to -0.44)	-0.44 (-0.88 to 0.01)
Hypotheses for carer sub-sample solely related to 'caring for' the patient (n=199)		
Hours of care provided	-0.21** (-0.34 to -0.07)	-0.21** (-0.34 to -0.06)
Carer Experience Scale	0.34***(0.19 to 0.47)	0.23** (0.08 to 0.38)
Shares house	-0.21 (-0.58 to 0.17)	-0.06 (-0.45 to 0.32)
Daily care	-0.04 (-0.39 to 0.32)	-0.43* (-0.80 to -0.06)
Main carer	0.07 (-0.29 to 0.43)	-0.50* (-0.87 to -0.12)
Provides majority of care	-0.08 (-0.37 to 0.23)	0.12 (-0.19 to 0.42)
Provides personal care	0.11 (-0.17 to 0.39)	0.14 (-0.13 to 0.42)
Impact of meningitis on work	-0.24 (-0.53 to 0.05)	-0.35* (-0.65 to -0.05)
Impact of meningitis on finances	-0.13 (-0.42 to 0.18)	-0.04 (-0.34 to 0.26)
Provides constant supervision	-0.10 (-0.40 to 0.20)	-0.20 (-0.51 to 0.10)

*p<0.05, **p<0.01, ***p<0.001

§VAS- visual analogue scale

§Spearman's rank correlation coefficients of > 0.3 are considered weak associations, > 0.5 moderate, > 0.7 strong associations. For Cohen's D effect sizes of > 0.2 are considered small, > 0.5 moderate and > 0.8 large. The same interpretations apply for negative correlation coefficients and effect sizes.

§ Spearman's rho reported for all constructs which are continuous variables (patient EQ-5D-5L, VAS, mobility, self-care, usual activity, anxiety, pain, hours of care provided, Carer Experience Scale). Cohen's D reported for all other variables.

§ Note: Higher score of patient EQ-5D-5L and VAS indicates better patient health, whereas higher score of the individual items of patient EQ-5D-5L indicates poorer patient health.

§ Note: Higher score on the Carer Experience Scale indicates a better experience, hence a positive association with family member index scores.

Table 4: Tests of responsiveness of the EQ-5D-5L and SF-6D in non-carers

*p<0.05, **p<0.01, ***p<0.001

Patient EQ-5D-5L	Non-carer EQ-5D-5L 2012 baseline (mean)	Non-carer EQ-5D-5L 2013 follow-up (mean)	Difference between follow-up and baseline EQ-5D-5L (95% CI)	Effect size (Cohen's D)	n
Improved	0.83	0.84	0.01 (-0.02, 0.04)	0.01	46
No change	0.91	0.88	-0.03*** (-0.04, -0.01)	-0.19	234
Worsened	0.84	0.81	-0.03** (-0.06, -0.01)	-0.14	115
Patient EQ-5D-5L	**Non-carer SF-6D 2012 baseline (mean)**	**Non-carer SF-6D 2013 follow-up (mean)**	**Difference between follow-up and baseline SF-6D (95% CI)**	**Effect size (Cohen's D)**	**n**
Improved	0.76	0.76	0.00 (-0.03, 0.03)	0.00	43
No change	0.81	0.79	-0.02** (-0.03, -0.01)	-0.17	210
Worsened	0.76	0.75	-0.01 (-0.03, 0.01)	-0.05	104

§ Cohen's D effect sizes of 0.2 to 0.5 are considered small, 0.5 moderate and > 0.8 large.

Table 5: Tests of responsiveness of the EQ-5D-5L and SF-6D in carers

	Carer EQ-5D-5L 2012 baseline (mean)	Carer EQ-5D-5L 2013 follow-up (mean)	Difference between follow-up and baseline EQ-5D-5L (95% CI)	Effect size (Cohen's D)	n
Patient EQ-5D-5L					
Improved	0.79	0.83	0.04 (-0.04, 0.13)	0.19	22
No change	0.84	0.83	0.00 (-0.03, 0.02)	-0.02	60
Worsened	0.80	0.73	-0.06** (-0.11, -0.02)	-0.27	41
Hours of care provided					
Less care	0.80	0.77	-0.03 (-0.08, 0.01)	-0.16	29
No change	0.81	0.82	0.01 (-0.04, 0.06)	0.05	30
More care	0.84	0.79	-0.05* (-0.10, 0.00)	-0.31	23

	Carer SF-6D 2012 baseline (mean)	Carer SF-6D 2013 follow-up (mean)	Difference between follow-up and baseline SF-6D (95% CI)	Effect size (Cohen's D)	n
Patient EQ-5D-5L					
Improved	0.71	0.70	-0.01 (-0.07, 0.06)	-0.04	22
No change	0.71	0.70	-0.01 (-0.04, 0.01)	-0.12	59
Worsened	0.69	0.65	-0.05* (-0.08, -0.01)	-0.36	39
Hours of care provided for patient					
Less care	0.68	0.66	-0.02 (-0.06, 0.03)	-0.12	27
No change	0.71	0.71	0.00 (-0.04, 0.04)	-0.02	31
More care	0.72	0.67	-0.05* (-0.10, -0.01)	-0.51	21

*p<0.05, **p<0.01, ***p<0.001

§ Cohen's D effect sizes of of 0.2 to 0.5 are considered small, 0.5 moderate and > 0.8 large.

Acknowledgements

We are grateful to the Meningitis Research Foundation and families affected by meningitis for assistance in generating the original dataset. We are also grateful to Rachel Meacock and participants at the 2nd staff-student EuHEA in Paris (2015) where an earlier version of the manuscript was discussed. This research described in this paper was funded through a University of Birmingham MDS doctoral studentship. Kate Jolly is part-funded by the National Institute for Health Research (NIHR) and Collaboration for Leadership in Applied Health Research and Care (CLAHRC). Hareth Al-Janabi is part-funded by the National Institute for Health Research (NIHR). The views expressed are those of the authors and not necessarily those of the NIHR, the NHS or the Department of Health.

Supplementary materials: This study is based on data derived from a study of the family impact of meningitis. Please contact the corresponding author for details about access to the data.

REFERENCES

1. Drummond F. Methods for the Economic Evaluation of Health Care Programmes: Oxford University Press; 2005.
2. Hoefman RJ, Van Exel J, Brouwer W. How to include informal care in economic evaluations. PharmacoEconomics. 2013;31(12):1105-19.
3. Lavelle TA, Wittenberg E, Lamarand K, Prosser LA. Variation in the spillover effects of illness on parents, spouses, and children of the chronically ill. Applied health economics and health policy. 2014;12(2):117-24.
4. Al-Janabi H, Van Exel J, Brouwer W, Coast J. A Framework for Including Family Health Spillovers in Economic Evaluation. Medical Decision Making. 2016;36(2):176-86.
5. Bobinac A, van Exel NJ, Rutten FF, Brouwer WB. Health effects in significant others: separating family and care-giving effects. Medical decision making : an international journal of the Society for Medical Decision Making. 2011;31(2):292-8.
6. Schulz R, Sherwood P. Physical and Mental Health Effects of Family Caregiving. 2008.
7. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. Jama. 2016;316(10):1093-103.
8. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2012;15(5):708-15.
9. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013 2013. Available from: http://www.nice.org.uk/article/pmg9/chapter/foreword.
10. Wittenberg E, Saada A, Prosser L. How Illness Affects Family Members: A Qualitative Interview Survey. Patient. 2013;6(4):257-68.
11. Wittenberg E, Ritter GA, Prosser LA. Evidence of spillover of illness among household members: EQ-5D scores from a US sample. Medical decision making : an international journal of the Society for Medical Decision Making. 2013;33(2):235-43.
12. University of Sheffield. SF-6D 2015 [cited 2015]. Available from: https://www.shef.ac.uk/scharr/sections/heds/mvh/sf-6d.
13. Streiner DL, Norman GR. Health Measurement Scales: A practical guide to their development and use: OUP Oxford; 2008.
14. Whitehurst DG, Bryan S, Lewis M. Systematic review and empirical comparison of contemporaneous EQ-5D and SF-6D group mean scores. Medical decision making : an international journal of the Society for Medical Decision Making. 2011;31(6):E34-44.
15. Yang Y, Longworth L, Brazier J. An assessment of validity and responsiveness of generic measures of health-related quality of life in hearing impairment. Qual Life Res. 2013;22(10):2813-28.
16. Kuspinar A, Mayo NE. A review of the psychometric properties of generic utility measures in multiple sclerosis. PharmacoEconomics. 2014;32(8):759-73.
17. Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. The European journal of health economics : HEPAC : health economics in prevention and care. 2015;16(9):927-39.
18. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess. 2014;18(34):vii-viii, xiii-xxv, 1-188.

19. Al-Janabi H, Van Exel J, Brouwer W, Trotter C, Glennie L, Hannigan L, et al. Measuring Health Spillovers for Economic Evaluation: A Case Study in Meningitis. Health Economics. 2015.
20. Davidson T, Krevers B, Levin LA. In pursuit of QALY weights for relatives: empirical estimates in relatives caring for older people. The European journal of health economics : HEPAC : health economics in prevention and care. 2008;9(3):285-92.
21. Payakachat N, Tilford JM, Brouwer WB, van Exel NJ, Grosse SD. Measuring health and well-being effects in family caregivers of children with craniofacial malformations. Qual Life Res. 2011;20(9):1487-95.
22. Khanna R, Jariwala K, Bentley J. Psychometric properties of the EuroQol Five Dimensional Questionnaire (EQ-5D-3L) in caregivers of autistic children. Qual Life Res. 2013;22(10):2909-20.
23. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res. 2013;22(7):1717-27.
24. Al-Janabi H, Flynn TN, Coast J. Estimation of a Preference-Based Carer Experience Scale. Medical Decision Making. 2011;31(3):458-68.
25. Schulz R, McGinnis KA, Zhang S, Martire LM, Hebert RS, Beach SR, et al. Dementia patient suffering and caregiver depression. Alzheimer disease and associated disorders. 2008;22(2):170-6.
26. Bobinac A, van Exel NJ, Rutten FF, Brouwer WB. Caring for and caring about: disentangling the caregiver effect and the family effect. Journal of health economics. 2010;29(4):549-56.
27. Spore E. Quality of life of caregivers of children with chronic conditions [Doctoral thesis]. Chicago: University of Illinois; 2012.
28. Raina P, O'Donnell M, Rosenbaum P, Brehaut J, Walter SD, Russell D, et al. The health and well-being of caregivers of children with cerebral palsy. Pediatrics. 2005;115(6):e626-36.
29. Arafa MA, Zaher SR, El-Dowaty AA, Moneeb DE. Quality of life among parents of children with heart disease. Health and quality of life outcomes. 2008;6:91.
30. Klassen AF, Klaassen R, Dix D, Pritchard S, Yanofsky R, O'Donnell M, et al. Impact of caring for a child with cancer on parents' health-related quality of life. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(36):5884-9.
31. Yamada T, Hatt SR, Leske DA, Holmes JM. Health-related quality of life in parents of children with intermittent exotropia. Journal of AAPOS : the official publication of the American Association for Pediatric Ophthalmology and Strabismus / American Association for Pediatric Ophthalmology and Strabismus. 2011;15(2):135-9.
32. Kespichayawattana J, VanLandingham M. Effects of coresidence and caregiving on health of Thai parents of adult children with AIDS. Journal of nursing scholarship : an official publication of Sigma Theta Tau International Honor Society of Nursing / Sigma Theta Tau. 2003;35(3):217-24.
33. Lawoko S, Soares JJ. Quality of life among parents of children with congenital heart disease, parents of children with other diseases and parents of healthy children. Qual Life Res. 2003;12(6):655-66.
34. Buysse CM, Raat H, Hazelzet JA, Hop WC, Maliepaard M, Joosten KF. Surviving meningococcal septic shock: health consequences and quality of life in children and their parents up to 2 years after pediatric intensive care unit discharge. Critical care medicine. 2008;36(2):596-602.
35. An KJ, Song MS, Sung KW, Joung YS. Health-related quality of life, activities of daily living and parenting stress in children with brain tumors. Psychiatry investigation. 2011;8(3):250-5.
36. Goldbeck L. The impact of newly diagnosed chronic paediatric conditions on parental quality of life. Qual Life Res. 2006;15(7):1121-31.
37. Poley MJ, Brouwer WBF, Exel NJA, Tibboel D. Assessing health-related quality-of-life changes in informal caregivers: an evaluation in parents of children with major congenital anomalies. Qual Life Res. 2011;21(5):849-61.

38. Lin JD, Hu J, Yen CF, Hsu SW, Lin LP, Loh CH, et al. Quality of life in caregivers of children and adolescents with intellectual disabilities: use of WHOQOL-BREF survey. Res Dev Disabil. 2009;30(6):1448-58.
39. Konstantareas MM, Papageorgiou V. Effects of temperament, symptom severity and level of functioning on maternal stress in Greek children and youth with ASD. Autism : the international journal of research and practice. 2006;10(6):593-607.
40. Dearden C, Becker S. Young carers in the UK. 2004.
41. Burton LC, Newsom JT, Schulz R, Hirsch CH, German PS. Preventive health behaviors among spousal caregivers. Preventive medicine. 1997;26(2):162-9.
42. Keeley T, Al-Janabi H, Nicholls E, Foster NE, Jowett S, Coast J. A longitudinal assessment of the responsiveness of the ICECAP-A in a randomised controlled trial of a knee pain intervention. Qual Life Res. 2015;24(10):2319-31.
43. Wyrwich KW, Norquist JM, Lenderking WR, Acaster S. Methods for interpreting change over time in patient-reported outcome measures. Qual Life Res. 2013;22(3):475-83.
44. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. Journal of clinical epidemiology. 2008;61(2):102-9.
45. Walters S, Brazier J. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res. 2005;14(6):1523-32.
46. Sims AL, Parsons N, Achten J, Griffin XL, Costa ML, Reed MR. The World Hip Trauma Evaluation Study 3: Hemiarthroplasty Evaluation by Multicentre Investigation - WHITE 3: HEMI - An Abridged Protocol. Bone & joint research. 2016;5(1):18-25.
47. Gibson PR, Vaizey C, Black CM, Nicholls R, Weston AR, Bampton P, et al. Relationship between disease severity and quality of life and assessment of health care utilization and cost for ulcerative colitis in Australia: a cross-sectional, observational study. Journal of Crohn's & colitis. 2014;8(7):598-606.
48. Chen P, Lin KC, Liing RJ, Wu CY, Chen CL, Chang KC. Validity, responsiveness, and minimal clinically important difference of EQ-5D-5L in stroke patients undergoing rehabilitation. Qual Life Res. 2016;25(6):1585-96.
49. Nolan CM, Longworth L, Lord J, Canavan JL, Jones SE, Kon SS, et al. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. Thorax. 2016;71(6):493-500.
50. Peacock JL, Kerry S. Presenting Medical Statistics from Proposal to Publication: A Step-by-step Guide: OUP Oxford; 2007.
51. Sullivan GM, Feinn R. Using Effect Size-or Why the P Value Is Not Enough. Journal of graduate medical education. 2012;4(3):279-82.
52. Schulz R, Hebert R, Boerner K. Bereavement after caregiving. Geriatrics. 2008;63(1):20-2.
53. Wittenberg E, Prosser L. Disutility of Illness for Caregivers and Families: A Systematic Review of the Literature. PharmacoEconomics. 2013;31(6):489-500.
54. Davidson T, Levin L-A. Is the Societal Approach Wide Enough to Include Relatives? Incorporating Relatives' Costs and Effects in a Cost-Effectiveness Analysis. Applied health economics and health policy. 2010;8(1):25-35.
55. Goodrich K, Kaambwa B, Al-Janabi H. The inclusion of informal care in applied economic evaluation: A review. Value in Health. 2012;15(6):975-81.
56. Jit M, Edmunds WJ. Evaluating rotavirus vaccination in England and Wales. Part II. The potential cost-effectiveness of vaccination. Vaccine. 2007;25(20):3971-9.