SYSTEMATIC REVIEWS

Sources and Characteristics of Utility Weights for Economic Evaluation of Pediatric Vaccines: A Systematic Review

Michael Herdman, MSc1,*, Amanda Cole, PhD1, Christopher K. Hoyle, PhD1, Victoria Coles, PhD2, Stuart Carroll, MSc, MBA2, Nancy Devlin, PhD2

1Office of Health Economics, London, UK; 2Sanofi-Pasteur MSD, Maidenhead, Berkshire, UK

ABSTRACT

Background: Cost-effectiveness analysis of pediatric vaccines for infectious diseases often requires quality-of-life (utility) weights. Objective: To investigate how utility weights have been elicited and used in this context. Methods: A systematic review was conducted of studies published between January 1990 and July 2013 that elicited or used utility weights in cost-effectiveness analyses of vaccines for pediatric populations. The review focused on vaccines for 17 infectious diseases and is presented following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology. Results: A total of 6410 titles and abstracts and 225 full-text articles were reviewed. Of those selected for inclusion (n = 101), 15 articles described the elicitation of utility weights and 86 described economic modeling studies using utilities. Various methods were used to generate utilities, including time trade-off, contingent valuation, and willingness to pay, as well as a preference-based measure with associated value sets, such as the EuroQol five-dimensional questionnaire or the Health Utilities Index. In modeling studies, the source of utilities used was often unclear, poorly reported, or based on weak underlying evidence. We found no articles that reported on the elicitation or use of utilities in diphtheria, polio, or tetanus. Conclusions: The scarcity of appropriate utility weights for vaccine-preventable infectious diseases in children and a lack of standardization in their use in economic assessments limit the ability to accurately assess the benefits associated with interventions to prevent infectious diseases. This is an issue that should be of concern to those making decisions regarding the prevention and treatment of infectious childhood illnesses. Keywords: infectious diseases, literature review, pediatric, utilities, vaccine-preventable.

Copyright © 2016, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

The cost-effectiveness of health care interventions is an important element in making decisions regarding access to and reimbursement of health care technologies. The quality-adjusted life-year (QALY), a measure of quantity and quality of life (QOL), is widely used as a measure of incremental effect in economic evaluations of medical interventions and is a stipulated outcome in the National Institute for Health and Care Excellence’s (NICE’s) reference case [1]. In the case of vaccines, UK’s Joint Committee on Vaccination and Immunisation (JCVI) also specifies that evaluation of new vaccines should take into account “cost effectiveness based on costs per QALY and as a function of vaccine price at different cost per QALY thresholds” [2].

Estimating QALYs requires that values (utilities or weights) be assigned to health states (HSs) that are relevant to the condition of interest. In some cases, values are elicited directly by using techniques such as time trade-off (TTO) and standard gamble (SG). Alternatively, values can be obtained indirectly by using a generic HS classification system, such as the EuroQol five-dimensional questionnaire (EQ-5D) [3] or the Health Utilities Index (HUI) [4,5], which is accompanied by value sets. In practice, various approaches have been used to elicit utility weights [6–8], although a recent review of how QALYs are estimated for pediatric patients in cost-utility analyses performed in the United Kingdom [9] found that QALYs were generated most frequently using existing preference-based instruments, particularly the EQ-5D and the HUI.

Conflict of interest: M. Herdman, A. Cole, C. K. Hoyle, and N. Devlin are or were all employees of the Office of Health Economics at the time this study was performed. V. Coles and S. Carroll are or were employees of Sanofi-Pasteur MSD at the time this study was performed.

* Address correspondence to: Michael Herdman, Office of Health Economics, Southside, 7th Floor, 105 Victoria Street, London SW1E 6QT, UK.

E-mail: mherdman@ohe.org.

1098-3015/$36.00 – see front matter Copyright © 2016, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.jval.2015.11.003
The methods used to obtain values for HSs can affect the conclusions drawn about the cost-effectiveness of health care interventions, including vaccines, because several studies have shown that different methods lead to different utility weights [10–12]. Although measuring and valuing health is a complex task in adults, it is perhaps even more conceptually and methodologically demanding in children [13–15], and the appraisal of vaccines in pediatric populations is particularly challenging, as highlighted recently by Bruggenjuren et al. [16]. Although previous reviews have examined the generation and use of utility weights in pediatric populations in general [6,9], none has focused specifically on the use of utility weights for HSs associated with infectious diseases and their use in economic evaluations of vaccines to prevent those diseases.

The Patient Reported Outcomes in Children with Infectious Diseases (PROCHID) study was a project to systematically review the development and use of patient-reported outcome (PRO) measures in pediatric populations with vaccine-preventable infectious diseases. As part of that project, we reviewed the use of utility measures and utility weights in that population, which is reported in this article. The methods and results of the PRO element of the PROCHID project are reported elsewhere [17].

Methods

Search Strategy, Data Sources, and Eligibility Criteria

The overall review for the PROCHID study, of which the review of utility generation and use formed a part, was conducted in compliance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [18]. The search period was from January 1, 1990, to July 31, 2013, and included MEDLINE and EMBASE using the SCOPUS search engine. Published studies that reported the generation and/or use of utility weights for the pediatric population with vaccine-preventable infectious diseases of interest were identified using the general study algorithm and the utility-specific algorithm presented in Appendix 1.

The aim was to identify studies that had generated or used utility weights for the economic evaluation of vaccines used in pediatric populations (defined for the purpose of this study as populations aged <18 years) to prevent any of the following 17 infectious diseases: anogenital warts (AGWs), diphtheria, hepatitis B, Haemophilus influenzae type B (HiB), influenza, measles, mumps, rubella, meningococcal B meningitis, meningococcal C meningitis, norovirus, pertussis, polio, pneumococcal disease, rotavirus, tetanus, or varicella.

The search focused on publications in English and Spanish, which were the languages spoken by members of the study team, and on studies performed in Europe, North America, Australia, and New Zealand, because we were primarily interested in the economic evaluation of vaccines in developed countries.

Study Selection

Two independent reviewers screened the titles and abstracts of all the studies identified by the search algorithms to determine whether they met selection criteria. Discrepancies in the selection of publications were resolved through discussions between reviewers and recourse to a third reviewer where necessary. Studies were included for full-text review if they

1. described the elicitation or generation of utility weights for HSs associated with any of the infectious diseases of interest;
2. described the use of utility weights in cost-effectiveness or cost-utility analysis of vaccines to prevent any of the infectious diseases of interest;
3. referred to pediatric populations (age <18 years);
4. were performed in the countries of interest;
5. were published in English or Spanish; and
6. were available as full-text publications.

The reference lists of all full-text articles retrieved were reviewed to further identify potentially relevant studies. We excluded gray literature, such as unpublished manuscripts, government reports, or conference proceedings, and studies that were performed only in adult populations.

Data Extraction and Synthesis

A predesigned data extraction form was used to extract information on the following: year(s) of research, country, study type (HS valuation, cost-effectiveness, cost-utility, burden of disease, modeling, others), study objectives, intervention evaluated, study population (particularly any information on age groups), strategy used to estimate utility decrements (including source of utility weights and method of elicitation), and utility weights generated or used in any modeling. We also recorded any study limitations noted by the authors in relation to the utilities derived or used. Data extraction was performed for each full-text article by two reviewers. If doubts arose, for example, about the methods or sources used to obtain utilities or about the values themselves, these were resolved through discussions between the two reviewers and, if need be, through recourse to a third reviewer. A thematic approach to data synthesis was adopted on the basis of the information retrieved. From an early stage in the review of the full-text articles, they were organized and analyzed according to whether they were primarily concerned with generating HS utilities or with applying utilities in economic models.

Results

The searches performed for the PROCHID project as a whole yielded 6410 journal article references covering both PRO and utility-based studies. Of these, 6301 were excluded because of duplication or a failure to meet inclusion criteria after title or abstract review (Fig. 1). Full-text articles were retrieved and reviewed for the remaining 107 references (two articles could not be obtained). The review of reference lists identified a further 118 journal articles for full-text review, giving a total of 225 articles. Of these, 101 were retained for data extraction—15 articles reported on the elicitation of utility weights and 86 referred to the application of utility weights in economic models involving pediatric populations. A summary of results from the two types of study is provided in Tables 1 and 2, respectively (for further information, see Appendix Tables 1 and 2 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2015.11.003). No full-text articles referring to either utility generation or modeling involving utilities were identified for diphtheria, polio, or tetanus.

Utility Generation Studies

Table 1 presents the key features of 15 studies reporting on the generation of utility weights. Utility generation studies were defined as those in which utility weights were elicited from the general population and/or patient samples either directly (using valuation methods such as the TTO or the SG) or indirectly (after collecting data from relevant patient and/or caregiver samples using a preference-based measure, such as the EQ-5D or the HUI, which has accompanying value sets). Modeling studies that used utility estimations based on author or expert opinion were not considered elicitation studies and are described later in the article. No studies were found that reported utility generation for children with diphtheria, hepatitis B, HiB, measles, mumps,
rubella, norovirus, or tetanus. The main findings for each condition are summarized hereafter. Further information on these studies can be found in Appendix Table 1 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2015.11.003.

Anogenital Warts
The two primary studies identified [19,20] administered the EQ-5D to patients with AGWs and used EQ-5D value sets to generate utilities. Woodhall et al. [20] used the value set provided by Dolan et al. [21], and results were compared with those of matched population controls from Kind et al. [22]; the source for the utilities used in the study by Brisson et al. [19] is unclear. The sample size in the study by Woodhall et al. was much larger (895 vs. 31), and the mean disutility for AGWs in the study was just over half of that presented by Brisson et al. Of note is the finding by Woodhall et al. that disutility was the highest in females aged 16 to 19 years.

Influenza
A wide range of methods were used to estimate utilities associated with influenza in pediatric populations in the five studies identified [23–27], ranging from direct elicitation using TTO and willingness-to-pay (WTP) methods to indirect elicitation through the administration of the EQ-5D [27] or the use of a visual analogue scale [24]. Some authors [23,26] investigated possible differences in utilities for pediatric and adult populations and found that survey respondents were willing to trade off more time to avoid influenza-associated HSs in young children, resulting in greater disutilities and QALY loss for younger age groups. Prosser et al. [26] also showed that values differed depending on whether the TTO or the WTP approach was used, and utilities and disutilities varied considerably between the studies identified or were not directly comparable because of differences in the way results were reported.

Meningitis
We identified only one primary study that met our inclusion criteria for meningitis. Koomen et al. [28] studied the health problems of school-age survivors of bacterial meningitis, using school-age siblings or friends as a reference group. The final cohort was quite large, but consisted of only 42% of the initially selected children, meaning that there may have been some selection bias. Parents assessed the health-related quality of life (HRQOL) of their children using the HUI2, and utility weights for the resulting HSs were generated using the Canadian value set, although the study was conducted in The Netherlands.

Pertussis
We found only one article that reported on the elicitation of utilities for use in pediatric patients with pertussis [29]. It was carried out in the United States in 2003 and involved obtaining TTO and WTP valuations via telephone interviews with adult patients with pertussis and with the parents of adolescents with pertussis in Massachusetts. As part of the study, the authors explored whether values for HSs were affected by whether they were experienced by infants, adolescents, or adults. They found that, overall, utilities were lower for HSs occurring in adolescents than for HSs occurring in adults and that utilities for HSs occurring in infants were the lowest of all, although the latter result may have been due, in part, to differences in the HSs valued for infants.

Pneumococcal Disease
Two studies had elicited utility values for pneumococcal disease, one in the United Kingdom [30] and the other in the United States [21]. The approaches to deriving utilities varied, with Prosser et al. [31] using the TTO and the WTP methods and Legood et al. [30] deriving utility scores indirectly, through the administration of the HUI3. Other differences included the fact that Legood et al. [30] provided a utility score for a “case” (i.e., a
Table 1 – Key features of studies on generation of utility weights.

<table>
<thead>
<tr>
<th>Study features</th>
<th>Total (n = 15)*</th>
<th>AGWs (n = 2)</th>
<th>Influenza (n = 5)</th>
<th>Meningitis (n = 1)</th>
<th>Pertussis (n = 1)</th>
<th>Pneumococcal disease (n = 2)</th>
<th>Rotavirus (n = 2)</th>
<th>Varicella (n = 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods used to generate utilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-5D†</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-5D-Y†</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HUI†</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTO (bespoke)</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTP (bespoke)</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS (bespoke)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General population</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient values</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parents/carers</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experts/clinicians</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement of pediatric population, either in completing questionnaire or in rating HSs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utilities (point estimate or range, depending on the number of studies)</td>
<td>U: 0.87</td>
<td>U: 0.29–0.609</td>
<td>DU: 0.05</td>
<td>U: 0.67–0.96</td>
<td>U: 0.774</td>
<td>U: 0.200–0.986</td>
<td>QALY loss: 0.001–0.010</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DU: 0.056</td>
<td>QLY: 0.0075–0.0475</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>QLY: 0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DU, disutility weight; EQ-5D, EuroQol five-dimensional questionnaire; EQ-5D-Y, youth version of EuroQol five-dimensional questionnaire; HS, health state; HUI, Health Utilities Index; QALY, quality-adjusted life-year; TTO, time trade-off; U, utility weight; VAS, visual analogue scale; WTP, willingness to pay.

* The sum is often greater than the total number of studies because of multiple studies using more than one method to elicit utilities.
† Questionnaire administered to a sample of the population of interest, or their proxies, and utilities calculated from associated algorithms.
‡ For one of the pneumococcal studies, the HUI was completed by patients if aged ≥11 years, and by parents if the patient was aged 5–10 years.
### Table 2 - Key findings from modeling studies by condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Anogenital warts     | - 17 modeling studies identified involving populations aged <18 y (published 2004–2013)  
- Variety of countries (four studies each in the United States and the United Kingdom)  
- Generally modeled costs and outcomes for girls aged 12 y and older, although one modeled costs and outcomes for boys and men aged 9–26 y  
- Independent of where they were performed, 11 studies used utility weights for AGWs derived by Myers et al. [49] from a sample of healthy college-aged students in the United States  
- Methodologies for obtaining utilities used in the models were variable, but were primarily based on the use of the VAS or the TTO to value HSs related with AGWs, or via application of the EQ-5D or the HUI and generation of utilities using the corresponding algorithms  |
| Influenza            | - Mean utility weights cited for a case of AGW varied from 0.91 to 0.944, and 0.98  
- 18 modeling studies identified involving populations aged <18 y (published 2003–2013)  
- Largest number of studies (six) carried out in the United States, followed by the United Kingdom (five)  
- The modeling approach varied, with some studies modeling a lifetime cohort and others concentrating on children and adolescents or very specific populations such as pregnant women aged 15–44 y  
- Utilities generated using a very wide variety of methods, from the conventional approach of generating utilities from self-rated scores on the EQ-5D or the HUI (occasionally but not usually administered in adolescents or children when appropriate) to the less conventional approaches such as recalibrating Likert scale ratings to VAS scores and then converting to the TTO [65]  |
| Hepatitis B          | - A number of assumptions were used, e.g., in one study, children were assumed to have utility weights equivalent to those of “otherwise healthy adults”  
- Occasionally unclear how utility weights are derived  
- Examples of utility weights used include influenza with mild symptoms 0.52 (0.28–0.93); severe symptoms 0.05 (–1.0–0.93) [70]; “influenza without hospitalization” 0.65 and “influenza with hospitalization” 0.5 [66]  |
| Haemophilus influenza type B | - 2 modeling studies identified involving populations aged <18 y (published 2010–2011)  
- One in the United Kingdom, one in The Netherlands  
- Modeled costs and outcomes for children from 0 mo and 12 y to 99 y (United Kingdom) and from <15 y to >65 y (The Netherlands)  
- Utilities generated using the TTO and the SC; utilities in the UK study largely based on expert estimates from the United States; utilities in the Dutch study were from an international survey of patients and general population in six countries  
- Mean utility weights used for chronic HBV and compensated cirrhosis were 0.96 and 0.93 in the United Kingdom and 0.68 and 0.69 in The Netherlands, respectively  |
| Measles              | - Only one modeling study identified involving populations aged <18 y (published 2001)  
- Calculated the benefits of a supplemental immunization effort to increase preschool immunization rates (US data)  
- Loss in utility associated with severe sequelae of measles expressed as difference in utility between average population health and health associated with disability  
- Utilities for average health estimated using data from the Beaver Dam Health Outcomes Study (BDHOS) and the Medical Outcomes Study (MOS); method for calculating utilities associated with measles-related states is not clear, although appears to be based also on the BDHOS and the MOS  |
| Meningitis           | - Utility decrement for slight disability due to measles: 0.1; moderate disability: 0.6; severe disability: 0.9  
- Six modeling studies identified involving populations aged <18 y (published 2002–2013)  
- Two in The Netherlands, two in Canada, one each in the United States and Switzerland  
- Costs and outcomes modeled for a range of immunization strategies and age groups, from a time horizon of 2–12 mo to lifetime  
- Range of approaches used to estimate utilities, from application of a percentage reduction (e.g., utility decrement for moderate sequelae: 20% reduction in QOL) used in the study by Jaccard Ruedin et al. [83] to estimates of utility weights or losses on the basis of a wide range of sources, e.g., by De Wals and Erickson [89], or a single source, e.g., by Hepkema et al. [90]  
- Some studies (e.g., Ortega-Sanchez et al. [86]) used utilities obtained from different preference-based measures for different sequelae, as if weights obtained using the EQ-5D and the HUI were captured on a similar scale  
- Methods used to estimate utilities occasionally unclear or not reported, or based on use of nonstandard versions of preference-based measures, e.g., use of the EQ-5D+ [136]  |
|                      | - Differences in utility weights or losses for same sequelae across studies, e.g., quality of life of patients with amputations or scars: 0.83 in the study by Bos et al. (2001) compared with utility weights of 1 for scarring and 0.70 for a single amputation in the study by Ortega-Sanchez et al. [86]  |

continued on next page
child below the age of 14 years who had been infected with pneumococcal meningitis), whereas Prosser et al. [31] focused on different outcomes stemming from the infection, for example, otitis media and meningitis. Legood et al. [30] appeared to have used Canadian values associated with the HUI3 to calculate utilities, even though the study was performed in the United Kingdom. However, the value set used was not made clear in the article.

**Rotavirus**

Two studies, one performed in Canada [32] and the other in the United Kingdom [33], elicited utilities for pediatric populations.

**Table 2 – continued**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Pertussis         | • None of the studies appeared to take into account possible differences in weights by age group  
• Five modeling studies identified involving populations aged <18 y (published between 2002 and 2010)  
• Four in The Netherlands, one in the United Kingdom  
• Costs and outcomes modeled for a range of immunization strategies and age groups  
• Utilities used in modeling were all based on two studies [29,66] performed in the United States using TTO and WTP in patients with pertussis and parents of children with the condition, and/or on authors’ own assumptions  
• The TTO and WTP studies by Lee et al. [29,66] provided age-specific utility weights or decrements  
• Despite using the same source for utilities, there was variation in weights used across studies because of authors’ assumptions  
• Examples of age-specific utility decrements corresponding to mild, moderate, and severe cough, respectively, were as follows [94]: 0 y, 0.42 (severe only); 1–3 y, 0.2, 0.28, 0.39; 4–9 y, 0.175, 0.25, 0.36; 10–17 y, 0.15, 0.22, 0.33; 18–59 y, 0.1, 0.15, 0.19; 60+ y, 0.1, 0.15, 0.19  
• Wide range of sources used for utility weights, although there was a tendency to rely heavily on a series of “core” publications for utilities. [133,134]  
• Utility weights or losses are thereby “transferred” from one article to another and across countries, e.g., Strutton et al. [107], when modeling results for Germany, Greece, and the Netherlands state that country-specific utilities are provided, but the utilities used are identical for the three countries and all were derived from Oostenbrink et al. and Maddigan et al. [133, 135]  
• Methods used to estimate utilities occasionally unclear or not reported, or based on use of nonstandard versions of preference-based measures, e.g., the use of EQ-5D+. [136]  
• One study [137] provided some cost per QALY results, but provided no information on how utility weights were obtained or which values were used  
• Some differences in the utility values used for the same sequelae, e.g., utility weight of 0.45 used for deafness in the study by Diéz-Domingo et al. [97] and utility decrement of 0.8 used for hearing loss in the study by Earnshaw et al. [98]  
| Pneumococcal disease | 14 modeling studies identified involving populations aged <18 y (published 2003–2013)  
• Studies covered 11 countries (three studies each in the United Kingdom and The Netherlands)  
• A range of modeling strategies used, e.g., from a time horizon of birth to 10 y [Bos et al.] to 2–12 mo or birth to 5 y [96] to use of lifetime cohorts and data  
• Wide range of sources used for utility weights, although there was a tendency to rely heavily on a series of “core” publications for utilities. [133,134]  
• Utility weights or losses are thereby “transferred” from one article to another and across countries, e.g., Strutton et al. [107], when modeling results for Germany, Greece, and the Netherlands state that country-specific utilities are provided, but the utilities used are identical for the three countries and all were derived from Oostenbrink et al. and Maddigan et al. [133, 135]  
• Methods used to estimate utilities occasionally unclear or not reported, or based on use of nonstandard versions of preference-based measures, e.g., the use of EQ-5D+. [136]  
• One study [137] provided some cost per QALY results, but provided no information on how utility weights were obtained or which values were used  
• Some differences in the utility values used for the same sequelae, e.g., utility weight of 0.45 used for deafness in the study by Diéz-Domingo et al. [97] and utility decrement of 0.8 used for hearing loss in the study by Earnshaw et al. [98]  
| Rotavirus | We identified 15 modeling studies involving populations aged <18 y (published between 2007 and 2013)  
• Studies performed across 11 countries (maximum number of studies per country was three, in The Netherlands)  
• Almost all studies modeled costs and outcomes for hypothetical cohorts of newborns from 0 to 5 y, although one study modeled up to 15 y  
• Utilities in most of the studies were derived from one of two publications [32,125], both of which reported the results of a Canadian study to derive utility weights for children and caregivers using the HUI2 and the EQ-5D  
• Canadian and UK population weights used to calculate utilities on the basis of self-rated health on the HUI2 and the EQ-5D, respectively  
• Examples of mean utility weights used were as follows: healthy child, 0.986; child with rotavirus, 0.927; healthy caregiver, 0.967; caregiver of child with rotavirus, 0.91 [113]; however, even when the same Canadian source publications were used for utility weights, there were occasionally differences in the weights used, e.g., in another study weights attributed to child and carer were 0.884 for “mild rotavirus diarrhea” and 0.816 for “severe diarrhea”  
| Varicella | We identified four modeling studies involving populations aged <18 y (published between 2002 and 2012)  
• Studies performed in the United Kingdom, the United States (two), and The Netherlands  
• The modeling approach varied, with two studies modeling a lifetime cohort and two studies less than a lifetime  
• All used utility weights or decrements or losses from a Canadian study to derive utility weights for children and caregivers using the HUI2 and the EQ-5D  
• Canadian and UK population weights used to calculate utilities on the basis of self-rated health on the HUI2 and the EQ-5D, respectively  
• Examples of mean utility weights used were as follows: healthy child, 0.986; child with rotavirus, 0.927; healthy caregiver, 0.967; caregiver of child with rotavirus, 0.91 [113]; however, even when the same Canadian source publications were used for utility weights, there were occasionally differences in the weights used, e.g., in another study weights attributed to child and carer were 0.884 for “mild rotavirus diarrhea” and 0.816 for “severe diarrhea”  
|  | AGW, anogenital wart; EQ-5D, EuroQol five-dimensional questionnaire; HBV, hepatitis B virus; HS, health state; HUI, Health Utilities Index; QALY: quality-adjusted life-year; SG, standard gamble; TTO, time trade-off; VAS, visual analogue scale; WTP, willingness to pay. |
that had experienced infection with rotavirus. Both relied on the use of preference-based measures rather than directly eliciting values. Brisson et al. [32] asked parents to describe their children’s health using the HUI3, whereas Martin et al. [33] relied on the judgment of clinical experts, using the EQ-5D. The two studies produced substantially different utility scores, with a range from 0.200 to 0.781 in the study by Martin et al. and 0.896 to 0.986 in the study by Brisson et al. [32]. The discrepancies in values are perhaps not surprising given the methodological differences between the studies, including the use of different instruments, different raters, and different value sets.

**Varicella**

The two studies that generated utilities for varicella-related HSs [34,35] both used an indirect approach via the administration of the EQ-5D [34] and the HUI2 [35]. Of note is the fact that in the study by Bilcke et al. [34] the instrument could be completed up to 6 months after the infection, leading to the possibility of recall bias. Sample sizes were relatively small in both studies, and neither reported which value sets were used to calculate utilities. Also of note is the fact that 26% of the sample in the study by Bilcke et al. did not complete the EQ-5D, with the authors assuming that this was due to the unsuitability of some EQ-5D content for this age group.

**Modeling Studies Using Utility Weights from Secondary Sources**

Table 2 summarizes findings related to the use of utility weights used in economic modeling studies in pediatric populations with any of the infectious diseases of interest. Further information on these studies can be found in Appendix Table 2 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2015.11.003.

**Anogenital Warts**

We identified 17 studies that modeled the cost-effectiveness of vaccines that could prevent AGWs and included individuals younger than 18 years in the modeling framework [36–52]. Only two studies [43,52], however, used utility weights generated from samples that included individuals younger than 18 years. The rest all used utility weights generated from adult samples even though it is not clear whether “adult” utility values are equally applicable to younger populations with AGWs. In fact, Woodhall et al. [20] suggested that the disutility for AGW-associated HSs may be greater in younger patients. Interestingly, the studies by Elbasha and Dasbach [41] and Jit et al. [42] both reported that the outcome of their cost effectiveness analysis was sensitive to the utility weights used. There were also doubts as to whether instruments such as the EQ-5D and the HUI would adequately capture the negative psychological effects of the condition.

**Haemophilus influenzae Type B**

We identified three studies that modeled the cost-effectiveness of vaccines for Hib, two for an Australian infant population [53,54] and one for France [55]. Because there were no primary studies reporting utilities in Hib, authors of modeling studies all relied on early generic HS classifiers such as the Rosser-Kind Index [56] or Torrance’s multiattribute utility model [57], and assumed profiles on the basis of their own judgments [53] or those of clinical experts [54,55].

**Hepatitis B**

Only two studies incorporated utility weights to calculate the cost-effectiveness of vaccination or screening programs for hepatitis B in populations including subjects in the pediatric age range. The study by Siddiqui et al. [58] was performed from a UK perspective but used utility weights from a study in the United States by health care professionals [59]. It was also difficult to identify the specific source of utility estimates in the study by Siddiqui et al.; although they referenced Shepherd et al. [60], the values used were in fact taken from one of the manufacturers’ evidence submissions to NICE. However, Veldhuijzen et al. [61] used HS valuations from Levy et al. [62], which were obtained in six countries. Veldhuijzen et al. [61] used the overall utility weights from the study despite the fact that there was significant variation between countries. There were considerable differences between the Siddiqui et al. and Veldhuijzen et al. studies in some of the utility weights applied.

**Influenza**

We identified 18 modeling studies for influenza in populations that included subjects in the pediatric age range [63–80]. As noted in other conditions, HS values used were occasionally derived from adults’ self-ratings on preference-based measures, and not from children, leading to an assumption that the impact of the disease would be the same in children as in adults [76,79]. Also of note was the use of identical utility weights for complicated and uncomplicated influenza or for severe and mild symptoms [70,78].

Many of the studies combined utility estimates from several sources, despite important methodological differences between those sources [64], and transparency was occasionally lacking about how utility values were derived or modified by the authors, making it difficult to verify or understand the values used [65]. Source data also appeared to be misinterpreted on occasion [67], and some studies relied on somewhat outdated literature. Lee et al. [66], for example, based their utility estimates for influenza on a value for “contagious disease” from a 1978 publication [81], and Tarride et al. [78] used a value originally published in 1972 in a study of tuberculosis [82].

**Measles**

Only one article [83], in which Zwanziger et al. estimated utility decrements for mild, moderate, and severe disability caused by the neurological sequelae of measles, was identified as being relevant to the present review. We found no studies reporting on QoL associated with the disease state itself. The utility decrements used were estimated on the basis of the authors’ assumptions around health parameters for measles characterized by the EQ-5D, and no adjustment was made to take into account possible differences in utility weights across age groups.

**Meningitis B and C**

Six studies [84–90] (two articles drew from the same study) referred to the use of utility weights in economic analyses of interventions for meningitis. In the study by Jaccard Ruedin et al. [85], decrements in utility scores associated with long-term complications of meningitis were derived in two steps, including the use of a (generic) valuation framework [57] to estimate a utility on the basis of the authors’ own assumptions, which were then submitted to an expert panel for review. Ortega-Sanchez et al. [86] and Shepard et al. [84] used estimates for utility weights based on multiple conditions closely resembling each of the long-term sequelae of meningococcal disease, although the sources were not specific to meningitis and were not always based on studies in children. Bos et al. [87] made use of utility values obtained using an unconventional six-dimensional version of the EQ-5D, although it is not clear whether utility values obtained with this extended version would be comparable to values obtained with the standard version of the EQ-5D. De Wals et al. [88] assumed an average life utility value reduction of 0.282 in
meningococcal disease survivors, but it is not clear how utility values were calculated or whose values were used to weight HSs.

Pertussis
For pertussis, we found five studies that met the inclusion criteria [91–95]. The models reported were all developed to evaluate the cost-effectiveness of vaccines in the Netherlands or the United Kingdom, yet most used utility values derived in a US study [29]. As in some of the studies for influenza, Lee et al. [29] (one of the source studies) used a TTO method in which adults were asked to specify how much time they would trade from the end of their own lives to avoid an HS in their child; potential implications of this method include the possibility that parents’ preferences may incorporate features of their relationship with the child, such as altruism [29].

Interestingly, the utility weights used in the studies by de Greeff et al. [91] and de Vries et al. [92] diverged somewhat, despite being drawn from the same source [29]. de Greeff et al. [91] used median rather than mean utility values. Stevenson et al. [93] provided no information on the source of their utility estimates.

Pneumococcal Disease
We identified 14 studies that used utilities in economic analyses of vaccines against pneumococcal disease, with many of the authors tending to resort to the same sources to obtain utility weights or QALY losses [96–109]. The values used were therefore rarely specific to the setting in which the economic evaluations were performed and the values used were a mix of utility weights, QALY loss estimates, and utility decrements, which were not always well described or easy to trace to the source studies. In one case [109], there was no information available on how the utility weights used were derived or what they were.

Rotavirus
Most of the 15 studies that used existing HS valuations to model the economic outcomes of vaccines against rotavirus [110–124] referenced the study by Briss et al. [32] or Sénecal et al. [125] as the source for utilities. Both of these reported the results of a Canadian study to derive utility values for children and for caregivers. Some of the modeling studies included the utility impact on caregivers, whereas others did not, and several of the studies incorporated untested assumptions into the analysis. For instance, Bickle et al. [110] assumed that QALY loss in children and caregivers would be halved in a situation in which no health care was received. Tilson et al. [124] noted that one of the potential problems in using the utility weights from Briss et al. [32] is that they do not distinguish between different severities of illness, which may lead to an underestimation of the QALY loss for hospitalized infants and an overestimation for home-treated cases. Another limitation is that the HUI2 was not designed for use in very young children or in acute health conditions [124].

We also found that utility values used differed even when drawn from the same source study [113,117] and that the source study was sometimes incorrectly referenced [111]. In two studies [119,122], the utility values for children and caregivers were switched.

Varicella
Four modeling studies for varicella met our inclusion criteria. They were performed in the Netherlands [126], the United Kingdom [127], and the United States [128,129], although all used utility weights from a study conducted in the United Kingdom [130], except the study by Rotberg et al. [128], which used expert assessments on the Index of Well-Being scale. Again, untested assumptions were applied in some cases, such as the assumption by Van Hoek et al. [127] that patients with less than 50 spots would have only 25% of the QALY loss of patients with more than 50 spots. In the three modeling studies that used US values, there were differences in the utility values used, despite being drawn from the same source. Finally, Brisson and Edmunds [130] made interesting comments regarding the difficulties of using the SG to capture health impacts for short-term and often mild disease states such as those associated with chickenpox and the very different utility values derived using SG and contingent valuation techniques. Table 3 summarizes the principal findings from the study overall.

Discussion
In this study, we found that there are considerable gaps and weaknesses in the current evidence base for utilities used in economic evaluations of pediatric vaccines for infectious diseases. For many of the diseases covered in this review, no utility weights were available, and in cases in which such weights had been generated, we found substantial variation in the methodology used to generate them. Most of those studies have been carried out in North America and the United Kingdom, and the extent to which utility weights are transferable geographically is not clear. The use of different methodologies and approaches undermines the comparability of results across studies, and the lack of standardization may cast doubt on the reliability of results in cost-effectiveness analysis in which such weights are used.

In part at least, these issues arise because there are considerable challenges in undertaking research in this area. For example, for diseases for which vaccines have been available for some time, and whose prevalence is low, it may be difficult to obtain relevant data on preferences regarding QOL. Furthermore, many of the relevant disease states are either short-duration acute states or very mild states, both of which pose challenges for the use of conventional methods such as the TTO and the SG. These difficulties in obtaining utilities for children’s HRQOL compound the difficulties noted by Herdman et al. (forthcoming [17]) in using PROs in children.
Furthermore, there are fundamental unresolved normative issues. It is not clear who should be valuing HRQOL states experienced by children, and there is considerable variation in this respect among the studies we reviewed. Some of the tasks involved in HS valuation might be too difficult for younger children, although it is not clear whether having HSs valued by children is even desirable on normative grounds. For example, it is often argued that utilities used in the estimation of QALYs should reflect the preferences of the general public, as taxpayers, rather than the preferences of the subgroup of the population affected by a disease or treatment [131,132]. The issue is further complicated in the case of children, though, because adult members of the general public, if asked to value states that clearly relate to children, will be potentially confounding both their views on how bad the states will be when experienced by children and views about the priority that should be accorded to the treatment of children. Furthermore, it is not clear what perspective an adult will or should take in valuing children's HSs: thinking of themselves as a child? A hypothetical child or a child they know, perhaps their own?

As well as the need for more and better studies to elicit utility weights for the conditions studied here, there is also arguably a need for better guidance on how economic evaluations for vaccines are carried out. As noted, UK's JCVI follows the NICE reference case and specifies that evaluation of new vaccines should take into account cost-effectiveness based on costs per QALY, but there are aspects that are specific to vaccines and warrant the development and implementation of a JCVI-specific reference case. A part of that guidance could focus on how utility weights are collected and used in the economic evaluation of vaccines for pediatric populations. The JCVI Working Group on Cost-Effectiveness, which has recently been formed to look at how the JCVI should be approaching economic evaluation, could be an important step in that direction. Further research into areas such as public preferences for prevention over cure could also help inform vaccination policy, more accurately capture the value of pediatric vaccination programs, and improve the transparency of decision making.

Without such initiatives, there is a danger that vaccines will be valued inappropriately. Inaccurate valuation of health interventions will undermine the principle of economic evaluation and will lead to inefficient allocation of resources. Undervaluation of both existing and novel vaccines could have serious consequences if it acts as a disincentive for future investment and innovation in vaccines, with detrimental effects on health at the population level.

In terms of study limitations, perhaps the most relevant is that we focused only on developed countries, which means that the results may not be generalizable to countries in the developing world. This was done to limit the scope of the review to ensure its feasibility and because the relevance of different conditions and the environment for the use of PRO instruments and cost-utility analysis are likely to differ substantially between developed and developing countries.

In conclusion, our results show that there is a paucity of studies providing utility weights for use in economic evaluation in the vaccine-preventable infectious conditions reviewed. This makes it extremely difficult to accurately assess the impact of these conditions on population health and limits the possibilities for reliably evaluating the benefits and value for money of their prevention or treatment.

Acknowledgments

We thank Antoni Parada of the Agency for Health Quality and Assessment of Catalonia (AQuAS) for his invaluable help in developing the search strategy and performing the searches in SCOPUS.

Source of financial support: Funding for this study was provided by Sanofi Pasteur MSD.

Appendix A. SCOPUS Search Algorithm to Identify Publications on the Use of Utilities in Diphtheria

Utilities in Diphtheria

((TITLE(newborn OR neonatal OR infant* OR child OR children OR kid OR kids OR pediatric* OR paediatric* OR teenage* OR young* OR youth* OR juvenile) OR KEY(newborn OR neonatal OR infant* OR child OR children OR kid OR kids OR pediatric* OR paediatric* OR adolescent* OR teenager* OR young* OR youth* OR juvenile)) AND ((TITLE(diphtheria)) OR ((TITLE(outcome* W/1 (assess* OR patient* OR measure*)) OR “quality of life” OR “health related quality of life” OR ((health OR functional) W/1 status) OR (QL OR QoL OR HRQL OR HRQoL)) OR KEY((outcome* W/1 (assess* OR patient* OR measure*)) OR “quality of life” OR “health related quality of life” OR ((health OR functional) W/1 status) OR (QL OR QoL OR HRQL OR HRQoL))) OR ((((status W/1 (psychological OR mental OR physical OR social OR psychosocial OR disability)) OR “activities of daily living” OR “daily activities” OR “usual activities” OR “common activities” OR wellbeing OR “well being” OR “Perceived health status”) OR KEY((status W/1 (psychological OR mental OR physical OR social OR psychosocial OR disability))) OR (((TITLE(Sickness Impact Profile) OR “attitude to health” OR “health status indicators”) OR SF-36 OR “self esteem” OR “self assessment” OR “self evaluation” OR “self report” OR “daily activities” OR “usual activities”) OR KEY (“Sickness Impact Profile” OR “attitude to health” OR “health status indicators”) OR SF-36 OR “self esteem” OR “self assessment” OR “self evaluation” OR “self report” OR “daily activities” OR “usual activities” OR “common activities” OR wellbeing OR “well being” OR “Perceived health status”))) OR (((TITLE(Survey OR Questionnaire OR Instrument)) OR KEY((symptom* OR sign*) AND (assess* OR index OR indices OR instrument* OR measur* OR scale* OR profile* OR rating* OR report* OR scale* OR schedule* OR score* OR survey* OR questionnaire*))) OR KEY(((symptom* OR sign*) AND (assess* OR index OR indices OR instrument* OR measur* OR profile* OR rating* OR report* OR scale* OR schedule* OR score* OR survey* OR questionnaire*))))).

Supplemental Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2015.11.003 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

References


