HHS Public Access

Author manuscript

Expert Rev Pharmacoecon Outcomes Res. Author manuscript; available in PMC 2016 April 01.

Published in final edited form as:

Expert Rev Pharmacoecon Outcomes Res. 2015 April ; 15(2): 267–283. doi: 10.1586/14737167.2015.1001372.

Estimates of utility weights in hemophilia: implications for cost-utility analysis of clotting factor prophylaxis

Scott D Grosse[*,1], **Shraddha S Chaugule**[2], and **Joel W Hay**[2]

[1]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, 1600 Clifton Rd. NE, Mail Stop E-64, Atlanta, GA 30333, USA

[2]Leonard Schaeffer Center for Health Policy and Economics, University of Southern California, VPD 214-L, Los Angeles, CA 90089-3333, USA

Abstract

Estimates of preference-weighted health outcomes or health state utilities are needed to assess improvements in health in terms of quality-adjusted life-years. Gains in quality-adjusted life-years are used to assess the cost–effectiveness of prophylactic use of clotting factor compared with on-demand treatment among people with hemophilia, a congenital bleeding disorder. Published estimates of health utilities for people with hemophilia vary, contributing to uncertainty in the estimates of cost–effectiveness of prophylaxis. Challenges in estimating utility weights for the purpose of evaluating hemophilia treatment include selection bias in observational data, difficulty in adjusting for predictors of health-related quality of life and lack of preference-based data comparing adults with lifetime or primary prophylaxis versus no prophylaxis living within the same country and healthcare system.

Keywords

health utilities; hemophilia; prophylaxis; quality-adjusted life-years; quality of life

Cost–effectiveness analyses assess the value of interventions, whether preventive or curative, in terms of expected changes in health and costs. If an intervention improves health outcomes and lowers total costs relative to its comparator (status quo or an alternative intervention), it is said to be dominant; otherwise the incremental cost–effectiveness ratio (ICER) is calculated as the ratio of the incremental direct cost to the incremental improvement in health [1,2]. A cost-utility analysis (CUA) assesses health gains using a quality-adjusted life-year (QALY), a preference-based summary measure that combines

[*]Author for correspondence: Tel.: +1 404 498 3074 Fax: +1 404 498 6799 sgrosse@cdc.gov.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending or royalties.

No writing assistance was utilized in the production of this manuscript.

information on morbidity, including quality of life and functional limitations, with information on premature mortality. To calculate QALYs, states of health are valued using QALY or 'utility' weights on a scale equal to 1 for perfect health and 0 for being dead.

Hemophilia A and B are hereditary bleeding disorders caused by deficiencies in clotting factors VIII and IX, respectively, and because they are inherited in an X-linked recessive pattern they mostly occur in males. Other, less common forms of hemophilia also occur. People with hemophilia can be categorized as having mild, moderate or severe disease based on the level of endogenous clotting factor activity; severe hemophilia is categorized as people with <1% clotting factor activity. People with severe hemophilia bleed longer than others after an injury and often experience internal bleeding, particularly in joints (knees, ankles and elbows). Conventional episodic or on-demand management of hemophilia involves the administration of factor concentrate to treat bleeding episodes. Routine prophylactic administration of factor concentrate can reduce the frequency of bleeds. Prophylaxis that is administered after someone with hemophilia has had associated bleeds is called secondary prophylaxis, whereas prophylaxis that is begun asymptomatically upon diagnosis in an infant or young child is called primary prophylaxis. A side effect of treatment with clotting factor is the development of an antibody inhibitor that can neutralize factor activity and make treatment ineffective. Treatments for inhibitors include immune tolerance therapy to eradicate the inhibitor and, if that is not successful, the use of bypassing agents to control bleeding.

In hemophilia management, one question has dominated the cost–effectiveness literature in recent years: is continuous prophylaxis cost-effective in comparison with episodic or on-demand treatment [3]? Six CUAs of clotting factor prophylaxis in males with hemophilia have been published. Four reported ICERs <$75,000 US dollars or <50,000 GB pounds or euros [4–7]. Two other CUAs concluded that prophylaxis is very unlikely to be cost-effective by conventional criteria, with ICERs in excess of 1 million dollars or euros [8,9]. A recent review by Miners of economic evaluations of hemophilia prophylaxis, consisting of five CEAs, five CUAs and one CBA, concluded that the diversity in estimates and conclusions reflected differences in local conditions and design features such as time horizons [3]. In addition, Miners noted that the two CUAs with the least favorable ICERs also assumed lower utility differences between prophylaxis and treatment on-demand than in three CUAs with favorable ICERs [3].

The purpose of this review is to shed light on these differences in utility estimates through a systematic review of the relevant literature on utility weights in patients with hemophilia. How sensitive are the results of CUAs of prophylaxis with clotting factor to assumptions about the incremental utility of prophylaxis relative to on-demand treatment? What are the potential biases in utility estimates that have been incorporated in CUAs of hemophilia? This paper does not ask whether prophylaxis is cost-effective, in part because the criteria for cost– effectiveness are variable and often arbitrary [10]. Some have suggested that different economic criteria be used, such as cost per bleed prevented or monetary benefit based on willingness to pay [11]. Instead, the focus of this paper is on the sources and implications of uncertainty in the utility scores used to assess cost–effectiveness in CUAs.

Overview of health utility scores

Health state utilities are preference-weighted health-related quality of life (HRQL) measures, which use assessments of the utility or value of hypothetical health states. The US expert guidance recommends using utility weights based on the preferences of the general population or decision-makers rather than the preferences of individuals affected by a condition [1]. Likewise, NICE in the UK specifies the use of population-based utility scores in estimating QALYs for the purpose of assessing the cost–effectiveness of new drugs and treatments and to inform policy decisions on coverage [12]. The argument for using community preferences is that societal resource allocations should reflect broad community values [1,13,14]. Some experts argue that patient preferences, or 'experienced utility' as opposed to ex ante 'decision utility', should inform economic evaluations of health interventions [14]. It is argued that members of the general public frequently overstate the negative attributes of living with a condition [15]. On the other hand, patient preferences often do not yield higher utility weights [13]. Further, if utility weights for a condition based on patient preferences are very high, using them will undervalue prevention or treatments [14].

The foundation of the QALY measure is expected utility theory. Expected utility is the sum of the product of expected number of life-years and the utility of time spent in each of those life-years. In theory, after adjusting for time preferences via discounting, decision-makers should be indifferent between equivalent expected utilities, for example, 10 discounted years lived in perfect health followed by death versus 20 discounted years lived in a health state with a utility weight of 0.5. Utility weights for health states can be either directly or indirectly assessed. Direct utility assessment is accomplished by asking people to rate health states relative to death using a variety of elicitation methods. Indirect utility elicitation entails asking people to assess their own health according to a questionnaire with specified domains of health and then valuing those health states through a separate mechanism in which a general population sample is asked to value hypothetical combinations of health domains (see below for details).

As seen in Table 1, three commonly used direct elicitation methods are the standard gamble (SG), time trade off (TTO) and visual analogue scale (VAS) [16,17]. The SG and TTO methods require respondents to explicitly trade off between health states and risk of death or life expectancy. SG requires respondents to choose between two profiles, one of which involves continuing to live with a suboptimal health state with a utility weight less than 1 and the other involves a hypothetical cure which will either restore one to perfect health or kill one immediately. The SG utility weight is 1 minus the probability of death at which the respondent is indifferent. TTO asks people how many years of life they would be willing to give up at the end of their lifespan to be restored to full health now. VAS is simpler to use because it only asks people to fill out a rating scale and is readily self-completed but VAS is not a choice-based measure and hence cardinal VAS preferences may be misleading [18]. Also, VAS estimates have a wider dispersion, with lower values for most conditions than is true for SG or TTO [17].

The SG yields weights that are generally higher than those that result from use of a TTO method [14,19]. This difference may reflect biases in how people respond to SG questions [20]. In addition, standard expected utility theory may not reflect how people value uncertain outcomes; rank-dependent expected utility theory and cumulative prospect theory allow for nonlinear weighting of probabilities and differential aversion for losses and gains. Evidence of risk perceptions calculated using a discrete choice experiment suggests that SG typically overstates utilities and that utility values are more consistent with a rank-dependent expected utility function [21]. Other researchers have also used different methods of pairwise comparisons based on discrete choice experiment to elicit preferences on health states [22,23].

More commonly, utility scores are assessed indirectly by asking people to fill out a questionnaire using a multi-attribute survey that describes health on multiple domains, each of which has multiple levels. The surveys are generic measures that can be used on their own to measure HRQL or in combination with preference-based valuation sets to generate health utilities. For the latter, numerical scores for each combination are assigned values that are derived from a population-based survey or exercise in which direct utility elicitation methods were used to assess trade offs among hypothetical health states. The combination of a descriptive questionnaire and the values assigned to each combination of score is known as a multi-attribute utility assessment instrument (MAUI). The choice of instrument in principle should depend on what aspects of health the potential researcher or policy body wishes to cover as well as the disease type and age group of patients being evaluated.

As seen in Table 2, six MAUIs are commonly used [24], of which three have been used in research with people with hemophilia. Two were developed as stand-alone MAUIs, the Euro-Qol EQ-5D and the Health Utilities Index (HUI), each of which has more than one version. The EQ-5D has five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression with three levels for each (now known as the EQ-5D-3L to distinguish it from a new version with five levels, the EQ-5D-5L, whose valuations are still being developed) [25]. The HUI2 [26] was developed specifically for use with children and has six dimensions (plus a fertility dimension that is usually excluded); the HUI3 is targeted for adults and adolescents. The HUI3 split the HUI2 sensory dimension into separate hearing, speech and vision dimensions [27,28].

Each MAUI has values assigned to health states based on surveys of members of the public. The HUI uses SG valuations derived from a Canadian sample taken to be representative of the general community. The SG utilities for the HUIs were derived from VAS values using a power transformation based on a subset of respondents and health states for which both SG and VAS valuations were available [18]. The use of power transformation methods has been debated [29,30]. However, the HUI3 scoring function was validated in a study which compared predicted and directly measured SG scores for 73 health states [27].

The EQ-5D has multiple sets of valuations for various countries, in addition to a VAS-based measure (EQ-5D-VAS). The EQ-5D valuations employ the TTO method, with a 10-year time horizon typically employed. However, the choice of a time horizon for TTO preference elicitation is arbitrary, and TTO values have been shown to be sensitive to the length of the

time horizon [31]. Valuations for the same EQ-5D health state using different national 'tariffs' (valuations assigned to EQ-5D combinations, the product of which yield utility weights) can differ appreciably, which has implications for cost–effectiveness analyses [32]. For example, Noyes *et al*. found the ICER for a particular drug therapy to be US$42,899 using the UK scoring function and US$108,498 using the US weights [33]. That may be a relatively extreme example, though.

Indirect preference-based measures of utility can be calculated using algorithms derived from non-preference-based HRQL instruments. To do so, a utility elicitation protocol is used with a representative sample of adults to derive valuation sets for the various health states represented by combinations of responses to the HRQL questions. For example, the Rand Short-Form HRQL instrument with either 36 (SF-36) or 12 (SF-12) questions produces separate component scores for physical (PCS) and mental (MCS) health. A group in Britain developed a preference-based MAUI with six domains, the SF-6D (Table 2) that through algorithms based on subsets of questions from the SF-36 or SF-12 instruments using values calculated through a SG exercise in a population sample can be used to calculate utility weights [34,35]. The US weights for the SF-6D have recently been developed using an innovative pairwise pivot method [23].

The different MAUIs can yield different estimates of utility weights for a given health state or disorder both as overall means and in terms of responsiveness to clinical differences. Although differences in values for health states between MAUIs often have relatively minor impacts on ICERs [29], examples of large discrepancies can be found. For example, a British study of hearing aid fitting in older adults reported that mean utilities calculated using the EQ-5D, SF-6D (both using UK weights) and the HUI3 were 0.80, 0.78 and 0.58, and the pre-post difference was 0.01, 0.01 and 0.06, respectively [36]. Unlike the EQ-5D or SF-6D, the HUI3 has a separate domain for hearing, which presumably accounts for the greater apparent sensitivity of that instrument to hearing loss. A Canadian study of adults with breast, colorectal or lung cancer reported that the difference in utility weights between advanced stage cancers and less advanced cancers was 0.01 using the EQ-5D and SF-6D and 0.05–0.06 using HUI2 and HUI3, a fivefold difference in the apparent impact of advanced cancer using HUI2 and HUI3 versus EQ-5D or SF-6D [37]. Teckle *et al*. note that the EQ-5D and SF-6D have domains of role and social function that can reflect adaptation to physical limitations, whereas the HUI instruments have domains for impairments, but do not have domains that reflect adaptation (Table 1) [37].

In addition to differences in the types of domains included, the commonly used MAUIs differ in terms of ceiling and floor effects. The EQ-5D has ceiling effects, with many values of 1.0 and no values between 0.88 and 1.0, whereas the SF-6D has values that are more tightly compressed, with a floor at 0.30 [29,38]. The HUI3 has the greatest dispersion in scores and tends to yield a larger estimate of difference in scores associated with interventions. That is particularly true for vision and hearing, which are distinct domains in the HUI3 (Table 2), but are not included in the EQ-5D or SF-6D [16].

A number of head-to-head comparisons of the EQ-5D and SF-6D have been published [29]. At the individual level, the correlation of the two measures is moderate, typically with a

correlation coefficient of about 0.7. The EQ-5D generally does a better job of discriminating among severe health states and the SF-6D is better at discriminating among people with relatively mild health states (Table 2). The two instruments often yield non-comparable mean estimates [39,40]. Consequently, there may be value in collecting both SF-6D and EQ-5D scores and using both to estimate QALY gains for CUAs.

Most MAUIs, except for the HUI2, were developed for use with adults or adolescents, although the EQ-5D has been used by some researchers with pediatric samples, and a variant, the EQ-5D-Y, was designed to be applied to children or adolescents [40]. The HUI3 can be taken directly by adolescents aged 13 years and above and can also be administered to parents of younger children [41]. The assessment of health utilities in children is a challenging field [32,42,43].

The choice of utility elicitation method to calculate utility scores may depend on the intended policy audience. Some organizations, notably NICE in the UK, use cost per QALY thresholds to make coverage decisions for prescription medications. NICE strongly recommends the use of the EQ-5D for valuing health outcomes in cost–effectiveness analyses for pharmaceuticals, whereas France appears to prefer direct utility elicitation methods and several European countries as well as Canada indicate method agnosticism. In some countries, this lack of preference may reflect lack of interest or support for QALY measures more generally [31,44]. The USA has no official guidance, but two influential panels of experts indicated a preference for use of MAUIs, although accepting direct methods such as SG and TTO [1,45].

Methods

One of the authors as a first step identified published studies that used empirical data to model the impact of prophylaxis on health utilities among adults with hemophilia. PubMed was initially searched in 2012 for articles referring to ‘cost–effectiveness’ ‘hemophilia’ and ‘prophylaxis’; a total of five cost-utility studies were identified. The search was repeated in April 2013 and one additional CUA was identified [5] as well as a review [3]; a hand search of references confirmed the exhaustiveness of the list of CUAs. Two CUAs of treatment for hemophilia inhibitors in adolescents and adults were also identified through a search of the Tufts CEA Registry. In a second step, CUAs were reviewed to identify the sources of estimates of utility weights. In addition, searches were conducted in PubMed and Google Scholar (which allows full-text searches) for studies that elicited health utilities in hemophilia patients by using the names of specific utility elicitation methods in combination with ‘hemophilia’. Those searches were repeated in February 2014 and May 2014.

Several different study designs have been used to assess health utilities among individuals with hemophilia. One approach is to ask hypothetical questions about different hemophilia treatment states among people familiar with hemophilia, with direct utility elicitation methods. Three other study designs use indirect utility elicitation and cross-sectional data. The simplest of these approaches is a comparison of mean utility weights for individuals using prophylaxis with those receiving on-demand treatment, holding constant the severity of hemophilia. However, this approach will likely understate the actual impact of

prophylaxis if people with severe hemophilia who experience bleeding events less commonly are less likely to use prophylaxis on a regular basis. Another approach presumes that the effect of prophylaxis is to turn the patient with severe hemophilia into the equivalent of someone who has moderate hemophilia and takes the difference in mean utility weights for patients with severe and mild/moderate hemophilia, few or none of whom are on prophylaxis, as equivalent to the expected impact of prophylaxis among males with severe hemophilia. A third approach is to use retrospective data to assess the cumulative lifetime use of prophylaxis as the independent predictor of utility weights.

Results

Utility weights for adults with hemophilia

As seen in Table 3, 17 sets of published estimates of utility weights for adult and adolescent males with hemophilia were identified. Two studies used direct utility elicitation with SG [46,47]. Fifteen studies used indirect elicitation, 11 of which used the EQ-5D [9,48–57]. HUI measures were used in two studies [58,59], and the SF-6D was used in another two studies [60,61].

SG estimates

A SG exercise with seven hypothetical health scenarios was conducted in Canada with 30 members of the public, 30 parents of children with hemophilia and 28 adults with hemophilia [46]. Adults with hemophilia were the least willing to accept a risk of death to avoid poorer health states, and members of the general public were most willing to accept a risk of death to achieve a hypothetical cure. Consequently, SG utility weights were higher for adults with hemophilia than in the general population and were also higher for parents of children with hemophilia. For each group, the least preferred scenario was on-demand therapy with frequent bleeding (≥3 bleeds in 3 months), with a median value of 0.825 in the general population and 0.895 among adult patients. The baseline scenario of on-demand therapy with less frequent bleeds (<3 in 3 months) had median SG weights of 0.905 and 0.940, respectively, compared with 0.950 and 0.955, respectively for highdose prophylaxis in the absence of a port [46]. In addition, the disutility of having a target joint was modeled, with the results used in a CUA [8].

A second SG study was conducted in the USA with 64 adults with hemophilia and 64 children, most of whom had parent proxy respondents (VAS scores were also assessed) [46]. The respondents chose among nine health states. The mean SG scores for mild hemophilia for the adult samples were 0.884; for moderate hemophilia, 0.868; for severe hemophilia on episodic treatment with no significant joint disease, 0.810 and for severe hemophilia on prophylaxis with no significant joint disease, 0.799. If one compares the mean scores for the two groups of hypothetical severe hemophilia patients without joint disease, episodic treatment versus prophylaxis was associated with a difference in health utility of +0.11 for the adult sample, indicating no reduction in utility with episodic treatment *per se*. Average utilities were lower for scenarios with severe joint disease, but no comparison was made for treatment types for those scenarios. In addition, presence of HIV or hepatitis C virus (HCV) infection and severe liver disease were each associated with major decrements in utility.

Having an inhibitor was associated with significantly lower scores, 0.745, equivalent to a decrement of roughly 0.06 relative to severe hemophilia without joint disease [47].

EQ-5D estimates

A cross-sectional UK sample of 249 adult males with hemophilia, none of whom received prophylaxis, were administered both the SF-36 and the EQ-5D; usable data were available for 166 subjects [43]. According to that study, 66 males with severe hemophilia had a median EQ-5D weight of 0.66, compared with a median weight of 0.85 for males with mild/moderate hemophilia. Miners *et al*. stated that neither EQ-5D nor SF-36 scores were significantly related to HIV seropositivity, history of bleeding or history of orthopedic surgery among the 66 subjects with severe hemophilia, but no results were reported. The EQ-5D score was most strongly associated with the pain component of the SF-36, followed by social functioning, general health perception and energy/vitality [48].

A smaller study of 56 patients with hemophilia in Italy, who were not on prophylaxis and did not have inhibitors, collected both EQ-5D and SF-36 information [49]. There was a marked decrement in EQ-5D scores with age: 0.81 at 15–30 years, 0.70 at 31–45 years and 0.49 at 45 years or over. HIV infection had a large, negative effect that was independent of age, a regression coefficient of –0.180 [49].

A similar study from Italy applied the EQ-5D and SF-36 to 52 males aged 15–64 years with inhibitors [50]. The mean EQ-5D utility score was 0.66 calculated using the UK weights and 0.69 using weights from Catalonia. The authors noted that this was comparable to the UK patients with severe hemophilia without inhibitors [48], but did not compare with the findings from an Italian study [49].

A multinational trial of a bypassing agent prophylaxis for a mix of pediatric and adult patients with inhibitors enrolled 37 patients and collected information prior to, during and post-prophylaxis for 22 who received prophylaxis [51]. Health utilities were elicited using EQ-5D index (referred to in the publication as TTO) scores. Mean EQ-5D scores increased from 0.56 pre-prophylaxis to 0.69 post-prophylaxis; median scores were 0.62 and 0.79, respectively.

Another trial of prophylaxis with a different bypassing agent collected EQ-5D index scores from 25 patients with inhibitors who completed the 12-month trial, 10 on prophylaxis and 15 on on-demand treatment [52]. Mean scores increased from 0.62 to 0.70 for the prophylaxis group and remained essentially unchanged for the on-demand group, decreasing slightly from 0.63 to 0.61. The difference for the prophylaxis group was considered clinically significant. Results on EQ-5D index scores from a previous 6-month trial of the same agent were less favorable [62].

Another inhibitor-specific longitudinal study enrolled 18 adult patients and 19 caregivers of pediatric patients with inhibitors to fill out the EQ-5D [53]. For 18 adult patients, mean EQ-5D index values were 0.54 on days with bleeds and 0.74 on days without bleeds, or 0.72 on average. Mean scores were higher for patients on prophylaxis, not controlling for age, but

pediatric patients were both more likely to receive prophylaxis and to have higher EQ-5D scores.

A sample of Dutch males with hemophilia, median age 62, who had undergone major joint surgery an average of 12 years earlier, responded to the EQ-5D [54]. Their mean EQ-5D score was 0.73, which compared with a Dutch norm of 0.84 for males aged 50–59 years.

Noone *et al*. conducted a telephone survey with 58 respondents with hemophilia living in four European countries: France, Ireland, Sweden and the UK [55]. The average EQ-5D weight was highest in Sweden, where prophylaxis was universal, 0.93, compared with 0.73–0.76 in the other three countries where on-demand treatment was the norm. In addition, they reported pooled data stratified by individual lifetime experience with prophylaxis: 0.88 for those on lifetime primary prophylaxis, 0.77 for those with over 50% prophylaxis, 0.79 for those with less than 50% prophylaxis and 0.72 for those who had never received prophylaxis. Mean scores were 0.88 for those on primary prophylaxis and 0.78 for those on secondary prophylaxis. No data were collected on comorbid conditions such as infection with HIV or hepatitis C [55].

Noone *et al*. conducted a similar survey with 124 respondents from Canada, France, Ireland, the Netherlands, Poland and the UK, of which responses were used for 116 subjects aged 18–35 years with severe hemophilia [56]. The use of on-demand treatment was 79% in Poland, 62% in France, 20% in Ireland, 13% in Canada and 8% in both the UK and the Netherlands. The mean health utility values for these countries were 0.629, 0.687, 0.786, 0.791, 0.768 and 0.915, respectively. Excluding 13 respondents with a history of an inhibitor who were analyzed separately, the remaining sample was split into four groups: always on-demand (n = 26), <50% of their life on prophylaxis (n = 26), ≥50% of their life on prophylaxis (n = 35) and always on prophylaxis (n = 15). Mean EQ-5D scores were 0.619 for on-demand, 0.755 for those with less than half their life on prophylaxis, 0.812 for those mostly on prophylaxis and 0.866 for those always on prophylaxis. Parallel curves were observed when the sample was stratified by age into two groups. The mean score for the inhibitor cohort was 0.798, which is similar to the average score of those who were mostly on prophylaxis [9].

A recent study from the Netherlands assessed EQ-5D scores in young adults (aged 20–33 years in 2007) with severe (n = 60) or moderate (n = 34) hemophilia seen at a hemophilia clinic in Utrecht, as well as unaffected controls (n = 105) [57]. All 60 severe patients had been on lifelong prophylaxis throughout childhood, beginning at a median age of 4.8 years, although 15% discontinued prophylaxis as adults. The authors reported that mean EQ-5D scores for patients with severe and moderate hemophilia were 0.80 and 0.92, respectively, an absolute difference of 0.12 (not statistically significant). The mean score for unaffected controls was 0.87. The authors concluded that quality of life was ‘similar’ for people with different levels of severity of hemophilia and with the general Dutch population [9].

A subsequent publication reported findings from EQ-5D data collected from males with severe hemophilia and no inhibitor [57]. The data were collected during regular outpatient visits between 2006 and 2009 from 78 Dutch adult patients and 50 Swedish patients seen at

participating hemophilia clinics, including a Utrecht clinic that was included in an earlier study [9]. All subjects with one exception in each country had been on prophylaxis since early childhood, although only 78% of the Dutch and 96% of the Swedish patients had remained on routine prophylaxis for the past 5 years. The mean EQ-5D scores, calculated using Dutch tariffs (valuation sets), were 0.88 for the Dutch sample and 0.86 for the Swedish sample, not significantly different from each other and 0.05 points lower than equivalent scores from the general male population aged 20–29 years in each country [57].

Other indirect utility elicitation estimates

A German study published in 1996 used an early version of the HUI instrument developed by Torrance to assess health utilities for 50 patients with hemophilia, 39 on episodic treatment and 11 on prophylaxis [58]. There was essentially no difference in mean utility values for the two groups, 0.60 and 0.59, respectively.

The impact of HIV status on the overall gap in utility scores is demonstrated in a Canadian study that reported HUI2 and HUI3 scores for 101 males aged 13 years or above stratified by HIV seropositivity among those with severe hemophilia. The absolute gap in scores between those with severe and mild/moderate hemophilia was 0.12–0.14, not taking HIV status into account (Table 3). Among HIV seronegative individuals with hemophilia, utility weights were lower among those with severe hemophilia by just 0.05–0.06, or by less than half as much as the overall difference [59].

A cross-national study collected SF-36 survey data from over 500 males aged 14–83 years with severe hemophilia and without an inhibitor in four European countries and used the SF-6D algorithm to generate utility scores that were included in a CUA of prophylaxis [60]. The mean predicted SF-6D utility scores for HIV-seronegative patients aged 30 years or under were 0.77 for those on prophylaxis and 0.73 for those treated on-demand, a difference of 0.04 [60]. Among subjects over age 30, mean scores did not differ significantly by treatment type. The multivariate regression used to predict scores controlled for age, disability and number of bleeds [60].

A Belgian study administered the SF-36 to 71 adult males seen at one treatment center [61]. The mean SF-6D utility score was 0.66, differing by severity: 0.63 for 44 with severe hemophilia, 0.66 for 15 with moderate hemophilia and 0.74 for 12 with mild hemophilia. The authors reported that adults with severe hemophilia were much more likely to have HIV and HCV infections than those with moderate hemophilia. Another publication from the same group of investigators reported that the decrements from the Belgian SF-6D norm for adult males were 0.197, 0.151 and 0.054, respectively [63].

Other studies have collected SF-36 or SF-12 data on patients with hemophilia [54,64–71]; those data could be used to generate SF-6D utility scores.

Expert commentary

Different approaches have been taken in the literature to estimate the disutility of on-demand treatment in severe hemophilia relative to prophylaxis. One approach entails the subtraction

of average utility weights for individuals with severe hemophilia from those with mild or moderate hemophilia on the assumption that prophylaxis will make people with severe hemophilia comparable to those with mild hemophilia [48]. Miners *et al*. reported an absolute difference of 0.19 points, equal to a relative difference of 20% in utility, between those with severe and mild/moderate hemophilia [48]. Three CUAs that used those data calculated that prophylaxis has a roughly 20% impact on QALYs and concluded that prophylaxis was cost-effective by usual criteria [4,6,7]. Another CUA of treatment for patients with inhibitors used the same utility scores [72].

The approach taken by Miners *et al*. has two limitations. First, much of the overall difference in utilities is likely to have been the result of confounding of severe hemophilia with HIV seropositivity. Barr *et al*. reported that when comparisons were restricted to HIV-seronegative individuals, the difference in utility scores for those with severe versus mild/moderate hemophilia A was only half as large as the overall difference [59]. Other investigators also reported significantly lower health utilities among patients with hemophilia with HIV infection [47,49].

Second, a Dutch study recently questioned the notion that lifelong prophylaxis can turn severe patients with hemophilia into the equivalent of moderate patients, because significant clinical differences remain [9]. Specifically, the median Haemophilia Joint Health Score measuring loss of joint function was four-times higher in the severe hemophilia group, 8 versus 2. In contrast, the authors stated that EQ-5D scores for 80 patients with severe hemophilia who had been on lifelong prophylaxis were ‘similar’ to 40 patients with moderate hemophilia, 25% of whom were on prophylaxis. By ‘similar’, the authors meant a statistically non-significant difference. In absolute terms, however, the mean EQ-5D score for patients with severe hemophilia was 0.80 and that for moderate patients with hemophilia was 0.92, a difference of 0.12. The latter difference exceeds estimates of the minimally important difference for EQ-5D scores of 0.03–0.07, which is used to assess clinical significance [73,74]. Indeed, the disutility of severe relative to moderate hemophilia in that study is close to differences observed in the pre-prophylaxis era [59].

The utility estimates in the two CUAs that estimated very high ICERs were also subject to limitations. First, Lippert *et al*. estimated the effect of prophylaxis through the comparison of mean health scores for people with hemophilia in a clinical sample pooled from four European countries who did or did not use prophylaxis [60]. The difference of 0.04 for males under age 30 is one-fifth as large as in the CUA as noted by Miners [3,6]. However, that approach ignores selection bias [5]. If individuals with few complications are less likely to initiate or persist with secondary prophylaxis, an inverse association of prophylaxis with health scores could result. Similar average utilities could mask a beneficial effect of prophylaxis on health offset by an adverse selection bias. Two studies reported that the comparison of adults on secondary prophylaxis versus on-demand treatment leads to virtually no difference in SF-12 scores [64,65]. Two other studies reported very small, statistically insignificant differences in mean SF-6D or HUI scores by prophylaxis versus on-demand therapy [58,60].

Second, the CUA by Risebrough *et al*. [8] used the SG results from Naraine *et al*. [46] using preferences from a general population sample with a direct effect of prophylaxis on utility scores of 0.045 and an indirect effect of target joint bleeds on utility scores of –0.08 [8]. It should be noted that the direct effect of prophylaxis on utility would have been one-third as large, 0.015, if adult patient preferences had been used, a phenomenon that has been noted previously [14]. Furthermore, SG results may overstate utility scores and understate the disutility of worse states of health [14]. On the other hand, Risebrough *et al*. modeled the effect of prophylaxis on the risk of developing a target joint in addition to a direct effect on health utility [8]. It appears that the difference in the expected number of target joints through age 6 between on-demand and primary prophylaxis in that model was 0.775, which implies a total utility gain from primary prophylaxis relative to on-demand treatment of 0.107 per patient-year.

A recent CUA by Farrugia *et al*. [5] used estimates from a cross-national study that compared EQ-5D utility scores among persons with severe hemophilia living in six countries with different patterns of prophylaxis use to model the cost-utility of prophylaxis in the USA and UK [56]. Specifically, it was assumed that the mean utility scores under prophylaxis and on-demand treatment for a 20-year-old male would be roughly 0.886 and 0.632, respectively, a difference of 0.25 [4]. Noone *et al*. reported that the mean utility score for individuals who reported always having been on prophylaxis was 0.866 and that of people always on-demand was 0.619 [56]. The comparison by Noone *et al*. of utility scores for reported history of primary prophylaxis, secondary prophylaxis and on-demand treatment reflects an important advance over previous studies, which compared current use of prophylaxis with on-demand therapy. The effect of secondary prophylaxis among people with established joint disease is likely to be very small [5,59]. One US study reported that six adults who had always used prophylaxis had higher scores on SF-12 dimensions than those who either used secondary prophylaxis (n = 26) or no prophylaxis (n = 32); the difference in one dimension, physical functioning, was statistically significant [64].

However, the mean scores by treatment type (always on-demand vs always on prophylaxis) in the study by Noone *et al*. appear to largely reflect between-country differences in mean utility scores, ranging from 0.629 in Poland to 0.915 in the Netherlands [56]. Consequently, the putative effect of prophylaxis on utility scores in that study and in the CUA by Farrugia *et al*. is confounded by cross-national differences in patterns of care, prevalence of chronic viral infections and socioeconomic conditions. It is implausible that the difference in mean utility scores between persons with severe hemophilia in Poland and the Netherlands could be due almost solely to differences in use of prophylaxis. Noone *et al*. could have estimated the association of treatment type with utility values within countries using a fixed effects (dummy variable) model.

Also, Noone *et al*. did not elicit HIV status from survey respondents, which has been shown to be an important con-founder in previous studies. Farrugia *et al*. [5] argue that the mean age of the sample, 18–35 at the time of the survey [56], was lower than the time elapsed since HIV infections occurred through contaminated blood products. However, the older half of the sample could have been subject to bias from unmeasured confounding with HIV

seropositivity. That issue could be resolved through an analysis restricted to the younger half of the sample.

The cross-national studies by Noone *et al*. were based on convenience samples of persons with hemophilia and hence by definition are of uncertain representativeness. For The Netherlands, the mean EQ-5D score for self-selected Dutch young adult patients with severe hemophilia of 0.915 [56] is higher than that reported in two published reports from the Netherlands with representative samples of the treated population at a single center, 0.80 [9] and 0.88 [57].

The mean difference in utility scores between the prophylaxis and on-demand treatment arms of the three CUAs that reported favorable ICERs ranged from 0.21 [4,6,7] to 0.25–0.27 at ages 0–20 years in the USA and UK and 0.33 in Sweden [5]. In contrast, the two CUAs with unfavorable ICERs assumed much lower differences in scores, 0.04 in one [60] and approximately 0.11 in the other [6]. The latter estimate is similar to the mean difference between patients with mild/moderate and severe hemophilia of roughly 0.12 reported in three studies [9,59,61].

None of the CUAs explicitly modeled the *difference* in utility associated with prophylaxis versus on-demand treatment. Each of the four analyses that reported relatively low ICERs included a probabilistic sensitivity analysis, but they modeled each utility score independently. For example, the analysis of Farrugia *et al*. preserved a difference of 0.25 in utility scores using either the lower or upper bound estimates for both scores [5]. In addition, a one-way sensitivity analysis in the same study varied each utility score by just 0.05, that is, a range of difference in utility scores from 0.20 to 0.30.

The choice of MAUI can be influential in estimates of utility scores in hemophilia. In particular, SF-6D utility scores in hemophilia appear to be lower than EQ-5D scores, although no direct comparisons have been published to date. The mean SF-6D score for patients with severe hemophilia in four European countries (Germany, Sweden, the UK and The Netherlands) was reported to be 0.70 [60], and for a sample of adults in Belgium it was 0.63 [61]. Mean EQ-5D scores for males with severe hemophilia in Sweden and the Netherlands appear to be about 0.86–0.88 [57]. Mean EQ-5D scores for males with severe hemophilia in the UK are about 0.75–0.78 [55,56]. Through the Universal Data Collection (UDC) system, the US hemophilia treatment centers collected patient-reported EQ-5D and SF-12 data from thousands of individuals with hemophilia and other bleeding disorders between 2005 and 2011; EQ-5D data were previously used to estimate disability-adjusted life-years [75]. A head-to-head comparison of EQ-5D and SF-6D scores for more than 1800 males with severe hemophilia A using these data is currently in preparation.

Generic MAUIs do not necessarily fully capture the adverse health effects of severe hemophilia. A recent study by den Uijl *et al*. [9] found a statistically insignificant difference in mean utility scores between individuals with severe and mild/moderate hemophilia. In contrast, that study found significant differences using the Haemophilia Joint Health Score instrument, which appears to be a more sensitive outcome measure in hemophilia. In studies with relatively small numbers of subjects, which is typical of rare disorders such as

hemophilia, it is challenging to find statistically significant differences in utility scores even if the differences are of clinically significant magnitude as was the case in the Dutch study [9].

Non-preference-based HRQL measures have also been used in studies of the benefits of secondary prophylaxis. Buchbinder and Ragni [76] identified nine studies of adult patients with hemophilia that compared measures of HRQL by prophylaxis status, of which six were said to report better HRQL among adult patients treated with prophylaxis compared with on-demand therapy. However, it is not clear how many differences were statistically significant or clinically meaningful. An observational retrospective study from Italy with data on 37 individuals reported that all EQ-5D dimensions were higher with prophylaxis; some (pain, mobility, usual activities) were statistically significant at the 0.01 level [77]. Pollmann *et al*. reported findings from 152 German adult subjects enrolled in the PASS study, of whom 103 completed the SF-36 questionnaire [78]. The authors reported no differences by treatment type (on-demand vs continuous prophylaxis) in the MCS of the SF-36 and slight, not statistically significant differences in the physical health component (PCS). In contrast, both the presence of a ‘target joint’, a joint where bleeds frequently occur, and HIV infection were associated with statistically significant and clinically meaningful declines in the PCS; hepatitis C infection was not associated. The association of PCS with target joints was particularly strong [78].

Although the focus of the hemophilia literature on HRQL has been on joint bleeds and arthropathy, an important influence on the SF-36 PCS measure among people with hemophilia is pain [69]. Pain is included as a single attribute or dimension in HUI2, HUI3 and EQ-5D. Among patients with hemophilia, self-reports of pain are associated with decreased EQ-5D scores [49]. Similarly, pain is one of two HUI3 attributes (along with ambulation) that significantly predicts overall HUI3 scores among individuals with hemophilia [59]. More attention should be paid to the role of pain management and self-management strategies in improving outcomes for people with hemophilia.

Finally, there is potential value in the development of condition-specific preference-based measures (CSPBMs) for hemophilia. Disease-specific HRQL instruments are typically reported to be more responsive or sensitive in detecting relatively small but statistically significant and clinically important differences in functioning [79]. For example, a cancer-specific preference-based instrument was able to detect significant differences that could not be picked up using either the EQ-5D or SF-6D, although HUI2 and HUI3 were found to be comparably sensitive to the cancer-specific instrument [37]. A CSPBM for rheumatoid arthritis was found to be only slightly superior to generic measures (HUI2, HUI3, EQ-5D and SF-6D), all of which were able to significantly discriminate differences in disease severity [80]. Despite the lack of clear evidence of utility of CSPBMs over generic measures, development of new measures continues to occur, motivated in large part by concerns over the insensitivity of generic measures [81,82]. An alternative is to abandon CUA and QALYs in favor of other metrics [11].

CSPBMs can be estimated in two ways, through statistical mapping in datasets where patients fill out both disease-specific instruments and generic preference-based measures or

through direct valuation in which people undergo direct utility elicitation for health states included in disease-specific instruments [83]. Utility weights for disease-specific measures derived from direct valuation are more likely to yield useful results than through statistical mapping algorithms, in part because the mapping algorithm is limited to the information content of the generic instrument such as the EQ-5D to which it is mapped [83]. For example, a preference-based instrument for osteoarthritis based on direct valuation was found to yield significant differences in utility weights for patients who had experienced significant improvement or deterioration in clinical measures whereas the HUI3 measure was not responsive [84].

As seen in Table 4, researchers from Germany developed hemophilia-specific HRQL questionnaires for children and adolescents [85] (Haemo-QoL) and for adults (Haem-A-QoL) [86]. Culturally adapted and linguistically validated translations were developed so that these instruments could be used internationally [86]. Another disease-specific instrument, the Haemo-QoL-A, has been shown to discriminate significantly between adults with hemophilia by severity and HIV status, and the physical functioning subscale to discriminate between patients receiving prophylactic or on-demand therapy [87]. CHO-KLAT is a disease-specific instrument developed by Canadian researchers with a focus on children's perspectives [88]. CHO-KLAT was found to be a reliable and valid measure of quality of life for boys with hemophilia [89]. In the future, validated utility weights could potentially be derived from disease-specific instruments by measuring the impact on specific functionalities affected by the disease.

Five-year view

Future studies of HRQL in hemophilia will focus on using patient registries for longitudinal assessments and on having a robust study design with statistical methods that control for confounding by comorbid conditions (e.g., HIV or HCV infection). In addition, it is important for researchers to model the effects of prophylaxis on various complications of hemophilia as well as the disutility associated with those complications. For example, one SG study suggested that prophylaxis has little direct benefit among patients who do not have joint disease, but joint disease itself has a large negative effect on utility [47]. Since primary prophylaxis is known to reduce the frequency of joint disease, the total effect of prophylaxis requires modeling the probabilities of progressing to joint disease for the two treatment groups [90].

Study designs will allow for an increased understanding of the impact of timing of the start of prophylaxis, schedule of prophylaxis, severity of the disease and type of treatment on HRQL. These developments could help researchers quantify the degree to which frequent, intravenous infusions given several times weekly during prophylaxis influence HRQL. Finally, this will help in facilitating an increased understanding of the role of pain management and self-management strategies in improving health outcomes for people with hemophilia.

It also is anticipated that disease-specific measures will be used more frequently for assessing HRQL in children and adults with hemophilia motivated by the relatively small

differences in utility scores found using generic measures such as the EQ-5D. This could result in increased interest in preference-based scoring algorithms that could be used with hemophilia-specific HRQL instruments. However, mapping algorithms based on regression analyses that can be used to impute EQ-5D utilities from non-preference-based HRQL measures may not necessarily be reliable at the individual level, as shown by Blome *et al*. in the case of psoriasis [91]. However, empirical evidence comparing CSPBMs and generic instruments is mixed [82]. If such hypothetical CSPBMs do show greater responsiveness, which is uncertain, their use could potentially make interventions appear cost-effective relative to other interventions valued using generic measures more than could be justified. In order to inform prioritization decisions, generic instruments such as the EQ-5D, HUI3 or SF-6D will continue to be used [92].

CUAs of hemophilia prophylaxis will increasingly follow cost–effectiveness analysis guidelines and include utility scores in sensitivity analyses. Because even probabilistic sensitivity analyses often do not adequately capture uncertainty [93], it is important for sensitivity or scenario analyses to reflect the full range of variation in the published literature of *differences* in utility scores [32]. In addition to hemophilia prophylaxis with regular clotting factor, CUAs will continue to be used to compare management options for individuals with inhibitors, including prophylaxis with bypassing agents [94–96].

References

Papers of special note have been highlighted as:

•• of considerable interest

1. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. 1996USAOxford University Press; NY •• Gold standard for understanding cost–effectiveness analysis.
2. Haddix, AC.; Teutsch, SM.; Corso, PS., editors. Prevention effectiveness: a guide to decision analysis and economic evaluation. 2nd edition. NY, USA: Oxford University Press; 2003.
3. Miners AH. Economic evaluations of prophylaxis with clotting factor for people with severe haemophilia: why do the results vary so much? Haemophilia. 2013; 19:174–180. [PubMed: 22989090]
4. Colombo GL, Di Matteo S, Mancuso ME, Santagostino E. Cost-utility analysis of prophylaxis versus treatment on demand in severe hemophilia A. Clinicoecon Outcomes Res. 2011; 3:55–61. [PubMed: 21935333]
5. Farrugia A, Cassar J, Kimber MC, et al. Treatment for life for severe haemophilia A- A cost-utility model for prophylaxis vs. on-demand treatment. Haemophilia. 2013; 19:e228–e238. [PubMed: 23534877]
6. Miners A. Revisiting the cost-effectiveness of primary prophylaxis with clotting factor for the treatment of severe haemophilia. Haemophilia. 2009; 15:881–887. [PubMed: 19473422]
7. Miners AH, Sabin CA, Tolley KH, et al. A cost-utility analysis of primary prophylaxis versus treatment on-demand for individuals with severe haemophilia. Pharmacoeconomics. 2002; 20:759–774. [PubMed: 12201795] •• Landmark example of cost-utility analysis of primary prophylaxis versus on-demand treatment for individuals with severe hemophilia.
8. Risebrough N, Oh P, Blanchette V, et al. Cost-utility analysis of Canadian tailored prophylaxis, primary prophylaxis and on-demand therapy in young children with severe haemophilia A. Haemophilia. 2008; 14:743–752. [PubMed: 18422610]
9. den Uijl I, Biesma D, Grobbee D, Fischer K. Turning severe into moderate haemophilia by prophylaxis: are we reaching our goal? Blood Transfus. 2013; 11:364–369. [PubMed: 23149144]

10. Grosse SD. Assessing cost effectiveness in health care: the history of the $50,000 per QALY threshold. Expert Rev Pharmacoecon Outcomes Res. 2008; 8:165–178. [PubMed: 20528406]
11. Farrugia A, O'Mahony B, Cassar J. Health technology assessment and haemophilia. Haemophilia. 2012; 18:152–157. [PubMed: 22335450]
12. Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. BMJ. 2004; 329:224–227. [PubMed: 15271836]
13. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. Annu Rev Public Health. 2000; 21:587–611. [PubMed: 10884966]
14. Nord E, Daniels N, Kamlet M. QALYs: some challenges. Value Health. 2009; 12:S10–S15. [PubMed: 19250125]
15. Ubel PA, Loewenstein G, Schwarz N, Smith D. Misimagining the unimaginable: the disability paradox and health care decision making. Health Psychol. 2005; 24:S57–S62. [PubMed: 16045420]
16. Bullinger M, Globe D, Wasserman J, et al. Challenges of patient-reported outcome assessment in hemophilia care - a state of the art review. Value Health. 2009; 12:808–820. [PubMed: 19490552]
17. Revicki DA, Lenderking WR. Methods and issues associated with the use of quality-adjusted life-years. Expert Rev Pharmacoecon Outcomes Res. 2012; 12:105–114. [PubMed: 22280200]
18. Torrance GW, Feeny D, Furlong W. Visual analog scales: do they have a role in the measurement of preferences for health states? Med Decis Making. 2001; 21:329–334. [PubMed: 11475389]
19. Carroll AE, Downs SM. Improving decision analyses: parent preferences (utility values) for pediatric health outcomes. J Pediatr. 2009; 155:21–25. [PubMed: 19394030]
20. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health Econ. 2002; 11:447–456. [PubMed: 12112493]
21. van Houtven G, Johnson FR, Kilambi V, Hauber AB. Eliciting benefit-risk preferences and probability-weighted utility using choice-format conjoint analysis. Med Decis Making. 2011; 31:469–480. [PubMed: 21310854]
22. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian valuation of EQ-5D health states: preliminary value set and considerations for future valuation studies. PLoS One. 2012; 7(2):e31115. [PubMed: 22328929]
23. Craig BM, Pickard AS, Stolk E, Brazier JE. US valuation of the SF-6D. Med Decis Making. 2013; 33:793–803. [PubMed: 23629865]
24. Richardson J, McKie J, Bariola E. Review and critique of health related multi attribute utility instruments. Centre for Health Economics, Monash University. Research Paper. 2011:64.Available from: www.buseco.monash.edu.au/centres/che/pubs/researchpaper64.pdf • Comprehensive overview of all six multi-attribute utility assessment instruments.
25. Devlin NJ, Krabbe PF. The development of new research methods for the valuation of EQ-5D-5L. Eur J Health Econ. 2013; 14(Suppl 1):1–3. [PubMed: 22699457]
26. Torrance GW, Feeny DH, Furlong WJ, et al. A multi-attribute utility function for a comprehensive health status classification system. Health Utilities Mark 2. Med Care. 1996; 34:702–722. [PubMed: 8676608]
27. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the Health Utilities Index Mark 3 system. Med Care. 2002; 40:113–128. [PubMed: 11802084]
28. Feeny D, Furlong W, Saigal S, et al. Comparing directly measured standard gamble scores to HUI2 and HUI3 utility scores: group- and individual-level comparisons. Soc Sci Med. 2004; 58:799–809. [PubMed: 14672594]
29. Stevens KJ, McCabe CJ, Brazier JE. Mapping between visual analogue and standard gamble data; results from the UK Health Utilities Index 2 valuation. Health Econ. 2005; 15:527–534. [PubMed: 16389651]
30. McDonough CM, Tosteson AN. Measuring preferences for cost-utility analysis: how choice of method may influence decision-making. Pharmacoeconomics. 2007; 25:93–106. [PubMed: 17249853]
31. Feeny D. Standardization and regulatory guidelines may inhibit science and reduce the usefulness of analyses based on the application of preference-based measures for policy decisions. Med Decis Making. 2013; 33:316–329. [PubMed: 23184461]

32. Grosse SD, Prosser LA, Asakawa K, Feeny D. QALY weights for neurosensory impairments in pediatric economic evaluations: case studies and a critique. Expert Rev Pharmacoecon Outcomes Res. 2010; 10:293–308. [PubMed: 20545594]
33. Noyes K, Dick AW, Holloway RG. The implications of using US-specific EQ-5D preference weights for cost-effectiveness evaluation. Med Decis Making. 2007; 27:327–334. [PubMed: 17502449]
34. Brazier J, Usherwood T, Harper R, et al. Deriving a preference-based single index from the UK SF-36 Health Survey. J Clin Epidemiol. 1998; 51:1115–1128. [PubMed: 9817129]
35. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care. 2004; 42:851–859. [PubMed: 15319610]
36. Barton GR, Bankart J, Davis AC, Summerfield QA. Comparing utility scores before and after hearing-aid provision: results according to the EQ-5D, HUI3 and SF-6D. Appl Health Econ Health Policy. 2004; 3:103–105. [PubMed: 15702947]
37. Teckle P, Peacock S, McTaggart-Cowan H, et al. The ability of cancer-specific and generic preference-based instruments to discriminate across clinical and self-reported measures of cancer severities. Health Qual Life Outcomes. 2011; 9:106. [PubMed: 22123196]
38. Fryback DG, Palta M, Cherepanov D, et al. Comparison of 5 health-related quality-of-life indexes using item response theory analysis. Med Decis Making. 2010; 30:5–15. [PubMed: 19843961]
39. Whitehurst DG, Bryan S, Lewis M. Systematic review and empirical comparison of contemporaneous EQ-5D and SF-6D group mean scores. Med Decis Making. 2011; 31:E34–E44. [PubMed: 21993430]
40. Gaujoux-Viala C, Rat AC, Guillemin F, et al. Responsiveness of EQ-5D and SF-6D in patients with early arthritis: results from the ESPOIR cohort. Ann Rheum Dis. 2012; 71:1478–1483. [PubMed: 22419775]
41. Noyes J, Edwards RT. EQ-5D for the assessment of health-related quality of life and resource allocation in children: a systematic methodological review. Value Health. 2011; 14:1117–1129. [PubMed: 22152182]
42. Ungar WJ. Challenges in health state valuation in paediatric economic evaluation: are QALYs contraindicated? Pharmacoeconomics. 2011; 29:641–652. [PubMed: 21604821]
43. Payakachat N, Tilford JM, Kovacs E, Kuhlthau K. Autism spectrum disorders: a review of measures for clinical, health services and cost-effectiveness applications. Expert Rev Pharmacoecon Outcomes Res. 2012; 12:485–503. [PubMed: 22971035]
44. Prosser LA, Grosse SD, Wittenberg E. Health utility elicitation: is there still a role for direct methods? Pharmacoeconomics. 2012; 30:83–86. [PubMed: 22185185]
45. Miller, W.; Robinson, LA.; Lawrence, RS., editors. Valuing health for regulatory cost-effectiveness analysis. Institute of Medicine. Washington, DC, USA: National Academies Press; 2006.
46. Naraine VS, Risebrough NA, Oh P, et al. Health-related quality-of-life treatments for severe haemophilia: utility measurements using the Standard Gamble technique. Haemophilia. 2002; 8:112–120. [PubMed: 11952846] • Used direct utility elicitation method for obtaining utility weights for males with hemophilia.
47. Wasserman J, Aday LA, Begley CE, et al. Measuring health state preferences for hemophilia: development of a disease-specific utility instrument. Haemophilia. 2005; 11:49–57. [PubMed: 15660989] • Used direct utility elicitation method for obtaining utility weights for males with hemophilia.
48. Miners AH, Sabin CA, Tolley KH, et al. Assessing health-related quality-of-life in individuals with haemophilia. Haemophilia. 1999; 5:378–385. [PubMed: 10583523]
49. Trippoli S, Vaiani M, Linari S, et al. Multivariate analysis of factors influencing quality of life and utility in patients with haemophilia. Haematologica. 2001; 86:722–728. [PubMed: 11454527]
50. Gringeri A, Mantovani LG, Scalone L, Mannucci PM. Cost of care and quality of life for patients with hemophilia complicated by inhibitors: the COCIS Study Group. Blood. 2003; 102:2356–2363.

51. Hoots WK, Ebbesen LS, Konkle BA, et al. Secondary prophylaxis with recombinant activated factor VII improves health-related quality of life of haemophilia patients with inhibitors. Haemophilia. 2008; 14:466–475. [PubMed: 18282155]
52. Stasyshyn O, Antunes S, Mamonov V, et al. Prophylaxis with anti-inhibitor coagulant complex improves health-related quality of life in haemophilia patients with inhibitors: results from FEIBA NF Prophylaxis Study. Haemophilia. 2014; 20:644–650. [PubMed: 24589084]
53. Neufeld EJ, Recht M, Sabio H, et al. Effect of acute bleeding on daily quality of life assessments in patients with congenital hemophilia with inhibitors and their families: observations from the dosing observational study in hemophilia. Value Health. 2012; 15:916–925. [PubMed: 22999142]
54. de Kleijn P, Sluiter D, Vogely HCh, et al. Long-term outcome of multiple joint procedures in haemophilia. Haemophilia. 2014; 20:276–281. [PubMed: 24533953]
55. Noone D, O'Mahony B, Prihodova L. A survey of the outcome of prophylaxis, on-demand or combined treatment in 20–35 year old men with severe haemophilia in four European countries. Haemophilia. 2011; 17:e831–e848. [PubMed: 21507150] • Multi-country survey of hemophilia patients stratified by life-time experience with prophylaxis.
56. Noone D, O'Mahony B, van Dijk JP, Prihodova L. A survey of the outcome of prophylaxis, on-demand treatment or combined treatment in 18–35-year old men with severe haemophilia in six countries. Haemophilia. 2013; 19:44–50. [PubMed: 22913831] • Multi-country survey of hemophilia patients stratified by life-time experience with prophylaxis.
57. Fischer K, Steen Carlsson K, Petrini P, et al. Intermediate-dose versus high-dose prophylaxis for severe hemophilia: comparing outcome and costs since the 1970s. Blood. 2013; 122:1129–1136. [PubMed: 23777770]
58. Szucs TD, Offner A, Schramm W. Socioeconomic impact of haemophilia care: results of a pilot study. Haemophilia. 1996; 2:211–221.
59. Barr RD, Saleh M, Furlong W, et al. Health status and health-related quality of life associated with hemophilia. Am J Hematol. 2002; 71:152–160. [PubMed: 12410568]
60. Lippert B, Berger K, Berntorp E, et al. Cost effectiveness of haemophilia treatment: a cross-national assessment. Blood Coagul Fibrinolysis. 2005; 16:477–485. [PubMed: 16175006] • Directly compared utility scores for patients on episodic treatment and prophylaxis.
61. Carvalhosa AM, Henrard S, Lambert C, Hermans C. Physical and mental quality of life in adult patients with haemophilia in Belgium: the impact of financial issues. Haemophilia. 2014; 20:479–485. [PubMed: 24354566]
62. Gringeri A, Leissinger C, Cortesi PA, et al. Health-related quality of life in patients with haemophilia and inhibitors on prophylaxis with anti-inhibitor complex concentrate: results from the Pro-FEIBA study. Haemophilia. 2013; 19:736–743. [PubMed: 23731246]
63. Henrard S, Devleesschauwer B, Beutels P, et al. The health and economic burden of haemophilia in Belgium: a rare, expensive and challenging disease. Orphanet J Rare Dis. 2014; 9:39. [PubMed: 24655371]
64. Duncan N, Shapiro A, Ye X, et al. Treatment patterns, health-related quality of life and adherence to prophylaxis among haemophilia A patients in the United States. Haemophilia. 2012; 18:760–765. [PubMed: 22548353]
65. Poon JL, Zhou ZY, Doctor JN, et al. Quality of life in haemophilia A. Hemophilia Utilization Group Study Va (HUGS- Va). Haemophilia. 2012; 18:699–707. [PubMed: 22507546]
66. Royal S, Schramm W, Berntorp E, et al. Quality-of-life differences between prophylactic and on-demand factor replacement therapy in European haemophilia patients. Haemophilia. 2002; 8:44–50. [PubMed: 11886464]
67. Szende A, Schramm W, Flood E, et al. Health-related quality of life assessment in adult haemophilia patients: a systematic review and evaluation of instruments. Haemophilia. 2003; 9:678–687. [PubMed: 14750933]
68. Posthouwer D, Plug I, van der Bom JG, et al. Hepatitis C and health-related quality of life among patients with hemophilia. Haematologica. 2005; 90:846–850. [PubMed: 15951299]
69. Elander J, Robinson G, Mitchell K, Morris J. An assessment of the relative influence of pain coping, negative thoughts about pain, and pain acceptance on health-related quality of life among people with hemophilia. Pain. 2009; 145:169–175. [PubMed: 19573989]

70. Molho P, Rolland N, Lebrun T, et al. Epidemiological survey of the orthopaedic status of severe haemophilia A and B patients in France. Haemophilia. 2000; 6:23–32. [PubMed: 10632738]
71. Lindvall K, Astermark J, Bjorkman S, et al. Daily dosing prophylaxis for haemophilia: a randomized crossover pilot study evaluating feasibility and efficacy. Haemophilia. 2012; 18:855–859. [PubMed: 22681244]
72. Knight C, Paisley S, Wight J, Lloyd-Jones M. Economic modelling of different treatment strategies for haemophilia A with high-responding inhibitors. Haemophilia. 2003; 9:521–540. [PubMed: 12828681]
73. Lubetkin E, Jia H, Franks P, Gold M. Relationship among sociodemographic factors, clinical conditions health-related quality of life: examining the EQ-5D in the U.S. general population. Qual Life Res. 2005; 14:2187–2196. [PubMed: 16328899]
74. 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; 4:1523–2532. [PubMed: 16110932]
75. Siddiqi AE, Ebrahim SH, Soucie JM, et al. Burden of disease resulting from hemophilia in the U.S. Am J Prev Med. 2010; 38:S482–S488. [PubMed: 20331947]
76. Buchbinder D, Ragni V. What is the role of prophylaxis in the improvement of health-related quality of life of patients with hemophilia? Hematology Am Soc Hematol Educ Program. 2013; 2013:52–55. [PubMed: 24319162] • Comprehensive evidence-based mini-review discussing the role of prophylaxis in improvement of health-related quality of life.
77. Tagliaferri A, Franchini M, Coppola A, et al. Effects of secondary prophylaxis started in adolescent and adult haemophiliacs. Haemophilia. 2008; 14:945–951. [PubMed: 18540895]
78. Pollmann H, Klamroth R, Vidovic N, et al. Prophylaxis and quality of life in patients with hemophilia A during routine treatment with ADVATE [antihemophilic factor (recombinant), plasma/albumin-free method] in Germany: a subgroup analysis of the ADVATE PASS post-approval, non-interventional study. Ann Hematol. 2013; 92:689–698. [PubMed: 23324986]
79. Wiebe S, Guyatt G, Weaver B, et al. Comparative responsiveness of generic and specific quality-of-life instruments. J Clin Epidemiol. 2003; 56:52–60. [PubMed: 12589870]
80. Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med. 2005; 60:1571–1582. [PubMed: 15652688]
81. Lloyd A, Kerr C, Breheny K, et al. Economic evaluation in short bowel syndrome (SBS): an algorithm to estimate utility scores for a patient-reported SBS-specific quality of life scale (SBS-QoL™). Qual Life Res. 2014; 23:449–458. [PubMed: 24026630]
82. Versteegh MM, Leunis A, Uyl-de Groot CA, Stolk EA. Condition-specific preference-based measures: benefit or burden? Value Health. 2012; 15:504–513. [PubMed: 22583461]
83. Lin FJ, Longworth L, Pickard AS. Evaluation of content on EQ-5D as compared to disease-specific utility measures. Qual Life Res. 2013; 22:853–874. [PubMed: 22729670]
84. Davis EM, Lynd LD, Grubisic M, et al. Responsiveness of health state utility values in knee osteoarthritis. J Rheumatol. 2013; 40:2075–2082. [PubMed: 24187098]
85. von Mackensen S, Bullinger M. Haemo-QoL Group. Development and testing of an instrument to assess the Quality of Life of Children with Haemophilia in Europe (Haemo-QoL). Haemophilia. 2004; 1:17–25. [PubMed: 14987245]
86. von Mackensen S, Campos IG, Acquadro C, Strandberg-Larsen M. Cross-cultural adaptation and linguistic validation of age-group-specific haemophilia patient-reported outcome (PRO) instruments for patients and parents. Haemophilia. 2013; 19:e73–e83. [PubMed: 23167873]
87. Rentz A, Flood E, Altisent C, et al. Cross-cultural development and psychometric evaluation of a patient-reported health-related quality of life questionnaire for adults with haemophilia. Haemophilia. 2008; 14:1023–1034. [PubMed: 18665853]
88. Young NL, Bradley CS, Blanchette V, et al. Development of a health-related quality of life measure for boys with haemophilia: the Canadian Haemophilia Outcomes-Kids Life Assessment Tool (CHO-KLAT). Haemophilia. 2004; 10:34–43. [PubMed: 14987247]
89. Young NL, Bradley CS, Wakefield CD, et al. How well does the Canadian Haemophilia Outcomes-Kids' Life Assessment Tool (CHO-KLAT) measure the quality of life of boys with haemophilia? Pediatr Blood Cancer. 2006; 47:305–311. [PubMed: 16206209]

90. Fischer K, Pouw ME, Lewandowski D, et al. A modeling approach to evaluate long-term outcome of prophylactic and on demand treatment strategies for severe hemophilia A. Haematologica. 2011; 96:738–743. [PubMed: 21273268]

91. Blome C, Beikert FC, Rustenbach SJ, Augustin M. Mapping DLQI on EQ-5D in psoriasis: transformation of skin-specific health-related quality of life into utilities. Arch Dermatol Res. 2013; 305:197–204. [PubMed: 23266843]

92. Mulhern B, Rowen D, Brazier J, et al. Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. Health Technol Assess. 2013; 17(5):1–140.

93. Groot Koerkamp B, Hunink MG, Stijnen T, et al. Limitations of acceptability curves for presenting uncertainty in cost-effectiveness analysis. Med Decis Making. 2007; 27:101–111. [PubMed: 17409361]

94. Hay JW, Zhou ZY. Systematic literature review of economics analysis on treatment of mild-to-moderate bleeds with aPCC versus rFVIIa. J Med Econ. 2011; 14:516–525. [PubMed: 21699366]

95. Rasekh HR, Imani A, Karimi M, Golestani M. Cost-utility analysis of immune tolerance induction therapy versus on-demand treatment with recombinant factor VII for hemophilia A with high titer inhibitors in Iran. Clinicoecon Outcomes Res. 2011; 3:207–212. [PubMed: 22163168]

96. Konkle BA, Ebbesen LS, Erhardtsen E, et al. Randomized, prospective clinical trial of recombinant factor VIIa for secondary prophylaxis in hemophilia patients with inhibitors. J Thromb Haemost. 2007; 5:1904–1913. [PubMed: 17723130]

Key issues

- The hemophilia cost–effectiveness literature has struggled to assess the value of continuous prophylaxis versus on-demand or episodic treatment.
- Accurate estimates of quality-adjusted life-years associated with different hemophilia management strategies, including different prophylaxis protocols, are needed to assess cost-utility.
- Although indirect generic preference-based instruments such as EQ-5D, SF-6D, Health Utilities Index-2 and Health Utilities Index-3 allow researchers to estimate utility weights, studies using different instruments often find substantially different estimates and it is not clear which ones are most useful.
- Studies have reported small cross-sectional differences in utility scores by hemophilia severity, particularly if one controls for confounding by HIV infection.
- Males in countries where primary prophylaxis has been long promoted have significantly better utility scores than males in countries with less use of prophylaxis.
- The timing of the start of prophylaxis, schedule of prophylaxis, severity of the disease, type of treatment and presence of complications such as joint disease or an inhibitor may play an important role in health-related quality of life.
- It could be helpful to estimate the association of treatment type with utility values using a fixed effects (dummy variable) model to control for country of residence because currently, there is no published evidence on difference in utility weights in severe hemophilia by timing or duration of use of prophylaxis within national populations.
- Future research should focus on longitudinal assessments using registry data and on including controls for confounding by co-morbid conditions (e.g., HIV or hepatitis C virus infection).
- Disease-specific health-related quality of life instruments may be more sensitive than generic instrument in detecting relatively small but clinically important differences in outcomes while measuring a narrow aspect of functioning related to a disease. The development of mapping algorithms to derive utility measures from condition-specific instruments require vetting to demonstrate reliability and value.

Table 1

Common direct elicitation methods (SG vs TTO vs VAS).

Technique	Type	Description	Strength	Weakness
SG	Choice based	Respondents explicitly trade off between health states and risk of death or life	• Rooted in rigorous theoretical foundation of expected utility theory (dominant theory of decision making) • Favored by many health economists	• Cognitively challenging • Valuations can be seen as containing biases (source of bias: probability weighting and loss aversion) • SG yields weights that are generally higher than those that result from use of a TTO method
TTO	Choice based	Respondents are asked the number of years of life they would be willing to give up at the end of lifespan to be restored to full health now	• Favored by many health economists and also recommended by NICE in its technology appraisals	• Cognitively challenging • Valuations can be seen as containing biases (source of bias: utility curvature, scale compatibility and loss aversion)
VAS	Non-choice based	Respondents are asked to fill out a rating scale comparing health states to death or perfect health	• Most feasible valuation technique • High response rate and high levels of completion	• Prone to context effects (average rating for items is influenced by the level of other items being valued) • Estimates have wider dispersion, with lower values for most conditions than estimated by SG or TTO

SG: Standard gamble; TTO: Time trade off; VAS: Visual analogue scale.

Table 2

Multi-attribute utility assessment instruments.

Study (year)	Instrument	Dimension	Country	Valuation technique	Strength	Weakness	Ref.
Devlin and Krabbe (2013)	EQ-5D (EuroQol - 5D)	Mobility, self-care, usual activities, pain/ discomfort, anxiety/ depression	Belgium, Denmark, Finland, Germany, Japan, The Netherlands, Slovenia, Spain, UK, USA, Zimbabwe	VAS, TTO, ranking	Preferred for populations with more severe problems	Suffers from ceiling effect. Hence, EQ-5D is less sensitive for milder problems Does not have dimensions for particular impairments	[25]
Feeny *et al* (2002)	HUI3	Vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain	Canada (Hamilton), France, the Netherlands, Spain	VAS transformed to SG	Performs better than many other instruments for people with sensory problems	Does not examine role or social function	[27]
Torrance *et al* (1996)	HUI2	Sensory, mobility, emotion, cognitive, self-care, pain, fertility	Canada (Hamilton), UK	VAS transformed to SG	Only generic instrument designed for use in children		[26]
Brazier *et al* (1998), Brazier and Roberts (2004)	SF-6D	Physical functioning, role limitation, social functioning, pain, energy, mental health	Hong Kong, UK, Japan, Australia, Brazil	SG, ranking	Preferred for populations with mild problems	Suffers from floor effect Does not have dimensions for particular impairments	[34,35]

HUI2: Health Utilities lndex-2; HUI3: Health Utilities lndex-3; TTO: Time trade off; VAS: Visual analogue scale.

Table 3

Review of quality-adjusted life-year weights in hemophilia.

Study (year)	Sample description	Country	Utility elicitation technique	On-demand	Prophylaxis	Inhibitor patients	Ref.
Naraine *et al* (2002)	30 members of general population aged 22–53 years, 28 adults aged 21–67 years with hemophilia, 30 parents aged 25–56 years of children with hemophilia	Canada	SG	Baseline: low bleeding frequency • General population – 0.905 • Adult patients – 0.940 • Parents – 0.985	• General population – 0.950 • Adult patients – 0.955 • Parents – 0.985	–	[46]
Wasserman *et al* (2005)	64 adults aged 19–81 years and 64 children aged 1–18 years with hemophilia (most of whom had parent proxy respondents)	USA	SG and VAS	• Mild patients Adults – 0.884 Pediatrics – 0.936 • Moderate patients Adults – 0.868 Pediatrics – 0.907 • Severe patients Adults – 0.810 Pediatrics – 0.868	• Severe patients Adults – 0.799 Pediatrics – 0.872	• Severe patients Adults – 0.745 Pediatrics – 0.799	[47]
Miners *et al* (1999)	66 patients with severe hemophilia (mean age 37.6) and 100 with mild/moderate hemophilia (mean age 46.1) (none received prophylaxis)	UK	EQ-5D and SF-36	• Mild/Moderate patients – 0.85 • Severe patients – 0.66	–	–	[48]
Trippoli *et al* (2001)	56 patients with hemophilia: 22 age 15–30, 16 age 31–45, 18 age 45 and over (none on prophylaxis or had inhibitor)	Italy	EQ-5D and SF-36	EuroQol Self Classifier • Age 1 5–30 years – 0.81 • Age 31–45 years – 0.70 • Age 45 and over – 0.49	–	–	[49]
Gringeri *et al* (2003)	52 patients age 15–64 with inhibitor	Italy	EQ-5D and SF-36			0.66 – UK weights 0.69 – Catalonia weights	[50]
Hoots *et al.* (2008)	22 patients with inhibitor	Multinational trial	EQ-5D Index (TTO)			Pre-prophylaxis Mean = 0.56 Median = 0.69 Post-prophylaxis	[51]

Study (year)	Sample description	Country	Utility elicitation technique	On-demand	Prophylaxis	Inhibitor patients	Ref.
						Mean = 0.62 Median = 0.79	
Stasyshyn *et al* (2014)	25 total patients with inhibitor 12-month trial before and after scores for 1 5 on-demand and 10 prophylaxis patients	16 HTCs in Europe and USA	EQ-5D			On-demand Before and after 12 months scores remain unchanged Prophylaxis EQ-5D scores changed from 0.62 to 0.70 after 12 months	[52]
de Kleijn *et al* (2014)	22 patients who had undergone surgery at least 12 months earlier, mean age 59 at assessment	Dutch	EQ-5D	Mean = 0.73 (EQ-5D)			[54]
Noone *et al* (2011)	58 patients aged 20–35 years	4 European countries (France, Ireland, Sweden and the UK)	EQ-5D	• 100% of lifetime on on-demand therapy – 0.72	• Lifetime primary prophylaxis – 0.88 • 50% of lifetime on prophylaxis – 0.77 • <50% of lifetime on prophylaxis – 0.72	–	[55]
Noone *et al* (2013)	124 patients aged 18–35 years	Canada, France, Ireland, the Netherlands, Poland and the UK	EQ-5D	• 100% of lifetime on on-demand therapy-0.619	• Lifetime primary prophylaxis – 0.866 • ≥50% of lifetime on prophylaxis – 0.812 • <50% of lifetime on prophylaxis – 0.755	0.798	[56]
den Uijl *et al* (2013)	40 moderate and 80 severe patients aged 20–33 years and 105 members of general Dutch population aged 20–31 years	The Netherlands	EQ-5D	–	• Moderate patients (lifetime prophylaxis) – 0.92 • Severe patients (lifetime	–	[9]

Study (year)	Sample description	Country	Utility elicitation technique	On-demand	Prophylaxis	Inhibitor patients	Ref.
					prophylaxis) – 0.8 prophylaxis) – 0.8		
Fischer *et al* (2013)	Severe patients (none had inhibitors) aged 18 years and over 78 Dutch patients 50 Swedish patients	The Netherlands and Sweden	EQ-5D	–	• Severe patients (lifetime prophylaxis) – 0.88 (Dutch) – 0.86 (Sweden)	–	[57]
Neufeld *et al* (2012)	18 adult patients (aged 19–61 years) 19 caregivers of pediatric patients (aged 1–18 years) with inhibitors	USA	EQ-5D	–	–	Bleed days – 0.66 Non-bleed days – 0.82	[53]
Barrefa/. (2002)	101 patients aged 13–87 years with severe (n = 47) or mild/moderate (n = 58) hemophilia stratified by HIV seropositivity	Canada	HUI2	• Mild/Moderate patients = 0.85 • Severe patients HIV+ = 0.67 HIV- = 0.80 All = 0.73	–		[59]
			HUI3	• Mild/Moderate patients = 0.77 • Severe patients HIV+ = 0.57 HIV- = 0.71 All = 0.63			
Szucs *et al.* (1996)	50 patients aged 18–60 years 11 on prophylaxis 39 on-demand treatment	Germany	HUI (early instrument developed by Torrance)	0.59	0.60		[58]
Lippert *et al* (2005)	500 male patients aged 14–83 years with severe hemophilia and no inhibitor 194 aged 14–30 years 126 on prophylaxis 68 on on-demand treatment 312 aged 31–83 years 150 on prophylaxis 162 on on-demand treatment	Germany, Sweden, UK, the Netherlands	SF-6D	• Age 30 or below – 0.73 • Age 30 or above – 0.66	• Age 30 or below – 0.76 • Age 30 or above – 0.68		[60]
Carvalhosa *et al* (2014)	71 adult males with hemophilia (mean age 45) 44 severe patients 15 moderate patients 12 mild patients	Belgium	SF-6D	• • Severe – 0.63 • Moderate – 0.66 • Mild – 0.74			[61]

EQ-5D: EuroQol 5-D; HUI: Health Utilities Index; SG: Standard gamble; TTO: Time trade off; VAS: Visual analogue scale.

Table 4

Hemophilia disease-specific instruments.

Study (year)	Instrument	Dimension	Country	Strength	Weakness	Ref.
Child specific						
Von Mackensen *et al* (2004) (long form), Haemo-QoL Group	Haemo-QoL	Physical health, feeling, attitude, family, friends, other people, sport and school, coping, treatment, future (for 1–16 years age group only), and relationship (for 1–16 years age group only)	6 European countries (Germany, Italy, France, Spain, the Netherlands, and the UK)	Three versions for different age groups: Version I for children 4–7 years of age, version II for children 8–12 years of age, version III for children 13–16 years of age. Uses psychosocial determinants of QoL, such as, coping, locus of control, life satisfaction and social support. Possesses sufficient internal consistency, validity and retest reliability. Age generic, eight-item QoL index short form Haemo-QoL instrument available for use in daily clinical routine	Reliability was lower in younger children	[85]
Young *et al* (2004), Young *et al* (2013)	(CHO-KLAT) Canadian Hemophilia Outcomes-Kids Life Assessment Tool 2004, CHO KLAT Version 2.0 2013	Nine domains somatic symptoms, physical functioning, sleep disturbance, stigma, social functioning, fear, resentment/ reaction, energy level, mood/ behavior and restrictions	Canada	Specifically developed for children. High construct, convergent and discriminant validity	Possible ceiling effects	[88,89]
Adult specific						
Rentz *et al.* (2008)	Haemo-QoL-A	Physical functioning, role functioning, worry, consequences of bleeding, emotional impact and treatment concerns	USA	High test-retest reproducibility. Able to discriminate patients based on severity. US English version adapted into German and Spanish. Available in subjects speaking French and Italian languages as well	Not validated extensively. Process ongoing	[87]
Von Mackensen *et al.* (2013)	Haem-A-QoL	Forty-six items in ten domains	Europe (Originally validated in Italian)	Developed for adults (>17 years)	Good validity and reliability. Linguistic validation done for 28 languages	[86]