Volume 12 • Number 5 • 2009
VALUE IN HEALTH

Challenges of Patient-Reported Outcome Assessment in Hemophilia Care—a State of the Art Review

Monika Bullinger, PhD,[1] Denise Globe, PhD,[2] Joan Wasserman, PhD,[3] Nancy L. Young, PhD,[4] Sylvia von Mackensen, PhD[1]*

[1]Institute for Medical Psychology, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany; [2]California State University Northridge, Department of Health Science, Northridge, CA, USA; [3]University of Texas Health Science Center at Houston, School of Nursing, Center on Aging, Houston, TX, USA; [4]Laurentian University, School of Medicine, Sudbury, ON, Canada

ABSTRACT

Introduction: One of the recent advances in assessing outcomes of medical care is the inclusion of the patient perspective. The term patient-reported outcome (PRO) is used to reflect the patient perceptions of disease and its consequences as well as of treatment and health-care provision. The development of PRO measures has advanced rapidly, and implementation in clinical research and practice is now underway. From an evaluation perspective, recommendations for the choice of PRO measures and an appraisal of the potential benefits of PRO data collection within specific health conditions are needed.

Methods: Hemophilia is a rare and clinically well-defined health condition with established and cost-intensive treatment strategies, in which PRO assessment is increasingly recognized as important. For this reason, measurement of PROs in hemophilia focusing on health-related quality of life (HRQoL) and patient preferences were reviewed to identify appropriate measures, to make recommendations for their choice, and to critically examine their impact in international hemophilia research and practice.

Results: Using literature searches and expert discussion strategies, generic and targeted measures for HRQoL and patient preferences in adults and children with hemophilia were screened, and 20 were reviewed on the basis of their psychometric properties and international availability. Only a few of the 20 measures have been used in clinical settings or research related to persons with hemophilia.

Conclusion: Consequently, an increased use of these measures is recommended to understand patient views on disease and treatment and to judge the impact of PROs for improvements in health care.

Keywords: assessment, hemophilia, patient-reported outcomes, patient preferences, quality of life.

Background

Recently, there has been a growing interest in including patient-reported outcomes (PROs) as secondary or primary end points in medical research. PROs are efficacy measures for evaluating the benefits of new and existing treatments. A PRO can be defined as any report coming directly from patients, without interpretation by physicians or persons, about how they function or feel in relation to a health condition and its treatment [1]. PROs include, but are not limited to, measures of quality of life, health status, treatment preferences, and patient satisfaction. They aim at capturing disease and treatment from the patient's perspective to better assess the value of the treatment received and to recommend strategies for improved care. This is especially important in chronic health conditions with high demands on care, often involving lifelong therapies, as is the case in hemophilia.

Hemophilia is a rare congenital disorder characterized by spontaneous and post-traumatic bleeding events due to low levels of blood-clotting factors. Hemorrhages are particularly frequent in joints and cause progressive destruction of articular structures, leading to impairment of joint function and chronic pain. Intracerebral and gastrointestinal hemorrhages are infrequent, but life-threatening. The major goal in treating patients with hemophilia is to reduce the frequency of bleeds, and consequently, mortality and joint damage to prevent future disability.

Address correspondence to: Monika Bullinger, Institute for Medical Psychology, Center for Psychosocial Medicine, University Medical Centre Hamburg-Eppendorf, Martinistr. 52, Haus S 35, D- 20246 Hamburg, Germany. E-mail: bullinger@uke.uni-hamburg.de

10.1111/j.1524-4733.2009.00523.x

*on behalf of the Health-Related Quality of Life Expert Working Group of the International Prophylaxis Study Group (IPSG)

Treatment is based on replacement of the missing clotting factor when a bleed occurs (on-demand treatment) or regularly and continuously (prophylactic treatment). The cost of therapy is extremely high mainly because of the costs of clotting factor concentrates [2]. In developed countries, it is estimated to be about €60,000 per patient year on average, but the costs can increase to €220,000 per patient year in patients with treatment complications such as the development of clotting factor-inhibiting antibodies. These costs are distributed differently between patients and payers in different countries. Dose, frequency, and type of replacement therapy in hemophilia vary depending not only on severity of the disease, patient age, lifestyle, and preferences and treatment schemes, but also on availability of factor concentrates and the reimbursement and the distribution system in the country in which it is being given. Reimbursement for prophylactic therapy also varies along a continuum of full to no reimbursement. On one end of the continuum, countries such as Germany and Canada provide reimbursement for prophylactic care for all hemophilic patients. In the middle are countries such as the United States where reimbursement is variable depending on individual insurer restrictions. On the other end of the continuum are countries such as India where no prophylactic care is provided for patients with hemophilia. Translating quality of life gains from prophylaxis into economic analyses is imperative for decision-makers who must decide whether or not to reimburse prophylactic treatment with factor concentrate for patients with hemophilia.

Bleeding frequency and joint damage have been the key outcome measures to evaluate treatment strategies in clinical studies and in routine practice. Assessment of PROs is increasingly considered necessary to understand which treatment strategy is clinically most effective and fits the patient's needs best [3]. Among the different PRO approaches, this current article focuses on two:

1098-3015/09/808 808–820

health-related quality of life (HRQoL) and patient preference measures (PPMs) [4]. In the subsequent review of measures, careful evaluation of conceptual background, methodological quality, clinical practicability, and international availability was carried out and information on the use of these measures in clinical research and practice was gathered.

A Brief Overview of HRQoL Assessment

Conceptually, HRQoL has been defined as reflecting a person's valuation of goals, expectations, and aspirations with regard to different areas of life. More specifically, HRQoL has been viewed as the subjective perception of well-being and function in different life domains, namely concerning physical, social, emotional, mental, and everyday life or role performance [5]. The inclusion of the term HRQoL in medical research reflects the need to address patients' perspectives of their health conditions and treatments. Because the focus on health constitutes a more specific aspect of quality of life, the term HRQoL has been coined to describe patient-perceived health status and treatment experience [6].

HRQoL instruments differ in aspects such as scope (generic or chronic generic vs. targeted), respondents (adults vs. children or adolescents), form of administration (self-completed, questionnaire, or computer administration vs. interview), mode of response (self-report vs. proxy report or observation), purpose of assessment (descriptive, evaluative, or predictive), and method of development (psychometric primarily based on psychological test theory vs. clinimetric primarily based on clinical epidemiology principles to integrate the perspectives of patients and clinicians) [7,8]. In psychometric and clinimetric paradigms, reliability refers to the precision of a measure (e.g., internal consistency, test–retest) and validity refers to the correct representation of the concept to be measured in the questionnaire (e.g., face, criterion, construct validity). Within either construct or criterion validity, discriminant validity denotes the ability of the measure to distinguish between different clinical subgroups, and convergent validity denotes the correspondence of the measure with other instruments measuring similar concepts.

In general, instruments also differ regarding their purpose of use, ranging from diagnosis of the individual patient in a specific clinical context over outcome assessment in prospective observational studies or controlled clinical trials to epidemiological studies [9,10].

Generic measures are designed to be applicable to persons with a variety of health conditions, including the general population where most subjects have no disease or defined health condition. Targeted—or disease-specific measures—are developed specifically for a defined clinical population, and hence focus on the most relevant health issues for that group. Both types of measures have been developed for adults and for children/adolescents.

The availability of measures for use in different cultural contexts is also of importance. Steps in the construction of an internationally usable instrument involve item development, translation, psychometric testing (e.g., examination of reliability, validity, and responsiveness), and norming (obtaining reference data for the questionnaire from large nonclinical or clinical populations) in the respective country. Guidelines on how to conduct these steps have been published [11,12].

A Brief Overview of PPM

Patient preferences (also referred to as utilities) can be characterized as numeric measures that represent the value placed on a particular health state [13]. The number reflects how much risk that individual would be willing to take in exchange for a specific improvement in this health state and provides standardized metrics to compare different outcomes across various diseases. PPMs may differ depending on whose values are applied—those of the individual vs. those of society. Preferences for a health state may be elicited either directly or indirectly.

Direct measurement, a theoretically based measure involving choice, requires respondents to value a prespecified health state [14], using one of three current approaches to assess preferences or utilities directly: standard gamble (SG), time-tradeoff (TTO), and visual analogue scale (VAS). These three methods most often yield different preference scores [15]. With these measures, a patient has to decide between a theoretical lifetime that he/she would be willing to give in exchange for a beneficial treatment (TTO), or which risk a patient would take for which probability of being cured (SG) or how desirable a health outcome is (VAS).

Indirect measurement involves using a health status instrument where the preferences have been assigned independent of the study, preferably by community-based judges. These preference-based measures with preassigned weights (preference scores) theoretically yield scores that are comparable with measures elicited directly, and the differences in methods are in the definition of health states [16].

Recently, patient preferences have been derived from existing quality of life scales by using items and response categories to construct health state scenarios that patients are then asked to rate in terms of preferences. If these states are then rated a priori by community-based judges, it is possible to assign utilities as well [17].

This article reviews the PRO assessment in the field of hemophilia with specific attention to different age groups (childhood, adolescence, and adulthood). The purpose is to identify valid and reliable HRQoL and PPMs measures for use in prophylaxis research, to critically examine the current use of these measures in clinical studies, and to evaluate the impact of PRO assessment in hemophilia.

Methods

A literature review on HRQoL or PPM was performed using the MEDLINE System with the aim to identify disease-specific or targeted measures for children and adults in the hemophilia field and to examine which of these measures in fact had been used in published hemophilia studies. An important aspect for such measures is that they reflect patient perceptions from the start of instrument development and that they provide information whether and how patient input was used to develop key domains and items. For the literature analysis, a PubMed search was conducted using the keywords "haemophilia" with "quality of life," and with "patient preferences" or "utilities" without limitation in publication date or language. Criteria for the inspection of a publication were:

1. description of approaches to measurement as well as information on the PRO instrument creation ranging from the generation of items to the finalization of the instrument;
2. description of an instrument with data about the methodological quality of the measure in terms of reliability, validity, and responsiveness;
3. information on the availability of language versions other than the original language, with measures preferably not only translated but also validated in these languages; and
4. the presence of empirical data reported in outcome studies about these measures from the literature including their actual use in research and practice.

The keywords (text words) "haemophilia" or "hemophilia" combined with "quality of life" resulted in 210 publications, combined with "patient preferences," 30 articles were found, and when combined with "utilities," 7 articles remained. In the second stage, a more detailed search was conducted; articles were inspected in terms of their design and psychometric characteristics. In addition, reviews of HRQoL and PPMs published within the area of hemophilia were included in the results presented here [18–24]. The current review attempts to add to these existing reviews by including the most recent information on instrument development and by conjointly addressing patient preference and quality of life measures.

Using this strategy, a total of 20 quality of life measures were inspected: five generic measures for adults, six generic measures for children, five targeted measures for adults, and four targeted measures for children. For patient preference assessment, four generic instruments for adults, two generic instruments for children, one targeted instrument for adults, and no targeted instrument for children were found. These measures were examined for the report of psychometric criteria (reliability, validity, responsiveness) and for available translations or validation in other than the original language. To examine targeted measures, the criteria specified for evaluating PRO instruments were derived from the US Food and Drug Administration (FDA) PRO Draft Guidance [25]; these include information on the conceptual framework and intended application, development of the instrument, psychometric properties, and the origin of the instruments.

The results of this review were then screened for the application of measures in actual studies. A second PubMed search was carried out in which the instruments' names, as one keyword, was connected with the term haemophilia or hemophilia. Both search strategies were performed as a basis for expert discussions within the HRQoL expert working group of the International Prophylaxis Study Group (IPSG), a subgroup of which met several times over a period of 4 years to arrive at a consensus about HRQoL and PPM assessment in hemophilia. The goal of the subgroup was to arrive at guidelines for clinicians and researchers in selecting relevant HRQoL measures and PPMs, and in implementing PRO measures in area of hemophilia.

Results: Quality of Life

Quality of Life Measures for Adults and for Children

Adult generic measures. A number of generic measures to assess HRQoL in adults have been developed over the past 30 years. The four most frequently used and internationally available measures were the Sickness Impact Profile (SIP) [26], Nottingham Health Profile (NHP) [27], the World Health Organization Quality of Life Assessment (WHOQOL) [28,29], and the SF-36 Health Survey [30,31]. In general, these instruments' psychometric properties have been examined in a wide variety of studies. The SIP and NHP were the first measures developed to capture the subjective representation of health—they were primarily oriented toward identifying major impairments in self-reported health, although newer measures such as the WHOQOL-100 and its short version, the WHOQOL-BREF [32], as well as the SF-36 Health Survey and its abbreviated version SF-12 [33], represent the positive and negative spectrum of functioning and well-being. Because of their generic nature and their known measurement properties, they are principally applicable for HRQoL assessment across health conditions including hemophilia. Information on these measures, as shown in Table 1, indicates good psychometric properties, but there are differences in item number and in the existence of short versions. With the exception of the internationally derived WHOQOL, the instruments have been developed originally in English. In addition, these instruments have been translated in a variety of languages and many have been validated in different countries. As an example, over 80 language versions exist for the SF-36, which has undergone validation and norming in many countries (http://www.qualitymetric.com).

Nevertheless, only a few of the generic measures have been included in hemophilia studies (SF-36, SF-12, SIP, and the Quality of Well-Being Index (QWB)). As already reported in earlier reviews, the most frequently included measure is the SF-36/SF-12 Health Survey. An (up-to-date) PubMed review (June 2007) confirmed this: 13 hemophilia articles were cited for the SF-36 [34,35] and 2 articles were cited for the SF-12. In comparison, the SIP has been cited in only one hemophilia study [36], and the NHP and the WHOQOL have not been used in adult patients with hemophilia according to the literature. Two other generic PRO instruments have been included in two hemophilia studies, namely the QWB and the Manchester Minneapolis QoL Index [37,38].

In addition, targeted measures for chronic diseases with consequences for joint function and movement have been developed such the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [39,40], and the Arthritis Impact Measurement Scales (AIMS-2 [41]). The AIMS-2 has been cited five times, and is the most prominent chronic generic measure used in the hemophilia field. By contrast, the WOMAC has been cited once [42] in this field. An older measure, the Stanford Health Assessment Questionnaire [43] has recently been recommended of functional impairment because of arthropathy [44]. Albeit developed for other chronic diseases, these targeted measures deserve further psychometric testing in the hemophilia patient population before they can be used in hemophilia to assess the level of specific functional impairments.

Child generic measures. In terms of children's quality of life assessment, efforts have recently been directed toward developing and testing scales measuring quality of life from the child's (as well as parents') perspective and in different age groups [45]. Several generic measures such as the Child Health Questionnaire (CHQ) by Landgraf [46] and the PedsQL by Varni [47] have been developed in the United States, while other measures such as the KINDL/KINDL-R questionnaire for child self-report and for parental proxy report [48,49] and TACQOL [50] have been developed in Europe. Most recently, two multinationally developed instruments are available for use: the KIDSCREEN [51,52] and the DISABKIDS measure [53,54]. Many of these measures have been developed during the past decade, and have been designed to be applicable to children's age and developmental status. Some of them already include self-report in children from age 4 on, although for younger children, only parent report is possible. Differences between age group versions of an instrument relate to the number of items, their formulation, and the type of domains included. The number of items tends to be low, and parents' report paralleling children's items is frequently available. Short forms are often developed in a way to make assessment across a number of items in age groups possible.

The CHQ is available in child and parent forms and has been extensively and successfully psychometrically tested. This is also true for the KINDL-R questionnaire, which includes three age group versions for children 4 to 7, 8 to 11, and 12 to 16. The TACQOL is again available for older children in parent and self-report. The PedsQL subdivides four age groups (2–4, 5–7, 8–12, and 13–18) with 23 items, is available in self- and parent-report in over 20 languages, and is one of the most widely used

Table I (a) Generic QoL measures for adults; (b) generic QoL measures for children

Instrument characteristics						Measurement properties	
(a)							
Instrument		Author	No. of Domains	No. of items	Original Language	Reliability	Validity
Generic							
SIP		Bergner et al., 1981 [26]	12	136	English (US)	$r = 0.92$ $\alpha = 0.96$	**Criterion**: concurrent **Construct**: convergent discriminant
NHP		Hunt et al., 1981 [27]	6 + 1	38 + 6	English (UK)	$r = 0.77–0.85$	**Criterion**: known groups
WHOQoL Family	WHOQoL-100	Power et al., 1998, 1999	6	100	International	$\alpha = 0.81$ (sum score)	**Construct**:discriminant
	WHOQoL-BREF	[28,29,32]	4	24		$\alpha = 0.66–0.84$ (subscales)	
SF Health Survey Family	SF-36 Health Survey	Ware et al., 1992 [30]	8	36	English (US)*	$\alpha = 0.68–0.96$ (subscales) ICC = 0.648–0.868	**Criterion**: known groups **Construct**: convergent discriminant
	SF-12 Health Survey	Ware et al., 1996 [33]	2	12		$\alpha = 0.63–0.91$ (subscales) ICC > 0.80	
Chronic generic							
WOMAC		Bellamy et al., 1988 [40]	24	3	English (US)	$\alpha = 0.80$ (sum score) $r = 0.80$	**Construct**: Discriminant content
AIMS2		Meenan et al., 1982 [41]	12	57	English (US)	$\alpha = 0.72–0.91$ $r = 0.78–0.94$	**Criterion**: known groups

(b)

Instrument: all in child-report (C) and parent-report (P)	Author	Age groups	No. of domains	No. of Items	Language	Reliability	Validity
Generic							
KINDL-R	Ravens-Sieberer et al. 1998 [49]	4–7	6	12	German†	$\alpha = 0.63–0.84$ (C)	**Criterion**: known groups
		8–11 12–16		24		$\alpha = 0.62–0.89$ (P)	**Construct**: convergent discriminant
CHQ	Landgraf et al., 1997 [46]	8–15	14	87	English (US)‡	$\alpha = 0.63–0.90$ (C) $\alpha = 0.59–0.94$ (P)	**Construct**: discriminant
TACQOL	Theunissen et al., 1998 [50]	8–15 (C) 5–15 (P)	7	56	Dutch§	$\alpha = 0.66–0.79$ (C) $\alpha = 0.67–0.84$ (P)	**Construct**: discriminant
PedsQl	Varni et al., 1999 [47]	5–7	4	23	English (US)‖	$\alpha = 0.88$ (C) $\alpha = 0.90$ (P)	**Construct**: discriminant
		8–12				ICC = 0.88(C)	
		13–18				ICC = 0.90 (P)	
KIDSCREEN	Ravens-Sieberer et al., 2001 [51]	8–18 KIDSCREEN-52	10	52	International¶	$\alpha = 0.76–0.89$	**Construct**: convergent discriminant
		8–18 KIDSCREEN-27	5	27		$\alpha = 0.79–0.84$	
		8–18 KIDSCREEN-10	1	10		$\alpha = 0.82$	
Chronic generic							
DISABKIDS	Bullinger et al., 2002 [53]	8–16 DCGM-37	6	37	International#	$\alpha = 0.70–0.93$ (C) $\alpha = 0.77–0.95$ (P)	**Construct**: convergent discriminant
		8–16 DCGM-12	3	12		$\alpha = 0.90$ (C) $\alpha = 0.93$ (P)	
		4–7 Smiley Measure	1	6		$\alpha = 0.69$ (C) $\alpha = 0.71$ (P)	

*Validated Languages
Afrikaans, Arabic (Qatar), Arabic (Saudi Arabia), Arabic (Tunisia), Armenian (Armenia), Bulgarian, Chinese (Hong Kong SAR), Chinese (Malaysia), Chinese (Simplified), Chinese (Singapore), Chinese (Taiwan), Chinese (United States), Croatian, Czech, Danish, Dutch (Belgium), Dutch (Netherlands), English (Australia), English (Canada), English (Great Britain), English (Hong Kong), English (Malaysia), English (New Zealand), English (Philippines), English (Singapore), English (South Africa), English (Taiwan), English (United States), Estonian, Finnish, French (Belgium), French (Canada), French (Standard), French (Switzerland), German (Austria), German (Standard), German (Switzerland), Greek, Hebrew, Hungarian, Icelandic, India (Bengali), India (English), India (Gujarati), India (Hindi), India (Kannada), India (Malayalam), India (Marathi), India (Punjabi), India (Tamil), India (Telugu), India (Urdu), Indonesian, Italian (Standard), Italian (Switzerland), Japanese, Korean, Latvian, Lithuanian, Malay, Malay (Malaysia), Maltese, Norwegian, Philippines (Cebuano), Philippines (Tagalog), Polish, Portuguese (Brazil), Portuguese (Portugal), Romanian, Russian, Serbia (Serbian), Singapore (Malaysia), Slovak, Slovenian, South Africa (hosa), South Africa (Zulu), Spanish (Argentina), Spanish (Chile), Spanish (Colombia), Spanish (Costa Rica), Spanish (Guatemala), Spanish (Honduras), Spanish (Mexico), Spanish (Peru), Spanish (Puerto Rico), Spanish (Spain), Spanish (United States), Spanish (Uruguay), Spanish (Venezuela), Swedish, Thai, Turkish, Ukrainian, Vietnamese.
†German, English, French, Spanish, Italian, Dutch, Greek, Turkish, Norwegian, Swedish, Polish, Portuguese, Taiwanese, Russian, Serbian, Serbo-Croatian, Vietnamese, Japanese, Arabian.
‡Chinese Traditional (Hong Kong), Danish, Dutch (Belgium), Dutch (Netherlands), English (Australia), English (Canada), English (UK-Eire), English (US), French (Canada), French (European), German (Austria), German (Germany), Greek Hebrew (Israel), Italian, Norwegian, Polish, Portuguese (Brazil), Slovakian, Spanish (Costa Rica), Spanish (Iberian), Spanish (Mexico), Spanish (Puerto Rico), Spanish (U.S.), Swedish.
§English, French, Russian, Turkish.
‖Spanish for Argentina, English for Australia, Belgium Dutch, Belgium French, Portuguese for Brazil, French for Canada, Spanish for Chile, Croatian, Czech, Danish, French, Finnish, German, Hungarian, Icelandic for Iceland, Arabic, Hebrew, Russian, Italian, Spanish for Mexico, Norwegian, Urdu for Pakistan, Arabic, Spanish for Peru, Polish, Portuguese, Romanian, Russian, Slovakian, Slovenian, Spanish, Swedish, Dutch, English for the UK, English for the USA.
¶German, German (Austria), Czech, French, Greek, Hungarian, English (Ireland), Dutch (Netherlands), Polish, Spanish, Swedish, German (Switzerland), English (UK).
#German, English, Dutch, French, Greek, Swedish.
α = internal consistency; AIMS2, arthritis impact measurement scales; C, child report; CHQ, child health questionnaire; ICC = concordance (intraclass correlation); P, parent report; NHP, Nottingham health profile; PedsQl, pediatric quality of life inventory; r = test–retest correlation; SIP, sickness impact profile, TACQOL, TNO AZL child quality of life; WHOQoL, World Health Organization quality of life; WOMAC, Western Ontario and McMaster Universities osteoarthritis index.

instruments for generic quality of life assessment in children. Very recently, the European KIDSCREEN instrument has been simultaneously developed in several languages and validated in over 20,000 children, but has not yet been applied in pediatric hemophilia research. The same is true for a new chronic generic measure: the DISABKIDS instrument. The DISABKIDS instrument focuses on the experience of having a chronic illness and originated, as did the KIDSCREEN, from a European research consortium [55,56]. Both instruments are age-group related, were derived from respondents' focus groups, and can be completed by children/adolescents from the age of 8 years and by parents. For the DISABKIDS, a smiley version for younger children (4–7 years) has been developed. Especially, the newer instruments provide assessments of core HRQoL concepts as well as modules for different diseases are brief and are psychometrically well tested. All of these instruments sufficiently fulfill psychometric criteria of reliability and validity and are available in various languages (see Table 1b). The PedsQL has so far been included in the hemophilia published literature [57], as have the KINDL Questionnaire [58] and the CHQ [59,69].

Targeted Quality of Life Measures for Children and Adults

Hemophilia-specific HRQoL measures for adults. In addition to generic scales, which had been used before, hemophilia-specific HRQoL measures for adults have only very recently been published according to the guidelines of international instrument construction. The HAEMO-QOL-A questionnaire was developed internationally [60]; initially, focus groups were conducted simultaneously in participating countries to derive items. The HAEMO-QOL-A consists of 41 items representing six relevant HRQoL dimensions: physical functioning, psychosocial issues, role functioning, fear/worry, positive affect, and treatment worry, and is available in several languages. Psychometric characteristics showed good values for reliability and satisfactory values for validity (concurrent, discriminant). The questionnaire was pilot-tested in the United States, translated, and was field-tested in the United States, Spain, and Germany, in 221 patients. The Haem-A-QoL [61] instrument was compiled from focus groups of patients with hemophilia in Italy who were interviewed about their perceptions of their condition and treatment. The Haem-A-QoL consists of 46 items pertaining to 10 dimensions with an average completion time of 14 minutes. The Haem-A-QoL has a core instrument consisting of 27 items that are identical to items in the pediatric Haemo-QoL questionnaire (see further discussion), and thus allows a comparison between children's and adults' hemophilia-specific quality of life and the ability to assess patients' HRQoL from early childhood to adulthood [62]. The Haem-A-QoL was validated in the Italian Cost of Care of Hemophilia Study (COCHE) in 233 Italian adult patients with hemophilia. The psychometric characteristics include good reliability (ranging from $\alpha = 0.74$–0.88), and high convergent (correlation with SF-36) and discriminant validity (e.g., severity and infections) [63]. The Haem-A-QoL has been translated into 21 languages using a quality-controlled forward/backward translation procedure; cross-cultural validity was determined with positive results by clinicians' reviews in 17 languages [64].

The Hemofilia-QoL [65] was developed in Spain and consists of 36 items reflecting to nine scales: physical health (eight items), daily activities (four items), joint damage (three items), pain (two items), treatment satisfaction (two items), treatment difficulties (four items), emotional functioning (five items), mental health (three items), and relationship and social activities (five items). Psychometric testing demonstrated acceptable reliability values with Cronbach's $\alpha = 0.95$ for the Hemofilia-QoL total score and adequate convergent validity with the SF-36 subscales. The questionnaire was field-tested in 121 hemophilic patients in Spain.

Also, the Hemolatin-QoL Questionnaire [66] from Latin America is a newer instrument available in Spanish as well as Portuguese for Brazil with current psychometric testing underway. A newer French instrument, the hemophilia age-group specific quality of life questionnaire assessing 11 domains with 89 items, is currently under validation [67].

The instruments reviewed are all designed to evaluate efficacy of treatment, assess symptoms/signs, functional status, and health perception, and include multiple items from multiple domains in multiple concepts. They are self-administered, all spanning a 4-week time frame, and yielding profile and composite scores. Patients are asked to fill out the questionnaires that can be administered in person or by mail. Face-to-face and phone surveys are possible, but have not been conducted so far. An article comparing the concepts of three adult measures (Haemo-QoL-A, Haem-A-QoL, Hemofilia-QoL) found that they all address five concepts of HRQoL in hemophilia, with, however, different items, namely physical, emotional, social, functional, mental, and treatment-related dimensions [63] (Table 2).

Hemophilia-specific HRQoL measures for children. A special problem with measuring HRQoL among children is the age dependence of quality of life domains as well as difficulties obtaining their self-report throughout the developmental stages. Two of the first measures for children with hemophilia were developed concurrently: the Haemo-QoL [68,69] and the Canadian Haemophilia Outcomes–Kids Life Assessment Tool (CHO-KLAT) [70]. Both are able to capture the perspectives of children's and parents' in separate versions.

The initial development of the Haemo-QoL questionnaire used parents' assessment of children's quality of life as well as clinical expert consensus on relevant dimensions. This input was used to construct a quality of life instrument for children. Statements were then modified for three age groups (4–7, 8–12, 13–16 years) of children and tested for acceptance using cognitive debriefing techniques. The age-specific versions differ in number of domains and items, but contain an identical core item set. The Haemo-QoL was pilot-tested in 58 children and their parents and field-tested and validated in 339 children and their parents in six European countries (Germany, Italy, Spain, France, the UK, and The Netherlands) [58]. These validated Haemo-QoL versions can be downloaded from the Haemo-QoL website (http://www.haemoqol.de). Moreover, the Haemo-QoL has been linguistically cross-cultural validated in 30 languages.

Psychometric analysis revealed good psychometric properties both for the interview-based form in small children (21 items), the self-report in children age 8–12 (64 items), and adolescents 13–16 years (77 items) as well as for the corresponding forms for parents. Cronbach's alpha for the total score varied from $\alpha = 0.85$ to 0.91. The convergent validity of the questionnaire was satisfactory.

In addition to the full version, a short version for small children (4–7 years) containing 16 items and a short version for older children (8–16 years) containing 35 items were developed and showed excellent psychometric properties. An eight-item index version was developed spanning all age groups, which is also available as a self- and parent-report (Haemo-QoL-Index) [71].

The CHO-KLAT was developed initially in Canada and a single cross-age version was sought [70]. The CHO-KLAT is unique in being developed primarily from first principals and from

Table 2 (a) Disease-specific QoL measures for adults with hemophilia; and (b) disease-specific QoL measures for children with hemophilia

Instrument characteristics					Language		Measurement properties	
(a) Adults								
Instrument	Author	No. of domains	No. of items	No. of patients	Original	Validated	Reliability	Validity
Haemo-QoL-A	Rentz et al., 2008 [59]	6	46	221	English (US)	English (US) German Spanish	$\alpha = 0.95$	**Construct**: convergent discriminant
Haem-A-QoL	Mackensen et al., 2004 [55]	10	46	233	Italian	Italian English (UK) German Hungarian	$\alpha = 0.96$	**Content** **Construct**: convergent discriminant
Hemofilia-QoL	Arranz et al., 2004 [64]	9	36	121	Spanish	Spanish	$\alpha = 0.95$	**Construct**: convergent discriminant
Hemolotain-QoL	Remor et al., 2005 [65]	9	47	231	Spanish (US)	Spanish (US) Portuguese (US)	$\alpha = 0.95$	**Content**
QUAL-HEMO	Trudeau et al., 2003 [66]	11	89	70	French	French	$\alpha = 0.35$–0.88	

(b) Children

Instrument	Author	Age range	No. of domains	No. of items	No. of patients	Version	Original	Validated	Reliability	Validity
Haemo-QoL	Bullinger et al. 2002 [68] & v. Mackensen et al., 2004 [58]	4–7 Haemo-QoL$_{Fullversion}$	9	21	339	CP	English (UK)	English (UK) German Spanish French Italian Dutch	$\alpha = 0.85$ (C) $\alpha = 0.85$ (P) $\alpha = 0.85$ (C) $\alpha = 0.93$ (P) $r = 0.90$ (C) $\alpha = 0.91$ (C) $\alpha = 0.95$ (P) $r = 0.92$ (C)	**Construct**: convergent discriminant
		8–12 Haemo-QoL$_{Fullversion}$	10	64						
		13–16 Haemo-QoL$_{Fullversion}$	12	77						
	v. Mackensen et al., 2004 [58]	4–7 Haemo-QoL$_{Shortversion}$	1	16					$\alpha = 0.92$ ICC = 0.33 (C-P)	
		8–16 Haemo-QoL$_{Shortversion}$	1	35					$\alpha = 0.921$ CC = 0.78 (C) ICC = 0.86 (P) ICC = 0.47 (C-P)	
	Pollak et al., 2006 [70]	4–16 Haemo-QoL$_{Index}$	1	8					$\alpha = 0.70$ (C) $\alpha = 0.78$ (P)	
CHO-KLAT	Young et al., 2004 [69]	4–18	1	35	52	C P	English (CA)	English (CA and UK) French (CA and France) Dutch Spanish German	ICC = 0.74 (C) ICC = 0.83 (P) $\alpha > 0.8$ (all) ICC = 0.84 (C-P)	**Concordant** **Construct**: Convergent discriminant
Quality of life for young patients	Manco Johnson et al., 2004 [72]	2–6	9	39	103	P	English (US)	English (US)	$\alpha = 0.73$–0.94	**Construct**: convergent discriminant

the perspectives of children with hemophilia of all ages. Priority at all stages of the development of the CHO-KLAT was given to children to ensure that the measure reflected the perspectives of children and could be used by self-report by children. The CHO-KLAT is available for children 4 to 18 years of age and includes parents- and self-report forms [70]. The CHO-KLAT consists of 35 items and has a single summary score representing overall quality of life based on questions regarding treatment, physical health, family, future, feeling, understanding of hemophilia, other people and friends, and control over your life. Child–parent concordance is very strong (intraclass correlation coefficient [ICC] = 0.75), as is consistency over time for both child self-report (ICC = 0.74) and parent proxy report (ICC = 0.83).

The instrument was pilot-tested in a group of 52 children and is currently field-tested in a larger group. The CHO-KLAT is available in English (both Canadian and UK versions), French (both Canadian and France versions), Spanish, Dutch, and German. Furthermore, the CHO-KLAT and Haemo-QoL have been used conjointly in a recent retrospective analysis that showed acceptable psychometric properties as well as strengths and weaknesses of both measures [72].

The instruments reviewed so far evaluated efficacy and assessed symptoms and signs, as well as functional status for multiple domains with multiple concepts. All measures are self-reported, relate to the past 4 weeks and are filled out by the children. In young children from the early age of 4 years onward, they are interview administered. Each of the measures is also available for parent-proxy report using the same items formulated from the parents' perspective.

Other targeted instruments for children with hemophilia have also been published. For example, Manco-Johnson et al. [73] developed and tested a parent-administered instrument to assess the HRQoL of very young children between 2 and 6 years. Some of the other more recent measures in young hemophilia patients such as the Dutch Hem-Dux [74], however, still have to undergo more extended psychometric testing.

Preference Measures

Generic PPM for Adults and Children

Within the indirect assessment approach, systematic generic patient-preference measurement has begun as early as 1980—and has since then been used in health economic studies of chronic diseases. Older measures such as the QWB [75] were complemented by newer approaches such as EUROQOL [76], now named EQ-5D [77], Health Utilities Index (HUI) [78], and the SF-6D [17].

The EQ-5D is a generic, utility-based instrument assessing five dimensions: mobility, self-care, usual activities, pain, and anxiety/depression [79]. The utility score from this measure is derived from community-based preference scores assigned to several health scenarios, created by the permutation and combination of the degrees of capacity or impairment across the five different health dimensions. The EQ-5D was psychometrically tested for reliability [80] and validity [81,82]. Recently, a patient-preference version of the SF-36, the SF-6D, has been developed, which can be administered directly or can be derived from the SF-36 items [83]. Validity information on the SF-6D, compared with the EQ-5D, has been reported by Petrou et al. [84]. In addition, the HUI [85,86] has been utilized with adults and is the only generic PPM instrument so far used with children. Reliability [87] and validity [88] have been reported for the general adult population.

Multiattribute health states assessed in the HUI Mark 3 include vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. In hemophilia, these measures have been included only very recently.

In an effort to begin to assess PPMs in adult patients with hemophilia A and B, Miners et al. [89] utilized both the SF-36 and the EQ-5D. The authors found that individuals with severe hemophilia have reduced levels of HRQoL compared with: 1) individuals with moderate or mild forms of the disease; and 2) with the UK male normative population.

Targeted PPM for Adults and Children

Newer developments have included efforts to derive disease-specific utilities, oriented at the target health issues associated with a given disease. As one of the first approaches in the use of utility-based measures in hemophilia [90], Wasserman et al. (2005) developed and validated a disease-specific utility instrument that directly measured patient preferences for nine unique hemophilia health states [91]. Health states ranged from mild to severe and included common morbidities experienced by persons with hemophilia. The VAS and the SG methods were both used to assess preferences. The authors found statistically significant differences for all health states combined between pediatric and adult participants ($P = 0.045$) as well as differences among adult and pediatric group preferences for the mild, severe (episodic treatment), and severe (prophylactic treatment) health states. These results indicated that age can influence patients' preferences regarding their state of health (Table 3).

Recent research demonstrated that both generic and disease-specific instruments can be used to reliably measure utilities in hemophilia. However, given the complex nature of the disease, its unique characteristics, and the comorbidities often associated with it, there is a need for multiattribute preference measurements. To date, no preference measure integrating two or more health states (multiattribute) has been developed for this population. Reliable and valid generic (e.g., EQ-5D) and disease-specific (e.g., TTO or algorithms of indirect methods) can be used to integrate multiple attributes into a single preference instrument for hemophilia.

Inclusion of HRQoL and PPM Measures in Hemophilia Research and Care

Given the short history of PROs, it is not surprising that only few studies have used HRQoL and PPM measures in hemophilia and that no randomized clinical trials have been published yet. Within the 210 citations identified from the literature search, many used respective keywords without having empirically addressed the topic. Examples of including quality of life measures are cross-sectional studies to better understand patients' quality of life as related to hemophilia [92] in specific cultural settings (e.g., the Chinese population) [93] in specific disease conditions (e.g., von Willebrand Disease) [94], regarding complications (HIV, inhibitors) [2] or impact of events (e.g., intracranial bleeds and arthropathy) [95,96].

Only recently, studies are underway combining quality of life and PPM assessments [97–99]. In comparison, PPM measures have primarily been used in health economic studies. Nevertheless, both measures are increasingly included in comprehensive health outcome studies [100], such as the European Study of Clinical, Health-Economic and Quality of Life outcomes in hemophilia (http://www.eschqol.lmu.de). In this multinational study, HRQoL and PPM measures are used conjointly and prospectively in more than 1400 patients from 21 European countries, namely adults and children as well as their parents.

Furthermore, PPM and HRQoL concepts are also interrelated. Patient preferences are an important factor in the

Table 3 Generic and targeted utility measures for adults and children

Name		Author of hemophilia-related article	Generic/disease specific Adult/children	No. of items/ No. of Domains	No. of patients	Method	Country scoring algorithm validated	Original language	Validated languages	Reliability	Validity
Generic EQ-5D (adult instrument)		Miners et al. (2002) [88]	Adult	5/1	168	VAS	Yes	Dutch, English, Finnish, Norwegian, Swedish	All major	$r = 0.63–0.80$ ICC = 0.86	**Construct:** convergent discriminant
SF-6D		Brazier et al. (1998) [17]	Adult	31/6	over 1000	VAS	Yes	see SF 36	English	not reported	**Construct:** discriminant
Health Utility Index	HUI–2	Barr et al. (2002) [91]	Adult and Children	7/1	115	VAS, TTO,	Yes	English	English, French, Spanish, Italian, German, Dutch, Japanese		**Criterion:** predicitve
	HUI-3			8/1		SG				ICC = 0.767	**Construct:** discriminant
Disease-specific Hemophilia Health States		Wasserman et al. (2005) [90]	Adult	9 health states	128	VAS, SG	No	English	English, Spanish	$r = 0.91$	**Content Construct:** discriminant

α = internal consistency; C, child report; ICC = concordance; QUAL-HEMO, quality of life questionnaire in hemophilia; P, parent report; r = test-retest correlation.

assessment of overall HRQoL and vice versa. Patient preferences for a particular health state can be associated with the likelihood of seeking care, the type of treatment selected, selection of provider specialty, and compliance with care [101].

Preferences along with individual and environmental characteristics shape an individual's perception of his/her general health. For example, high utilization of health-care services is potentially associated with poorer self-reported health status. Because patient preferences are related to an individual's perception of health status, it is not surprising that results vary with respondent characteristics. Initial investigations have supported the existence of a relationship between health-state preferences and patient health-seeking behavior. For instance, Fifer et al. [102] demonstrated that prescriptions for psychotropic medications for patients with anxiety were related to patient self-assessment of the desirability of their current state. The association between patient preferences and treatment decisions of patients with hemophilia has not been explored, and it is not yet known how severity of hemophilia, treatment of hemophilia, patient preferences, and quality of life might be related [100].

Choosing PRO Measures in Hemophilia

The selection of the HRQoL or PPM measures should be guided by the study-related and instrument-related aspects [62]. Instrument-related characteristics include the concepts assessed, the psychometric properties and feasibility of the instrument. Because study objectives vary, relevant concepts will have to reflect the dimensions under study. As concerns comparisons of dimensions or domains assessed across instruments, this constitutes a difficult task. The reason is that although domain names might seem similar, close inspection of their content shows a wide variation. Beyond a classification of very broad domains (physical, emotional, social), a specific comparison makes a close inspection of the items included in subscales of an instrument necessary. As concerns the study purpose, there is a wide range of study designs and objectives including a description of the patient population (epidemiological, cross sectional), a description of change over time in the patient cohort (longitudinal prospective), a comparative evaluation of treatment outcomes (randomized clinical trials), an evaluation of the quality of care (longitudinal quality assurance studies), or an assessment of treatment benefit (health economic studies).

Study designs for research specific to hemophilia care including HRQoL and PPM measures may serve different purposes. For instance, a cross-sectional or epidemiological description of patients' quality of life that is linked to clinical sociodemographic or other psychosocial characteristics is important for making funding and access decisions. If the comparison is meant to concern hemophilia patients only, the choice of a hemophilia-specific measure might be sufficient, especially if one is interested in subdividing patient populations according to clinical characteristics. Nevertheless, if the purpose is to compare hemophilia patients with populations suffering from other chronic conditions, the use of a generic measure is a prerequisite. Studies focusing on following patients' HRQoL over time in terms of a naturalistic setting or in terms of longitudinal cohort studies, which might or might not involve an intervention, usually benefit from both types of assessments: targeted and disease-specific.

Several studies have postulated that changes can be detected most sensitively with disease-specific measures, but other studies have reported that such changes over time are also discernable with generic scales. A solution is to combine generic and targeted measures in a modular fashion. Some measures—especially the pediatric ones—have been designed that way, for example the

PEDsQL, the KINDL-R, and the DISABKIDS have a generic core and disease-specific modules. In hemophilia literature, the head-to-head comparison or combination of both generic and disease-specific scales for assessment of HRQoL or PPM has yet to be completed.

Controlled, randomized clinical trials include PRO measures as primary or secondary end points. If the measures are to be used as primary end points, the criteria for selecting the measure should include previous knowledge of the measures performance in other trials in the same population and psychometric evidence. Prior estimates of the sample size necessary to detect changes over time and differences between groups should be utilized.

If the clinical trial is multinational, it is important that the measure has been translated and psychometrically tested in the relevant cultural and linguistic contexts. In hemophilia, multinational patient recruitment is frequently necessary to accumulate a sufficient number of patients in studies. Thus, it is extremely helpful if psychometric analysis of measures have already been performed in different countries, so that an instrument can be used in hemophilic patients around the world.

This is, for example, the case for the recently developed quality of life measures in hemophilia, both for children (e.g., Haemo-QoL for children) and for adults (e.g., Haemo-QoL-A, Haem-A-QoL), and it is also true for a range of generic and targeted measures.

In children, specific problems arise regarding the age from which children can respond to a questionnaire. Usually, if appropriate forms of questioning and graphic reports are used, children from an age as early as 4 years can give information on their quality of life. Nevertheless, for PPM measures, good-level reading and writing and a generally well-developed cognitive functioning is expected. Another question pertains to whether parents (as proxies) should be included. Research on correlation between parent and child report shows, depending on the area of quality of life assessed, high correlations (for example for the physical function), but the emerging picture is complex and does not easily conclude that one mode of report may be substituted by the other.

Regarding the choice of measures, recommendations based on the literature review and intensive discussion within the Expert HRQoL Working Group of the IPSG expert group, are:

1. reflect on aim, design, patient, and scientific reasons for HRQoL assessment
2. always seek the opinion of the patient; observers provide valuable additional, but not identical information. Parent ratings for very young children are acceptable
3. use methodologically sound instruments providing interpretable scores
4. implement generic and targeted instruments conjointly as both pick up intervention effects; generic scales reflect HRQoL impact on the population level
5. if instrument alternatives are available, combine the short form of one with the long form of the other, to be used for item banking and further instrument development
6. use HRQoL and health economic instruments conjointly, if possible

Instruments can be used to derive utilities (SF-6D, EQ-5D, HUI, QWB), but so far, little is known about their responsiveness in hemophilia. The most important overall recommendation (was) is to include PRO instruments and to examine their properties in future studies. The reason for this general recommendation is the paucity of empirical results regarding the use of measures in clinical studies and trials so far. As has been demonstrated, generic and targeted measures for adults and children to assess PROs in terms of HRQoL and patient preferences are available. Nevertheless, they have not yet been sufficiently included in hemophilia studies, so that confirmation of psychometric properties in a specific study context, especially regarding specificity and responsiveness, is still lacking. Only the proactive and timely inclusion of these measures will aid solving urgent problems in prophylactic research over the patient's life span, such as when to start, when to stop, and when to modify the course of treatment recognizing its impact on patients' lives [103].

Discussion

In the area of PROs, generic and disease-specific instruments are available to assess HRQoL and patient preferences in adults and in children with hemophilia. These instruments were recently developed from patient input and can be therefore expected to assess relevant HRQoL domains in hemophilia. The studies were designed to psychometrically test these measures before they can be used in clinical studies. Such use will help determine which generic/disease-specific instrument is more sensitive to measure PROs in hemophilia across severity and over time to evaluate change. The majority of these measures were developed before the dissemination of the FDA PRO draft guidance. The most recently developed instrument (Haemo-QoL, CHO-KLAT, and Haem-A-QoL) have incorporated a patient-centric approach along with rigorous psychometric analyses throughout the development process, consistent with the FDA guidance. Comparison of the existing measures (such as the SF-36) with the FDA guidance has been completed by other investigators and is beyond the scope of this article. The current article sets out to present measures that can be used in a wide range of studies from clinical trials to observational studies, and is not specifically focused on integrating PRO impact into labels approved by regulatory agencies such as the European Medicines Agency (EMEA) or FDA.

Choosing between instruments can pose challenges to the individual researcher or clinician. The existence of different PRO instruments and their scarce use in research and clinical settings makes it difficult to obtain enough empirical data about specific instruments' performance, thereby preventing comparisons. Nevertheless, the availability of alternative measures is often necessary to determine those approaches that are conceptually appropriate, methodologically sound, and practically applicable. For example, a recent publication addresses the comparison of disease-specific measures HRQoL for children with hemophilia [72]. One of these measures, the CHO-KLAT, was developed in Canada with emphasis on the perspectives of children. Another, the Haemo-QoL, was developed in Europe, with emphasis on the perspectives of clinicians. Although these two measures are unique and independent, researchers from both studies were collaboratively linked throughout development and testing. The strengths, limitations, and unique contributions of these two measures have been compared using a retrospective analysis of data from field-testing of both measures [72]. The analysis included a comparative assessment of the basic validity, reliability, and items used in each measure. Overall, the results showed that the CHO-KLAT and the Haemo-QoL are promising and valuable measures of HRQoL for children with hemophilia. Analyses confirmed the basic psychometric properties of both tools, but identified some discrepancies between them. Additional data is expected to allow for greater understanding of these discrepancies and lend clarity as to how the tools should be used in clinical studies (separately or merged). The present recommendation is that the measures be run independently, but preferably concurrently in studies of children with hemophilia.

Combination of measures (i.e., a generic with a disease-specific HRQoL measure) might be helpful for researchers and clinicians, but patient burden has to be considered. To minimize respondent burden, one strategy to achieve optimal information on measurement performance would be to use core sets of instruments across studies. The development of short forms, which reflects the need to assess quality of life or patient preference concisely and parsimoniously, may help in this development. Clearly, it is beneficial to use common strategies for quality of life assessment in a rare disease such as hemophilia. Patient preferences enable investigators to quantify the level of impairment of an individual due to a specific disease state. The degree to which a patient values the improvement in HRQoL from a given treatment can also be assessed through preference (utility) scores. Utility measures have been integrated into outcomes and economic analysis across a myriad of clinical conditions including coronary artery disease, depression, and asthma [4,104]. Information concerning the relative costs of alternative interventions and their impact on outcomes of value to the patient are important for accurate estimates from these analyses. The evaluation of treatment approaches such as prophylaxis is a potential application of utility measurement in hemophilia.

Frequently, patient preference and quality of life measures are used independently. However, research, such as economic analyses, can be enriched by incorporating both methods. One possibility is to use the HRQoL instruments in health economic research. The second possibility is to use quality of life measures that in themselves make it possible to derive utility values. These include the SF-6D, derived from the SF-36 Health survey, as well as the EQ-5D. A third possibility is to further develop hemophilia-specific patient preference methodology. Wasserman et al. [91] developed an instrument to measure patient preferences for specific hemophilia health states. For clinical practice, the combination of both methods is critical because reimbursement policies are often based on both economic and quality of life data.

PPMs are a useful way to translate the benefits of prophylactic care into economic terms. This information is imperative to justify the high up front costs of prophylactic care from the perspective of the hemophilic patient. To date, completion of economic models has been difficult because of two issues. First, few studies have integrated a PPM, so utilities for various hemophilia health states at different periods of time are scarce. Research, such as that conducted by Wasserman et al. [91], as well as general measures used in current prospective studies in both the United States and Europe, will help to fill this gap. Second, although some cost studies have been conducted in the past decade, the detailed information needed to include the models is also limited. Recent developments in economic modeling techniques, as well as cost studies in the United States (e.g., Haemophilia Utilization Study Group [HUGS] V) [105,106] and in Europe (COCIS is the Italian version of cost of Care in Haemophilia [and COCHE]) [2,107], will provide the breadth of detail needed for these models. Measuring HRQoL in hemophilia clinical research will provide important data to integrate the benefit of economic treatment to the patient into economic analyses that are needed to justify reimbursement for prophylactic care throughout the world.

As repeatedly stated by proponents of evidence-based medicine, treatment choices should be based as much as possible on an empirical foundation, obtained via high-quality studies, focusing on results of well-designed prospective clinical trials and the meta-analysis or pooling of data. Informed decision-making, however, is only possible if such data exist, and in persons with hemophilia for quality of life as well as patient preference, this is not yet the case. Specific treatment scenarios such as implanting a port-a-cath in young children, identifying the best prophylaxis strategy, and changing treatment strategies in later life, are examples of hemophilia care delivery that would greatly benefit from quality of life and patient preference evaluation. These measures should be incorporated as much as possible into research to augment the data upon which to base the best choice of treatment strategies for patients with hemophilia. It is thus hoped that the use of PROs will lead to a better understanding of the benefits and limitations of prophylaxis strategies in persons with hemophilia, and would ultimately aid in clinicians and policymakers in delivering appropriate care to hemophilic patients by assessing the outcomes of treatment.

We would like to express our recognition to the IPSG Steering Committee: Dr. Victor S. Blanchette (chair), Dr. Louis M. Aledort (cochair), Dr. Rolf Ljung (cochair), Dr. Marijke van den Berg, Dr. Brian M. Feldman, Dr. Alessandro Gringeri, Dr. Marilyn Manco-Johnson, Dr. Pia Petrini, Dr. Georges E. Rivard, and Dr. Wolfgang Schramm, for their continuous scientific incentive and advice. As members of the IPSG HRQoL expert group, we want to thank our fellow members: Dr. Victor Blanchette for leading the IPSG and being responsible for bringing us together and providing the context for the ideas for the articles to be created; Dr. Lou Aledort for his guidance in enhancing the relevance of the citations to clinicians and policymakers; Dr. Wolfgang Schramm for serving as co-chair of the HRQoL Expert Group, and Dr. Hilary Llewellyn-Thomas for her helpful advice as a HRQoL group member. We would like to acknowledge the help of Audrey Abad and Marjorie McClimont for facilitating and coordinating our international collaboration and Anna Levke Brütt for helping to prepare the article. The work of the IPSG group is supported by external funding from Bayer HealthCare LLC, Biological Products Division, Baxter BioScience, CSL Behring, and Wyeth Pharmaceuticals administered through the Hospital for Sick Children Foundation, Toronto, Canada.

Source of financial support: The work of the IPSG group is supported by external funding from Bayer HealthCare LLC, Biological Products Division, Baxter BioScience, CSL Behring, and Wyeth Pharmaceuticals administered through the Hospital for Sick Children Foundation, Toronto, Canada. Members of the IPSG have not received any remuneration for work on this article.

References

1 Revicki DA, Gnanasakthy A, Weinfurt K. Documenting the rationale and psychometric characteristics of patient reported outcomes for labeling and promotional claims: the PRO evidence dossier. Qual Life Res 2007;16:717–23.
2 Gringeri A, Mantovani LG, Scalone L, Mannucci PM COCIS Study Group. Cost of care and quality of life in hemophilia complicated by inhibitors: the COCIS Study Group. Blood 2003;102:2358–63.
3 Patrick DL, Chiang YP. Measurement of health outcomes in treatment effectiveness evaluations: conceptual and methodological challenges. Med Care 2000;38(9 Suppl. 2):S14–25.
4 Patrick DL, Erickson P. Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation. New York: Oxford University Press, 1993.
5 Bullinger M. Quality of life—definition, conceptualization and implications. A methodologists view. Theor Surg 1991;6:143–8.
6 Spilker B. Revicki, D. Taxonomy of quality of life. In: Quality of Life Assessment in Clinical Trails. New York: Raven Press, 1996.
7 Feinstein AR. Clinimetrics. New Haven, CT: Yale University Press, 1987.
8 Marx RG, Bombardier C, Hogg-Johnson S, Wright JG. Clinimetric and psychometric strategies for development of a health measurement scale. J Clin Epidemiol 1999;52:105–11.
9 Spilker B. Quality of Life Assessment in Clinical Trails. New York: Raven Press, 1996.

10 Saleck S. Compendium of Quality of Life Instruments. Toronto, ON, Canada: John Wiley & Sons, 2004.
11 Bullinger M, Anderson R, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. Qual Life Res 1993;2:451–9.
12 Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186–91.
13 Torrance GW. Measurement of health state utilities for economic appraisal. J Health Econ 1986;5:1–30.
14 Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions. Health Utilities Index. Pharmacoeconomics 1995;7:503–20.
15 Arnesen T. Are TROMMALD M. QALYs based on time trade-off comparable?—A systematic review of TTO methodologies. J Health Econ 2005;14:39–53.
16 Lamers LM. The transformation of utilities for health states worse than death: consequences for the estimation of EQ-5D value sets. Med Care 2007;45:238–44.
17 Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. J Clin Epidemiol 1998;51:1115–28.
18 Beeton K. Evaluation of outcome of care in patients with haemophilia. Haemophilia 2002;8:428–34.
19 De Kleijn P, Heijnen L, Van Meeteren NL. Clinimetric instruments to assess functional health status in patients with haemophilia: a literature review. Haemophilia 2002;8:419–27.
20 Fischer K, van der Bom JG, van den Berg HM. Health-related quality of life as outcome parameter in haemophilia treatment. Haemophilia 2003;9(Suppl. 1):S75–81.
21 Bullinger M, von Mackensen S. Quality of life assessment in haemophilia. Haemophilia 2004;10(Suppl. 1):S9–16.
22 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;9:678–87.
23 Remor E, Young NL, von Mackensen S, Lopatina EG. Disease-specific quality-of-life measurement tools for haemophilia patients. Haemophilia 2004;10(Suppl. 4):S30–4.
24 Gringeri A, Mantovani L, Mackensen SV. Quality of life assessment in clinical practice in haemophilia treatment. Haemophilia 2006;12(Suppl. 3):S22–9.
25 Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims. Draft Guidance. February 2006 Available from: http://www.fda.gov/Cder/Guidance/5460dft.pdf [Accessed July 2, 2008].
26 Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: development and final revision of a health status measure. Med Care 1981;19:787–805.
27 Hunt SM, McKenna SP, McEwen J, et al. The Nottingham Health Profile: subjective health status and medical consultations. Soc Sci Med 1981;15:221–9.
28 Power M, Harper A, Bullinger M. The World Health Organization WHOQOL-100: tests of the universality of Quality of Life in 15 different cultural groups worldwide. Health Psychol 1999;18:495–505.
29 WHO QOL Group. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. Soc Sci Med 1998;46:1569–85.
30 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–83.
31 Ware JE. The SF-36 Health Survey. In: Spilker B, ed., Quality of Life and Pharmaeconomics in Clinical Trials. Philadelphia: Lippincott-Raven, 1996.
32 The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med 1998;28:551–8.
33 Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–33.
34 Rodriguez-Merchan EC. Health outcome assessment: SF-36 in English and Spanish. Haemophilia 2003;9:343–4.
35 Miners A, Sabin C, Tolley K, et al. Assessing health-related quality-of-life in individuals with haemophilia. Haemophilia 1999;5:378–85.
36 de-Joode E, van-Meeteren N, van-den-Berg H, et al. Validity of health status measurement with the Dutch Arthritis Impact Measurement Scale 2 in individuals with severe haemophilia. Haemophilia 2001;7:190–7.
37 Djulbegovic B, Goldsmith G, Vaughn D, et al. Comparison of the quality of life between HIV-positive haemophilia patients and HIV-negative haemophilia patients. Haemophilia 1996;2:166–72.
38 Khair K. Quality of life in children with haemophilia. Haemophilia 2000;6:419.
39 Bellamy N. Womac Osteoarthritis Index: A User's Guide. London, Ontario, Canada: University of Western Ontario, 1996.
40 Bellamy N, Buchanan WW, Goldsmith CH, et al. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
41 Meenan RF. The AIMS approach to health status measurement: conceptual background and measurement properties. J Rheumatol 1982;9:785–8.
42 Schick M, Stucki G, Rodriguez M, et al. Haemophilic; arthropathy: assessment of quality of life after total knee arthroplasty. Clin Rheumatol 1999;18:468–72.
43 Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982;9:789–93.
44 Poonnoose PM, Srivastava A. Functional assessment of arthropathy—an international perspective. Semin Hematol 2006;43(Suppl. 1):S27–32.
45 Ravens-Sieberer U, Erhart M, Wille N, et al. Generic health-related quality-of-life assessment in children and adolescents: methodological considerations. Pharmacoeconomics 2006;24:1199–220.
46 Landgraf JM, Abetz L, Ware JE. Child Health Questionnaire (CHQ): A User's Manual. Boston: The Health Institute Press, 1997.
47 Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care 1999;37:126–39.
48 Bullinger M, von Mackensen S, Kirchberger I. KINDL—ein fragebogen zur erfassung der lebensqualität von kindern. Zeitschrift für Gesundheitspsychologie 1994;2:64–77.
49 Ravens-Sieberer U, Bullinger M. Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results. Qual Life Res 1998;7:399–407.
50 Theunissen NC, Vogels TG, Koopman HM, et al. The proxy problem: child report versus parent report in health-related quality of life research. Qual Life Res 1998;7:387–97.
51 Ravens-Sieberer U, Gosch A, Abel T, et al. Quality of life in children and adolescents: a European public health perspective. Soz Präventivmed 2001;46:294–302.
52 The KIDSCREEN Group Europe. The KIDSCREEN Questionnaires- Quality of life questionnaires for children and adolescents. Lengerich, Germany: Pabst Science Publishers, 2006.
53 Bullinger M, Schmidt S, Petersen C. Assessing quality of life of children with chronic health conditions and disabilities: a European approach. Int J Rehabil Res 2002;25:197–206.
54 The DISABKIDS Group Europe. The DISABKIDS Questionnaires- Quality of Life Questionnaires for Children with Chronic Conditions. Lengerich, Germany: Pabst Science Publishers, 2006.
55 Ravens-Sieberer U, Erhart M, Bullinger M. The kidscreen and disabkids questionnaires- two measures for children and adolescent's health related Quality of life. Patient Reported Outcomes Newsl 2006;37:9–11.

56 Baars RM, Atherton CI, Koopman HM, et al. The European DISABKIDS project: development of seven condition-specific modules to measure health related quality of life in children and adolescents. Health Qual Life Outcomes 2005;3:70.

57 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;47:305–11.

58 Mackensen S, Bullinger M. The 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;10(Suppl. 1):S17–25.

59 Ekert H, Brewin T, Boey W, et al. Cost-utility analysis of recombinant factor VIIa (NovoSeven) in six children with long-standing inhibitors to factor VIII or IX. Haemophilia 2001;7: 279–85.

60 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;14:1023–34.

61 v. Mackensen S, Gringeri A, Santoni L, et al. Development and Pilot testing of a Disease-Specific Quality of Life Questionnaire for Adult Patients with Haemophilia (Haem-A-QoL). Blood 2004;104, abstract number 2214.

62 Gringeri A, v. Mackensen S. Quality of life in haemophilia. Haemophilia 2008;14(Suppl. 3):19–25.

63 Gringeri A, Mantovani L, von Mackensen S. Quality of life assessment in clinical practice in haemophilia treatment. Haemophilia 2006;12:1–8.

64 Chevallet L, Weatherall J, von Mackensen S. Linguistic validation of the Haemo-QoL and Haem-A-QoL for use in international studies. Value Health 2008;11:PSY44:A165.

65 Arranz P, Remor E, Quintana M, et al. Development of a new disease-specific quality-of-life questionnaire to adults living with haemophilia. Haemophilia 2004;10:376–82.

66 Remor E, The Hemolatin-QoL Group. Hemolatin-QoL: Desarrollo de una medida específica para la evaluación de la calidad de vida en pacientes adultos con hemofilia en America-Latina: el Hemolatin-QoL. Rev Interam Psicolog 2005;39:211–20.

67 Trudeau E, Dias C, Lambert T, et al. Development of the haemophilia age-group specific quality of life questionnaire. Value Health 2003;6:A744.

68 Mackensen S, Bullinger M, Ravens-Sieberer U, Negrini C. Development of the new HAEMO-QoL questionnaire for the assessment of health-related quality of life in children with haemophilia. Qual Life Res 2000;9:302.

69 Bullinger M, von Mackensen S, Fischer K, et al. Pilottesting of the Haemo-QoL quality of life questionnaire for haemophiliac children in six European countries. Haemophilia 2002;8(Suppl. 2):S47–54.

70 Young NL, Bradley CS, Blanchette V, et al. Development of a health related quality of life measure for boys with haemophilia: the Canadian Hemophilia Outcomes – Kids Life Assessment Tool (CHO-KLAT). Haemophilia 2004;10(Suppl. 1):S34–43.

71 Pollak E, Muhlan H, von Mackensen S, et al. The Haemo-QoL Index: developing a short measure for health-related quality of life assessment in children and adolescents with haemophilia. Haemophilia 2006;12:384–92.

72 Bradley CS, Bullinger M, McCusker PJ, et al. Comparing two measures of quality of life for children with haemophilia: the CHO-KLAT and the Haemo-QoL. Haemophilia 2006;12:643–53.

73 Manco-Johnson M, Morrissey-Harding G, Edelman-Lewis B, et al. Development and validation of a measure of disease-specific quality of life in young children with haemophilia. Haemophilia 2004;10:34–41.

74 Robben W, Plug I, DeGoede-Bolder A, et al. The development and evaluation of a quality of life questionnaire for children with hemophilia. J Thromb Haemost 2003;1(Suppl.1): SP0634.

75 Anderson JP, Kaplan RM, Berry CC, et al. Interday reliability of function assessment for a health status measure. The quality of well-being scale. Med Care 1989;27:1076–83.

76 The EuroQoL Group. EuroQoL—a new facility for the measurement of health-related quality of life. Health Policy 1990;16: 199–208.

77 Brooks R. EQ-5D, the current state of play. Health Policy 1996;37:53–72.

78 Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Health Utilities Index. Pharmacoeconomics 1995;7:490–502.

79 EuroQoL Group. EuroQo L—a new facility for the measurement of health-related quality of life. Health Policy 1990;16: 199–208.

80 van Agt HME, Essink-Bot ML, Krabbe PFM, Bonsel GJ. Test-retest reliability of health state valuations collected with the EUROQOL questionnaire. Soc Sci Med 1994;39:1537–44.

81 Brazier J, Jones N, Kind P. Testing the validity of the EuroQoL and comparing with the SF-36 health survey. Qual Life Res 1993;2:169–80.

82 Hurst NP, Jobanputra P, Hunter M, et al. Validity of EuroQoL—a generic health status instrument—in patients with rheumatoid arthritis. Economic and Health Outcomes Research Group. Br J Rheumatol 1994;33:655–62.

83 Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. J Health Econ 2004;13:873–84.

84 Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. J Health Econ 2005;14:1169–89.

85 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;40:113–28.

86 Feeny D, Wu L, Eng K. Comparing short form 6D, standard gamble, and Health Utilities Index Mark 2 and Mark 3 utility scores: results from total hip arthroplasty patients. Qual Life Res 2004;13:1659–70.

87 Boyle MH, Furlong W, Feeny D, et al. Reliability of the Health Utilities Index—Mark III used in the 1991 cycle 6 Canadian General Social Survery Health Questionnaire. Qual Life Res 1995;4:249–57.

88 Franks P, Gold M, Erickson P. Do utility-based measures of health-related quality of life predict future health states? Longitudinal evidence from a nationally representative cohort. Presented at the Society of Medical Decision Making. Portland, 1992. University of Rochester, Department of Family Medicine, Technical Report.

89 Miners AH, Sabin CA, Tolley KH, Lee CA. Cost-utility analysis of primary prophylaxis versus treatment on-demand for individuals with severe haemophilia. Pharmacoeconomics 2002;20: 759–74.

90 Sung L, Young NL, Greenberg ML, et al. Health-related quality of life (HRQL) scores reported from parents and their children with chronic illness differed depending on utility elicitation method. J Clin Epidemiol 2004;57:1161–6.

91 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.

92 Barr RD, Saleh M, Furlong W, et al. Health Status and health-related quality of life associated with haemophilia. Am J Hematol 2002;71:152–60.

93 Wang T, Zhang L, Li H, et al. Assessing health-related quality-of-life in individuals with haemophilia in China. Haemophilia 2004;10:370–5.

94 Barr RD, Sek J, Horsman J, et al. Health status and health-related quality of life associated with von Willebrand disease. Am J Hematol 2003;73:108–14.

95 Revel-Vilk S, Golomb MR, Achonu C, et al. Effect of intracranial bleeds on the health and quality of life of boys with hemophilia. J Pediatr 2004;144:490–5.

96 Fischer K, Astermark J, Van-Der-Bom J, et al. Prophylactic treatment for severe haemophilia: Comparison of an intermediate-

dose to a high dose regimen. Haemophilia 2002;8:753–60.
97 Szucs TD, Offner A, Kroner B, et al. Resource utilisation in haemophiliacs treated in Europe: results from the European Study on Socioeconomic Aspects of Haemophilia Care. The European Socioeconomic Study Group. Haemophilia 1998;4: 498–501.
98 Schramm W, Royal S, Kroner B, et al. Clinical outcomes and resource utilization associated with haemophilia care in Europe. Haemophilia 2002;8:33–43.
99 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.
100 Schramm W, Berger K. Linking medicine and economics: Health economics and quality of life in haemophilia care. Haemophilia 2002;8:217–20.
101 Patrick DL, Ferketich SL, Frame PS, et al. National institutes of health state-of-the-science panel. national institutes of health state-of-the-science conference statement: symptom management in cancer: pain, depression, and fatigue, July 15–17, 2002. J Natl Cancer Inst 2003;95:1110–17.
102 Fifer SK, Buesching DP, Henke CJ, et al. Functional status and somatization as predictors of medical offset in anxious and depressed patients. Value Health 2003;6:40–50.
103 Berntorp E, Astermark J, Bjorkman S, et al. Consensus perspectives on prophylactic therapy for haemophilia: summary statement. Haemophilia 2003;9(Suppl. 1):S1–4.
104 Say RE, Thomson R. The importance of patient preferences in treatment decisions—challenges for doctors. Br Med J 2003; 327:542–5.
105 Wu J, Nichol M, Globe D, et al. Estimating utility values for self-reported joint pain and motion limitation outcomes in adult hemophilia patients-the Hemophilia Utilization Group Study-Part V (HUGS-V). Value Health 2007;10:PHM8.
106 Globe DR, Curtis R, Koerper M. HUGS Steering Committee. Utilization of care in hemophilia: a resource-based method for cost analysis from the hemophilia utilization group study (HUGS). Haemophilia 2004;10(Suppl. 1):S63–70.
107 Scalone L, Gringeri A., Mackensen S, Mantovani L. Cost and quality of life in hemophilic patients without inhibitors: The COCHE Study (Poster). Haematologica 2006;91(Suppl. 1): abstract number 0412.