ORIGINAL ARTICLE *Clinical haemophilia*

The association of haemophilic arthropathy with Health-Related Quality of Life: a *post hoc* analysis

K. FISCHER,* P. DE KLEIJN,* C. NEGRIER,†‡ E. P. MAUSER-BUNSCHOTEN,* P. R. VAN DER VALK,* K. P. M. VAN GALEN,* A. WILLEMZE* and R. SCHUTGENS*

Van Creveldkliniek; †Department of Rehabilitation, Nursing Science and Sports, University Medical Center Utrecht, Utrecht, the Netherlands; and ‡Unité d'Hémostase Clinique, Centre Régional de traitement de l'Hémophilie, Lyon, France

Background: The aim of replacement therapy in haemophilia is to improve Health-Related Quality of Life (HRQoL) by preventing bleeding and arthropathy. However, the association of arthropathy with HRQoL is unknown. **Aim:** To explore the association of haemophilic arthropathy with HRQoL. **Methods:** A *post hoc* analysis on patients with severe/moderate haemophilia with SF36 questionnaire (SF36) and X-rays of ankles, knees and elbows made within 2.5-years. The SF36 scores of 'physical functioning' (SF36-PF, range 0–100, optimum 100) and Utility (SF6D-Utility, range 0–1, optimum 1) and radiological Pettersson scores (PS, range 0–78, optimum 0) were calculated. The association of PS with reduced SF6D-Utility and SF36-PF (<age-specific normal values) was analysed using ROC analyses and multivariable logistic regression. **Results:** Overall, 176 assessments were analysed: 130 from the Van Creveldkliniek and 46 from a French multicentre study. Most patients had severe haemophilia (89.9%), evaluated at 26.6 years, and with a range 15.7–65.8. Overall PS median (interquartile range) was 16 (7–34), SF6D-Utility was 0.76 (0.64–0.86) and SF36-PF was 85 (60–95). Receiver operating curve analysis identified a threshold PS of 21 points for both SF6D-Utility (AUC 0.65) and SF36-PF (AUC 0.76). In patients with PS > 21 points, the risk of reduced SF6D-Utility was stable (OR 4.16; 95% CI: 2.03–8.51) but SF36-PF continued to decrease: compared to lowest PS, OR for reduced SF36-PF was 5.69 (95% CI: 1.62–20.06) for PS 22–39 and 25.15 (95% CI: 6.53–96.81) for PS 40–78. **Conclusion:** Health-Related Quality of Life only showed a significant deterioration in patients with a Pettersson score of >21 points. This suggests that HRQoL is relatively insensitive to early joint changes.

Keywords: arthropathy, haemarthrosis, Quality of Life, radiograph, utility

Introduction

Repeated spontaneous or traumatic bleeding is the hallmark of haemophilia. In the absence of prophylaxis, patients with severe haemophilia (FVIII/IX < 0.01 IU mL^{-1}) experience repeated joint bleeds, most frequently in the ankle, knee and elbow joints [1]. Repeated joint bleeding ultimately results in chronic pain, functional limitations and crippling deformities causing loss of quality of life. Prophylactic replacement therapy is very effective in preventing bleeds and reduces or even prevents subsequent haemophilic arthropathy [2,3]. Due to the high costs of

concentrates, prophylaxis is associated with annual costs ranging from 180×10^3 to 298×10^3 USD in adult patients [4]. High costs have limited the widespread introduction of prophylactic replacement therapy, and have induced queries into the cost-effectiveness of prophylactic replacement therapy in countries where prophylaxis is widely used [3,5]. Health-economic evaluation of prophylactic treatment in severe haemophilia is notoriously difficult due to the low number of patients, variability in treatment regimens and the lifelong perspective [6,7]. For health-economic evaluation from the societal perspective, the method of choice is a cost-utility analysis calculating the cost per Quality Adjusted Life Year (QALY) gained [8]. To obtain QALYs, Health-Related Quality of Life (HRQoL) is expressed as a utility value between zero (death) and one (perfect health). Utilities can be measured using several questionnaires, of which the Euroqol has been most widely used [9,10]. Alternatively, utilities can be calculated from the

Correspondence: Kathelijn Fischer, Van Creveldkliniek, University Medical Center Utrecht, Room HP C01.425, PO Box 85500, 3508 GA Utrecht, the Netherlands.
Tel.: +31 88 755 8450; fax: +31 88 755 5438;
e-mail: k.fischer@umcutrecht.nl

widely used generic Short Form 36 (SF36) questionnaire [11,12]. In patients with haemophilia, utilities have been associated with age, diagnosis, bleeding frequency and treatment, but not yet with other parameters of long-term outcome, such as radiologic arthropathy [13–15]. Establishing the association of HRQoL, including utility, with a measure of arthropathy would provide important information for the interpretation of reported treatment results as well as provide input for health-economic models [16–19].

The aim of the present project is to study the association of HRQoL, specifically utility, with haemophilic arthropathy on conventional radiology.

Methods and scope of the analysis

The present analysis is a *post hoc* analysis of five completed studies including X-rays, and SF36 questionnaires administered to patients with severe [4,20–23] or moderate [24] haemophilia. Data on patients treated at the Van Creveldkliniek in the University Medical Center Utrecht, the Netherlands, were combined with data from a French cross-sectional multicentre study [22,23]. Ethical approval was obtained for all individual studies, as well as separate approval of the ethics committee of the University Medical Center Utrecht for this *post hoc* analysis. Inclusion criteria for this study were: severe or moderate haemophilia A or B, minimum age of 16 years at the time of completing the SF36, and X-rays of all six main joints taken within 2.5 years of completing the SF36. All X-rays were scored according to Pettersson [25]. This is a validated radiological scoring system assessing osteochondral changes in knees, elbows and ankles [26]. Pettersson scores at patient level represent the sum of the six joints (range 0–78; higher scores indicate more joint damage). Joints with arthrodesis or arthroplasty were given the maximum score of 13 points. All Dutch X-rays were scored by one radiologist, and French X-rays were scored centrally by two French radiologists [22]. The SF36 questionnaire assesses HRQoL over eight domains: physical functioning (SF36-PF), physical role limitations, bodily pain, general health, social functioning, emotional role limitations, mental health and vitality [11,27]. Each domain is determined by 1–10 questions and is scored from 0 to the optimum of 100 points. The SF36-PF measures physical activities: i.e. walking stairs, walking distances and carrying groceries. An algorithm for conversion SF36 scores into utility values (SF6D, range 0.0–1.0) has been developed by Brazier *et al.* [12]. Patient and treatment characteristics were extracted from the databases [4,20,21,24]. All patient characteristics were collected to reflect the time point of assessment of the SF36. Hepatitis C (HCV) infection and human immunodeficiency virus (HIV) infection were considered present in case of current positive viral

antigen. Current prophylaxis was defined as regular infusions with FVIII, with a minimum of twice weekly infusions for haemophilia A, and once weekly for haemophilia B, for a minimum of 46 weeks year^{-1}. Disability allowances (i.e. compensation for the inability to perform paid work) were considered positive if any allowance was received at the time of completing the SF36.

Statistical analysis

For descriptive analyses, medians and interquartile ranges (IQR: P25–P75) or proportions were calculated. In addition, means and standard deviations were calculated for the SF6D-Utility and SF36-PF scores.

First, SF6D-Utility scores were classified as 'lower than normal' (below the lower limit of the 95% confidence interval (CI) of the reference value according to age [28]) or 'normal' (higher or equal to the lower limit of this 95% CI). Thresholds used were: <0.852 for age 16–19; <0.830 for age 20–24; <0.825 for age 25–29; <0.819 for age 30–34; <0.818 for age 35–39; <0.812 for age 40–44; <0.802 for age 45–49; <0.783 for age 50–54; <0.793 for age 55–59; <0.771 for age 60–64; <0.784 for age 65–69 [28]. Second, SF36-PF scores were classified as 'lower than normal' (below age-specific mean values minus one standard deviation for Dutch men [29]) or 'normal' (for all values equal or higher than the age-specific mean-1SD). Threshold used were <76.8 for age 16–40; <62.4 for age 41–60; <27.1 for age 61–70 [29]. The association of total Pettersson scores (i.e. the sum score for all six joints) with SF36-derived utility (SF6D-Utility) was studied. Graphical exploration suggested the presence of a 'threshold' in Pettersson scores associated with lower SF6D-Utility. Subsequently, this threshold was studied using receiver operating curve (ROC) analysis. This was performed for both SF6D-Utility 'below normal' and the SF36-PF 'below normal'. The threshold value of the Pettersson score was identified as the score with the maximum sum of sensitivity and specificity for both curves combined, when considering this score as the threshold. For descriptive analyses, the Pettersson scores were divided into five groups of approximately equal size, around the threshold value. Subsequently, univariable and multivariable logistical regression analysis was performed to assess the association of a Pettersson score (in two or five categories) with SF6D-Utility, independent of age at SF36 assessment, current use of prophylaxis, presence of inhibitory antibodies, presence of chronic HCV infection, HIV status, haemophilia type, haemophilia severity, and country of origin. All multivariable regression analyses were performed in a stepwise backwards manner (entry criterion $P < 0.016$; exit criterion $P > 0.10$). To provide more detailed information on the association of

Pettersson scores with SF6D-Utility, multivariable linear regression was performed using SF6D-Utility as dependent (Tables S4, S5).

To assess the association of Pettersson scores with physical functioning, multivariable logistic regression was repeated using the SF36 domain score for 'physical functioning' (SF36-PF) 'below normal' as dependent variable, and Pettersson scores in two and five categories.

To assess the potential effect of repeated observations within patients, regression analyses were repeated using only the most recent observation for each patient. To address the potential effects of including data from two different countries using different primary treatment strategies [23], the ROC analyses were repeated for Dutch and French patients only, and multivariable regression analyses included adjustment for 'French origin'.

Results

Data sources and patient characteristics

In total, data on 176 Pettersson scores were extracted from five databases. For 130 assessments in Dutch patients, Pettersson scores originated from routine assessments, questionnaires were administered between 1997 and 2006 for a large European study ($n = 60$) [20], during a nationwide survey ($n = 24$) [21], for a Dutch-Swedish comparative study ($n = 37$) [4] and during a study on moderate haemophilia ($n = 14$) [24]. All studies included either consecutive patients recruited during clinic visits [20,22] or aimed to include full birth cohorts [4,21,24], patient selection was based on availability of SF36 and routinely performed X-rays of the six main joints. In total, 116 patients with severe haemophilia (median age 30.1 years), and 14 patients with moderate haemophilia (median age 38.1 years) from the Van Creveldkliniek and 46 French patients (41 severe, median age 21.7 years) [22] were studied. The eventual dataset included 176 assessments in 159 patients; 9 patients had two assessments, 4 patients had three.

Patient- and treatment characteristics, as well as outcome according to country and severity are presented in Table 1. The median age at the time of assessment was 26.6 years, ranging from 15.7 to 65.8 years. The majority of assessments were in patients with severe haemophilia ($n = 158$, 89.9%), and 84.7% with haemophilia A. The median year of birth was 1973, (range 1941–97); only 25% of patients were born before 1965, when clotting factor replacement became available. A slight majority of 56.3% was receiving prophylactic replacement therapy at the time of evaluation. Prophylaxis was only used in 14% of Dutch patients with moderate haemophilia and 15% of French patients, compared to 78% of Dutch severe patients. Data included 10 assessments in Dutch severe patients who discontinued prophylaxis, and only three in patients with a current inhibitor.

The majority of patients (68.8%) had chronic HCV, and 19.3% were HIV positive. About one-third of patients (37.5%) had a history of any joint surgery, and 19.9% currently received a disability allowance.

Radiological joint outcome in the patients studied included the full range of haemophilic arthropathy (range 0–78), but the majority had limited arthropathy, with a median Pettersson score of 16 points. The interval between completing the SF36 questionnaire and X-ray evaluation was 0.1 years (IQR 0.0–1.0). HRQoL measurement with the SF36 showed a median Utility of 0.76 (range 0.47–1.00) with 61% of patients reporting SF6D-Utility values below the age-specific reference value, especially in French patients (87%, vs. 53% and 43% in Dutch patients). Reduction in 'physical functioning' (SF36-PF) scores was more limited with a median score of 85 points, and 37% below the age-specific reference value and similar proportions between Dutch and French patients. Details of other SF36 domain scores are provided in Supplemental material (Table S1).

Receiver operating curve analysis showed that the predictive value of Pettersson scores on SF6D-Utility (AUC 0.653; 95% confidence interval (CI) 0.570–0.736) was lower than the predictive value of SF36-PF (AUC 0.759; CI: 0.683–0.835). A Pettersson score of 21.5 points had the highest sensitivity and specificity and was therefore used as threshold value in the analyses. Analyses performed separately in Dutch and French data yielded the same threshold value.

Several patient characteristics were associated with increasing Pettersson scores (Table 2). Especially in the categories with Pettersson scores over 21 points, patients were older and had more disability allowances. Use of prophylaxis, HCV- and HIV serology showed less association with arthropathy. SF6D-Utility showed a steady but limited deterioration with increasing Pettersson scores, from a median of 0.83 to a median of 0.72 (i.e. −0.11/0.83, a reduction of 13%). The association of the SF36 domain scores for 'physical functioning', however, showed a much stronger deterioration from a median of 100 in the group with lowest Pettersson scores to 50 points in the group with highest Pettersson scores (a reduction of 50%). For the other SF36 domain scores, only the domain of physical role limitations showed a considerable change from median scores of 100 to 63 points.

The association of Pettersson scores with SF6D-Utility

Figure 1 shows that SF6D-Utility only showed limited reduction with increasing Pettersson scores with the strongest reduction in patient with Pettersson scores

Table 1. Patient characteristics and outcome according to country and severity.

	All patients	Dutch-severe	Dutch-moderate	French
No. of patients	176	116	14	46
Hem A (%)	149 (84.7%)	84%	78%	89%
Severe haemophilia	159 (90.3%)	100%	0%	42 (91%)
Age (yrs)	26.6 (20.7–36.0)	30.1 (23.4–39.5)	38.1 (26.4–53.4)	21.7 (18.7–25.3)
Born < 1965	44 (25%)	33%	43%	0%
Current prophylaxis	99 (56.3%)	78%	14%	15%
Chronic HCV	121 (68.8%)	67%	36%	83%
HIV positive	34 (19.3%)	10%	0%	50%
Disability allowance	37 (21.0%)	10%	0%	11%
Pettersson score.	16 (7–34)	19 (8–44)	5 (1–13)	17 (8–28)
SF6D-Utility (max 1.00)	0.76 (0.64–0.86)	0.80 (0.68–0.88)	0.83 (0.74–0.93)	0.69 (0.59–0.79)
	0.76 (0.13)*	0.78 (0.12)*	0.81 (0.12)*	0.69 (0.12)*
SF6D-Utility < reference	108 (61.4%)	53%	43%	87%
SF36-physical	85 (60–95)	85 (50–95)	83 (64–100)	80 (65–90)
functioning (SF36-PF)	75 (24)*	74 (26)*	79 (18)*	75 (21)*
SF36-PF < reference	65 (36.9%)	37%	21%	43%

Numbers (%) or medians (P25–P75) and *means (standard deviation).

Table 2. Patient characteristics and outcome, overall and according to Pettersson scores.

		According to Pettersson scores				
	All assessments	0–4	5–12	13–21	22–39	40–78
No. of assessments	176	33	35	36	37	35
Hem A (%)	149 (84.7%)	82%	86%	81%	89%	86%
Severe haemophilia	158 (89.8%)	76%	86%	89%	97%	100%
Age (yrs)	26.6 (20.7–36.0)	20.9 (18.2–26.5)	20.6 (18.5–27.3)	26.1 (22.8–30.1)	29.1 (24.4–35.0)	43.1 (37.7–47.4)
Born < 1965	44 (25%)	6%	0%	14%	27%	77%
Current prophylaxis	99 (56.3%)	49%	60%	56%	51%	66%
Chronic HCV	121 (68.8%)	52%	57%	67%	89%	77%
HIV positive	34 (19.3%)	3%	23%	25%	30%	14%
Disability allowance	37 (21.0%)	3%	3%	14%	30%	54%
Pettersson score	16 (7–34)	2 (1–3)	8 (7–11)	16 (14–18)	31 (26–34)	56 (45–67)
SF6D-Utility	0.76 (0.64–0.86)	0.83 (0.68–0.90)	0.83 (0.74–0.89)	0.79 (0.70–0.83)	0.73 (0.60–0.82)	0.72 (0.60–0.81)
	0.76 (0.13)*	0.80 (0.13)*	0.80 (0.12)*	0.77 (0.13)*	0.70 (0.12)*	0.71 (0.11)*
SF6D-Utility < reference	108 (61.4%)	48%	49%	53%	78%	77%
SF36-physical functioning	85 (60–95)	95 (85–100)	90 (74–100)	90 (80–95)	75 (43–85)	50 (39–66)
(SF36-PF)	75 (24)*	90 (14)*	84 (19)*	83 (20)*	67 (22)*	51 (22)*
SF36-PF < reference	65 (36.9%)	12%	26%	19%	50%	79%

Numbers (%) or medians (IQR = P25–P75); Values marked with *represent mean (standard deviation).

over 21 points. To study the association Pettersson scores with SF6D-Utility in more detail, the Pettersson scores were analysed in the same five categories (Table 3, Model 1). The proportion of patients with an SF6D-Utility below the age-specific reference value was significantly increased in patients with Pettersson scores over 21 points (appr 78% vs. appr 50%, $P < 0.001$). Multivariable regression showed that Pettersson scores of over 21 points were associated with a significant decrease in SF6D-Utility compared to the lowest category of 0–4 points, with similar Odds Ratios of 3.05 (95% CI: 1.02–9.18) and 4.50 (95% CI: 1.53–13.28) for Pettersson scores of 22–39 and 40–78 points respectively. Although all patient characteristics were included in the multivariable model, besides Pettersson scores only French origin showed an independent statistically significant association with SF6D-Utility. Details of the univariable and multivariable regression analyses as well as analyses of the

Pettersson scores in two categories are provided in the Supplemental data (Table S2–S5).

The association of Pettersson scores with the SF36 domain of 'physical functioning'

The SF36 domain of Physical functioning (SF36-PF) represents the ability to perform activities such as walking distances, walking stairs and carrying objects. Figure 2 shows SF36-PF scores according to Pettersson scores. Again, the SF36-PF appeared to decrease more steeply in patients with a Pettersson score of over 21 points, and continued to decrease in patients with Pettersson scores over 39 points. This was also reflected in the steady increase of the proportion of patients with SF36-PF below the age-specific reference value (Table 2). Multivariate regression analysis of the association of arthropathy with SF36-PF score according to categories of the Pettersson score, showed a
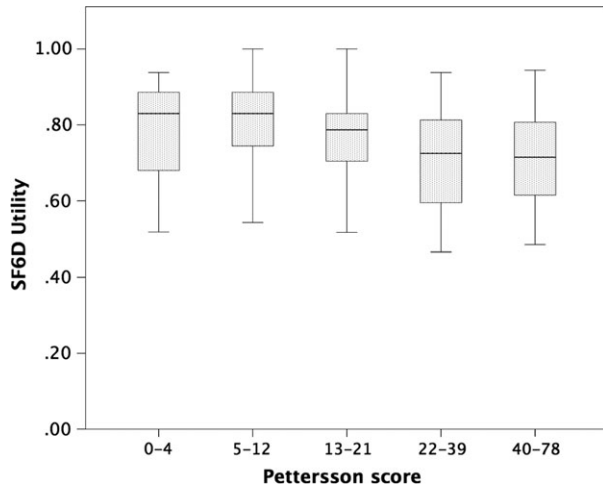
Fig. 1. SF6D-Utility according to Pettersson scores. Subscript to Figure 1: Multivariable regression showed a significantly lower SF6D-Utility in patients with Pettersson scores > 21 points, but no additional decrease in Utility in patients with Pettersson scores of 40–78 points.
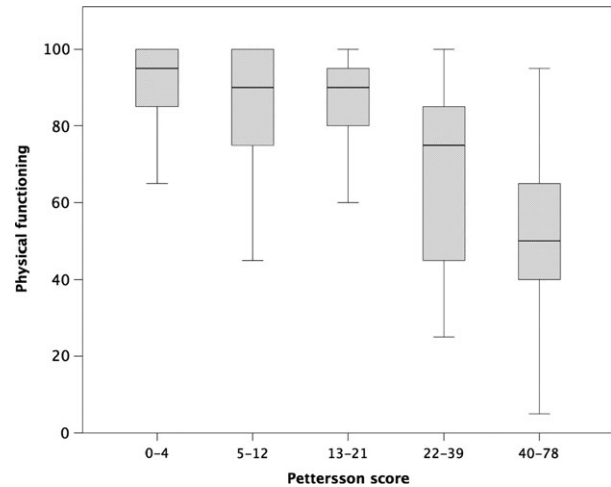
Fig. 2. SF36 Physical Functioning according to Pettersson scores. Subscript to Figure 2: Multivariable regression showed a significantly lower SF36-PF in patients with Pettersson scores > 21 points and a progressive reduction of SF36-PF scores with increasing Pettersson scores.

progressive and statistically significant decrease in SF36-PF scores with odds ratios of 5.69 (95% CI: 1.62–20.06) for PS 22–39 and 25.15 (95% CI: 6.53–96.81) for PS 40–78 (Table 3, Model 2). Although all patient characteristics were included in the multivariable model, we observed that besides Pettersson scores, only current HCV infection showed an independent statistically significant association with SF36-PF; details on the regression analysis are provided in the supplemental data (Table S3).

Sensitivity analyses consisting of repeating the analyses of both models after exclusion of repeated assessments (159 unique patients) yielded similar findings (data on request).

Discussion

This *post hoc* analysis of 176 assessments in adults with severe and moderate haemophilia reports on the association of HRQoL with haemophilic arthropathy. SF6D-Utility showed a limited deterioration with increasing Pettersson scores, from a median of 0.83 to a median of 0.72 (−0.11, i.e. −13%), while the SF36 domain score for 'physical functioning' showed a much stronger deterioration from a median of 100 to 50 points (−50%). Multivariable regression analyses showed that Health-Related Quality of Life (SF6D-Utility and SF36-PF) was significantly lower in patients with Pettersson scores over 21 points only.

Study design

The present analysis is a *post hoc* analysis, which is associated with an increased risk of measurement errors and/or selection bias. Dutch X-rays scores were scored by a single radiologist, while French X-rays were scored by two radiologists. This may have added some inter-rater discrepancies to the analysis. As more

Table 3. Multivariate logistic regression models describing the association of arthropathy with HRQoL.

Parameter	Univariable regression			Multivariable regression		
	Odds ratio	95% confidence interval	*P*-value	Odds ratio	95% confidence interval	*P*-value
Model 1:						
Pettersson scores and SF6D-Utility below age-specific reference value						
Pettersson score						
5–12 points vs. 0–4 points	1.00	0.39–2.60	0.99	0.66	0.23–1.91	0.44
13–21 points vs. 0–4 points	0.72	0.46–3.06	0.72	1.06	0.39–2.93	0.91
22–39 points vs. 0–4 points	3.85	1.36–10.88	0.01	3.05	1.02–9.18	0.05
40–78 points vs. 0–4 points	3.59	1.26–10.18	0.02	4.50	1.53–13.28	<0.01
Model 2:						
Pettersson scores and SF36 domain 'physical functioning' below age-specific reference value						
Pettersson score						
5–12 points vs. 0–4 points	2.61	0.76–9.51	0.15	2.52	0.68–9.30	0.17
13–21 points vs. 0–4 points	1.75	0.46–6.63	0.41	1.58	0.41–6.06	0.51
22–39 points vs. 0–4 points	7.25	2.11–24.87	0.002	5.69	1.62–20.06	<0.01
40–78 points vs. 0–4 points	27.96	7.35–106.33	<0.001	25.15	6.53–96.81	<0.001

Multivariable regression adjusted for age, HCV, HIV prophylaxis, inhibitor status, haemophilia, severity and country of origin.

SF36 than Euroqol (EQ5D) questionnaires were available in our cohort, SF6D-Utility values were used in this analysis. In general, the SF6D-Utility is better at discriminating between people with relatively mild health impairment, and the EQ5D is better at discriminating between those with more severe morbidity [30]. In this population with limited joint disease, the SF6D-Utility may be therefore the most appropriate tool for Utility assessment.

Selection bias was avoided by the use of SF36 data from several cohort studies aiming at including either full cohorts or consecutive patients visiting the clinic. The selection of patients for the present analysis was dependent on the availability of X-rays within 2.5 years from SF36 assessment. This was random, as the present analysis was not planned at the time and X-rays are performed at 5-year intervals in all Dutch patients with severe haemophilia and 7-year intervals in those with moderate haemophilia at the Van Creveldkliniek. The cohorts included both haemophilia A and B, but (ex-) inhibitors were under-represented as they were excluded for the majority of the original studies. The cohort of patients with severe haemophilia born between 1965 and 94 participated in two studies, resulting in repeated assessments in 13/159 patients (8.2%) and 17/176 assessments (9.7%). However, additional analyses excluding the repeated assessments showed similar results.

In the multivariable analysis, the association of HRQoL with arthropathy was independent of haemophilia severity and age. The effect of age was excluded by the use of age-specific reference values for both logistic regression analyses. Although the full range of Pettersson scores was included in this study, the patients with extensive arthropathy were under-represented. However, for the present population, the number of observations above the threshold value of 21 points was sufficient to perform multivariable regression analyses and to pick up the deterioration with higher Pettersson scores. Another striking characteristic of the population studied is the low proportion of patients on prophylaxis. Especially in the younger patient group this was unexpected, as early prophylaxis is the standard of care at the Van Creveldkliniek. This can be explained by the inclusion of patients with moderate haemophilia, French patients who were mostly treated on demand, and some patients who stopped prophylaxis in adulthood. Again, multivariable regression showed that the association of HRQoL with arthropathy was independent of prophylactic treatment.

Comparison with others

The data of this study partly overlap with those from a previous report on the association of Pettersson scores with SF36 in 96 patients from the Van Creveldkliniek. That study showed that a Pettersson score of ≥28 points was associated with decreased SF36 scores and loss of labour force participation [31]. This study includes 80 additional assessments (total 176), presents Utility values, and provides more extensive analyses including adjustment for other parameters than age only. The current results suggest that the 'threshold' Pettersson score associated with a steep decrease in HRQoL is >21 points. We could identify no other reports on the association of radiological arthropathy with Utility, but the association of clinical joint scores with the SF36-PF has been reported in 70 Spanish patients [32]. Although a deterioration of SF36-PF with increasing joint scores was reported, a threshold was not explored. A Dutch nationwide study ($n = 603$) reported univariate associations of SF36-PF with self-reported joint health, HIV, HCV and age, but objective parameters of joint health were not included [33].

So far, only two studies have presented SF6D-Utility in haemophilia; both showed lower Utilities than the overall value of 0.76 reported in this study. In a large European study ($n = 500$, age 14–83), Lippert *et al.* [13] reported a Utility of 0.73–0.76 for patients <30 years, and 0.66–0.68 for patients aged 30 years and over. In Belgium, Carvalhosa *et al.* [34] reported SF6D-Utility of 0.63 in patients with severe haemophilia (mean age 47 years), which showed a strong association with arthropathy.

The association of HRQoL scores with age, HIV, HCV, haemophilia type, prophylaxis and haemophilia severity is well established [13,14,20,30], and was corroborated by this study.

Clinical relevance and future research

Although the direct clinical relevance of radiological joint changes is limited, its association with HRQoL is highly relevant, especially in the context of choosing the intensity of prophylactic treatment; should the aim be to prevent all bleeding or can a limited number of joint bleeds be accepted [35]. In addition, it corroborates the observation that HRQoL is impacted by joint disease, rather than treatment received [36]. Both regression analyses showed that HRQoL was only significantly decreased in patients with a Pettersson score over 21 points. Repeating the present analysis in a larger number of patients will not only allow the evaluation of smaller increments in Pettersson score, but also enable detection of smaller reductions in HRQoL. However, the mean SF6D-Utility reduction of 0.03 for patients with Pettersson scores of 13–21 points was still below the minimally important difference of 0.041 established for the SF6D [37], while the reduction for patients with higher Pettersson scores was twice as high. The reduction of SF36-PF scores was more pronounced, and although not statistically

significant, patients with Pettersson scores of 5–21 points showed a SF36-PF reduction of 6–7 points, which is over the minimally important difference of 5 points for the SF36 [27]. However, in the current analysis, a very pronounced and statistically significant deterioration was observed for those with Pettersson scores >21 points only. This is very interesting from a clinical point of view: it reflects that although patients were limited in some of their physical activities (measured by SF36-PF), this does not immediately affect activities of daily living and/or role limitations, which are included in the SF6D-Utility.

The limited ability of Utility measures to pick up changes in health; in this case, radiological joint health or physical limitations (SF36-PF) has been observed in other conditions like cancer. It can be explained by the fact that SF6D-Utility includes domains that can reflect adaptation to physical limitations as well as response shift [38]. Even though utility values will always underestimate the effect of certain interventions, these data are used in cost-effectiveness analyses across all areas of healthcare [6,39]. The present data emphasize the potential pitfalls of economic models for haemophilia treatment, which are usually based on estimations of Utility values [16–19]. To better appreciate differences between treatment regimens, however, we recommend including assessment of joint structure and function, using objective tools such as X-rays or standardized physical examination, combined with self-reported assessment in the SF36-PF. For multicentre studies, however, standardization of X-ray scoring as well as physical examination is mandatory [40,41].

In conclusion, in adults with severe and moderate haemophilia, only quite advanced haemophilic arthropathy (Pettersson scores > 21 points) was significantly and independently associated with SF6D-Utility and SF36 domain scores for 'physical functioning'. Overall, HRQoL, especially SF6D-Utility was not affected by early joint disease.

Acknowledgements

Author contributions

This study was designed by KF and PdK, in collaboration with all authors. Data collection was performed by all authors. The analysis was performed by KF, with interpretation of data by all authors. The first draft of the manuscript was written by KF, and further developed by all authors.

Disclosures

The authors stated that they had no interests which might be perceived as posing a conflict or bias.

References

1 Ahlberg A. Haemophilia in Sweden VII. Incidence, treatment and prophylaxis of arthropathy and other musculo-skeletal manifestations of haemophilia A and B. *Acta Orthop Scand* 1965; **77**(suppl): 5–99.

2 Iorio A, Marchesini E, Marcucci M, Stobart K, Chan AK. Clotting factor concentrates given to prevent bleeding and bleeding-related complications in people with hemophilia A or B. *Cochrane Database Syst Rev* 2011; doi: 10.1002/14651858.CD003429.pub4.

3 Berntorp E, Astermark J, Baghaei F *et al.* Treatment of haemophilia A and B and von Willebrand's disease: summary and conclusions of a systematic review as part of a Swedish health-technology assessment. *Haemophilia* 2012; **18**: 158–65.

4 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**: 1129–36.

5 Hay CR. Purchasing factor concentrates in the 21st century through competitive tendering. *Haemophilia* 2013; **19**: 660–7.

6 Nicholson A, Berger K, Bohn R *et al.* Recommendations for reporting economic evaluations of haemophilia prophylaxis: a nominal groups consensus statement on behalf of the Economics Expert Working Group of The International Prophylaxis Study Group. *Haemophilia* 2008; **14**: 127–32.

7 Miners AH. Economic evaluations of prophylaxis with clotting factor for people with severe haemophilia: why do the results vary so much? *Haemophilia* 2013; **19**: 174–80.

8 Robinson R. Cost-utility analysis. *BMJ* 1993; **307**: 859–62.

9 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.

10 EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy (New York)* 1990; **16**: 199–208.

11 Ware JE. The SF36 Health Survey, Manual Interpretation Guide 1993. Boston: New England Medical Center, 1993.

12 Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002; **21**: 271–92.

13 Lippert B, Berger K, Berntorp E *et al.* Cost effectiveness of haemophilia treatment: a cross-national assessment. *Blood Coagul Fibrinolysis* 2005; **16**: 477–85.

14 Trippoli S, Vaiani M, Linari S, Longo G, Morfini M, Messori A. Multivariate analysis of factors influencing quality of life and utility in patients with haemophilia. *Haematologica* 2001; **86**: 722–8.

15 O'Mahony B, Noone D. 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**: e842–3.

16 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.

17 Martínez-Murillo C, Quintana S, Ambriz R *et al.* An economic model of haemophilia in Mexico. *Haemophilia* 2004; **10**: 9–17.

18 Fischer K, Pouw ME, Lewandowski D, Janssen MP, Van den Berg HM, Van Hout BA. A modeling approach to evaluate long-term outcome of prophylactic and on demand treatment strategies for severe hemophilia A. *Haematologica* 2011; **96**: 738–43.

19 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**: e228–38.

20 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.

21 Plug I, Van Der Bom JG, Peters M *et al.* . Thirty years of hemophilia treatment in the Netherlands, 1972–2001. *Blood* 2004; **104**: 3494–500.

22 Molho P, Rolland N, Lebrun T *et al;* GROUP and THEFS. Epidemiological survey of the orthopedic status of severe haemophilia A and B patients in France. *Haemophilia* 2000; **6**: 23–32.

23 Fischer K, van der Bom JG, Molho P *et al.* Prophylactic versus on-demand treatment strategies for severe haemophilia: a comparison of costs and long-term outcome. *Haemophilia* 2002; **8**: 745–52.

24 den Uijl I, Biesma D, Grobbee D, Fischer K. Outcome in moderate haemophilia. *Blood Transfus* 2014; **12**(Suppl 1): s330–6.

25 Pettersson H, Nilsson IM, Hedner U, Norehn K, Ahlberg A. Radiologic evaluation of prophylaxis in severe haemophilia. *Acta Paediatr Scand* 1981; **70**: 565–70.

26 Erlemann R, Rosenthal H, Walthers EM, Almeida P, Calleja R. Reproducibility of the Pettersson scoring system–an interobserver study. *Acta Radiol* 1989; **30**: 147–51.

27 Ware JE, Sherbourne CD. The MOS 36-item short form health survey (SF36). Conceptual framework and item selection. *Med Care* 1992; **30**: 473–83.

28 van den Berg B. SF-6D Population Norms. *Health Econ* 2012; **21**: 1508–12.

29 Aaronson NK, Muller M, Cohen PD *et al.* Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol* 1998; **51**: 1055–68.

30 Grosse SD, Chaugule SS, Hay JW. Estimates of utility weights in hemophilia: implications for cost-utility analysis of clotting factor prophylaxis. *Expert Rev Pharmacoecon Outcomes Res* 2015; **15**: 267–83.

31 Fischer K, Van der Bom JG, Mauser-Bunschoten EP, Roosendaal G, Berg HM. Effects of haemophilic arthropathy on health-related quality of life and socio-economic parameters. *Haemophilia* 2005; **11**: 43–8.

32 Gorina E, Alberca I, Aznar JA *et al.* . Results of an orthopaedic survey in young patients with severe haemophilia in Spain. *Haemophilia* 2002; **8**: 38–42.

33 Posthouwer D, Plug I, Van der Bom JG, Fischer K, Rosendaal FR, Mauser-Bunschoten EP. Hepatitis C and health-related quality of life among patients with haemophilia. *Haematologica* 2005; **90**: 846–50.

34 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**: 479–85.

35 Fischer K, Lewandowski D, Janssen MP. Modelling lifelong effects of different prophylactic treatment strategies for severe haemophilia A. *Haemophilia* 2016; **22**: e375–82. epub ahead of print June 28. Available at http://www.ncbi.nlm.nih.gov/pubmed/27353496.

36 Wasserman J, Aday L a, Begley CE, Ahn C, Lairson DR. Measuring health state preferences for hemophilia: development of a disease-specific utility instrument. *Haemophilia* 2005; **11**: 49–57.

37 Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005; **14**: 1523–32.

38 Sprangers MA, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med* 1999; **48**: 1507–15.

39 Drummond MJ, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddard GL. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press, 2005.

40 Foppen W, van der Schaaf IC, Beek FJA, Verkooijen HM, Fischer K. Scoring haemophilic arthropathy on X-rays: improving inter- and intra-observer reliability and agreement using a consensus atlas. *Eur Radiol* 2016; **26**: 1963–70.

41 Nijdam A, Bladen M, Hubert N *et al.* Using routine Haemophilia Joint Health Score for international comparisons of haemophilia outcome: standardization is needed. *Haemophilia* 2016; **22**: 142–7.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. SF36 domain scores according to Pettersson scores.

Table S2. Multivariate logistic regression models describing the association of arthropathy with SF6D-Utility.

Table S3. Multivariate logistic regression models describing the association of arthropathy with SF36-physical functioning.

Table S4. Linear multivariable regression on the association of Pettersson scores (in two categories) with SF6D-Utility.

Table S5. Linear multivariable regression on the association of Pettersson scores (in five categories) with SF6D-Utility.