

Test-Retest Reliability of Ukrainian Version of the Haemo-QoL Questionnaire

Markin Andriy* and Dubei Leonid

Danylo Halytskyi Lviv National Medical University, Lviv, Ukraine

*Corresponding author: Markin Andriy, Danylo Halytskyi Lviv National Medical University, Lviv, Ukraine; E-mail: markinandrij@gmail.com

Abstract

Background: Today psychometric characteristics of Ukrainian version of the questionnaire Haemo-QoL remain insufficiently researched. This publication describes the procedure and results of the test-retest reliability of this questionnaire in children with hemophilia A.

Objectives: Study the test-retest reliability of the Ukrainian version of the Haemo-QoL questionnaire for assessing the health-related quality of life (HRQoL) in children with hemophilia A, and determine the possibility of its use in practical and theoretical medicine in terms of time reliability.

Methods: The quality of life assessment (QoLA) was performed using the correspondence method, by interviewing 32 children with hemophilia A (self-report) and 32 parents of the children (proxy-report). The questioning was conducted twice at intervals of 4-6 weeks (primary test and re-test). The Haemo-QoL questionnaire version of age group I (4-7 years old), II (8-12 years old) or age group III (13-16 years) was used to measure the quality of life in children with hemophilia; and three versions were used for parents of children of corresponding age groups. These interviews were analyzed, the correlation between the data of primary test and re-test was determined using Spearman's rho, Pearson correlation coefficients, Wilcoxon's W criterion, Cohen's d effect size was determined for separate comparisons.

Results: HRQoL indices are not statistically different between tests on all scales ($p > 0.05$), except for the "Family" and "Others" scales ($p < 0.05$). Total HRQoL is statistically different ($p = 0.0013$), however, with a median difference of 0.25 only and an average absolute difference variability of 1.67 ± 1.51 ($5.42 \pm 2.83\%$). Total HRQoL in parent versions did not change statistically, unlike the children's versions, but the difference between the tests was only 1.32 and Cohen's d ES - 0.08. For versions Ip, IIp, IIIc/p there was no statistical difference, the versions Ic and II statistically differed ($p = 0.038$; $t = -2.39$ and $p = 0.0022$; $t = -3.98$, respectively) with a mean difference of 2.0 and 1.6, respectively.

Conclusion: The Ukrainian version of the Haemo-QoL questionnaire has sufficient test-retest reliability for quantitative dynamic reflection HRQoL in patients with haemophilia A.

Keywords: *Haemophilia; Health-related quality of life; Haemo-QoL; Test-retest reliability; Psychometric analysis; Questionnaire*

Received Date: April 23, 2018; **Accepted Date:** May 27, 2018; **Published Date:** June 4, 2018

Citation: Markin Andriy, Test-Retest Reliability of Ukrainian Version of the Haemo-QoL Questionnaire. J Clin Cases Rep 2018; 1(2): 74-83. DOI: <https://doi.org/10.46619/joccr.2018.1-1016>

Introduction

In today's medicine, which increasingly shifts towards the biopsychosocial model year after year, monitoring and improving the health-related quality of life (HRQoL) becomes one of the main components of managing children with chronic diseases; and scientific and practical developments on HRQoL research issues remain a priority direction of modern medicine [1-4]. The dynamic control of psychosocial health in patients with hemophilia, with HRQoL being one of the main indicators, is now recommended by the World Federation of Hemophilia (WFH) and the European Haemophilia Therapy Standardization Board (EHTSB) as one of the main management tasks for these patients [5-7]. The so-called disease-specific questionnaires are the most optimal tool for determining HRQoL in patients with chronic diseases [8,9]. Among these, to date, the most proven and widely used in children with hemophilia is the Haemo-QoL questionnaire [7,10-18]. This questionnaire has sufficient psychometric characteristics, which is proved in numerous studies [7,16,19-24]. To date, there is no validated Ukrainian version of this questionnaire, that significantly limits the full-fledged management of patients with hemophilia in Ukraine. Cultural and linguistic adaptation of the Ukrainian test version of this questionnaire, study of its validity and sensitivity has already been carried out in a separate research paper. Taking into account that, according to numerous recommendations, for the legitimate use of the instrument in the new ethnolinguistic environment the necessary step is to carry out a comprehensive psychometric analysis with confirmation of its validity, reliability and sensitivity, the issue of determining the reliability of the Ukrainian version of the questionnaire remains topical [3,23,25,26]. These three main characteristics of the questionnaires: sensitivity, validity and reliability Middel et al. [27] aptly called "holy trinity" of psychometric analysis.

The reliability of the questionnaire is understood as its ability to give constant and accurate indicators when evaluating HRQoL. Today, the main components of reliability of questionnaires are the following: equivalence, stability (or time constancy, reproducibility, test-retest reliability) and internal consistency (homogeneity) [28]. In practice, stability and internal consistency are most often determined. Stability is the ability of the questionnaire to steadily in terms of time equivalent and accurately reflect the target function in quantitative dimension. The stability in the majority of adaptation studies is determined using the test-retest method, therefore, the term "test-retest reliability" (TRR) is often synonymous with reliability in time. This method of determining the sensitivity is to establish a correlation between the quantitative assessment of the target function in the primary and repeated questionnaires of the same respondent (group of respondents), provided that there are no objective changes in the health condition between the examinations that can potentially affect the assessment of their own well-being expressed in quantitative terms [28-31]. Although TRR study during adaptation of questionnaires is recommended by most guidelines, however, the study by Solans et al. [32] demonstrates rather rare use of this method in adaptation processes. Thus, among 94 questionnaires for assessing the quality of life (including 30 general and 64 disease-specific) analyzed by researchers within 6 years, the test-retest method was used in validating questionnaires in only 20% of such adaptation processes [32].

The reliability of the Haemo-QoL questionnaire has been demonstrated in a number of studies (Bullinger et al. [16], von Mackensen et al. [19], Mercan et al. [14]), however, TRR studies [14,16,19] remain extremely limited. In our opinion, the study of the TRR questionnaire, although it is labour-intensive in the time and emotional terms, is a key (but not unique) one for characterizing a specific instrument with a sufficiently reliable for its application in practical and theoretical medicine in terms of reliability. We fully share the opinion of Limperg et al. [22] that today there is no need to develop new tools for assessing HRQoL for patients with hemophilia, and the study of characteristics of the existing instruments remains a priority.

Objectives

Study the test-retest reliability of the Ukrainian version of the questionnaire for assessing HRQoL in children with hemophilia, Haemo-QoL; determine the suitability for using this version of the questionnaire in terms of the reliability in time in the new ethnolinguistic environment.

Methodology

Assessment of HRQoL: The ethnolinguistically adapted pilot Ukrainian long version of the questionnaire was used for Haemo-QoL assessment. The questionnaire has six versions: three versions of the questionnaire for children allowing for age groups: the Ic version - for children aged 4-7 years, IIc - 8-12 years, IIIc - 13- 16 years; three analogous versions of the questionnaire for parents of children of the corresponding age groups (Ip, Iip, IIip). The versions for children of different age groups differed both in terms of the number of scales of the questionnaire and the number of questions in the corresponding scales. The number of questions in the scales varies from 2 to 10 questions. The structural load of the questionnaire is the following: Ic/p - 8 scales/21 questions, IIc/p - 10 scales/64 questions, IIIc/p - 12 scales/77 questions. The study includes the results of 64 forms with answers of respondents-participants of the primary questionnaire and 64 forms with answers of respondents-participants of the repeated questionnaire. Scores for the scale were taken into account provided more than 50% of answers to the statement. The HRQoL obtained by counting the scores of the questionnaire is presented in the form of a transmuted scale score (hereinafter – “TSS”) according to the well-known formula [33]. At the same time, higher TSS values indicate a worse quality of life, and smaller ones – a better one, with an interval of acceptable values from “0” to “100” scores, where “0” is the best HRQoL, and “100” is the worst HRQoL. Total HRQoL for each respondent is determined as the mean of scores (TSS) of all scales of the questionnaire. Taking into account that the majority of indicators for the characteristics of distribution differed from normal, the results of the study were presented as Me (25%, 75%), where Me is median, 25% - 1st quartile (25th percentile), 75% - third quartile (75th percentile).

Procedure

A study of the reliability of the Ukrainian version of the Haemo-QoL questionnaire was conducted on the basis of the Children Thrombosis and Hemostasis Center in of the Western Ukrainian Specialized Children’s Medical Center, Lviv. Reliability in time of the questionnaire was determined using the test-retest method by determining the correlation between the results of the primary questioning and re-questioning of 64 respondents participating in the study. Duration of the study was 2 months. Sex and age as well as clinical and nosological structure of respondents is presented in Table 1. The questioning was conducted by the correspondent method. Joint hemorrhages were registered using the general clinical method, the method of retrospective analysis of patients’ clinical records, ultrasonography (US) of target joints. All participants in the study were fluent in Ukrainian, were well informed about the purpose and the course of the study.

Criteria of selection in the study: Prior to repeated HRQoL questioning, all respondents had to answer the question on the cover of the questionnaire: “How has your health condition changed over the last month?” with the choice options: “no changes”, “improved”, “became worse”. The criteria for inclusion in the study were as follows:

1. Children with hemophilia A (factor VIII concentration in plasma <30 IU/dl), aged 4 to 16 years, and their parents (one parent).
2. Patients who answered “no changes” to the above question.

3. Lack of dynamic changes in the clinical characteristics of the disease in the repeated questioning of patients surveyed in the primary questioning. Hemorrhages in the joints for 4 weeks before the re-questioning were chosen as indicators of changes in the clinical characteristics.

Name	Absolute number of respondents	%
Total number of respondents	64	100
Respondents - children	32	50
Severe hemophilia A*	15	46.9
Moderate hemophilia A**	13	40.6
Mild hemophilia A***	4	12.5
Boys	32	100
Girls	0	0
Respondents - parents	32	50
Mother	24	75
Father	8	25
Age groups of respondents according to versions of the questionnaire		
I children (4-7 years)	11	17
I parents	11	17
II children (8-12 years)	12	19
II parents	12	19
III children (13-16 years)	9	14
III parents	9	14
Joint hemorrhages		
Without hemorrhages	44	69
With hemorrhages	20	31

Table 1: Sex and age as well as clinical and nosological structure of respondents

Note: *Factor VIII concentration in plasma ≤ 1 MO/dl; ** Factor VIII concentration in plasma 1-5 IU/dl; *** Factor VIII concentration in plasma 5 - 30 IU/dl.

To identify the subjective severity of joint hemorrhages, the respondents' answers to the following question of the questionnaire were evaluated: "How much have you been disturbed by hemorrhages during the last 4 weeks?" Patients should have been assessed for the choice options for respondents Ic: "not at all", "somewhat", "very much", for IIc, IIIc: "not at all", "somewhat", "moderately", "quite a bit", for Ip, IIp, IIIp: "not at all", "somewhat", "moderately", "quite a bit", "very much". According to answer to this question, the criterion of inclusion in the study was the same answer in the primary and repeated questionnaires. Some respondents who did not respond identically, cognitive debriefing was conducted to determine the understanding and objectivity of such answers. According to the indicator of joint hemorrhages, the structure of the respondents was as follows: 44 respondents did not have any hemorrhages in the primary and repeated questionnaires (22 children and 22 parents of the corresponding children), 20 had joint re-hemorrhages (1-2 hemorrhages between the examinations) in the repeated questioning (10 children with hemorrhages and their 10 parents), however, indicated a subjective assessment of their severity in the questionnaire, similar to that given in the primary survey.

Lack of psychosocial factors that could affect the objectivity of answers: Such factors include: divorce of the patient's parents, relocation, death of relatives, change of educational institution and others that could potentially be found out in an individual interview. To this end, the parents of patients were required to fill in the appropriate item in the supplementary

form to the questionnaire on the availability or absence of such factors. Such information was also collected through personal interviewing.

Statistical data processing: For the purpose of statistical research, a preliminary analysis of the groups was conducted for the normality of distribution, using the analysis of excess and asymmetry, Kolmogorov-Smirnov and Shapiro-Vilka tests, as well as distribution alignment chart and histogram analysis. For correlation analysis, the Pearson correlation coefficient for normal distribution data and the Spearman's rank correlation coefficient for data that differed from the normal, nonparametric Wilcoxon's W criterion for related groups were used. Any difference with P value <0.05 was considered statistically significant. To determine the standardized effect size (ES) of the difference between the HRQoL of the primary questioning and re-questioning the Cohen's d ES for nonparametric data was used. The obtained Cohen's d value was further estimated by a scale of 0.1 to 0.3 as a "slight difference", 0.3 to 0.5 as an "average difference", 0.5 and higher – a "big difference" [34-36].

Results

The results of evaluation of HRQoL in the primary (hereinafter – "TSS1") and repeated (hereinafter – "TSS2") survey for all respondents without age grouping are presented in Table 2. As we can see from this table, we evaluated the reliability of the difference between TSS1 and TSS2 both for each scale separately and by the overall HRQoL. A separate comparison of the results of TSS1 and TSS2 was carried out for each of the six versions of the questionnaire (the result is presented in Table 3).

Relation of the total HRQoL between the surveys of all respondents: Total HRQoL of all groups in the primary survey was 32.69 (18.01, 39.81), in the repeated survey the total HRQoL slightly increased to 32.94 (19.68; 39.68). At the same time, the increase in the indicator was statistically reliable ($p=0.0013$, $W=478$). Taking into account the statistically reliable difference of the total HRQoL between TSS1 and TSS2, the difference of medians for these indicators is also estimated to be only 0.25, which is $0.77 \pm 1.09\%$ from the value of TSS1. The mean absolute value of difference variation of the total HRQoL between TSS1 and TSS2 is 1.67 ± 1.51 , which is $5.42 \pm 2.83\%$ of the total HRQoL value for the primary survey, the minimum difference value = 0, the maximum is 8.33. The value of the difference in each patient separately for the total HRQoL between the primary and the repeated survey to the value of TSS1 ("total HRQoL difference, %") is 4.95 (1.89; 9.46)%, and its maximum value in some cases is 66.67%.

Relation of HRQoL scales of the questionnaire between surveys of all respondents: In the analysis of the difference between the values of TSS1 and TSS2, the statistically reliable difference was not found on all scales of the questionnaire, except for the "Family" scale ($p=0.014$, $W=201$) and "Others" ($p=0.0097$, $W=53$), in which the difference between the primary and secondary questionnaires was statistically reliable. In this case, we conducted paired comparisons without differentiation of respondents into groups by age and versions of the questionnaire.

Relation of the total HRQoL between surveys by age groups: Comparing total HRQoL in three age groups (Ic/p, IIc/p, IIIc/p), no statistically reliable difference between TSS 1 and TSS 2 groups Ic/p ($p=0.085$, $t=-1.81$) and IIIc/p ($p=0.47$, $t=-0.75$) was found, whereas in the group IIc/p, in the primary questionnaire, the total HRQoL was 28.98 ± 13.59 and in the repeated questionnaire 30.00 ± 13.35 , and this increase was statistically reliable ($p=0.0016$, $t=-3.58$). At the same time, the mean difference is 1.02 and Cohen's d ES is only 0.08.

HRQoL Scales	N	TSS 1 Mean (n=64)	TSS 1 Me (n=64)	TSS 2 Mean (n=64)	TSS 2 Me (n=64)	P _{1,2}	W
Physical health	64	36.67	38.39(12.50; 56.25)	37.57	38.39(12.50; 54.91)	0.18	218.5
Feelings	64	23.37	19.79(6.70; 39.29)	24.25	21.13(7.29; 42.86)	0.10	68.5
View	64	24.07	25.00(9.72; 31.25)	24.48	25.00(11.80; 30.00)	0.80	206
Family	64	44.18	46.88(23.44; 65.31)	45.72	50.00(25.00; 63.75)	0.014	201
Friend	64	31.91	31.25(12.50; 50.00)	31.61	31.25(16.67; 50.00)	0.72	126
Perceived support	42	27.93	25.00(12.50; 43.75)	29.27	25.00(12.50; 43.75)	0.063	32
Others	64	20.38	16.67(0.00; 31.25)	22.27	20.83(4.17; 29.17)	0.0097	53
Sports and school*	64	40.78	39.76(23.44; 58.33)	42.03	43.75(28.13; 58.33)	0.16	164
Dealing	42	29.29	26.79(14.29; 42.86)	30.41	26.79(14.29; 42.86)	0.31	77.5
Treatment	64	38.72	34.38(17.86; 53.57)	37.98	34.38(18.75; 53.57)	0.50	70
Future	18	24.77	18.75(12.50; 31.32)	24.08	18.75(12.50; 31.32)	0.53	13.5
Relationship	18	18.06	18.75(0.00; 25.00)	19.44	12.50(0.00; 37.50)	0.46	7
Total HRQoL	64	30.83	32.69(18.01; 39.81)	31.68	32.94(19.68; 39.68)	0.0013	478

Table 2: Comparison of HRQoL in the primary and repeated questionnaires in Children and their Parents with Haemophilia A, all age groups

Note: *"Sports and school" is "Nursery school/Kindergarten" for version I ch/p, N – number of respondents surveyed for HRQoL for the relevant scale, TSS1 Me – median of scores in the primary questionnaire, TSS2 Me – median of score in the repeated questionnaire, P_{1,2} – probability null hypothesis between the "TSS1 Me" and "TSS2 Me", "TSS1 Mean" and "TSS2 Mean" - mean points for primary and repeated questionnaires; W - nonparametric Wilcoxon's W criterion.

Relation of the total HRQoL between the surveys in the parent/children questionnaire versions: Comparing the difference between the total HRQoL of the parent questionnaire versions (Ip, Iip, IIIp) and children versions separately (Ip, Iip, IIIp), without division into corresponding age groups, it was found that the indicators obtained in the analysis of parent questionnaire version in the primary and repeated questionnaires were as close as possible and no significant difference was found (p=0.28, t=-1.1). At the same time, the total HRQoL in the group of children in the repeated questionnaire was slightly different from the results of the primary questionnaire, and this difference was statistically reliable (p=0.0016, t=-3.45). It should be noted that for these groups the mean difference is 1.32, and the Cohen's d ES is only 0.08.

Relation of the total HRQoL between surveys in each group: We also compared the results of survey by the total HRQoL for each of the six versions of the questionnaire separately. For this purpose, the children and their parents were divided into age groups in accordance with questionnaire versions (6 groups). This was done in order to characterize the test-retest reliability of the age version of the questionnaire. The results are presented in Table 3.

Age group	N	TSS 1 Mean (SD)	TSS 2 Mean (SD)	P _{1,2}	t
Age group	N	TSS 1			
Mean (SD)	TSS 2				
Mean (SD)	P _{1,2}	t			
Ic	11	29.83 (20.14)	31.83 (20.16)	0.038	-2.39
Ip	11	36.87 (21.50)	37.02 (20.01)	0.85	-0.197
Iic	12	28.81 (14.09)	30.41 (13.71)	0.0022	-3.98

Table 3: Comparison of HRQoL in the primary and repeated questionnaires in Children and their Parents with Haemophilia A, different age groups

Note: N – number of respondents, TSS1 and TSS2 - mean points for primary and repeated questionnaires, P_{1,2} – probability null hypothesis between the TSS 1 and TSS2, t –t-test.

As can be seen from Table 3, the results of parents for all versions of the questionnaire showed no statistically reliable difference between the primary and secondary questionnaires. The mean difference in this case was from 0.15 to 0.56. As for the results of the HRQoL evaluation in children's questionnaires, no statistically reliable difference was found only in the

oldest children (version IIIc) with a mean difference of only 0.10. With regard to Ic and IIc versions, the reliable difference of the total HRQoL between TSS1 and TSS2 is statistically verified. In addition, it should be noted that the mean difference for Ic and IIc versions was 2.00 and 1.6, respectively.

Discussion

Before discussing the results obtained in the HRQoL survey in 32 patients with hemophilia A and their 32 parents the following should be noted. To date, there are no, and probably there can be no, single criteria for evaluating the results of the TRR survey. The investigated psychological component of hemophilia and its influence on the feeling of own well-being, is dynamic and labile in itself, and it is impossible to exclude all factors that can affect it, and therefore completely isolate the results of the primary and repeated questionnaires according to inclusion criteria. Therefore, the existing difference between the results of the primary and the repeated questionnaires does not always indicate the lack of reliability of the questionnaire. A kind of “residual” sensitivity of this questionnaire to factors that are not excluded and not taken into account in the study may have an impact, too. Therefore, the conclusion on the degree of reliability of the questionnaire is made by the method of complex data analysis, including internal consistency. To assess the ability of this questionnaire to reflect accurate and constant results as objectively as possible, we also tried to determine the difference between the indicators and the “size effect” of this difference, differentiated the indicators by the scales under study. It should be noted that we also investigated the internal consistency of the Ukrainian version of Haemo-QoL, but the results will be presented in another publication. The intervals between the surveys should also be taken into account in TRR studies. In the study of TRR by S. von Mackensen et al. [21] when developing and implementing the Haemo-QoL questionnaire, the interval between the surveys was 1 week, in a study by E. Pollak et al. [18] - 1-2 weeks, while in our survey it varied from 4 to 6 weeks. When studying TRR, time intervals between surveys should always be taken into account and analyzed when drawing conclusions about the reliability of the questionnaire because of the so-called “memory effect”. In the study of reliability, as Oladimeji A and Bolarinwa [28] noted, ‘... the time period is long enough yet short in time that respondents’ memories of taking the test at time 1 do not affect their scores at time 2 and subsequent test administrations ...’ However, there are no optimal commonly accepted boundaries for time between surveys to determine reliability in time so far.

When comparing the HRQoL indicators for each scale of the questionnaire, it was found that similar results were obtained for most scales, however, for the “Family” and “Others” scales the difference between the primary and secondary questionnaires was found, but it was negligible (3.12 for “Family” and 4.16 for “Others”). This is likely to show a greater degree of sensitivity of these scales to the factors that are not taken into account (and therefore not excluded) in the study, or low reliability of the inclusion criteria used in this study. In general, taking into account the size of the difference in indicators, such data can be interpreted as sufficient reliability for all scales of the questionnaire. Analyzing the total HRQoL, its increase was found in the repeated questionnaire. However, a small median difference (0.25) makes it possible to interpret these indicators as similar, especially mentioning the range of possible indicators from 0 to 100. The most similar indicators were found in the 1st and 3rd age group, while in 2nd they differed between the surveys. But, again, given the small mean difference and less than the “slight difference” of Cohen’s d ES, data between the surveys can be considered similar. Interestingly, comparing the TRR in parents and children, it has been found that parents give more similar scores than children, which may indicate a greater reliability of the parent questionnaire. This also applies to analysis in different age groups. For all versions of the questionnaire for parents (Ip, Iip, and IIIp) there was no significant difference in the indicators of the quality of life between the surveys. These data are similar to those of E. Pollak et al. [18], in which the higher

reliability of parental forms was demonstrated, but the conclusion was made on the basis of an analysis of internal consistency. The TRR study was conducted by the author only in age groups II and III where a lower intra-class correlation coefficient in parent versions between surveys is demonstrated. At the same time, in our study, versions for children showed lower reliability (with the exception of older children aged 13-16, where no statistical difference was found between the surveys), however, the minimum difference between the surveys and a “slight difference” in Cohen’s d ES allows us to assert that the indicators obtained in these groups are similar. It is difficult to compare data obtained with known researches due to a number of reasons, in particular due to the difference in age groups included in the study, the methods used and often the complex reliability profile. This comparison can be made after a complete psychometric analysis of this questionnaire.

It is also worth noting that in most studies of validating questionnaires, the authors make conclusion on the reliability of the newly created versions solely on the basis of an analysis of internal consistency. In our opinion, such a conclusion can be only complex, taking into account both the internal consistency and test-retest stability in time.

Conclusion

Taking into account that the main objective of this study was to investigate the test-retest reliability of the Ukrainian long version of the questionnaire to evaluate the HRQoL of children with haemophilia, Haemo-QoL, we can draw the following conclusions:

1. It is proved that the Ukrainian version of the Haemo-QoL questionnaire has sufficient reliability to changes in time to quantitatively reflect the change in HRQoL in patients with hemophilia A.
2. For a comprehensive description of the reliability of this version of the questionnaire, it is necessary to conduct a study of internal stability.

References

1. Antunes B, Harding R, Higginson IJ, et al. (2014) Implementing patient-reported outcome measures in palliative care clinical practice: a systematic review of facilitators and barriers. *Palliative Medicine* 28(2): 158-175.
2. Berg HM, Feldman BM, Fischer K, et al. (2015) Assessments of outcome in haemophilia—what is the added value of QoL tools?. *Haemophilia* 21(4): 430-435.
3. Bullinger M and Quitmann J (2014) Quality of life as patient-reported outcomes: principles of assessment. *Dialogues in Clinical Neuroscience* 16(2): 137.
4. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, et al. (2016) Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. *Journal of Comparative Effectiveness Research* 5(5): 507-519.
5. de Moerloose P, Fischer K, Lambert T, et al. (2012) Recommendations for assessment, monitoring and follow-up of patients with haemophilia. *Haemophilia*, 18(3): 319-325.
6. Srivastava A, Brewer AK, Mauser-Bunschoten EP (2013) Guidelines for the management of hemophilia. *Haemophilia* 19(1).
7. Novik A, Salek S, Ionova T (2012) Guidelines patient - reported outcomes in hematology. Genoa: Litoprint.
8. Megari K (2013) Quality of life in chronic disease patients. *Health Psychology Research* 1(3): 141-148.
9. Snyder CF, Aaronson NK, Choucair A, et al. (2012) Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Quality Life Research* 21(8): 1305-1314.

10. Deshbhratar P (2016) Comparison of Quality of Life in Adults and Children with Haemophilia. *Imperial Journal of Interdisciplinary Research* 2(7): 733-735.
11. Gringeri AV, Von Mackensen S, Auerswald G, et al. (2004) Health status and health-related quality of life of children with haemophilia from six West European countries. *Haemophilia* 10(S1): 26-33.
12. Oladapo AO, Epstein JD, Williams E, et al. (2015) Health-related quality of life assessment in haemophilia patients on prophylaxis therapy: a systematic review of results from prospective clinical trials. *Haemophilia* 21(5): 344-358.
13. Tantawy AA, Mackensen SV, El-Laboudy MA, et al. (2011) Health-related quality of life in Egyptian children and adolescents with hemophilia A. *Pediatric Hematology and Oncology* 28(3): 222-229.
14. Mercan A, Sarper N, Inanir M, et al. (2010) Hemophilia-Specific Quality of Life Index (Haemo-QoL and Haem-A-QoL questionnaires) of children and adults: result of a single center from Turkey. *Pediatric Hematology and Oncology* 27(6): 449-461.
15. Williams VK, Antoniou G, Jackson A, et al. (2016) Parents' perception of quality of life in their sons with haemophilia. *Journal of Pediatrics and Child Health* 52(12): 1095-1098.
16. Bullinger M, Mackensen S von, Fischer K (2002) Pilot testing of the 'Haemo-QoL' quality of life questionnaire for haemophiliac children in six European countries. *Haemophilia* 8(2): 47-54.
17. Lee CA, Berntorp EE, Hoots WK (2014) *Textbook of Hemophilia*. John Wiley & Sons, NJ, USA.
18. Pollak E, Mühlhan H, Von Mackensen S, et al. (2006) The Haemo-QoL Index: developing a short measure for health-related quality of life assessment in children and adolescents with haemophilia. *Haemophilia* 12(4): 384-392.
19. Remor E, Young NL, Von Mackensen S, et al. (2004) Disease-specific quality-of-life measurement tools for haemophilia patients. *Haemophilia* 10(4): 30-34.
20. Chevallet L, Weatherall JH, Von Mackensen S (2008) PSY44 Linguistic validation of the Haemo-QoL and Haem-A-QoL for use in international studies. *Value in Health* 11(3): A165.
21. Mackensen SV, Bullinger M, Haemo-QoL Group (2004) Development and testing of an instrument to assess the Quality of Life of Children with Haemophilia in Europe (Haemo-QoL). *Haemophilia* 10: 17-25.
22. Limperg PF, Terwee CB, Young NL, et al. (2017) Health-related quality of life questionnaires in individuals with haemophilia: a systematic review of their measurement properties. *Haemophilia* 23(4): 497-510.
23. Mackensen S, Campos IG, Acquadro C, et al. (2013). Cross-cultural adaptation and linguistic validation of age-group-specific haemophilia patient-reported outcome (PRO) instruments for patients and parents. *Haemophilia* 19(2): 73-83.
24. Young NL, ST-LOUIS J, Burke T, et al. (2012) Cross-cultural validation of the CHO-KLAT and HAEMO-QoL-A in Canadian French. *Haemophilia* 18(3): 353-357.
25. Epstein J, Santo RM, Guillemin F (2015) A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology* 68(4): 435-441.
26. Cassis FY, Querol F, Forsyth A, et al. (2012) Psychosocial aspects of haemophilia: a systematic review of methodologies and findings. *Haemophilia* 18(3): e101-e114.
27. Middel B and Sonderen E van (2002) Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *Integrated Journal of Integrated Care* 2: 1-21.
28. Bolarinwa OA (2015) Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal* 22(4): 195-201.

29. Erdvik IB, Øverby NC, Haugen T (2015) Translating, reliability testing, and validating a norwegian questionnaire to assess adolescents' intentions to be physically active after high school graduation. *Sage Open* 5(2): 1-6.
30. Pedisic Z, Bennie JA, Timperio AF, et al.(2014) Workplace Sitting Breaks Questionnaire (SITBRQ): an assessment of concurrent validity and test-retest reliability. *BMC Public Health* 14(1): 1249.
31. Rejano-Campo M, Ferrer-Peña R, Urraca-Gesto MA, et al. (2017) Transcultural adaptation and psychometric validation of a Spanish-language version of the “Pelvic Girdle Questionnaire”. *Health and Quality of Life Outcomes* 15(1): 30.
32. De Boer AS, Meuffels DE, Van der Vlies CH, et al. (2017) Validation of the American Orthopaedic Foot and Ankle Society Ankle-Hindfoot Scale Dutch language version in patients with hindfoot fractures. *BMJ Open* 7(11): e018314.
33. Solans M, Pane S, Estrada MD, et al. (2008) Health-Related Quality of Life Measurement in Children and Adolescents: A Systematic Review of Generic and Disease-Specific Instruments. *Value in health* 11(4): 742-764.
34. <http://haemoqol.de/scoring/manual/>.
35. Cohen J (1977) *Statistical power analysis for the behavioral sciences* 2nd edn. Erlbaum Associates, Hillsdale.
36. <https://www.pdfdrive.net/research-methods-and-statistics-in-psychology-e38313045.html>.
37. Fritz CO, Morris PE, Richler JJ (2012) Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General* 141(1): 2-18.
38. http://ccsg.isr.umich.edu/images/PDFs/CCSG_Full_Guidelines_2016_Version.pdf.