

ORIGINAL ARTICLE

## Health Related Quality of Life in stem cell transplantation: Clinical and psychometric validation of the questionnaire module, High Dose Chemotherapy (HDC-19)

I. ANDERSSON<sup>1,4</sup>, M. HJERMSTAD<sup>2,3</sup>, D. STOCKELBERG<sup>1</sup> & L. O. PERSSON<sup>4</sup>

<sup>1</sup>Section of Haematology and Coagulation, Sahlgrenska University Hospital, Göteborg, Sweden, <sup>2</sup>Department of Behavioural Sciences in medicine, University of Oslo, Norway, <sup>3</sup>The Norwegian Cancer Society, Oslo, Norway and <sup>4</sup>Institution for Caring and Health Science, Sahlgrenska Academy, Göteborg University, Sweden

### Abstract

The objective of this study was to assess the psychometric properties of the HDC-19, a module questionnaire for assessing symptoms and problems of patients undergoing stem cell transplantation (SCT) following high-dose chemotherapy (HDC). It consists of 19 questions and was developed for use in conjunction with EORTC QLQ-C30. Psychometric evaluations were performed according to guidelines recommended by the EORTC. The principal component analyses suggested that nine of the HDC-19 items could be reduced to four components (sexual functioning, future health perspectives, skin irritations and joint/muscle pain). Multitrait scaling analysis showed that most item-scale correlation coefficients met the standards of convergent ( $>0.40$ ) and discriminant validity. Test-retest reliability coefficients between assessments at inclusion and admission were high, indicating that perceived health status remained virtually unchanged during this period. As expected, correlations between admission and one month after transplantation were considerably lower. The internal consistency of the multi-item scales was also satisfactory, (Cronbach's alpha 0.59–0.87). Overall, the known-groups comparisons showed smaller differences between designated groups than expected. As expected, changes in the HDC-19 mirrored changes in QLQ-C30 'global quality of life'. These results lend support to the validity of the HDC-19 as a supplementary questionnaire for assessing specific health-related quality of life (HRQOL) issues relevant for SCT patients.

High-dose chemotherapy (HDC) followed by autologous or allogeneic stem cell transplantation (SCT) is today a well-established treatment for hematologic malignancies, certain solid tumours and immunologic disorders [1,2]. Although this procedure has become more sophisticated, there are several reports of acute and late medical side effects after HDC and SCT [3–5]. The complications range from minimal to life-threatening, and include nausea, vomiting, mucositis, loss of appetite and weight, fatigue, bleeding and frequent infections [3]. Patients receiving allogeneic SCT also experience additional complications, such as infections related to extended immunosuppression [4,5], risk for venoocclusive disease (VOD) [5] and graft-versus-host disease (GvHD) [4,5].

In oncology, it is now generally accepted that the outcomes of illness and treatment cannot be fully

described solely by medical measures [6,7]. Today such measures are commonly supplemented by assessments of Health Related Quality of Life (HRQOL), i.e. the various effects that disease and treatments have on patients' perceptions of functioning, symptoms and well-being [7,8].

Several instruments are now available to assess HRQOL in cancer. Among the most extensively used questionnaires are those developed by the European Organization for Research and Treatment of Cancer (EORTC) [6]. The 30-item EORTC Quality of Life Questionnaire (EORTC QLQ-C30) provides information about problems commonly experienced by cancer patients in general. To obtain specific information in relation to different cancer diagnoses, several supplemental diagnosis- and/or treatment-specific modules have been developed. The EORTC Head and Neck Cancer [9],

Oesophageal Cancer [10], Breast Cancer [11], Colorectal Cancer [12] and Gastric Cancer [13] modules are some of the official and internationally validated modules available today. These may be obtained at the EORTC website [www.eortc.be/home/qol](http://www.eortc.be/home/qol). Detailed guidelines have been established by Sprangers and colleagues [14] on behalf of the EORTC for the development of such modules to supplement the core instrument EORTC QLQ-C30.

As no specific questionnaires for evaluating HRQOL after SCT were currently available, Hjermstad and colleagues [15] developed a 19-item High-Dose-Chemotherapy module (HDC-19) for this purpose in their prospective study of autologous and allogeneic SCT patients [15]. This ad-hoc module was developed according to the guidelines set forth by the EORTC [14] and focuses on physical symptoms and psychological problems related to SCT not covered by the core questionnaire, EORTC QLQ-C30. In phase I of the development of the HDC-19, relevant HRQOL issues were generated by means of extensive literature searches and interviews with patients and health care providers. In phase II, items were operationalised and formatted to the same four-point response categories as in the core questionnaire. In phase III, the provisional module was first pilot-tested in ten patients and subsequently in a larger study with 177 patients [15]. The increased use of stem cell transplantation and the raised awareness of the need for more thorough assessments of its physical and psychological side effects, engendered in part by findings from the unofficial HDC-19, contributed to the development of an official EORTC HDC module, the HDC29. The HDC29 is targeted for the pre-transplant period, the time during treatment and up to 6 months after treatment. However, this module still awaits testing in a larger (international) group of patients in order to determine its reliability, validity and cross-cultural applicability [16].

In 2000, we initiated a project aimed at assessing HRQOL in patients treated with HDC followed by SCT. The present study is part of that project. As the work in the EORTC HDC group was in its infancy and no official treatment-specific module was available at study start, it was decided to use the HDC-19 developed by Hjermstad and colleagues [15] to supplement the EORTC QLQ-C30.

The aim of the present study was to field test a Swedish version of HDC-19 to examine its validity, reliability and other psychometric properties in a group of patients undergoing SCT. EORTC guidelines were followed in performing these evaluations [14].

## Patients and methods

### *Study sample*

Subjects included in the study consisted of patients who had been referred to the bone marrow transplantation unit at Sahlgrenska University Hospital in Göteborg and who had been accepted for SCT between the end of November 2000 and 2004. Inclusion criteria were age over 18 years, ability to speak and understand Swedish and no mental illness.

During this period, 335 patients were accepted for SCT (254 for autologous SCT and 81 for allogeneic SCT). Three patients were excluded because they were under 18 years, eight did not speak or understand Swedish and three had mental illness. Of the remaining 321 patients meeting the inclusion criteria, 202 agreed to participate in the study. Of these, 171 completed and returned the first questionnaires, giving a response rate of 84% (see Figure 1).

After informed consent the patients were asked to complete two questionnaires, the EORTC QLQ-C30 (version 3.0) and the HDC-19, at six occasions. The questionnaires were administered at inclusion (mean 49 days before transplantation), at admission to hospital (mean 8 days before transplantation) and at 30 days, 100 days, 6 and 12 months after transplantation. They were administered to the patients as inpatients when applicable or by mail after telephone contact. Although the patients were not requested to explain their reasons for discontinuing the study, 13 of the 31 who failed to complete the questionnaires at inclusion did not proceed to SCT due to progressive disease ( $n=10$ ) or death ( $n=3$ ). Questionnaires at admission were completed by 141 patients (4 did not proceed to SCT due to progressive disease and 18 failed to answer). Questionnaires at the 30-day follow-up were completed by 158 patients (17 of those missing admission questionnaires completed 30-day questionnaires). The 100-day questionnaires were completed by 149 patients (2 had died and 7 failed to answer) and 139 answered the 6-month questionnaires (3 had died and 7 failed to answer). The 12-month questionnaires were completed by 132 patients (5 had died and 2 failed to answer). Demographic data and other relevant clinical information were obtained from medical records.

### *Questionnaires*

The EORTC QLQ-C30 (version 3.0) consists of 30 items (6), covering six scales measuring physical, emotional, social, cognitive role functioning, and global QOL. Three symptom scales measure fatigue,

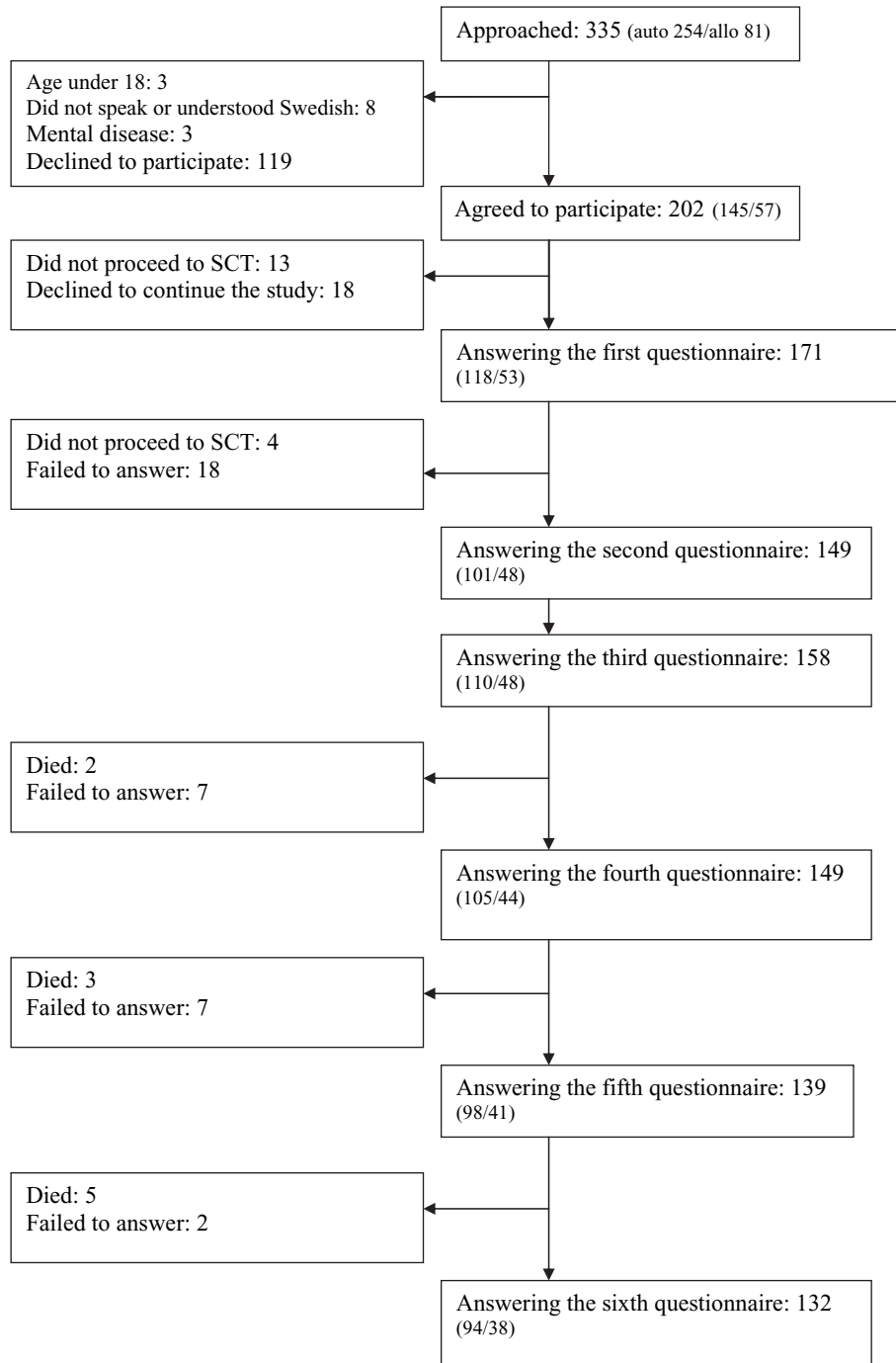


Figure 1. Attrition of patients from approach to one year follow-up.

pain, and nausea/vomiting, and six single items measure financial impact and the symptoms dyspnoea, sleep disturbance, appetite, diarrhoea, and constipation. The time frame is the past week. All items are rated on a 4-point response scale ranging from 1 (not at all) to 4 (very much), except the two items measuring general health and global QOL, which have visual-analogue scales from 1–7. All scales and single items are transformed to a percentile scale from 0–100. Higher scores represent better

functioning, general health and global QOL, whereas high scores on the symptom scales and single items indicate a higher level of symptoms and/or problems [17].

The HDC-19, which is the main outcome of the present study, consists of 19 questions covering circulation, mouth and skin problems, pain, anxiety, sexuality, social support and perspectives regarding future health. It is also rated using 4-point response scales and is scored according to EORTC guidelines

[6,14,15]. Items focusing on sexuality, future health perspective and social support constitute functional scales and items focusing on symptoms form symptom scales. As in the EORTC QLQ-C30, functional scales are calculated so that higher scores represent better functioning, whereas higher symptom scale scores represent higher levels of symptoms. All scales were transformed to a 0–100 scale.

#### *Translation procedure*

The HDC-19 was originally developed in Norway [15]. As Swedish and Norwegian are very similar, a traditional forward-backward translation [14] was judged to be unnecessary. Instead, three persons independently translated the complete questionnaire from Norwegian to Swedish and thereafter agreed on one version. Later, this translation was examined and approved by the author of the original version (M. Hjermstad).

#### *Statistical methods*

*Tests of internal structure between items.* Principal component analyses with varimax rotation were performed to examine for potential multi-item scales among the 19 items in the HDC-19 module. Listwise exclusion of missing values was utilised. The analyses were performed separately on each of the six assessments. Three strategies were used to determine the number of tentative factors to retain—Cattell's scree plot, absorption of variance and face validity of factors [18]. The tentative multi-item factor structure that best fitted a compromise of these criteria across all assessments was then tested by means of multi-trait scaling analysis, using the Multi-trait analysis program [19]. This technique is based on item-scale correlations and examines convergent and discriminatory validity. Evidence of convergent validity was defined as a correlation larger than 0.40, corrected for overlap, between an item and its expected scale. Discriminatory validity was tested by computing the proportion of items that correlated significantly lower, lower, higher and significantly higher with their expected scale, corrected for overlap, than with the other scales. The scaling success rate in these analyses was defined as the proportion of items that correlated higher or significantly higher with its expected scale than with competing scales.

*Reliability.* Test-retest reliability was examined with intra-class correlations (ICC) between the assessments at inclusion and baseline, since relatively stable health could be expected during this period [20]. The limits of agreement for this period were

also computed. This was defined as the mean of the individual differences between the two assessments  $\pm 1.96$  standard deviations ( $M_d \pm 1.96 \times SD_d$ ) and designates the interval comprising 95% of the differences between the two measurements in stable individuals [21]. The internal consistency of multi-item scales was assessed by Cronbach's alpha.

*Score distributions.* Score distributions of both HDC-19 multi-item scales and single items were examined. These analyses also included proportions of missing values, proportions of respondents scoring at maximum (ceiling) or minimum (floor) levels and the extent that the full range of possible scores was used.

*Clinical validity.* Clinical validity was examined with two methods. Firstly, known-groups comparisons were performed to assess the ability of the HDC-19 to discriminate between mutually exclusive groups of patients expected to differ in clinical status. Secondly, changes in HDC-19 scores over time were compared with changes in EORTC QLQ-C30 'global QOL'. Three categories of mutually exclusive groups were elaborated based on clinical experience: 1) patients with multiple myeloma versus lymphoma; 2) patients in complete remission (CR) versus partial remission (PR); 3) patients receiving stem cells from related donors (RD) versus patients receiving stem cells from unrelated donors (URD). Lymphoma patients were expected to have more symptoms compared to multiple myeloma patients because they had received more chemotherapy (BCNU, Etoposid, ARA-C and Melphalan (BEAM)) in their conditioning before SCT, whereas multiple myeloma patients had only received Melphalan. Patients in CR before SCT were expected to experience fewer symptoms compared to patients in PR because of a better starting-point, i.e. the latter had often received more chemotherapy of higher intensity over a longer period. RD patients were expected to have fewer problems during the SCT than URD patients because of a higher risk for GvHD and infections in this latter group. The known-group comparisons were performed on the assessments one month after transplantation because a higher level of symptoms and problems were expected at this time than at the other assessments.

*Responsiveness.* This was examined by computing dependent t-tests (p-values) and Standardised Response Means (SRM) between baseline and the subsequent assessments. The SRM is defined as the ratio of the mean change to the SD of that

change [20]. Commonly the thresholds suggested by Cohen [22] are used to define what is regarded to be small (0.2–0.5), moderate (0.5–0.8) and large (above 0.8). Most changes in health perceptions were expected to occur between baseline and the two subsequent assessments (one and three months after transplantation, respectively).

## Results

### *Patient characteristics*

Demographic and clinical characteristics of the 202 included patients are shown in Table I. Mean age was 51 (19–70) years and 60% of the patients were male. The largest diagnostic groups were lymphoma (33%) and multiple myeloma (28%). Other diagnoses were acute leukaemia (15%) and chronic leukaemia (7%). One hundred and forty-five (72%) of the patients underwent autologous SCT and 57 (28%) underwent allogeneic SCT. Clinical status at transplantation was complete remission (CR) for 33% and partial remission (PR) for 48%.

The demographic differences between the study group and the 119 patients who declined to participate in the study were small and insignificant. In the latter group, the mean age was 50 (18–69) years and 68% were male. Among clinical variables, the only significant differences were that a greater proportion of the non-participants had multiple myeloma (41% vs. 28% in the study group) and fewer had lymphoma (25% vs. 33% in the study group).

### *Tests of internal structure between items*

The principal component analyses suggested that ten of the 19 items in the HDC19 could be reduced to four components (data not shown). These components (and potential multi-item scales) were named future health perspective (items 11, 12, 13), sexual functioning (items 17, 18, 19), joint and muscle pain (items 9, 10) and skin irritations (items 7, 8). Together these four components accounted for 77–82% of the variance of these ten items (depending on time of measurement). All had eigenvalues above 1 and the Cattell's scree plot flattened after the 4<sup>th</sup> component [18]. Similar patterns were observed in all six assessments. Although the principal component analyses showed that item 19 (degree of sexual enjoyment) was highly related to the two items on sexual functioning (items 17 and 18), this item was treated as a single item in subsequent analyses because respondents were requested to answer this question only if they were sexually active. These three questions are also

Table I. Clinical and demographic data at inclusion (n = 202).

	N (%)
Sex	
Male	121 (60%)
Female	81 (40%)
Age Mean (SD)	51.5 (11.2)
Median(min-max)	54 (19–70)
Employment	
Employed	42 (21%)
Retired	24 (12%)
On sick-list	49 (24%)
Student	2 (1%)
Unknown	85 (42%)
Marital status	
Married	146 (72%)
Single	35 (17%)
Unknown	21 (10%)
Diagnosis <sup>1</sup>	
Lymphoma	66 (32%)
Multiple myeloma	58 (28%)
AL	31 (15%)
CL	14 (7%)
Solid tumors	14 (7%)
Other	19 (9%)
Transplantation	
Autologous	145 (72%)
Allogeneic	57 (28%)
Clinical status <sup>2</sup>	
CR I–II	67 (33%)
PR I–III	91 (48%)
CP	9 (5%)
Not available	35 (14%)
Conditioning <sup>3</sup>	
Melphalan	62 (31%)
BEAM	62 (31%)
Cy/TBI	16 (8%)
Flu/Bu	16 (8%)
Cy/Bu	14 (7%)
Other	23 (11%)

<sup>1</sup>AL = Acute Leukaemia; CL = Chronic Leukemia; Other = Amyloidosis (8 patients), Myelodysplastic Syndrome (5 patients), Mb Waldenström (1 patient), Myelofibrosis (3 patients), Aplastic Anaemia (1 patient), SLE (1 patient).

<sup>2</sup>CR = Complete Remission; PR = Partial Remission; CP = Chronic Phase.

<sup>3</sup>BEAM = BCNU, Etoposid, ARA-C and Melphalan; Cy/TBI = Cyclophosphamide/Total Body Radiation; Flu/Bu = Fludarabine/Busulphan; Cy/Bu = Cyclophosphamide/Busulphan.

treated in this manner in other EORTC modules, such as the breast cancer module [11].

The results of the multi-trait scaling analysis of the four potential scales are listed in Table II. The criterion for item convergent validity (correlations larger than 0.40, corrected for overlap) was reached for all scales at all assessments, except for 'skin irritations' at 1 month and 3 months follow-ups. This supports the item convergent validity. Item/scale discriminatory validity was also satisfactory.

Table II. Item convergent validity, Cronbach's alpha and item/scale discriminatory validity of the four potential multi-item scales in HDC19.

	Health worries 3 items	Sexual functioning 2 items	Joint and muscle pain 2 items	Skin irritations 2 items
Time 1 inclusion (n = 171)				
Mean(Range) convergent validity <sup>a</sup>	.73 (.64–.82)	.76 (.76–.76)	.42 (.42–.42)	.46 (.46–.46)
Cronbach's alpha	.86	.86	.59	.62
Percent success rate discrimin. validity <sup>b</sup>	0/0/0/100	0/0/0/100	0/0/50/50	0/0/0/100
Time 2 baseline (n = 149)				
Mean(Range) convergent validity <sup>a</sup>	.76 (.67–.84)	.78 (.78–.78)	.51 (.51–.51)	.49 (.49–.49)
Cronbach's alpha	.87	.88	.67	.65
Percent success rate discrimin. validity <sup>b</sup>	0/0/0/100	0/0/0/100	0/0/13/87	0/0/0/100
Time 3 1 month after tx (n = 158)				
Mean(Range) convergent validity <sup>a</sup>	.75 (.63–.84)	.73 (.73–.73)	.46 (.46–.46)	.36 (.36–.36)
Cronbach's alpha	.87	.83	.63	.49
Percent success rate discrimin. validity <sup>b</sup>	0/0/0/100	0/0/0/100	0/0/0/100	0/37/37/25
Time 4 3 months after tx (n = 149)				
Mean(Range) convergent validity <sup>a</sup>	.75 (.68–.83)	.76 (.76–.76)	.56 (.56–.56)	.38 (.38–.38)
Cronbach's alpha	.87	.91	.72	.55
Percent success rate discrimin. validity <sup>b</sup>	0/0/0/100	0/0/0/100	0/0/0/100	0/0/37/63
Time 5 6 months after tx (n = 139)				
Mean(Range) convergent validity <sup>a</sup>	.78 (.72–.82)	.71 (.71–.71)	.63 (.63–.63)	.48 (.48–.48)
Cronbach's alpha	.89	.83	.77	.65
Percent success rate discrimin. validity <sup>b</sup>	0/0/0/100	0/0/0/100	0/0/0/100	0/0/0/100
Time 6 12 months after tx (n = 132)				
Mean(Range) convergent validity <sup>a</sup>	.81 (.72–.86)	.77 (.77–.77)	.55 (.55–.55)	.53 (.53–.53)
Cronbach's alpha	.90	.87	.71	.69
Percent success rate discrimin. validity <sup>b</sup>	0/0/0/100	0/0/0/100	0/0/0/63	0/0/25/75

<sup>a</sup>Correlations between items and hypothesized scale corrected for overlap.

<sup>b</sup>Percent correlations that are 'significantly lower/lower/higher/significantly higher' with hypothesized scale compared to other scales.

The proportion of items correlating higher or significantly higher with their hypothesized scale than with competing scales (success rate) was 100% for all four scales at all assessments, except for 'skin irritations' at 1 month after tx.

#### Score distributions

Table III lists the means and SD's for the total sample at all six assessments. The most common symptoms were 'feeling cold more easily', 'change of taste', 'dry mouth' and 'potency problems' (among men). 'Sexual functioning' was also low, but 'future health perspective' was relatively high. The lowest levels of functioning and most symptoms were found one and three months after transplantation.

The proportion of missing values was low for all scales/ items (data not shown). The highest proportion missing was observed for 'fear of sterility' (6–7%), 'someone to talk with' (5–6%) and 'interested in sex' (3–4%); otherwise the proportions of missing values were around 1% or less.

Table IV shows the percent of patients responding at floor and ceiling, respectively. As can be seen all symptom scales were highly skewed with a large

percent responding at floor, particularly 'skin irritations', 'increased mucous production', 'soreness in the mouth' and 'worries about sterility'. This indicates that relatively few patients experienced high levels of symptoms/problems. However, the distribution of symptoms was less skewed at the assessments 1 month and 3 months after tx, i.e. when the influence of treatment was at maximum. The functional scales were more symmetrical, except for the sexual functioning scale. Most respondents scored at the floor level for this scale, indicating that the level of sexual functioning was relatively low during the whole study period.

#### Reliability tests

Table V shows the test-retest reliability coefficients (intra-class correlations) between the assessments at inclusion and baseline. These were very high for all scales, ranging from 0.96–1.00. The means of the individual differences were all around zero and the 95% limits of agreement of the observations were narrow. Cronbach's alpha was also good for 'sexual functioning' and 'future health perspectives' (above 0.70), except for 'joint and muscle pain'

Table III. Means (SD) of the HDC19.

	No. of items	Inclusion	Baseline	1 month after tx	3 months after tx	6 months after tx	12 months after tx
<b>Functioning scales<sup>a</sup></b>							
Sexual functioning	2	23.5 (23.9)	24.8 (24.5)	14.4 (20.2)	24.3 (22.2)	27.3 (23.6)	30.5 (26.1)
Sexual enjoyment	1	62.4 (26.0)	64.9 (25.0)	53.7 (26.7)	56.1 (23.2)	64.4 (25.7)	61.3 (24.8)
Future health perspect.	3	64.3 (27.6)	66.2 (26.9)	65.3 (27.2)	68.3 (26.4)	69.4 (26.1)	66.1 (27.7)
Someone to talk to	1	57.4 (33.4)	58.8 (32.3)	52.3 (35.0)	52.2 (36.2)	54.6 (33.9)	52.9 (36.1)
<b>Symptom scales<sup>b</sup></b>							
Joint and muscle pain	2	24.4 (23.6)	21.8 (23.1)	23.5 (21.5)	32.5 (25.3)	26.9 (25.5)	28.6 (23.7)
Skin irritations	2	17.4 (18.7)	16.3 (17.7)	24.5 (21.7)	19.5 (20.2)	20.0 (21.4)	20.0 (22.2)
Felt cold more easily	1	24.9 (27.2)	21.1 (25.2)	44.1 (28.8)	30.1 (29.3)	20.2 (26.4)	19.8 (23.4)
Been dizzy	1	13.7 (8.24)	10.7 (18.4)	20.7 (25.1)	12.3 (18.7)	11.0 (18.4)	11.8 (19.2)
Increased mucous prod.	1	8.2 (17.9)	5.9 (14.6)	18.5 (30.3)	10.2 (18.8)	5.3 (15.6)	8.5 (20.6)
Soreness in the mouth	1	2.4 (10.6)	1.6 (8.32)	8.7 (20.5)	3.3 (12.5)	5.5 (15.7)	5.1 (15.0)
Dry mouth	1	16.9 (25.9)	14.1 (22.6)	33.1 (32.0)	29.2 (32.1)	24.6 (30.9)	20.0 (26.0)
Change of taste	1	19.3 (27.6)	15.8 (25.7)	51.6 (32.5)	21.6 (28.1)	14.6 (26.9)	10.6 (22.4)
Worries about sterility	1	7.1 (21.2)	8.7 (23.5)	7.5 (21.3)	6.3 (18.4)	.1 (21.6)	6.6 (21.3)
Potency problems	1	37.2 (36.2)	33.3 (35.5)	48.9 (37.5)	36.1 (38.1)	35.9 (33.0)	36.4 (38.1)

<sup>a</sup>Scores range from 0 to 100, with higher scores representing higher level of functioning.

<sup>b</sup>Scores range from 0 to 100, with higher scores representing higher levels of symptoms.

and 'skin irritations' (Table II). However, it should be noted that these latter scales consisted of only two items.

#### Clinical validity

The known-groups comparisons are shown in Table VI. Multiple myeloma patients reported significantly higher levels on 'joint and muscle pain' and a more negative 'future health perspective' than the lymphoma patients. However, the lymphoma

patients less frequently reported that they had 'someone to talk with'. The CR patients reported significantly lower 'sexual functioning', 'felt cold more easily' and also had more problems with 'dry mouth'. However, no significant differences were found between the RD and URD groups.

As expected, the changes in the HDC-19 mirrored those in the EORTC QLQ-C30 'global QOL' (data not shown), indicating that changes in the specific functions/problems addressed in HDC-19 also corresponded to changes in global QOL. All

Table IV. Percent at floor and ceiling of the HDC-19.

	Inclusion		Baseline		1 month after tx		3 months after tx		6 months after tx		12 months after tx	
	% floor	% ceil.	% floor	% ceil.	% floor	% ceil.	% floor	% ceil.	% floor	% ceil.	% floor	% ceil.
<b>Functioning scales<sup>a</sup></b>												
Sexual functioning	37	1	37	1	56	1	35	0	32	0	27	3
Sexual enjoyment <sup>b</sup>	1	23	2	24	5	15	3	11	1	25	0	21
Future health perspect.	6	14	7	15	5	17	5	20	6	18	5	22
Someone to talk to	14	26	13	27	19	23	21	25	16	24	19	27
<b>Symptom scales<sup>c</sup></b>												
Joint and muscle pain	33	1	37	1	31	1	20	2	30	2	24	1
Skin irritations	40	0	42	0	26	2	36	1	36	1	39	0
Felt cold more easily	45	4	49	4	15	11	39	5	55	3	52	1
Been dizzy	68	1	68	1	52	2	67	0	70	1	70	0
Increased mucous prod.	80	1	80	1	68	6	75	0	88	0	83	1
Soreness in the mouth	94	0	96	0	81	2	92	1	87	1	88	1
Dry mouth	63	3	65	3	37	9	44	9	51	8	54	4
Change of taste	60	5	61	5	16	19	55	4	72	5	78	1
Worries about sterility	88	3	85	3	87	3	87	2	88	4	89	3
Potency problems	39	14	40	15	26	24	43	18	32	13	45	17

<sup>a</sup>Higher scores represent higher level of functioning.

<sup>b</sup>Note that respondents were requested to answer this question only if they were sexually active.

<sup>c</sup>Higher scores represent higher levels of symptoms.

Table V. Test-retest agreement between the assessments at inclusion and baseline. Intraclass correlations (ICC) and limits of agreement ( $M_d \pm 1.96 \times SD_d$ ).

	ICC	Limits of agreement
Functioning scales		
Sexual functioning	0.99	0.11 $\pm$ 4.6
Sexual enjoyment	0.98	0.54 $\pm$ 8.3
Future health perspective	0.99	-0.15 $\pm$ 5.0
Someone to talk to	0.98	0.22 $\pm$ 9.3
Symptom scales		
Joint and muscle pain	0.96	0.00 $\pm$ 10.5
Skin irritations	0.96	0.64 $\pm$ 9.0
Felt cold more easily	0.98	0.43 $\pm$ 7.4
Been dizzy	0.96	-0.21 $\pm$ 9.1
Increased mucous production	0.98	-0.21 $\pm$ 5.3
Soreness in the mouth	1.00	0.00 $\pm$ 0.0
Dry mouth	0.97	0.22 $\pm$ 9.1
Change of taste	0.99	-0.21 $\pm$ 5.2
Worries about sterility	0.97	-0.68 $\pm$ 9.3
Potency problems	0.99	-0.39 $\pm$ 7.2

correlations were significant except for 'someone to talk to', 'worries of sterility', 'skin irritations' and 'soreness in the mouth'.

### Responsiveness

Mean changes, p-values for dependent t-tests and standardised response means (SRM) between baseline and subsequent assessments are shown in Table VII. As expected, most changes were found between the assessments at baseline and one month after transplantation. This was particularly true for 'feeling more easily cold', 'dry mouth' and 'change of taste'. Moderate changes were noted for

'sexual functioning', sexual enjoyment', 'skin irritations', 'dizziness', 'increased mucous production, 'soreness in the mouth' and 'potency problems'. Smaller changes were observed between baseline and the assessment three months after transplantation. Although, functions and symptoms largely tended to improve after 3 months, significant impairments were still found up to 12 months after transplantation for 'sexual functioning', 'joint and muscle pain' and 'change of taste'. However, the changes between the 3<sup>rd</sup> month to the 6<sup>th</sup> and 12<sup>th</sup>, month, respectively, were small (data not shown).

### Discussion

The HDC-19 was developed to complement the EORTC QLQ-C30 in order to provide additional information specific for SCT [15]. No previous studies have sought to determine the psychometric properties for the HDC-19, other than the distribution of means reported by Hjermstad et al. [15].

The original HDC-19 consisted of 19 single items. Principal component analysis suggested four multi-item scales (future health perspective, sexual functioning, joint and muscle pain and skin irritations) and ten single items (felt cold more easily, been dizzy, increased mucus production, sore mouth, dry mouth, change of taste, anxiety, worries about sterility, potency problems and sexual enjoyment). The internal consistencies for the multi-item scales, as assessed by Cronbach's alpha coefficients, were well above the preferred level of 0.70, except for joint and muscle pain and skin irritations. However, these

Table VI. Known-groups comparisons. Patients with myeloma (n = 47) versus lymphoma (n = 47); Patients with complete remission (CR) (n = 61) versus partial remission (PR) (n = 44); Patients with related donor (RD) (n = 20) versus unrelated donor (URD) (n = 26). Mean values (SD) and p-values from independent t-tests one month after transplantation are presented.

	Myeloma – Lymphoma	p	CR – PR	p	RD – URD	p
Functioning scales						
Sexual functioning	17.7 (23.9)–13.1 (17.0)	0.28	11.9 (18.4)–19.5 (20.0)	<b>0.05</b>	20.9 (26.9)–9.6 (16.4)	0.09
Sexual enjoyment	51.3 (25.9)–48.5 (22.9)	0.78	57.1 (30.5)–48.9 (21.3)	0.40	61.9 (35.6)–50.0 (43.0)	0.63
Future health perspective	56.2 (29.7)–68.3 (26.2)	<b>0.04</b>	66.8 (23.2)–70.2 (26.1)	0.49	67.2 (28.7)–72.2 (20.7)	0.49
Someone to talk to	55.8 (39.2)–39.4 (28.1)	<b>0.03</b>	57.6 (32.7)–40.0 (34.8)	<b>0.01</b>	57.9 (34.9)–64.1 (32.6)	0.54
Symptom scales						
Joint and muscle pain	27.6 (21.9)–19.4 (19.9)	<b>0.05</b>	23.7 (23.5)–20.1 (18.5)	0.39	20.9 (20.9)–26.3 (23.7)	0.42
Skin irritations	24.8 (26.4)–20.4 (18.1)	0.34	25.4 (19.2)–22.3 (21.2)	0.44	26.7 (23.2)–25.6 (17.2)	0.86
Felt cold more easily	38.8 (29.1)–44.2 (28.3)	0.35	49.2 (25.5)–36.4 (26.7)	0.01	53.3 (27.4)–46.2 (29.9)	0.41
Been dizzy	14.9 (22.6)–23.1 (25.6)	0.10	23.5 (27.4)–21.2 (23.9)	0.65	20.0 (27.4)–24.6 (29.2)	0.61
Increased mucous production	23.8 (31.9)–23.1 (33.5)	0.92	19.1 (31.9)–18.2 (27.3)	0.87	20.0 (34.9)–7.7 (19.6)	0.14
Soreness in the mouth	12.2 (24.2)–6.8 (15.2)	0.19	9.2 (22.9)–7.6 (20.2)	0.62	10.0 (26.7)–7.7 (21.7)	0.75
Dry mouth	36.1 (31.8)–38.9 (31.7)	0.66	37.2 (33.9)–24.0 (27.5)	<b>0.04</b>	30.0 (35.7)–20.5 (28.4)	0.32
Change of taste	57.1 (33.3)–51.0 (34.1)	0.37	54.6 (32.8)–46.9 (28.1)	0.21	48.3 (29.6)–44.9 (32.6)	0.71
Worries about sterility	2.8 (9.40)–8.5 (22.5)	0.11	5.7 (16.7)–9.8 (23.4)	0.30	8.7 (26.9)–14.7 (30.6)	0.51
Potency problems	60.5 (37.0)–54.8 (35.0)	0.55	45.6 (35.5)–54.8 (34.2)	0.32	33.3 (30.9)–40.0 (44.0)	0.71

Table VII. Mean (SD) changes from baseline, p-values and Standardised Response Means (SRM).

	Baseline – 1 month			Baseline – 3 months			Baseline – 6 months			Baseline – 12 months		
	Mean (SD) change from baseline	p-value	SRM <sup>c</sup>	Mean (SD) change from baseline	p-value	SRM <sup>c</sup>	Mean (SD) change from baseline	p-value	SRM <sup>c</sup>	Mean (SD) change from baseline	p-value	SRM <sup>c</sup>
Functioning scales <sup>a</sup>												
Sexual functioning	-8.6 (23.9)	<b>0.00</b>	0.38	-0.1 (21.6)	0.31	0.09	+3.9 (24.3)	0.07	0.16	+6.5 (26.3)	<b>0.01</b>	0.25
Sexual enjoyment	-13.8 (30.2)	<b>0.02</b>	0.48	-4.2 (24.4)	0.16	0.21	+3.8 (21.8)	0.26	0.18	+0.0 (25.0)	0.99	0.00
Future health perspec.	+0.8 (21.7)	0.96	0.00	+4.4 (28.5)	0.11	0.15	+5.3 (22.3)	<b>0.01</b>	0.24	+3.1 (22.7)	0.13	0.14
Someone to talk to	-4.8 (29.6)	0.18	0.13	-6.2 (28.5)	<b>0.04</b>	0.20	-4.4 (30.7)	0.12	0.14	-6.5 (25.4)	0.07	0.18
Symptom scales <sup>b</sup>												
Joint and muscle pain	+0.2 (21.1)	0.63	0.04	+9.6 (23.9)	<b>0.00</b>	0.45	+5.3 (28.2)	<b>0.03</b>	0.19	+6.5 (25.4)	<b>0.01</b>	0.26
Skin irritations	+8.1 (23.3)	<b>0.00</b>	0.36	+3.7 (21.8)	<b>0.03</b>	0.20	+4.0 (24.1)	0.06	0.17	+4.1 (26.0)	0.08	0.16
Felt cold more easily	+20.5 (32.4)	<b>0.00</b>	0.74	+5.1 (29.6)	<b>0.01</b>	0.24	-2.0 (30.5)	0.45	0.06	-2.9 (27.8)	0.24	0.11
Been dizzy	+6.2 (29.5)	<b>0.00</b>	0.31	+2.1 (22.8)	0.42	0.07	-2.8 (23.3)	0.17	0.12	-1.9 (26.4)	0.43	0.07
Increased mucous prod.	+10.0 (30.2)	<b>0.00</b>	0.39	+2.2 (22.8)	<b>0.02</b>	0.21	-3.03 (18.2)	0.06	0.17	-0.0 (23.1)	0.99	0.00
Soreness in the mouth	+7.3 (21.6)	<b>0.00</b>	0.32	+0.7 (12.3)	0.37	0.08	+4.5 (15.8)	<b>0.00</b>	0.29	+2.6 (15.5)	0.06	0.17
Dry mouth	+17.5 (32.7)	<b>0.00</b>	0.60	+11.8 (20.3)	<b>0.00</b>	0.46	+9.1 (30.6)	<b>0.00</b>	0.30	+4.5 (31.7)	0.11	0.14
Change of taste	+31.9 (36.9)	<b>0.00</b>	0.94	+1.7 (30.3)	0.22	0.11	-3.8 (34.9)	0.21	0.11	-7.4 (32.9)	<b>0.01</b>	0.23
Worries about sterility	-1.5 (15.6)	0.36	0.08	-2.3 (17.1)	0.17	0.13	-1.4 (16.2)	0.36	0.08	-1.72 (18.6)	0.32	0.09
Potency problems	+10.7 (35.7)	<b>0.01</b>	0.34	+4.4 (35.2)	0.72	0.05	1.9 (32.3)	0.06	0.06	-1.5 (41.2)	0.77	0.04

<sup>a</sup>Higher scores indicate better functioning.

<sup>b</sup>Higher scores indicate more symptoms.

<sup>c</sup>The thresholds suggested by Cohen (1988) are; small (0.2–0.5), moderate (0.5–0.8) and large (above 0.8).

scales consist of only two items each. Multitrait scaling analysis also showed that most item-scale correlation coefficients met the standards of convergent and discriminatory validity.

The distribution of means was comparable with that reported by Hjermstad and colleagues [15]. Mean values ranged from 52.2 to 69.4 on the functional scales indicating that our patients experienced their functional status as satisfactory throughout the SCT period. An important exception was sexual functioning, which had considerably lower means ranging from 14.4–30.5. It is not surprising that sexual functioning was impaired in this patient group and may be attributed to late effects related to chemotherapy, fatigue, GvHD and problems with potency among men. This result is also consistent with other studies [23,24].

Results from the symptom scales indicated that our patients experienced most of their symptoms one month after SCT, at which point the highest mean values were registered for 'change of taste', 'dry mouth', 'feeling cold more easily' and men experienced most potency problems. However, only for 'change of taste' this change could be considered as large according to the SRM criteria suggested by Cohen [22]. Otherwise, the changes were moderate to small. Consistent with other studies on HRQOL after SCT [23,25–27], our patients still experienced some degree of symptoms as long as one year after SCT. Also consistent with these studies, our results indicated that most patients seemed to reach an acceptable level of functioning during the first year after SCT and reported their QOL to be good or excellent. Nevertheless, when specifically assessed, some areas remain impaired, such as physical functioning [25], sexual problems [23] and fatigue [27].

The test-retest reliability between assessments at inclusion and baseline were very high for all scales and single items, indicating that perceived health status remained virtually unchanged during this period. Similar correlations were also computed between assessments at baseline and one month after transplantation when most changes in perceived problems and symptoms related to high-dose chemotherapy and SCT were hypothesised to occur. As expected, the correlations were considerably lower between these two assessments.

Despite considerable heterogeneity in the patient sample regarding diagnosis, disease status, chemotherapy regimens and source of stem cells, our SCT patients showed remarkably little variation in HRQOL and transplant-related problems. The ability of the HDC-19 to discriminate between subgroups in the known-groups comparisons was

relatively poor, but the problems to find mutually exclusive groups should be noted. There are many factors that may influence the outcome for this group of patients and that the effect of a SCT is highly individual. However, lymphoma patients supposed to have more symptoms than myeloma patients depending on different conditioning regimens. Instead, multiple myeloma patients reported higher levels on 'joint and muscle pain' compared to the lymphoma patients. They also reported a more negative future health perspective. One plausible reason for this is that multiple myeloma patients know that they cannot be cured with SCT, but rather that the most they can hope for is a longer period of remission. Patients in PR were expected to have more symptoms and problems during the SCT period, but instead patients in CR reported significantly lower 'sexual functioning' and more 'dry mouth' and 'change of taste'. We also compared patients receiving stem cells from related and unrelated donors because the latter are known to have a higher risk of GvHD and infections. The HDC-19 was not, however, able to discriminate between these groups. This could be due to the fact that the patients did not have GvHD or any symptoms of GvHD at this measurement point.

Supporting the clinical validity of the HDC-19 was, however, that most of the scales mirrored changes in EORTC QLQ-C30 global QOL item, except skin irritation and three single items (anxiety, sore mouth and sterility). This result lends support to the clinical validity of the HDC-19.

## Conclusion

The results of this study give support to the validity of the HDC-19. The Swedish version of HDC-19 was found to be a reliable and valid HRQOL measure for patients receiving SCT after HDC, which indicates that it can be used to assess specific problems and symptoms related to SCT. This is one of the first studies to test the validity and reliability of HDC-19 in a larger group of patients receiving SCT after HDC. Additional studies are recommended to further evaluate the psychometric properties of the HDC-19 and possibly to also compare this ad hoc module with the EORTC HDC29 module, which is a continuation of HDC-19 extended with questions concerning body image, impact on family and in-patient issues. During the development (phase I–III), HDC-29 was tested in four countries and a larger sample was used compared to the development of the original HDC-19. However the final psychometric validation (Phase IV) is still ongoing and HDC29 is not yet an official EORTC module.

## References

- [1] Gratwohl A, Baldomero H, Frauendorfer K, Urbano-Ispizua A. EBMT activity survey 2004 and changes in the disease indication over the past 15 years. *Bone Marrow Transplant* 2006;37:1069–85.
- [2] Ljungman P, Urbano-Ispizua A, Cavazzana-Calvo M, Demirer T, Dini G, Einsele H, et al. Allogeneic and autologous transplantation for haematological diseases, solid tumors and immune disorders: Definition and current practice in Europe. *Bone Marrow Transplant* 2006;37:439–49.
- [3] Bakitas WM. In: Wujcik D, editor. Blood and marrow stem-cell transplantation. Indications, procedure, process. Sudbury: Jones & Bartlett Publishers; 1997.
- [4] Sanders JE. Chronic graft-versus-host disease and late effects after hematopoietic stem cell transplantation. *Int J Hematol* 2002;76(Suppl 2):15–28.
- [5] Tabbara IA, Zimmermann K, Morgan C, Nahleh Z. Allogeneic hematopoietic stem cell transplantation: Complications and result. *Arch Intern Med* 2002;162:1558–66.
- [6] Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez N, et al. The European organisation for research and treatment of cancer: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365–76.
- [7] Muldon M, Barger J, Manuck S. What are quality of life measurements measuring? *BMJ* 1998;316:542–5.
- [8] Sprangers M. Quality of life assessment in oncology. *Acta Oncol* 2002;41:229–37.
- [9] Bjordal K, Ahlner-Elmqvist M, Tolleson EA, Razavi D, Maher E, Kaasa S. Development of a European organisation for research and treatment of cancer (EORTC) questionnaire module to be used in quality of life assessments in head and neck cancer patients. *Acta Oncol* 1994;33:879–85.
- [10] Blazeby JM, Alderson D, Winstone K, Steyn R, Hammerlid E, Arraras J, et al. Development of an EORTC questionnaire module to be used in quality of life assessment for patients with oesophageal cancer. *Eur J Cancer* 1996;32:1912–7.
- [11] Sprangers M, Groenvold M, Arras J, Franklin J, te Velde A, Muller M, et al. The EORTC breast-cancer specific quality-of-life questionnaire (QLQ-BR23): First results from a three-country field study. *J Clin Oncol* 1996;14:2756–68.
- [12] Sprangers M, te Velde A, Aaronson A. The construction and testing of the EORTC colorectal cancer-specific quality of life questionnaire module (QLQ-QR38). *Eur J Cancer* 1999;35:238–47.
- [13] Blazeby JM, Conroy T, Bottomley A, Vickery C, Arraras J, Sezer O, et al. on behalf of the European organisation for research and treatment of cancer gastrointestinal and quality of life groups. Clinical and psychometric validation of a questionnaire module, the EORTC QLQ-STO22, to assess quality of life in patients with gastric cancer. *Eur J Cancer* 2004;40:2260–8.
- [14] Sprangers M, Cull A, Bjordal K, Groenvold M, Aaronson NK. The European organization for research and treatment of cancer approach to quality of life assessment: guidelines for developing questionnaire modules. *Qual Life Res* 1993;2:287–95.
- [15] Hjermstad M, Holte H, Evensen S, Fayers P, Kaasa S. Do patients who are treated with stem cell transplantation have a health related quality of life comparable to the general population after 1 year? *Bone Marrow Transplant* 1999a;24:911–8.
- [16] Velikova G, Weis J, Hjermstad M, Kopp M, Morris P, Watson M, et al. The EORTC QLQ-HDC29: A supplementary module assessing the quality of life during and after high-dose chemotherapy and stem cell transplantation. *Eur J Cancer* 2007;43:87–94.
- [17] EORTC Quality of life study group. EORTC QLQ-C30 scoring manual, 3rd ed. Brussels: EORTC; 2001.
- [18] Gorsuch RL. Factor analysis. Hillsdale NJ: Lawrence Erlbaum Associates; 1983.
- [19] Hays RD, Hayashi T, Carson S, Ware JE. User's guide for the multi-trait analysis program (MAP). The Rand Publication series; 1988.
- [20] Fayers PM, Machin D. Quality of life: Assessment, analysis and interpretation. Chichester UK: John Wiley & Sons; 2001.
- [21] Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
- [22] Cohen J. Statistical power analysis for the behavioural sciences. Hillsdale NJ: Lawrence Erlbaum; 1988.
- [23] Edman L, Larsen J, Häggglund H, Gardulf A. Health-related quality of life, symptom distress and sense of coherence in adult survivors of allogeneic stem-cell transplantation. *Eur Cancer Care* 2001;10:124–30.
- [24] Molassiotis A, Morris P. Quality of life in patients with chronic myeloid leukaemia after unrelated donor bone marrow transplantation. *Cancer Nurs* 1999;22:340–9.
- [25] Broers S, Kaptein A, Saskia L, Fibbe W, Hengeveld M. Psychological functioning and quality of life following bone marrow transplantation: A 3-year follow-up study. *J Psychosom Res* 2000;48:11–21.
- [26] Hjermstad M, Evensen S, Kvaloy S, Fayers P, Kaasa S. Health-related quality of life 1 year after allogeneic or autologous stem-cell transplantation: A prospective study. *J Clin Oncol* 1999b;17:706–18.
- [27] Hjermstad M, Knobel H, Brinch L, Fayers PM, Loge JH, Holte H, et al. A prospective study of health-related quality of life, fatigue, anxiety and depression 3–5 years after stem cell transplantation. *Bone Marrow Transplant* 2004;34:257–66.