

A validated decision model for treating the anaemia of myelodysplastic syndromes with erythropoietin + granulocyte colony-stimulating factor: significant effects on quality of life

EVA HELLSTRÖM-LINDBERG,¹ NINA GULBRANDSEN,² GREGER LINDBERG,¹ TOMAS AHLGREN,³ INGER MARIE S. DAHL,⁴ INGUNN DYBEDAL,⁵ GUNNAR GRIMFORS,⁶ EVA HESSE-SUNDIN,⁷ MARTIN HJORTH,⁸ LENA KANTER-LEWENSOHN,¹⁶ OLLE LINDER,⁹ MICHAELA LUTHMAN,¹⁰ EVA LÖFVENBERG,¹¹ GUNNAR ÖBERG,¹² ANJA PORWIT-MACDONALD,¹⁶ ANDERS RÅDLUND,¹³ JAN SAMUELSSON,¹⁴ JON MAGNUS TANGEN,² INGEMAR WINQUIST¹⁵ AND FINN WISLOFF² FOR THE SCANDINAVIAN MDS GROUP ¹Department of Medicine, Karolinska Institutet at Huddinge University Hospital, Stockholm, Sweden, ²Department of Haematology, Ullevål Hospital, Oslo, Norway, Departments of Medicine/Haematology at the Hospitals in ³Malmö, Sweden, ⁴Tromsø and ⁵Trondheim Norway, ⁶Karolinska Hospital, Stockholm, ⁷Eskilstuna, ⁸Lidköping, ⁹Örebro, ¹⁰St Görans Hospital, Stockholm, ¹¹Umeå, ¹²Uppsala, ¹³Linköping, ¹⁴Southern Hospital, Stockholm, ¹⁵Lund, and ¹⁶Department of Pathology, Karolinska Hospital, Sweden

Received 20 June 2002; accepted 23 September 2002

Summary. We have published previously a prototype of a decision model for anaemic patients with myelodysplastic syndromes (MDS), in which transfusion need and serum erythropoietin (S-Epo) were used to define three groups with different probabilities of erythroid response to treatment with granulocyte colony-stimulating factor (G-CSF) + Epo. S-Epo \leq 500 U/l and a transfusion need of $<$ 2 units/month predicted a high probability of response to treatment, S-Epo $>$ 500 U/l and \geq 2 units/month for a poor response, whereas the presence of only one negative prognostic marker predicted an intermediate response. A total of 53 patients from a prospective study were included in our evaluation sample. Patients with good or intermediate probability of response were treated with G-CSF + Epo. The overall response rate was 42% with 28.3% achieving a complete and 13.2% a partial response to treatment. The

response rates were 61% and 14% in the good and intermediate predictive groups respectively. The model retained a significant predictive value in the evaluation sample ($P < 0.001$). Median duration of response was 23 months. Scores for global health and quality of life (QOL) were significantly lower in MDS patients than in a reference population, and fatigue and dyspnoea was significantly more prominent. Global QOL improved in patients responding to treatment ($P = 0.01$). The validated decision model defined a subgroup of patients with a response rate of 61% (95% confidence interval 48–74%) to treatment with G-CSF + Epo. The majority of these patients have shown complete and durable responses.

Keywords: anaemia, erythropoietin, granulocyte colony-stimulating factor, myelodysplasia.

The concept of low-risk myelodysplastic syndromes (MDS) may be defined as MDS with a low probability of progression to acute myelogenous leukaemia (AML) and with a relatively favourable survival. These patients mostly belong to the French–American–British (FAB) groups refractory anaemia (RA), RA with ringed sideroblasts (RARS) or a subgroup of

patients with RA with excess of blasts (RAEB) (Bennett *et al.*, 1982). Using the new World Health Organization (WHO) system, they are classified as RA and RARS, with or without multilineage dysplasia, or RAEB I (Jaffe *et al.*, 2001). According to the International Prognostic Scoring System (IPSS) for risk stratification, the majority of these patients have low or intermediate-1 risk, but some patients could belong to the intermediate-2 group because of advanced karyotypic abnormalities (Greenberg *et al.*, 1997). Anaemia is the major clinical problem for these patients, and repeated transfusions may lead to secondary haemochromatosis (Greenberg, 1994). Quality of life has been reported to be

Correspondence: Eva Hellström-Lindberg, MD, PhD, Department of Medicine, Division of Haematology, Karolinska Institute at Huddinge University Hospital, 141 86 Stockholm, Sweden. E-mail: Eva.Hellstrom-Lindberg@medhs.ki.se

reduced (Thomas, 2001; Kornblith et al, 2002), but there are no previous studies comparing patients with age-matched healthy individuals. The bone marrow is mainly characterized by ineffective haematopoiesis, reduced clonogenic growth *in vitro* and an increased proportion of apoptotic bone marrow precursors (Raza et al, 1995; Hellström-Lindberg et al, 1997a; Van de Loosdrecht & Vellenga, 2000). Colony culture studies of myelodysplastic bone marrow have shown that the impaired erythroid growth may respond to erythropoietin (Epo) alone, but that it may be improved further if Epo is combined with other growth factors (Backx et al, 1992; Schmidt-Mende et al, 2001).

Treatment with Epo may improve the anaemia in about 15–20% of patients with MDS, and responses may extend over several years (Hellstrom-Lindberg, 1995; Hast et al, 2001). A placebo-controlled randomized study showed that Epo improved haemoglobin levels in MDS: however, significant differences were observed only in patients with refractory anaemia (RA) and in patients without transfusion need before treatment (Italian Cooperative Study Group for rHuEpo in Myelodysplastic Syndromes, 1998).

The combination of Epo + granulocyte colony-stimulating factor (G-CSF) may improve haemoglobin levels in around 40% of patients with MDS (Negrin et al, 1996; Hellström-Lindberg et al, 1998). Clear evidence of a synergistic effect between the two drugs *in vivo* has been demonstrated in several studies. The synergistic effect seems to be most pronounced in patients with RA with ringed sideroblasts (RARS), ≈50% of whom respond to the combination. A recently published study, including 28 patients with good prognostic factors and short disease duration, showed a response rate of 61% and demonstrated several late responses (12–36 weeks after start of treatment; Mantovani et al, 2000). Finally, a preliminary report of a prospective randomized study of treatment with G-CSF + Epo versus supportive care only indicated a significant effect of treatment on haemoglobin levels and a clear synergistic effect of the combination compared with Epo alone (Casadevall et al, 2001).

We have identified previously serum erythropoietin (S-Epo) and pretreatment transfusion need as significant

variables for a response to treatment of the anaemia in MDS with G-CSF + Epo and used them to construct a prototype predictive model (Fig 1; Hellström-Lindberg et al, 1997b). The cut-off values for pretreatment transfusion need were <2 or ≥2 units per month and, for S-Epo, <100, 100–500 and >500 U/l. In the model, patients with S-Epo ≤500 U/l and a transfusion need of <2 units/month had a high probability of response to treatment. S-Epo >500 U/l together with a transfusion need of 2 or more units/month predicted a poor response to treatment, whereas the presence of one but not the other negative prognostic marker (S-Epo >500 U/l or transfusion need ≥2 units/month) predicted an intermediate response. Although patients with S-Epo <100 U/l and those with S-Epo 100–500 U/l were allocated to the same predictive group in the decision model, the ternary variable was more informative in the logistic regression analysis than the binary variable. Therefore, we maintained the ternary variable in the model to improve its usefulness in future clinical studies with larger patient cohorts.

Many predictive models should be defined as prototype models, as they constitute the result of a first regression analysis. For a final model, the prototype model has to be tested and validated in a prospective study and, if possible, it should also be tested by another study group with independent patient material (Spiegelhalter, 1983). Our previously published model has been validated in one study (Remacha et al, 1999). It was also evaluated by Mantovani et al (2000) with less significant results, possibly because of small study numbers (28 patients) that included a selection of more benign cases with short disease duration.

The present study was designed to test prospectively the predictive value of the previously published model, and to compare the predictive value of the model with other pretreatment variables. The clinical usefulness of the treatment was evaluated by analyses of quality of life (QOL) and duration of response. The study was also designed to study optimal dosing during the maintenance phase, but these results are only reported briefly here.

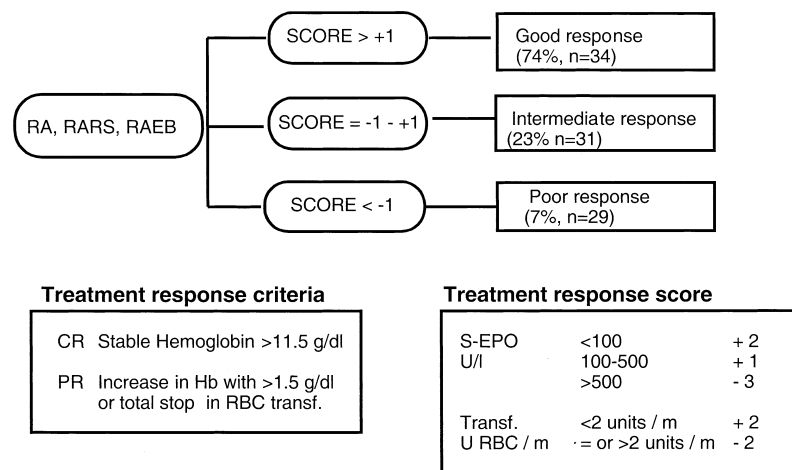


Fig 1. The prototype model (Hellström-Lindberg et al, 1997b) was used for selecting patients from the good and intermediate risk groups for the present study. The model retained a significant predictive value in the evaluation sample ($P = 0.001$), and the observed response rates did not differ from those expected by the model, thus validating the model.

PATIENTS AND METHODS

Patients. Patients were included in a Scandinavian phase IV multicentre study between November 1996 and March 1999. The diagnostic process and the morphological validation were performed as described previously (Hellström-Lindberg *et al*, 1998). A central clinical and morphological analysis was performed between inclusion in the study and the start of treatment to verify correct inclusion criteria. All patients signed consent forms, and the studies followed the guidelines of the investigation review boards of Sweden and Norway.

Patients with RA, RARS or RA with excess of blasts (RAEB) with either haemoglobin levels <10 g/dl or transfusion-dependent anaemia were included. The patients belonged to the good or intermediate groups, according to the prototype model, whereas those patients belonging to the poor responsive group were not eligible and were given other treatment. Patients were transfused when the haemoglobin level fell to 8.0 g/dl, unless a higher transfusion level was indicated by other medical conditions. The cut-off point for transfusion was constant during the study in each individual case. Patients with active ongoing bleeding or transfusion-dependent thrombocytopenia were excluded.

Sixty-three patients were included in the study. Nineteen had RA, 24 had RARS, and 20 had RAEB. A validated analysis of the number of dysplasias in each case was not performed at inclusion, and we did not reclassify patients according to the new WHO classification (Jaffe *et al*, 2001). However, of the 20 patients with RAEB, only five were of the RAEB II type. Thirty-four, 27 and two patients belonged to the good, intermediate and poor predictive groups respectively. The last two were reallocated to the poor response group at the review before the start of treatment and were defined as dropouts. Another three dropouts were defined: two patients deteriorated rapidly before starting treatment (one infection in combination with old age, one progressing towards AML), and one developed a psychiatric disorder and refused to start treatment. Five were withdrawn from treatment within 6 weeks of the start of treatment because of complications caused by advanced age and concomitant disease. The criteria for withdrawal/dropout were the same as those reported by Hellström-Lindberg *et al* (1998).

Patients were investigated before and after treatment with blood cell counts, C-reactive protein, red cell sedimentation rate, ferritin, cobalamin, folic acid, creatinine, liver enzymes, S-Epo, S-electrophoresis, direct antiglobulin test and bone marrow analysis. Cytogenetic analysis was performed before treatment.

Methodological considerations. The optimal design would have been to use the same inclusion criteria as Hellström-Lindberg *et al* (1998), thus including the patients assigned to the poor response group according to the prototype model. This would have allowed a verification of the total response rate and a validation of the response rates in the three predictive groups respectively. However, ethical considerations made us exclude the poor response group.

Within the Scandinavian MDS group trials, 80 patients had been treated with G-CSF + Epo, with a total response rate of 38% (Hellström-Lindberg *et al*, 1998). The response rate in the poor response group was only 7%, and the majority of side-effects were found in this group, the most serious being treatment-induced thrombocytopenia in already thrombocytopenic patients. Thus, we chose not to include a patient group that did not benefit from the treatment (Hellström-Lindberg *et al*, 1997b, 1998).

The validation process consisted of two parts. First, we analysed the performance of our model in predicting response to therapy in the present series of patients (evaluation sample). This was done by comparing actual response rates with the expected response rates, calculated from our model. Secondly, we analysed our evaluation sample for factors that could be used to identify patients with response to therapy. If the same factors were identified in the evaluation sample, this would strengthen the hypothesis that the predictive factors identified in our previous study (and used for constructing our predictive model) were externally valid.

Response criteria. These were identical to those in our previous studies (Hellström-Lindberg *et al*, 1998) and in a French randomized study (Casadevall *et al*, 2001). A complete erythroid response (CR) was defined as an increase in haemoglobin to at least 11.5 g/dl. A partial response (PR) was defined as an increase in haemoglobin of 1.5 g/dl or more in patients with non-transfused anaemia and a 100% reduction of transfusion need in combination with a stable haemoglobin level for ≥ 4 weeks in those with pretreatment transfusion need. As the effect of treatment on neutrophil and platelet counts has been described in detail (Hellström-Lindberg *et al*, 1998), we did not include these results in the present report.

Treatment. Treatment with G-CSF (filgrastim, Roche; and later Amgen Pharmaceutical) and Epo (erythropoietin β , Boehringer Mannheim; and later Roche) was started simultaneously. This was an academic study to which some research support, but neither free drug nor full financing, was given by the pharmaceutical companies. The dose of G-CSF was 75, 150 or 300 $\mu\text{g/d}$ subcutaneously (s.c.), given three times per week, and the dose was adjusted to obtain a neutrophil count of $6\text{--}10 \times 10^9/\text{l}$. The dosing regimen was changed from daily injection to three per week. This concept had been tried during the maintenance phase of the previous studies (unpublished observations). The starting dose of Epo was 10 000 U, 5 d/week, s.c. and, in case of a CR, the dose was reduced every 8 weeks as long as the response was maintained. The Epo dose was lower than in the previous study, which had a starting dose of 10 000 U, 7 d/week. At the initiation of the study, 10 000 U was the maximum content per ampoule, and there were practical reasons not to have injections during the weekend. Treatment was given for a minimum of 12 weeks. If no response was observed, treatment was stopped. Where there was some improvement in haemoglobin levels or reduction in transfusion need, treatment was continued for 16 weeks. Only patients fulfilling the criteria for PR at 16 weeks continued the first phase of the

study up to 20 weeks. Thus, responding patients were evaluated at 20 weeks, whereas a decision to define a patient as a non-responder was taken after 12 or 16 weeks. A patient was considered evaluable for a response to G-CSF and Epo if the treatment was given for 6 weeks or more.

Follow-up and duration of response. All evaluable patients treated within the Scandinavian studies (Hellström-Lindberg *et al*, 1998) plus the patients in the present study were followed-up with regard to duration of response. This analysis included 124 patients, 26 CR and 21 PR. Median follow-up times were similar; 43 months (Hellström-Lindberg *et al*, 1998) and 41 months (minimum 18 months; the present study).

Quality of life (QOL) questionnaire. The European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 is a health-related quality of life (HRQOL) questionnaire with 30 items constructed to be cancer specific and multidimensional. The questionnaire incorporates five functional scales (physical, role, emotional, cognitive and social functioning), three symptom scales (pain, fatigue and nausea/vomiting), a global health and quality of life scale and a number of single items (appetite loss, dyspnoea, diarrhoea, constipation, sleep disturbances and financial impact). All scale and item scores were linearly transformed so that the results range from 0 to 100. For the five functioning scales and the single global health/quality of life scale, higher scores represent higher levels of functioning. For the three symptom scales and the single symptom items, higher scores represent higher levels of symptoms. The psychometric properties and validity of the questionnaire have been found to be satisfactory (Aaronson *et al*, 1993; Kaasa *et al*, 1995; Wisloff *et al*, 1996) and the test-retest reliability to be high (Hjermstad *et al*, 1995). QLQ C-30 version 2.0 was used in this trial.

Design of the QOL study. The questionnaires were circulated to the patients before treatment and 12 weeks after the start of treatment. Responding patients were also analysed on later occasions, but these data are not included in the present report. Patients received the first questionnaire from the physician, and patients who did not return the first questionnaire did not receive the second form. The second questionnaire was mailed directly to the patients from the study secretariat.

Patients participating in the HRQOL study. Thirty-six (14 CR + PR, 22 NR) of the 53 (67.9%) evaluable patients answered the first questionnaire (Scandinavian versions). Thirty-two (12 CR + PR, 20 NR; 60.4%) answered the questionnaire 12 weeks after the start of treatment. The first six patients (11.3%) entering the clinical study were not included in the QOL analyses for logistic reasons. The reasons for the remaining 11 dropouts were: decreased cognitive/social function 4, language problems 2, unknown reasons 5.

The reference population. The HRQOL in the general population in Norway has been assessed by the EORTC QLQ-C30 questionnaire (Hjermstad *et al*, 1998a). In general, men reported better functioning and fewer symptoms on all scales. There was a consistent gradual decrease with

increasing age for both sexes. The population was divided into sex and 10-year age groups.

Adjustment for age and sex differences should be made for comparisons between a patient population and a reference population. We chose an approach (Hjermstad *et al*, 1998b) based upon the concept of obtaining expected mean scores using the population reference values and calculating the expected scores that would be observed for subjects of the same age and sex distribution as in the disease group.

Statistical analysis in the HRQOL study. The scores at different time points were compared using the Mann-Whitney test for related samples. Significance tests for comparison with the reference population were performed with multiple linear regression with stepwise forward selection.

Differences or changes of 10 or more on the 0–100 scale are usually regarded as clinically significant (Osoba *et al*, 1998). As a large number of comparisons were performed, a *P*-value of <0.01 was considered to be necessary for statistical significance. The Statistical Package for the Social Sciences (SPSS) software for Windows version 8.0 was used for all statistical analyses.

Statistical evaluation of the predictive model. The prototype model was tested in the evaluation sample by comparing the actual response rates with expected response rates in the two groups with predicted good response and intermediate response. Expected response rates were calculated from our prototype model, which predicted a response rate of 74% in the good response group and 23% in the intermediate response group (Hellström-Lindberg *et al*, 1997b). The difference between actual and expected response rates was tested for the null hypothesis of no difference using the appropriate test statistics described by Hilden *et al* (1978). Logistic regression was used for the multivariate analysis.

Other statistical methods. Student's *t*-test, the Mann-Whitney U-test and analysis of variance (ANOVA) were used for comparison of continuous variables, whenever appropriate. Chi-squared analysis (with continuity correction) was used to compare categories. Kaplan-Meier plots were used to determine the duration of response. Confidence intervals (CI) of 95% were used throughout the study.

RESULTS

Patients

Fifty-three patients were evaluable for a response to treatment. Baseline characteristics for evaluable patients, and for those reported by Hellström-Lindberg *et al* (1998) are described in Table I. The proportion of patients with pretreatment transfusion need and the proportion of patients with low-, intermediate- and high-risk karyotype, respectively, did not differ between the two studies. The last follow-up was 19 months from the inclusion of the last patient in the study.

Results from treatment; single variables

Twenty-two of the 53 patients (42%, CI 28–55%) responded to treatment: 15 (28%) showed a complete and seven (13%)

Table I. Baseline characteristics for 53 evaluable patients in the present study (evaluation sample) and 98 patients reported by Hellström-Lindberg *et al* (1998).

Variable	Evaluation sample (n = 53)	Construction sample (n = 98)
Age (years)	72 ± 10	70 ± 11
Sex (M/F)	29/24	61/37
RA/RARS/RAEB	15/21/17	30/31/32*
IPSS (low/Int-1/Int-2/high)†	16/26/5/0	
Disease duration (months)	14.3 ± 24.8	18 ± 25
Prior RBC transfusions (yes/no)	36/17	71/27
RBC transfusion units/month‡	2.5 ± 1.5	2.8 ± 1.5
Haemoglobin level (g/dl)	79.5 ± 30	86.0 ± 12.0
ANC count (× 10 ⁹ /l)	2.5 ± 2.4	2.2 ± 1.6
Platelet count (× 10 ⁹ /l)	243 ± 170	217 ± 160
Bone marrow blasts (%)	4.4 ± 3.3	5.3 ± 4.7
Karyotype (low, int, high)§	30/12/8	51/38
Serum erythropoietin (U/l)	83 (8–2800)	233 (6–5921)

*Also includes five patients with RAEB-t.

†IPSS, International prognostic scoring system (Greenberg *et al.*, 1997). Low, low risk; int-1, intermediate 1 risk; int-2, intermediate 2 risk; high, high risk. The construction sample was published before the IPSS risk score.

‡Transfusion need in patients *with* pretreatment RBC transfusions.

§Karyotype risk group according to IPSS; low, intermediate, high. In the previous study (Hellström-Lindberg *et al.*, 1998), only normal/abnormal karyotypes were used.

Table II. Category variables in responding and non-responding patients (univariate analysis).

Variable	Subgroups*	Response rates (CR + PR)	P-value
Sex (F/M)†	24 female/29 male	46%/38%	0.76
Diagnosis	15 RA/21 RARS/17 RAEB	47%/52%/24%‡	0.18
Diagnosis	36 RARS + RA/17 RAEB	50%/24%‡	0.07
Karyotype risk§	30 low/12 int/8 high	47%/58%/13%	0.12
IPSS¶	16 low/26 int-1/5 int-2/3 high	56%/46%/0%/33%	0.17
Transfusions (yes/no)	36 no/17 yes	71%/28%	0.008
Transfusion need**	33 no need or < 2/20 ≥ 2	58%/15%	0.006
Transfusion need	17 0/16 < 2/20 ≥ 2	71%/44%/15%	0.003
Serum Epo‡	30 < 100 U/23 ≥ 100 U/l	56%/22%	0.02
Serum Epo	40 < 200 U/11 ≥ 200 U/l	45%/18%	0.16
Serum Epo‡	12 100–200/8 200–500	25%/25%	1.0
Predictive group††	31 good/22 intermediate	61%/14%	0.001

*No. of patients per subgroup.

†Male/female.

‡U/l. Three patients had S-Epo > 500 U/l.

§Karyotype risk group according to IPSS; low, intermediate, high.

¶IPSS, International Prognostic Scoring System (Greenberg *et al.*, 1997). Low, low risk; int-1, intermediate 1 risk; int-2, intermediate 2 risk; high, high risk. Note, only five patients in Int-2 and -3 in high-risk group. see Table I.

**Units of RBC/month.

††Predictive group according to Hellström-Lindberg *et al* (1997b).

a partial response. Response rates in different subgroups of patients are shown in Table II. Only S-Epo, pretreatment transfusion need and group in the predictive model showed a significant impact on response to treatment.

Validation of the predictive model

Group in the prototype model was the most significant predictor of a response to treatment according to the univariate analysis ($P = 0.001$, Table II). No other combination

of variables or other cut-off levels for transfusion need or S-Epo showed higher significance. Multivariate analysis using logistic regression (not including the predictive model) retained only transfusion need (< 2 units/month, $P = 0.009$) and S-Epo (< 100 U/l, $P = 0.035$) as significant variables. However, no variable showed additional value beyond that of the model ($P = 0.0007$).

The response rates in the good and intermediate response groups were 19/31 (61%) and 3/22 (14%) respectively. The overall response rate was 41.5%, whereas the overall expected response rate was 52.8%. This difference was not significant. Figure 1 shows the validated model.

Duration of response and dosing in maintenance phase

The duration of a total of 48 responses (27 CR, 21 PR) was analysed. These included seven patients who, for various reasons, did not continue with maintenance treatment (Hellström-Lindberg et al, 1998). Median duration of all responses was 23 months, 29 months for CR and 5.5 months for PR. CRs were significantly longer than PRs, $P = 0.002$ (Fig 2). The doses of Epo and G-CSF were reduced successively and then stopped without a decline in the haemoglobin level in two patients with CR (one RA, one RARS). All other patients needed continuous treatment to maintain the response. Generally, patients with PR needed

the high initial Epo doses in order maintain their response. In CR patients, however, Epo dose reduction was achieved in ≈ 50% of the patients. Of these, the majority needed 30 000 or 40 000 U/week to maintain a CR, but doses as low as 10 000 U/week were achieved in a few patients.

Results of the HRQOL analysis

Compared with an age- and sex-adjusted reference population, MDS patients had a statistically highly significant impairment of most of the functioning scores except for the cognitive function, as shown in Fig 3. Regarding the symptom scales and single items, there were statistically and clinically significant increases in fatigue and dyspnoea.

Comparing the patients achieving either CR or PR with those achieving no response, there was clinically significant improvement in some of the important scores for the responding patients, i.e. global QOL, social functioning, fatigue and dyspnoea (Fig 4). The improvement was statistically significant only for global QOL ($P = 0.01$), with a borderline significance for fatigue ($P < 0.05$).

DISCUSSION

We have tested the clinical usefulness of a previously developed decision model for treating the anaemia of MDS

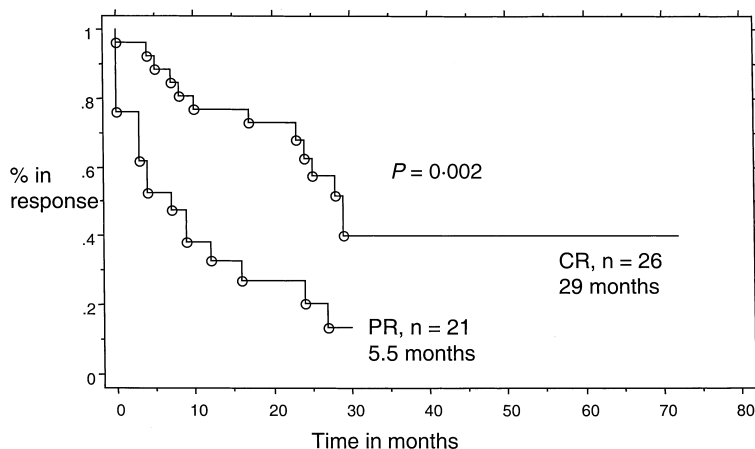


Fig 2. Duration of response in responding patients. This analysis was performed on the total number of evaluable patients treated within the Scandinavian studies; 124 patients, 26 CR and 21 PR. Kaplan-Meier plot.

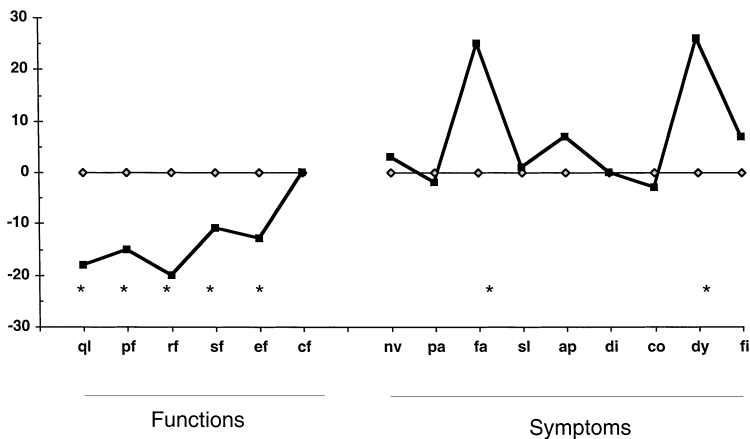


Fig 3. HRQOL scores from the MDS patients before the start of treatment compared with the age- and sex-adjusted reference population. ql, global health and quality of life; pf, physical functioning; rf, role functioning; sf, social functioning; ef, emotional functioning; cf, cognitive functioning; nv, nausea/vomiting; pa, pain; fa, fatigue; sl, sleep disturbances; ap, appetite loss; di, diarrhoea; co, constipation; dy, dyspnoea; fi, financial impact. * $P < 0.01$

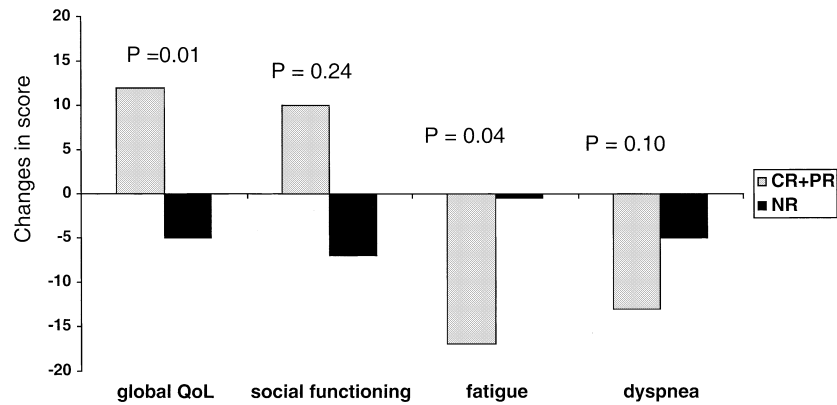


Fig 4. Changes in HRQoL scores from the start of treatment to 12 weeks after the start of treatment in patients achieving complete or partial response (CR + PR) (with corresponding *P*-values) compared with those with no response (NR). Improvement in functions = positive change; improvement in symptoms = negative change.

with a combination of G-CSF and Epo. We also present new results, which showed that patients with MDS and anaemia have a significantly decreased quality of life compared with age- and sex-matched healthy individuals, and that quality of life improves when anaemia is corrected by treatment.

Several independent studies have demonstrated a response rate of 40% or higher to treatment with G-CSF + Epo and also a clear synergistic effect between the two drugs. In the majority of these studies, the proportion of patients achieving CR, i.e. normalized haemoglobin levels, has been higher than that achieving PR. Both French and Scandinavian trials (Hellström-Lindberg *et al*, 1998; Casadevall *et al*, 2001) used more stringent response criteria than those suggested by Cheson *et al* (2000). A partial response in our studies is similar to the definition of good response, according to Cheson *et al* (2000), which they defined as a total cessation of transfusion need or an increase in haemoglobin of > 2.0 g/dl. The definition of CR, according to our studies, is a normalized haemoglobin level. Using these more precise criteria, the chance is higher that a response not only reflects a biological effect, but also a clear clinical value for the patient.

The importance of obtaining a complete erythroid response is supported by the significantly longer duration of response between CR and PR in our studies. Patients achieving CR had a median duration of 29 months, which is longer than for most other treatment alternatives for MDS (Hellström-Lindberg *et al*, 2000).

A recent study described a positive effect of treatment with 5-azacytidine on HRQoL in a group of patients, of whom the majority had intermediate- or high-risk MDS (Kornblith *et al*, 2002). However, there are no studies showing the HRQoL profile of elderly MDS patients compared with that of healthy individuals. We do not describe an unselected population-based cohort of patients in our study, but a group of MDS patients for whom anaemia is the main problem and the majority of whom have blast counts < 10%. We show that this patient group has a significantly lower HRQoL than an age- and sex-matched population, and that fatigue and dyspnoea constitute the main symptoms. The large reference population was collected from the same region as the patients, and

has been used in comparative studies (Hjermstad *et al*, 1998a,b). The analysis of HRQoL at 12 weeks, including a somewhat limited patient sample, showed a significant effect of treatment on global quality of life ($P = 0.01$), with borderline significance for fatigue ($P = 0.04$). This suggests that correcting the anaemia will improve the quality of life in elderly MDS patients, but larger studies are needed to verify this conclusion.

The clinical value of treatment with G-CSF + Epo is thus equally supported by the quality (completeness) and long duration of the response, and by the effect of the responses on the quality of life. However, both predictive groups (good and intermediate) in the present study showed somewhat lower response rates than in the previous study. There may be several reasons for this. The present phase IV study aimed to treat all patients who fulfilled the inclusion criteria, and the lower response rates could be a result of the wider range of unselected patient material. Another explanation could be the lower starting dose of Epo. Both the German (Mantovani *et al*, 2000) and the American (Negrin *et al*, 1996) studies, for example, used higher Epo doses than those administered in our study. It is therefore possible that the Epo dose should be 70 000 U/week, as in the first study.

The prototype model was evaluated prospectively and found to be valid in the present study. The overall response rate in our evaluation sample (42%) was lower than the corresponding rate (49% in the two groups good and intermediate response combined) in our previous sample. The lower response rate in the evaluation sample probably explains why the model overestimated the response rate. Despite this bias, our model remained valid. The value of our model was further supported by the fact that no other variables or combination of variables proved more significant in predicting a response to treatment in the evaluation sample. A similar model for the use of Epo alone does not exist, even if univariate analyses suggest that response rates to Epo are higher in patients with RA, lower serum Epo levels and no transfusion need.

The required probability for a response before initiating therapy with G-CSF + Epo in an anaemic patient with MDS depends on several factors and may vary between countries. In Europe, transfusions are still significantly cheaper than

both Epo and Epo + G-CSF, as long as the interval between transfusions is ≈ 4 weeks and no chelation therapy is needed. However, very little is known about the morbidity caused by anaemia in an elderly population and what secondary costs that may lead to. If treatment could prevent or postpone the high transfusion need that often develops in patients with RARS and 5q- syndrome, it is possible that both morbidity and costs would be decreased. We have shown that treatment is durable and improves quality of life. Thus, from a European point of view, it would be logical to recommend treatment for patients belonging to the good responsive group, whereas treatment for patients in the intermediate responsive group would be based on the merits of each case and not generally recommended. Approximately 30% of low-risk MDS patients eligible for treatment with Epo + G-CSF belong to the good predictive group, according to the combined US and Scandinavian study (Hellström-Lindberg *et al*, 1997b). In the US, with significantly higher costs for transfusions, decision-making may take different priorities into account. However, it is likely that patients would want to be informed of their chances of a response before choosing a therapy. We did not include the poor responsive group in our study but, from our previous experience, there is little evidence that patients from this group would be good candidates for therapy with G-CSF + Epo.

The original logistic regression analysis that led to the prototype model defined two levels of transfusion need (no need or < 2 units/month versus > 2 units/month) in combination with three Epo-levels, < 100 , 100–500 and > 500 U/l, as the optimal selection of pretreatment variables. The cut-off levels were verified in the present study. A difference in response rate between Epo < 100 and 100–500 U/l (Table II) was again observed, and we conclude that future studies will need to assess the value of S-Epo in predicting response to this treatment. However, the exact level of S-Epo, when < 500 U/l, will not make a difference when the model is used clinically, as a value of < 100 U/l will not allocate a patient to a different group. Thus, for practical use, the model could be expressed as two binary variables: transfusion need and serum level of Epo. In those cases in which both variables are favourable (transfusion need < 2 units/month and serum level of Epo < 500 U/l), the predicted response is good. In those cases with one favourable and one unfavourable variable, the predicted response is intermediate, whereas if both parameters are unfavourable, the predicted response is poor.

Four separate studies showed clear evidence of a synergistic effect of G-CSF and Epo *in vivo* (Negrin *et al*, 1996; Hellström-Lindberg *et al*, 1998; Remacha *et al*, 1999; Mantovani *et al*, 2000). Moreover, the proportion of CR patients seems to be higher after treatment with G-CSF + Epo than with Epo alone (Hellström-Lindberg, 1995; Italian Cooperative Study Group for rHuEpo in Myelodysplastic Syndromes, 1998). The combination also gives an acceptable response rate in patients with transfusion need. The response rate in patients with a low (< 2

units/month) transfusion need was 50%, compared with 82% in non-transfused patients and 14% in those with a high transfusion need (Table II). The synergistic effect seems to be most pronounced in RARS, but may also be observed in the other FAB groups. In a recent study, we showed that incubation of RARS CD34⁺ cells with G-CSF could improve erythroid colony growth (Schmidt-Mende *et al*, 2001), and that G-CSF significantly inhibited the increased caspase activity observed in bone marrow cells from these patients. In another study, G-CSF significantly reduced apoptosis in RARS but not in RA (Tehranchi *et al*, 2003). We conclude that G-CSF has additive or synergistic effects on erythropoiesis, particularly in RARS patients. However, these studies included patients who did not have RARS, but lost their response when G-CSF was withdrawn and regained it when it was reintroduced. Therefore, the role for G-CSF in non-RARS subtypes remains to be investigated further.

A practical recommendation for the use of treatment with G-CSF + Epo is to start therapy with Epo alone in non-RARS patients and in non-transfused RARS patients. A dose of 50 000 U/week is probably sufficient for the majority of patients. However, based on the results of our previous study, and the studies of Negrin *et al* (1996) and Mantovani *et al* (2000), an increase in Epo dose up to 70 000 U or more could be considered, especially in cases with transfusion need and higher serum Epo levels. If there was no response, G-CSF could be added in a dose adjusted to produce a clear rise in neutrophil count. RARS patients with a pretreatment transfusion need should probably start with the combined treatment from the beginning. Finally, it is possible that the addition of G-CSF may enable the Epo doses to be lowered in some patients. As Epo is more expensive than G-CSF at these doses, this could lower the cost of the whole treatment.

In this report, we present a validated, clinically useful decision model for treatment with G-CSF + Epo for the anaemia of MDS. The second phase of the study will include a detailed long-term follow-up and will focus on dosing regimens and on optimal doses of Epo and G-CSF for maintaining a long-term response.

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Cancer Society (3689-B99-05XBB) and from the Cancer Society in Stockholm (01:164).

REFERENCES

- Aaronson, N.K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N.J., Filiberti, A., Flechtner, H., Fleishman, S.B., de Haes, J.C., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P.B., Schraub, S., Sneeuw, K., Sullivan, M. & Takeda, F. (1993) The European Organization for Research and Treatment of Cancer QLQ C-30: a quality of life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, **85**, 365–376.
- Backx, B., Broeders, L. & Löwenberg, B. (1992) Kit ligand improves *in vitro* erythropoiesis in myelodysplastic syndromes. *Blood*, **80**, 1213–1217.

- Bennett, J.M., Catovsky, D., Daniel, M.T., Flandrin, G., Galton, D.A.G., Gralnick, H.R., Sultan, C. & The French-American-British (FAB) Group (1982) Proposals for the classification of the myelodysplastic syndromes. *British Journal of Haematology*, **51**, 189–199.
- Casadevall, N., Durieux, P., Dubois, S., Lepage, E., Hemery, F., Quarre, M.C., Fenaux, P., Damaj, G., Giraudier, S., Guerci, A., Laurent, G., Dombret, H., Stamatoullas, A., Vekhoff, A., Ribrag, V. & Dreyfus, F. pour le Groupe Francais des Myelodysplasies (2001) Erythropoietin (rHuEpo) plus G-CSF in the treatment of anemia in myelodysplastic syndromes (MDS): results of a randomized trial with impact on quality of life and costs. *Blood*, **98**, 3522 (abstract).
- Cheson, B., Bennett, J.M., Kantarijan, Pinto, A., Schiffer, C.A., Nimer, S.D., Löwenberg, B., Beral, M., de Witte, T., Stone, R., Mittelman, M., Sanz, G.F., Wijermans, P., Gore, S. & Greenberg, P. (2000) Report of an international working group to standardize response criteria for myelodysplastic syndromes. *Blood*, **96**, 3671–3674.
- Greenberg, P.L. (1994) The myelodysplastic syndromes. In: *Haematology: Basic Principles and Practice*, 2nd edn (ed. by R. Hoffman, E. Benz, S. Shattil, B. Furie & H. Cohen), pp. 1098–1121. Churchill Livingstone, New York.
- Greenberg, P.L., Cox, C., LeBeau, M., Fenaux, P., Morel, P., Sanz, G., Sanz, M., Vallespi, T., Hamblin, T., Oscier, D., Ohyashiki, K., Toyama, K., Aul, C., Mufti, G. & Bennett, J. (1997) International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*, **89**, 2079–2088.
- Hast, R., Wallvik, J., Folin, A., Bernell, P. & Stenke, L. (2001) Long-term follow-up of 18 patients with myelodysplastic syndromes with myelodysplastic syndromes responding to recombinant erythropoietin treatment. *Leukemia Research*, **25**, 19–21.
- Hellström-Lindberg, E. (1995) Efficacy of erythropoietin in the myelodysplastic syndromes. An analysis of 205 patients in 17 studies. *British Journal of Haematology*, **89**, 67–71.
- Hellström-Lindberg, E., Kanter-Lewensohn, L. & Öst, Å. (1997a) Morphological changes and apoptosis in bone marrow from patients treated with granulocyte-CSF and erythropoietin. *Leukemia Research*, **21**, 415–455.
- Hellström-Lindberg, E., Negrin, R., Stein, R., Krantz, S., Lindberg, G., Vardiman, J., Öst, Å. & Greenberg, P. (1997b) Erythroid response to treatment with G-CSF plus erythropoietin for the anemia of patients with myelodysplastic syndromes: proposal for a predictive model. *British Journal of Haematology*, **99**, 344–351.
- Hellström-Lindberg, E., Ahlgren, T., Beguin, Y., Carlsson, M., Carneskog, J., Dahl, I.M., Dybedal, I., Grimfors, G., Kanter-Lewensohn, L., Linder, O., Luthman, M., Lofvenberg, E., Nilsson-Ehle, H., Samuelsson, J., Tangen, J.M., Winqvist, I., Oberg, G., Osterborg, A. & Öst, Å. (1998) Treatment of the anemia in myelodysplastic syndromes with granulocyte-CSF plus erythropoietin: results from a randomized phase II study and long-term follow-up on 71 patients. *Blood*, **92**, 68–75.
- Hellström-Lindberg, E., Willman, C., Barrett, A.J. & Sauntharajah, Y. (2000) *Achievements in Understanding and Treatment of Myelodysplastic Syndromes*. *Haematology 2000*, pp. 110–132. American Society of Haematology, Washington, DC.
- Hilden, J., Habbema, J.D.F. & Bjerregaard, B. (1978) The measurement of performance in probabilistic diagnosis. II. Trustworthiness of the exact values of the diagnostic probabilities. *Methods in Infectious Medicine*, **17**, 227–237.
- Hjermstad, M.J., Fossa, S.D., Bjordal, K. & Kaasa, S. (1995) Test/retest study of the European Organization for Research and Treatment of Cancer core quality of life questionnaire. *Journal of Clinical Oncology*, **13**, 1249–1254.
- Hjermstad, M.J., Fayers, P.M., Bjordal, K. & Kaasa, S. (1998a) Health Related Quality of Life in the general population assessed by the EORTC Core Quality of Life Questionnaire – the EORTC QLQ-C30 (+3) *Journal of Clinical Oncology*, **16**, 1188–1196.
- Hjermstad, M.J., Fayers, P.M., Bjordal, K. & Kaasa, S. (1998b) Using reference data on quality of life – the importance of adjusting for age and sex, exemplified by the EORTC QLQ-C30 (+3). *European Journal of Cancer*, **34**, 1381–1389.
- Italian Cooperative Study Group for rHuEpo in Myelodysplastic Syndromes (1998) A randomized double-blind placebo-controlled study with subcutaneous recombinant human erythropoietin in patients with low-risk myelodysplastic syndromes. *British Journal of Haematology*, **103**, 1070–1074.
- Jaffe, E.S., Harris, N.L., Stein, H. & Vardiman, J.W. (eds) (2001). *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haemopoietic and Lymphoid Tissues*. IARC Press, Lyon.
- Kaasa, S., Bjordal, K., Aaronson, N., Moum, T., Wist, E., Hagen, S. & Kvikstad, A. (1995) The EORTC Core Quality of Life Questionnaire (QLQ C-30): validity and reliability when analyzed with patients treated with palliative radiotherapy. *European Journal of Cancer*, **31A**, 2260–2263.
- Kornblith, A.B., Herndon, I.I.J.E., Silverman, L.R., Demakos, E.P., Odchimar-Reissig, R., Holland, J.F., Powell, B.L., DeCastro, C., Ellerton, J., Larson, R.A., Schiffer, C. & Holland, J.C. (2002) Impact of azacytidine on the quality of life of patients with the myelodysplastic syndrome treated in a randomized phase III trial: a Cancer and Leukemia group B study. *Journal of Clinical Oncology*, **20**, 2441–2452.
- Mantovani, L., Lentini, G., Hentschel, B., Wickramanayake, P.D., Loeffler, M., Diehl, V. & Tesch, H. (2000) Treatment of anaemia in myelodysplastic syndromes with prolonged administration of recombinant human granulocyte colony-stimulating factor. *British Journal of Haematology*, **109**, 367–375.
- Negrin, R.S., Stein, R., Doherty, K., Cornwell, J., Vardiman, J., Krantz, S. & Greenberg, P.L. (1996) Maintenance treatment of the anemia of myelodysplastic syndromes with recombinant human granulocyte colony-stimulating factor and erythropoietin: evidence for in vivo synergy. *Blood*, **87**, 4076–4081.
- Osoba, D., Rodrigues, G., Myles, J., Zee, B. & Pater, J. (1998) Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology*, **16**, 139–144.
- Raza, A., Gezer, S., Mundle, S., Gao, X.Z., Alvi, S., Borok, R., Rifkin, S., Iftikhar, A., Shetty, V., Parcharidou, A., Loew, J.M., Marcus, B., Khan, Z., Chaney, C., Showel, J., Gregory, S. & Preisler, H. (1995) Apoptosis in bone marrow biopsy samples involving stromal and hematopoietic cells in 50 patients with myelodysplastic syndromes. *Blood*, **86**, 268–276.
- Remacha, A.F., Arrizabalaga, B., Villegas, A., Manteiga, R., Calvo, T., Julia, A., Fernandez Fuertes, I., Gonzalez, F.A., Font, L., Junca, J., del Arco, A., Malcorra, J.J., Equiza, E.P., de Mendiguren, B.P. & Romero, M. (1999) Erythropoietin plus granulocyte colony-stimulating factor in the treatment of myelodysplastic syndromes. Identification of a subgroup of responders. The Spanish Erythropathology Group. *Haematologica*, **84**, 1058–1064.
- Schmidt-Mende, J., Tehranchi, R., Forsblom, A.M., Joseph, B., Christensson, B., Fadeel, B., Zhivotovsky, B., Hellström-Lindberg, E. & Hellström-Lindberg, E. (2001) Granulocyte colony-stimulating factor inhibits Fas-triggered apoptosis in bone marrow cells isolated from patients with refractory anemia with ringed sideroblasts. *Leukemia*, **15**, 742–751.

- Spiegelhalter, D.J. (1983) Evaluation of clinical decision-aids, with application to a system for dyspepsia. *Statistics in Medicine*, **2**, 207–216.
- Tehranchi, R., Fadeel, B., Forsblom, A.M., Christensson, B., Samuelsson, J., Zhivotovsky, B. & Hellstrom-Lindberg, E. (2003) Granulocyte colony-stimulating factor inhibits spontaneous cytochrome c release and mitochondria-dependent apoptosis of myelodysplastic syndrome hematopoietic progenitors. *Blood*, **101**, 1080–1086.
- Thomas, M. (2001) Quality of life in myelodysplastic syndromes: Measurement issues in research and in clinical practice. *Leukemia Research*, Suppl. 1, Or28 (abstract).
- Van de Loosdrecht, A.A. & Vellenga, E. (2000) Myelodysplasia and apoptosis: new insights into ineffective erythropoiesis. *Medical Oncology*, **17**, 573–578.
- Wisloff, F., Hjorth, M., Kaasa, S. & Westin, J. (1996) Measurement of health-related quality of life in multiple myeloma. *British Journal of Haematology*, **92**, 604–613.