cialization, and the significance of morphology as part of this process is changing with it. So, what are the clinical hematologists going to do with their microscopes: use them or sell them?

Mars van 't Veer is a consultant hematologist at Addenbrooke's University Hospital, Cambridge, UK. Torsten Haferlach works at the MLL Munich Leukemia Laboratory, Germany.

Financial and other disclosures provided by the author using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are available with the full text of this paper at www.haematologica.org.

References

1. Romanowsky D. Zur Frage der Parasitologie und Therapie der Malaria. St Petersburg Med Wochenschrift 1891;16:307-15.
2. Giemsa G. Eine Vereinfachung und Vervollkommnung meiner Methylenazur-Methylenblau-Eosin-Färbemethode zur Erzielung der Romanowsky-Nochtschen Chromatinfärbung. Centralbladt für Bakteriologie 1904;37:308-11.
3. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DAG, Gralnick HR, et al. Proposals for the classification of acute leukaemias (FAB Cooperative Group). Br J Haematol. 1976;33(4): 451-8.
4. Bennet JM, Catovsky D, Daniel MT, et al. Proposed revised criteria for the classification of acute myeloid leukaemia. A report of the French-American-British Cooperative Group. Ann Intern Med., 1985a;103:620-5.
5. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, Sultan C. Criteria for the diagnosis of acute leukemia of megakaryocyte lineage (M7). A report of the French-American-British Cooperative Group. Ann Intern Med. 1985;103(3):460-2.
6. Nowell P, Hungerford D. A minute chromosome in chronic granulocytic leukemia. Science 1960;132:1497.
7. Rowley JD. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature 1973;243:290-3.
8. Jaffe ES, Harris NL, Stein H, Vardiman JW (eds.) WHO classification of tumours. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. IARC press: Lyon 2001.
9. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. (eds.). WHO Classification of Tumours of Hematopoietic and Lymphoid Tissue. IARC press, Lyon 2008.
10. Welch JS, Link DC. Genomics of AML: Clinical applications of next-generation sequencing. Hematology Am Soc Hematol Educ Program. 2011:30-5.

Disease severity in chronic graft-*versus*-host disease: doctors' gut feeling *versus* biostatistics?

Gerard Socié

Head of the Hematology / Immunology /Oncology Division, Head of the Transplant Program, Inserm U1160, University Paris VII & AP-HP, Hospital Saint Louis, 1 Avenue Claude Vellefaux, Paris, France

E-mail: gerard.socie@sls.aphp.fr

doi:10.3324/haematol.2014.113878

Chronic graft-*versus*-host disease (GVHD) is a frequent and potentially life-threatening complication of allogeneic hematopoietic stem cell transplantation. An increase in transplants in older patients and the more frequent use of unrelated donors have led to greater numbers of patients with this painful complication. Recent advances have been made in understanding the pathophysiology of chronic GVHD as well as in establishing precise criteria for the diagnosis and classification of disease manifestations. These advances will, it is hoped, pave the way to improving both the prophylaxis and treatment of chronic GVHD. We recently reviewed current issues in chronic GVHD with Dr. Ritz and readers of Haematologica interested in this field can find detailed information in this article.[1]

In the past, chronic GVHD included any clinical manifestations of GVHD that occurred beyond 100 days after transplantation. This definition was clearly imprecise and became inadequate. In 2005 a group of experts met under the auspices of the National Institutes of Health (NIH), USA in a consensus meeting. The goals of this NIH consensus working group on the diagnosis and staging of GVHD were: (i) to establish criteria for diagnosis of the disease, emphasizing the distinction between acute and chronic GVHD; (ii) to define criteria for scoring the severity of clinical manifestations in affected organs; and (iii) to propose categories describing the overall severity of the disease and the indications for treatment.[2]

The NIH consensus conference recognized two main categories of GVHD, each with two subcategories. The broad category of acute GVHD includes classic acute GVHD (maculopapular erythematous rash, gastrointestinal symptoms, or cholestatic hepatitis). The broad category of chronic GVHD includes classic chronic GVHD, presenting with manifestations that can be ascribed only to chronic GVHD. Chronic GVHD also includes an overlap syndrome, which has diagnostic or distinctive manifestations of chronic GVHD together with features typical of acute GVHD.

Numerous prognostic indices in chronic GVHD have been described.[1] Thrombocytopenia (platelet count $<100x10^9$/L) is the first reported and most reproducible prognostic factor even when using NIH criteria. Other factors, such as diarrhea, might be prognostic only due to the older definition of the disease or to the worse prognosis of the overlap syndrome. The NIH consensus conference proposed a new global chronic severity score establishing mild, moderate and severe forms of chronic GVHD based on a numerical scoring system for individual organs to calculate a summary scale.[2] Although the NIH global score was developed through expert opinion, several studies have shown that the global score at onset of chronic GVHD is associated with risk of subsequent mortality.[3-6] However as nicely described by Inamoto *et al*. in this issue of Haematologica,[7] several issues remain to be elucidated. Firstly, since the NIH global score was based on expert opinion and was not originally intended to predict mortality, does this score provide an optimal model for predicting mortality risk in patients with chronic GVHD? The authors hypothesized that empirically derived estimates of overall mortality risk incorporating the relative importance of different organ involvement might be more accurate

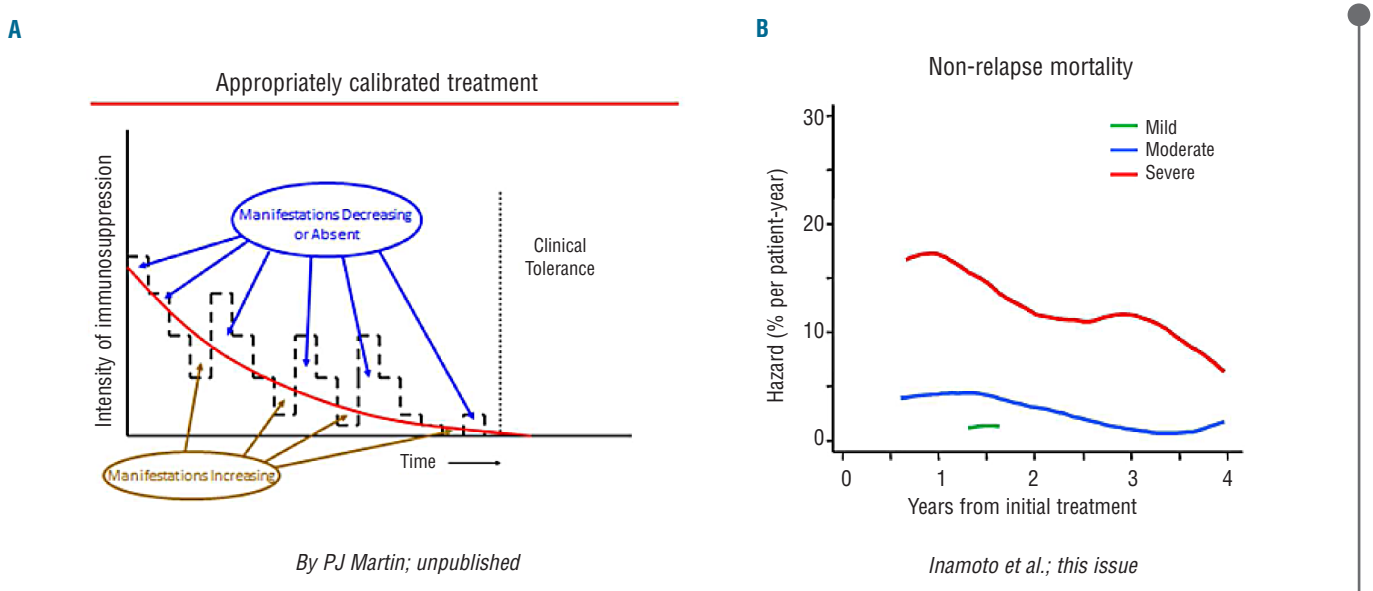

Figure 1. The waxing and waning manifestations of chronic GVHD.

than estimates derived from the opinion-based global score. Secondly, does the NIH global score predict mortality risk when it is applied at time points after the onset of treatment for chronic GVHD? Thirdly, does the NIH global score correlate with risk of recurrent malignancy? To address these important clinical questions, Inamoto *et al.* performed an impressive set of analyses on data collected in a prospective, multicenter, longitudinal, observational study of patients with chronic GVHD. The authors analyzed 574 adult patients with chronic GVHD, using multivariate time-varying analysis accounting for serial changes in severity of disease in eight individual organ sites over time. They randomly divided the cohort into a training set (482 patients who underwent 1602 visits) and a validation set (including 150 patients with 600 visit ratings). In the training set, severity of skin, mouth, gastrointestinal tract, liver and lung involvement were independently associated with the risk of non-relapse mortality. Weighted mortality points were assigned to individual organs based on the hazard ratios. The population was divided into three risk groups based on the total mortality points. The three new risk groups were validated in an independent validation set, but did not show better discriminative performance than the NIH global score. As compared to moderate or mild global score, severe global score was associated with increased risks of non-relapse mortality and overall mortality across time but not with a decreased risk of recurrent malignancy. Thus at a first glance one might have the impression that doctors' gut feeling (expert opinion) might discriminate overall chronic GVHD severity better than sophisticated biostatistics! Is this true? The response, as usual in good medical science, is both yes and no!

First the statistical analysis performed by Inamoto *et al.* must be applauded. In my view it is one of very few examples of how a statistical analysis can reproduce the true life in chronic GVHD. Every physician with expertise in the field will easily recognize that one of the hallmarks of chronic GVHD is its evolution; with flares and waves of manifestations increasing and decreasing sometimes due to new treatment but sometimes also in the absence of treatment modification (which is what PJ Martin often refers to as the waxing and waning of chronic GVHD symptoms). Some years ago Dr. Martin represented this evolution graphically and drew a cartoon (Figure 1A) which he kindly allowed to be published here in Haematologica. The real *tour de force* of the work by Inamoto *et al.* was to take advantage of a large cohort of patients who underwent hundreds of visit ratings to model the erratic evolution of chronic GVHD (Figure 1B). The only other study that aimed to model this evolution was one reporting the long-term results from a randomized trial on GVHD prophylaxis with or without anti-T-cell globulin[8] in which Dr. Schmoor used a similar (but not identical) methodology. Both of these models[7,8] will be of major interest in the analysis of future long-term studies of chronic GVHD.

The second issue is the predominance of clinical manifestations in estimating disease severity both in single organs and overall. Once again, it should be remembered that the current scoring system for each organ and the consequent overall grading was derived from experts' discussion and not from analysis of real patients' data. In the cohort analyzed in this paper, the patients were basically scored by the same senior authors who described the scoring system in 2005. Furthermore the inherent caveat regarding the Chronic GVHD Consortium analysis has always been that it is a mixture of incident cases (for which organ severity can be studied prospectively without problem) and of prevalent cases (for which prognostic studies could be the subject of controversies since by definition patients have to be alive long enough to be able to be scored after months, or even years of evolution).

It should be noted that patients with chronic GVHD now tend to survive for quite a long time and the median follow-up time of survivors in this study is rather short (for chronic GVHD): i.e. just a little longer than 3 years. Thus, longer follow-up of prevalent cases and studies including

larger numbers of patients might lead to different conclusions.

Finally, the authors found no relationship between chronic GVHD severity and relapse; this might be of interest but does not rule out that the graft-*versus*-leukemia effect is mainly restricted to NIH-defined chronic GVHD (and not to acute GVHD).[9]

Gerard Socie, MD, PhD, is Head of the Hematology-Immunology-Oncology Division at the Hospital St Louis, Paris; Professor of Hematology at the University Paris 7; and current President of the French Society of Hematology

Financial and other disclosures provided by the author using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are available with the full text of this paper at www.haematologica.org.

References

1. Socie G, Ritz J. Current issues in chronic graft-versus-host disease. Blood. 2014;124(3):374-84.
2. Filipovich AH, Weisdorf D, Pavletic S, Socie G, Wingard JR, Lee SJ, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. Biol Blood Marrow Transplant. 2005;11(12):945-56.
3. Perez-Simon JA, Encinas C, Silva F, Arcos MJ, Diez-Campelo M, Sanchez-Guijo FM, et al. Prognostic factors of chronic graft-versus-host disease following allogeneic peripheral blood stem cell transplantation: the national institutes health scale plus the type of onset can predict survival rates and the duration of immunosuppressive therapy. Biol Blood Marrow Transplant. 2008;14(10):1163-71.
4. Cho BS, Min CK, Eom KS, Kim YJ, Kim HJ, Lee S, et al. Feasibility of NIH consensus criteria for chronic graft-versus-host disease. Leukemia. 2009;23(1):78-84.
5. Arai S, Jagasia M, Storer B, Chai X, Pidala J, Cutler C, et al. Global and organ-specific chronic graft-versus-host disease severity according to the 2005 NIH Consensus Criteria. Blood. 2011;118(15):4242-9.
6. Kuzmina Z, Eder S, Bohm A, Pernicka E, Vormittag L, Kalhs P, et al. Significantly worse survival of patients with NIH-defined chronic graft-versus-host disease and thrombocytopenia or progressive onset type: results of a prospective study. Leukemia. 2012;26(4):746-56.
7. Inamoto Y, Martin PJ, Storer BE, Palmer J, Weisdorf DJ, Pidala J, et al. Association of severity of organ involvement with mortality and recurrent malignancy in patients with chronic gravt-versus-host disease. Haematologica. 2014;99(10):1618-23.
8. Socie G, Schmoor C, Bethge WA, Ottinger HD, Stelljes M, Zander AR, et al. Chronic graft-versus-host disease: long-term results from a randomized trial on graft-versus-host disease prophylaxis with or without anti-T-cell globulin ATG-Fresenius. Blood. 2011;117(23):6375-82.
9. Thepot S, Zhou J, Perrot A, Robin M, Xhaard A, de Latour RP, et al. The graft-versus leukemia effect is mainly restricted to NIH-defined chronic graft-versus-host disease after reduced intensity conditioning before allogeneic stem cell transplantation. Leukemia. 2010;24(11):1852-8.

Stem cell transplants for myelodysplastic syndromes: refining the outcome predictions

Austin John Barrett

Stem Cell Allotransplantation Section, Hematology Branch, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA

E-mail: barrettj@nhlbi.nih.gov *doi:10.3324/haematol.2014.113555*

Myelodysplastic syndromes (MDS) are best regarded as a spectrum of diseases with distinct underlying biologies (as characterized by chromosomal abnormalities) and different prognoses. Disease characteristics and outcome of MDS are defined by two well-established classifications. Firstly, the FAB that segregates MDS by the degree of progression from relatively benign conditions with no excess of blasts (RA/RARS), excess blasts (RAEB), and the poor prognosis excess blasts in transformation to leukemia (RAEBt); secondly the IPSS score,[1] or its more recent modification the IPSS-R score,[2] which combines marrow blast cell counts with chromosomal abnormalities to identify low, standard, poor and very poor prognosis MDS. The very poor risk category is separated from the poor risk by the inclusion of monosomies 7 and 5.

These classifications are derived from the natural evolution of MDS in the setting of supportive care and non-curative therapy. The question, therefore, arises as to how applicable these scores are to MDS recipients of HSCT, which represents the only curative treatment for MDS. As age barriers fall away and transplant outcomes for older patients improve, more patients with MDS receive allogeneic SCT and it becomes increasingly important to define MDS risk categories to aid transplant selection.[3]

Outcome after HSCT for MDS is determined by patient characteristics and factors under the control of the transplant physician. Patient characteristics segregate into factors which impact on transplant-related mortality, such as age and general health (using an adapted Charleson score[4]), and those that impact the curative potential of the transplant. For this latter, transplanters have applied the FAB and IPSS score with or without inclusion of monosomy as a marker of very high relapse risk.

In this issue of the Journal, Oneda and colleages, reporting for the European Group for Blood and Marrow Transplantation (EBMT) on over 500 MDS patients undergoing matched sibling HSCT, explore the relevance of chromosomal abnormalities on predicting transplant outcome.[5] As predicted, the IPSS and FAB scores correlated broadly with outcome. They went on to show that presence or absence of karyotypic abnormalities had no impact on the outcome of the good risk FAB group (RA/RARS). In contrast, RAEB and RAEBt patients with poor risk karyotypes by IPSS criteria had an almost 2-fold increase in the risk of relapse and a correspondingly lower overall survival. Into this matrix they then factored the impact of chemotherapy given prior to transplant and the subsequent marrow blast percentage and remission status at transplant to compare four groups: RA/RARS *versus* untreated RAEB/RAEBt CMML *versus* non-RA/RARS treated and achieving first complete remission (CR1) *versus* non-RA/RARS not in CR1 at transplant. Best outcome (survival) was seen in the good risk RA/RARS group but treated patients in CR1 fared equally well (5-year survival 55%). In contrast, untreated