Plasma cell quantification in bone marrow by computer-assisted image analysis

P. Went, S. Mayer, M. Oberholzer and S. Dirnhofer
Institute for Pathology, University of Basel, Switzerland

Summary. Background: Minor and major criteria for the diagnosis of multiple meloma according to the definition of the WHO classification include different categories of the bone marrow plasma cell count: a shift from the 10-30% group to the >30% group equals a shift from a minor to a major criterion, while the <10% group does not contribute to the diagnosis. Plasma cell fraction in the bone marrow is therefore critical for the classification and optimal clinical management of patients with plasma cell dyscrasias. The aim of this study was (i) to establish a digital image analysis system able to quantify bone marrow plasma cells and (ii) to evaluate two quantification techniques in bone marrow trephines i.e. computer-assisted digital image analysis and conventional light-microscopic evaluation. The results were compared regarding inter-observer variation of the obtained results.

Material and methods: Eighty-seven patients, 28 with multiple myeloma, 29 with monoclonal gammopathy of undetermined significance, and 30 with reactive plasmocytoisis were included in the study. Plasma cells in H&E- and CD138-stained slides were quantified by two investigators using light-microscopic estimation and computer-assisted digital analysis. The sets of results were correlated with rank correlation coefficients. Patients were categorized according to WHO criteria addressing the plasma cell content of the bone marrow (group 1: 0-10%, group 2: 11-30%, group 3: >30%), and the results compared by kappa statistics.

Results: The degree of agreement in CD138-stained slides was higher for results obtained using the computer-assisted image analysis system compared to light microscopic evaluation (corr.coeff.=0.782), as was seen in the intra- (corr.coeff.=0.960) and inter-individual results correlations (corr.coeff.=0.899). Inter-observer agreement for categorized results (SM/PW: kappa 0.833) was in a high range.

Conclusions: Computer-assisted image analysis demonstrated a higher reproducibility of bone marrow plasma cell quantification. This might be of critical importance for diagnosis, clinical management and prognostics when plasma cell numbers are low, which makes exact quantifications difficult.

Key words: CD138, Multiple myeloma, Plasma cell, Image analysis

Introduction

The WHO definition of plasma cell dyscrasias comprises plasma cell myeloma and variants such as smouldering myeloma and its precursor lesion, called monoclonal gammopathy of unknown significance (MGUS). The diagnosis is based on criteria such as the quantity of bone marrow plasma cells, the quantity of the immunoglobulin produced by the tumor cells (M-component) and clinico-radiologic findings such as lytic bone lesions, infections or serum calcium. Although the tumor burden can be reduced in multiple myeloma (MM), complete and sustained disease remission is rare, and only 35% of patients survive long-term. As new and aggressive therapies like stem cell transplantation arise (Barlogie et al., 2004), the diagnosis of multiple myeloma has to fulfill the highest standards.

But the quantity of bone marrow plasma cells is also of critical significance for the timing of therapy (Patriarca et al., 1999; Rajkumar et al., 2001). Visualization of plasma cells in histological bone marrow slides can be improved if the cells are marked with antibodies against immunoglobulin (Markey et al., 2003) or CD138 (Syndecan-1), a transmembrane protein expressed specifically on normal and neoplastic lymphoid cells with plasmocytic differentiation.
Quantification of bone marrow plasma cells

(Sebestyen et al., 1999). Quantification of bone marrow plasma cells is usually done by simple visual approximation of morphologic findings, a method subject to human fallibility (Furness and Lauder, 1997) and considerable intra- and interobserver variability: pathologists differed in counting mitotic figures between 2 and 34 mitoses per 10 high power fields on an identical slide (Silverberg, 1976). Computer-assisted image analysis contrarily should yield highly reproducible results, still, the improvement of interobserver concordance by digital image analysis has to be shown in every single application. In the bone marrow, morphometry was initially used to differentiate subtypes of myeloproliferative diseases with planimetric analysis of megakaryocyte morphology (Thiele et al., 1982) and its application was expanded since then to the detection of minimal residual disease in chronic lymphocytic leukemia (Gala et al., 1999). Nevertheless, only very few studies exist in which digital image analysis was used to quantify cells in the bone marrow (Thiele et al., 1988, 2001) We provide and compare for the first time data of plasma cell fractions in bone marrow obtained by a computer-assisted digital image analyzing system to conventional light-microscopic evaluation.

Material and methods

Patients

We retrospectively reviewed bone marrow biopsy slides of 87 patients with MM (n=28), MGUS (n=29) and reactive plasmocytosis (rP, n=30). All biopsies were taken between 2001 and 2003 and diagnosed by one of two surgical pathologists (SD, PW) specialized in hematopathology. Retrieval of tissue and clinical data was performed according to the regulations of the local institutional review board and data safety laws of the federal institutions (Basel-Stadt, Switzerland). Stained slides were retrieved from the archives of the Institute for Pathology, University of Basel. Combined clinical, radiological and laboratory data according to the established criteria of the WHO (Jaffe et al., 2001) was required, including a bone marrow biopsy with immunohistochemical labelling of the immunoglobulin light chains κ and λ, radiographic pictures of the axial skeleton and analyses of blood and urine. Monoclonal plasma cells were considered to be present in the biopsy sample when the ratio of κ+ve or λ+ve plasma cells fell outside the range of >10:1 for one of the immunoglobulin light chains.

In all cases, the biopsies contained at least ten mark cavities to be regarded as sufficiently representative.

Immunohistochemistry

All 87 bone marrow biopsies were SU SA-fixed (Bancroft and Gamble, 2001), decalcified, and paraffin-embedded. Five micron thick serial sections were cut and transferred to slides (Menzel-Gläser, Braunschweig, Germany). Routine hematoxylin and eosin (H&E) staining was performed on an automated device (tissue stainer cot 20, Medite, Nunningen, Switzerland). CD138 (mouse anti-human CD138, Clone M115, DAKO, Glostrup, Denmark, 1:25 diluted in Antibody Diluent [Cat No 251-018, Ventana Medical Systems, Tucson, AZ, USA]) immunostaining was accomplished using an automated modified streptavidin-biotin system (Ventana NexES, Ventana Medical Systems, USA) after heat-induced antigen retrieval in a microwave (100°C, 15 min) in EDTA buffer (pH=8.0). Osmium-enhanced diaminobenzidine was used as chromogen. For negative controls, the primary antibody was omitted, as positive controls served plasma cells normally present in bone marrow biopsies.

Light-microscopic analysis

The percentage of plasma cells relative to total bone marrow cells was estimated independently on H&E- and CD138-immunostained slides using a conventional microscope. The two investigators (SM, PW) were blinded for clinical diagnoses and each other’s results. A fraction of the cases (9 MM, 9 MGUS and 10 rP) were randomly selected and analysed a second time by one of the investigators (SM) to establish intra-observer variability.

Computer-assisted image analysis

Quantitative analysis of CD138-expressing plasma cells in bone marrow was assessed using an image processing system containing a microscope (Axioplan 2, Carl Zeiss AG, Feldbach, Switzerland), a color video camera (C 5810, Hamamatsu Photonics, Hamamatsu, Japan), a PC with a frame grabber (Meteor 2/MC, Matrox Electronic Systems, Quebec, Canada) and image processing software (KS 300, Carl Zeiss AG, Feldbach, Switzerland). A light source with fixed luminance (3200K) and a fixed exposure time (0.04sec) was used at x16 magnification. Only fields without bone formation were digitized. After grabbing and digitizing the selected field, the image was imported into the image processing system, where the CD138-positive areas and the cellular component without bone marrow fat were extracted. The immunopositive area fraction (IAF%) was calculated as the fraction of these two areas. In the first step of this process, the diaminobenzidine color signal was extracted using a command based on the hue, lightness, and saturation (HLS) color parameter. The resulting binary picture was optimized using the specific image postprocessing commands erosion, close, and elimination of small areas (Serra, 1982) until the binary signal precisely covered the plasma cells (Fig. 1). In a second analogous step, the process focused on the haematopoietic bone marrow, discarding fat cells from the analysis. All steps were established by trial and error, after which the protocol was fixed and all measurements...
were executed using exactly the same conditions to enable comparative results. The quantification unit was pixel (1 pixel = 0.297 mm).

In a first analysis, the standard error of the system was defined by measuring the same field 20 times. Variables that could add to the standard error include luminance instability, sensitivity of the camera, focus, and image processing variability. Secondly, bone marrow trephines were analysed methodically in as many fields as possible per biopsy. The final result is shown as immunopositive area fraction (IAF%) and was calculated as the median of the measured haematopoietic bone marrow and plasma cell fractions, respectively. The respective upper (75 percentile = Q3) and lower quartiles (25 percentile = Q1) illustrate the range of values. The two investigators analysed the cases analogous to the light microscopic evaluation.

**Statistical analysis**

The analysis was performed in Excel (Office 2000, Microsoft Corp., Redmond USA) and SPSS (v11.0, SPSS Inc., Chicago, USA). The strength of the correlation was evaluated by a Spearman correlation test, corrected for ties. In addition, weighted kappa statistics were performed using StatXact (Cytel Software Co., Cambridge MA, USA). To calculate Kappa-statistics, three groups of IAF% were formed: group 1 = 0-9%, group 2 = 10-29%, and group 3 = 30-100% in accordance with major and minor WHO criteria defining MGUS and MM (Jaffe et al., 2001), where bone marrow plasmocytosis of <10% is considered diagnostic for MGUS, and plasmocytosis of >30% specifies MM. This method allows us to verify if placement in these diagnostic relevant groups exceeds chance levels. P-Values < 0.05 were considered to be significant, two-sided tests were used in all calculations. For weighted kappa, discordance between groups 1 and 3 was considered more important than between groups 1 and 2 or 2 and 3. Weights were calculated according to Cicchetti-Allison (Cicchetti, 1972). Kappa values of 0.41-0.60 were considered moderate, and 0.61-0.80 as substantial (Landis and Koch, 1977). The following items were specifically addressed: inter-method variation designates the variation in results of different assessment methods independent of the investigator. Intra-observer variation designates variation of results from different assessments by a single investigator, whereas inter-observer variation designates variations between the results of the two investigators. Intra- and inter-observer variations are, therefore, independent of the methodologies.

**Results**

**Observer-dependent, H&E-based evaluation**

Twenty-eight specimens (9 MM, 10 rP, 9 MGUS) were used to conventionally assess plasma cell fractions by two investigators (SM, PW). Spearman correlation coefficient was low (0.472), and inter-observer agreement was insufficient (weighted kappa: 0.192, 95%CI -0.119-0.503, p value = 0.108).

**Observer-dependent, CD138-based evaluation**

CD138-stained bone marrow biopsies were first subjected to conventional assessments (Table 1). In 28 of the biopsies, the correlation coefficients between the first and second CD138-based estimation by SM was 0.782. The correlation coefficients of the CD138-based estimations by SM and PW were 0.876 for the first estimation, and 0.803 for the second estimation, respectively.

**Observer-independent, CD138-based assessment**

To evaluate the methodical accuracy of the

---

**Table 1. Results of estimation in CD138 stained slides.**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>n</th>
<th>SM mean</th>
<th>Q1a- Q3b</th>
<th>PW mean</th>
<th>Q1a- Q3b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Myeloma</td>
<td>28</td>
<td>25</td>
<td>17.5-35</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>MGUSc</td>
<td>9</td>
<td>25</td>
<td>17.75-30</td>
<td>30</td>
<td>25-31</td>
</tr>
<tr>
<td>Reactive Plasmocytosis</td>
<td>10</td>
<td>29</td>
<td>9.5-12.75</td>
<td>10</td>
<td>6.5-12.25</td>
</tr>
<tr>
<td>all</td>
<td>8</td>
<td>10</td>
<td>6-15</td>
<td>10</td>
<td>7-18</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8</td>
<td>7.25-21.5</td>
<td>16</td>
<td>8-30</td>
</tr>
</tbody>
</table>

a: 25 percentile; b: 75 percentile; c: monoclonal gammapathy of unknown significance; n.a. not available. Numbers indicate the median percentage of CD138-positive cells in the red marrow. Estimation was performed twice by investigator SM.
computer-assisted digital image analysis system and to

Quantification of bone marrow plasma cells

test the sensitivity to various parameters (luminance, camera sensitivity, focus), we first measured the same field 20 times and determined a mean haematopoietic area of 297738 pixels (min 297681, max 297782), and a mean plasma cell area of 277009 pixels (min 276252, max 277833). This resulted in a mean plasma cell fraction of 93.038% (min 92.776%, max 93.314%) in this specific field, establishing an acceptable value range of 0.538%. As many areas as possible in each bone marrow biopsy (mean 19.5 fields [4 min, 32 max]) were then assessed with the computer-assisted digital image analysis system by the two investigators. The results are presented in Table 2.

In 87 biopsies, the correlation coefficient was 0.899 for the two investigators (SM, PW). In the 28 cases measured twice by SM, the correlation coefficient was 0.881.

Intra-observer agreement for one of the investigators (SM) was substantial when the plasma cell count was estimated (kappa 0.596, 95% CI 0.343-0.849, p value<0.001), and even higher if measured digitally (kappa 0.704, 95% CI 0.476-0.933, p value<0.0001). Inter-observer estimations correlated better and were considered substantial (SM/PW, kappa 0.789, 95% CI 0.646-0.932, p value<0.0001), being near perfect for repeated measurements (SM/PW, kappa 0.849, 95% CI 0.687-1.000, p value<0.0001). Inter-method agreement was substantial (SM: kappa 0.658, 95% CI 0.537-0.780, p value<0.0001, PW: kappa 0.677, 95% CI 0.534-0.820, p value<0.0001) for both investigators. Table 3 shows the intra-observer and inter-method variability for the investigator SM and for inter-observer variability (SM/PW).

### Discussion

In this study, we found that CD18-immunostained plasma cells can be quantified using a computer-assisted digital image analysis system on histological bone marrow slides. The quantification of immunostained plasma cells obtained by image analysis-based measurements correlated with the corresponding CD18-based estimation. The computer-assisted digital image analysis system proved substantially superior to simple estimation of the plasma cell content since inter- and intra-observer variability was minimized. As presumed, H&E-based estimation of the plasma cells alone is insufficient for therapeutic design.

We focused the image analysis system in this study on plasma cell quantification. Qualitative morphologic features such as bizarre multinucleated plasma cells and immature plasma cells with nuclear cytoplasmic asynchrony, enlarged nucleoli or dispersed chromatin, which can be helpful in establishing final diagnosis (Barth et al., 1982) and prognosis (Goasguen et al., 1999), were not considered as they are subdued to individual consideration and therefore contribute to result variability. For example, agreement between three observers in grading dysplasia in colorectal adenomas was seen in only 47.8% of cases (Jensen et al., 1995), which indicates a high variation for qualitative histologic parameters. The bone marrow plasma cell content therefore, especially in the presence of microaggregates, is a more accessible tool in the diagnosis of plasma cell dyscrasias (Buss et al., 1986). Nevertheless, plasma cell nodules contribute to discrepancies in the plasma cell counts assessed by flowcytometry, bone marrow smears and bone marrow histology due to either the presence of fibrosis in bone marrow (Terpstra et al., 1992), or their patchy distribution. While the former complicates plasma cell detection with tissue-derived methods of bone marrow aspirates, the latter can only be overcome with representative bone marrow sampling and is independent of the method used to examine the tissue.

We compared our light-microscopic results with those obtained by semi-automated image analysis, and show that the computer-assisted image analyses of plasma cell contents were generally lower compared to simple estimations. Since the image analysis post-

### Table 2. Results of digital image analysis in CD138- stained slides.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>n</th>
<th>SM</th>
<th>PW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>mean IA F%</td>
<td>Q1a - Q3b</td>
</tr>
<tr>
<td>Multiple Myeloma</td>
<td>28</td>
<td>19.5-39.55</td>
<td>24.1-44.45</td>
</tr>
<tr>
<td>MGUS</td>
<td>29</td>
<td>6.3-11.13</td>
<td>6.7-8.68</td>
</tr>
<tr>
<td>Reactive Plasmocytosis</td>
<td>30</td>
<td>5.7-12.48</td>
<td>6.35-8.3</td>
</tr>
<tr>
<td>all</td>
<td>10</td>
<td>5.3-9.80</td>
<td>n.a.</td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>7.5-16.55</td>
<td>7.7-15.58</td>
</tr>
</tbody>
</table>

a: 25 percentile; b: 75 percentile; c: immuno-positive area fraction; d: monoclonal gammopathy of unknown significance; n.a. not available.

Numbers indicate the median percentage of CD138-positive cells in the red marrow. Measurement was performed twice by investigator SM.

### Table 3. Comparison of the different methods to assess plasma cell numbers in CD138-stained bone marrow slides.

<table>
<thead>
<tr>
<th>Comparison of differing methods</th>
<th>n</th>
<th>weighted kappa</th>
<th>Corr.coeff. a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraobserver-variability of conventional assessment</td>
<td>28</td>
<td>0.596</td>
<td>0.782</td>
</tr>
<tr>
<td>Intraobserver-variability of digital assessment</td>
<td>28</td>
<td>0.704</td>
<td>0.960</td>
</tr>
<tr>
<td>Interobserver-variability of conventional assessment</td>
<td>28</td>
<td>0.789</td>
<td>0.876</td>
</tr>
<tr>
<td>Interobserver-variability of digital assessment</td>
<td>87</td>
<td>0.833</td>
<td>0.899</td>
</tr>
<tr>
<td>Intermethod-variability</td>
<td>87</td>
<td>0.658</td>
<td>0.776</td>
</tr>
</tbody>
</table>

a: Spearman's rank correlation coefficient.
processing procedure was optimized to detect small clusters and single plasma cells, parts of single plasma cells could be missed, especially if technical artefacts are added, indicating that systematic underestimation of results by image analysis cannot be completely excluded. As the correlation of the conventionally and digitally assessed results shows, variation in cell size in individual cases does not influence the result. The detection of large clusters of plasma cells would fulfill as major criterion for the diagnosis of multiple myeloma according to the WHO criteria and was therefore regarded as less critical in this analysis.

The results of the intra-observer agreement assessments showed that experience in the use of the methods generally improved the outcome, e.g. results variability was reduced when more measurements were performed. Since plasma cell numbers in bone marrow is a critical value in patient management (Patriarca et al., 1999; Rawstron et al., 2002), high reproducibility and accuracy of the results is crucial, particularly in the management of patients requiring repeat bone marrow examinations, such as MGUS patients (Riccardi et al., 1990), or in assessment of residual disease after treatment (Blade et al., 1998). In chronic lymphatic leukemia, digital image analysis of the bone marrow was found to be of equal sensitivity in the detection of minimal residual disease if compared to semiquantitative polymerase chain reaction (Gala et al., 1999). Only very few other applications quantifying cells by the use of digital image analysis in the bone marrow exist: CD34 positive progenitor cells were successfully quantified to provide prognostic information in chronic myelogenous leukemia (Thiele et al., 2001) and the plasma cell content was corresponding well with the secreted serum M-component (Thiele et al., 1988). We can now add image analysis as a diagnostic tool in the diagnosis of plasma cell dyscrasias.

Decisions that are profoundly critical to patient management sometimes rely on assessments of very low reproducibility. Still, many health professionals advise a second expert opinion instead of observer-independent procedures such as digital morphometry before aggressive therapy regimens are executed. Furthermore, digital analysis has several undeniable advantages, such as low initial costs and, once established, ease of use and cost-effectiveness, since technicians perform the measurements. Moreover, the high reproducibility suggests that physicians involved in patient management would welcome the results. Important obstacles, however, remain to be overcome before this system can be successfully integrated into routine clinical use: Slides need to be of equally high quality, which can be difficult with bone marrow samples. This quality improvement is achieved through standardization of methods, including fixation and decalcification times as well as immunostaining protocols (Bancroft and Gamble, 2001; Fend et al., 2005). Antibody specificity must be high, since variation in a single factor, such as fixation time, can compromise the results. Nevertheless, standardization of procedures is a goal that will be beneficial to other aspects of laboratory work as well.

In summary, we have shown that computer-assisted digital image analysis can be a valuable tool in quantifying bone marrow plasma cells, particularly in cases where a plasma cell content of 10-30% can delay diagnosis because other criteria, which could be time-consuming to establish, are needed. Digital image analysis therefore has the potential to optimize clinical management and prognosis.

Acknowledgements. We thank H. Weisskopf for excellent technical assistance, and M. K. Occhipinti-Bender for editorial assistance.

References

Quantification of bone marrow plasma cells


Accepted March 23, 2006