Guidelines for risk stratification in multiple myeloma: report of the International Myeloma Workshop Consensus Panel 2

Nikhil C. Munshi¹,², Kenneth C. Anderson¹, P. Leif Bergsagel³, John Shaughnessy⁴, Antonio Palumbo⁵, Brian Durie⁶, Rafael Fonseca³, A. Keith Stewart³, Jean-Luc Harousseau⁷, Meletios Dimopoulos⁸, Sundar Jagannath⁹, Roman Hajek¹⁰, Orhan Sezer¹¹, Robert Kyle¹², Pieter Sonneveld¹³, Michele Cavo¹⁴, S. Vincent Rajkumar¹², Jesus San Miguel¹⁵, John Crowley¹⁶ and Hervé Avet-Loiseau⁷, and on behalf of the International Myeloma Workshop Consensus Panel 2

¹Dana-Farber Cancer Institute Boston, ²Boston VA Healthcare System, West Roxbury, MA; ³Mayo Clinic, Scottsdale, AZ; ⁴University of Arkansas for Medical Sciences, Little Rock, AR; ⁵University of Torino, Torino, Italy; ⁶Cedars-Sinai Medical Center, Los Angeles, CA; ⁷Universitaire de Nantes, Nantes, France; ⁸University of Athens, Athens, Greece; ⁹Mt. Sinai Medical Center, New York, NY; ¹⁰University of Brno, Czech Republic; ¹¹University Medical Center, Hamburg Germany; ¹²Mayo Clinic, Rochester, MN; ¹³Erasmus MC, Rotterdam, The Netherlands; ¹⁴University of Bologna, Bologna, Italy; ¹⁵University of Salamanca, Salamanca, Spain; ¹⁶Cancer Research & Biostatistics, Seattle, WA

Correspondence: Nikhil Munshi, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115; email: nikhil_munshi@dfci.harvard.edu. and Herve Avet-Loiseau, 42, Rue Jules SIMON, 44000 Nantes, France e-mail: herve.avetloiseau@chu-nantes.fr.
Abstract

A panel of members of the 2009 International Myeloma Workshop developed guidelines for risk stratification in multiple myeloma (MM). The purpose of risk stratification is not to decide time of therapy but to prognosticate. There is general consensus that risk stratification is applicable to newly-diagnosed patients; however, same genetic abnormalities characteristic of poor outcome at diagnosis may suggest poor outcome if detected at the time of relapse. Thus in good risk patients it is necessary to evaluate for high-risk features at relapse. Although detection of any cytogenetic abnormality is considered to suggest higher risk disease, the specific abnormalities considered as poor risk are: cytogenetically detected Chromosomal 13 or 13q deletion, t(4;14) and del17p; and detection by FISH of t(4;14); t(14;16) and del17p. Detection of 13q deletion by FISH only, in absence of other abnormalities, is not considered a high-risk feature. High serum β2M level and ISS stage II and III incorporating high β2-M and low albumin are considered to predict higher risk disease. There was a consensus that the high risk features will change in the future, with introduction of other new agents or possibly new combinations.
**Background**

Multiple myeloma is a heterogeneous disease with variable disease courses, response to therapy and survival outcome that ranges from less than 1 year in patients with aggressive disease to more than 10 years in patients with indolent disease presentation. Various patient, disease and therapy-related characteristics have been identified to predict the disease course as well as outcome in patients with myeloma. Such evaluation of prognostic factors and risk stratification is important to define treatment strategies, compare outcome of therapeutic trials, and predict survival from diagnosis. This consensus panel report has evaluated various aspects of risk stratification including its purpose and timing, especially at diagnosis and at relapse, its relationship to therapy and defined clinical, and laboratory features including genomic changes that may be utilized to stratify patients and predict outcome at the present time.

**Purpose of Risk stratification**

The general purpose of risk stratification is not to decide whether to treat a patient or not but to prognosticate.

Decision to treat is based on the criteria set for in the diagnosis of symptomatic myeloma which includes the CRAB (hypercalcemia, renal dysfunction, anemia and bone lesions) criteria \(^1\). Patients with clearly defined MGUS or SMM do not need initiation of therapy irrespective of any associated risk factors except on
specifically targeted protocols. For example, if a patient with clear diagnosis of SMM has 17p- by FISH or del13 on cytogenetics analysis, it does not constitute an indication to start therapy. The risk stratification being described here is only for determining prognosis and stratification of treatment, rather than to decide whether to initiate treatment.

There has been general consensus on the risk factors that help stratify patients receiving conventional therapeutic modalities. However, there are studies that suggest that Bortezomib, and to an extent Lenalidomide, may be able to overcome some of the poor risk features and achieve clinical benefit. Further studies are needed to decide on the ability of these agents to overcome the poor risk features. At the present time, it is important to stratify, but the available information does not indicate selection of therapies; e.g. if patient has t(4:14), it does not suggest that we should use a specific therapy or novel agent.

Currently, to mandate definitive treatment according to cytogenetic abnormalities is premature, although there is emerging data suggesting that some of the novel agents could overcome the negative prognosis of the cytogenetics abnormalities. It is important to continue to assess the impact of risk factors with novel therapies and combinations. Clinical trials should be done based on risk stratification, to try and test whether certain patients benefit more or less from certain therapeutic agents or strategies.
Timing of Risk stratification

At Diagnosis: There is consensus that the current risk stratification is applicable to newly-diagnosed patients. All current efforts are directed at stratification at diagnosis using the parameters obtained at diagnosis. The suggested investigations are listed in Table 1.

At Relapse: There has been documentation of change in risk factors at relapse. For example, in one study the mean labeling index increased from 1% at diagnosis to 2.5% at relapse. If patients are followed individually, it is always higher at relapse versus at diagnosis. Similar data exists for detection of p53 deletion at relapse or disease progression compared to time of diagnosis.

The evolving consensus is that if a patient acquires high risk features at relapse or progression, then that patient should be reclassified as having high-risk disease. For example if a patient was not detected to have del(17p) at diagnosis but at relapse 60% cells show this change, then it is our consensus that this patient now should be reclassified as having high-risk disease.

There is general consensus that same genetic abnormalities characteristic of poor outcome at diagnosis may suggest poor outcome if detected at the time of relapse.

Amongst the non-genetic risk factors, re-determination of factors such as β-2M or ISS at relapse or at follow up is not currently considered as predictive of change in risk stratification. Role of level of serum LDH at relapse is less clear. A very
high LDH is considered to represent proliferative disease. High LDH levels are uncommon in myeloma, but carry poor prognosis at diagnosis or relapse\textsuperscript{10-12}.

Moving forward, an important goal for relapse trials would be to evaluate these and other risk factors at relapse, keeping in consideration the type of therapy used at relapse. This is important as novel therapies are becoming available and patients are living longer.

In patients with relapsed disease, additional risk stratification criteria include type of response and length of response to prior therapy. Therapy-related poor risk features include progression while on therapy, and short duration of response\textsuperscript{6,13}. Unlike in the past, speed of response does not suggest overall outcome with newer agents.

If a patient already has an identified high risk feature at diagnosis, then there is no need to look for the same feature again. For example, if a patient at diagnosis has t(4;14), then one does not need to look for it again at relapse with same FISH probe. However, cytogenetic and FISH investigation should be performed at relapse to look for additional changes. If a patient is in a low risk group, then it is recommended that cytogenetics and FISH studies be performed at relapse for risk re-stratification as if not detected at the presentation, at relapse, due to selection of a pre-existing clone they may attain a detectable level.

\textbf{Is risk stratification specific for specific treatment}
There is general agreement that the risk stratification should be a global stratification, and not stratification for old versus new therapy or risk stratification for particular one treatment. We recognize that the risk features may be relevant to a given therapy. For example, when patients with del13 are considered to have poor prognosis, it is based on a large number of studies focused on outcome following high-dose therapy and transplantation. However with the use of novel agents, for example bortezomib, del13 does not seem to be predictive of high-risk.\textsuperscript{14,15} Thus risk factors for individual novel therapies are to be determined on an ongoing trial.

There was a consensus that the high risk features will change in the future, with introduction of other new agents or possibly new combinations.

It is unclear whether risk stratification should change after patients receive certain treatments. For example, Bortezomib is able to overcome, at least in part, the poor risk associated with t(4;14);\textsuperscript{2} do we need to identify different risk factors for patients who are post-bortezomib treatment. The general opinion was that there is not adequate information to implement such a recommendation.

**What risk factors to look for**

There is a consensus that both cytogenetics and FISH play an important and independent role in risk stratification. Both FISH with specific markers and cytogenetics with specific abnormalities needs to be performed on bone marrow
samples.

Although detection of any cytogenetic abnormality is considered to suggest higher risk disease, the specific abnormalities considered as poor risk are: cytogenetically detected Chromosomal 13 or 13q deletion, t(4;14) and del17p; and detection by FISH of t(4;14); t(14;16) and del17p.\textsuperscript{16}

High serum $\beta_2$M level and ISS stage II and III incorporating high $\beta_2$-M and low albumin are considered to predict higher risk disease.\textsuperscript{17}

What additional risk factors to look for.

A number of individual risk factors have been identified. However, there is in general emphasis to use a system that combines multiple factors, such as ISS. Some of these factors were considered in developing the ISS risk stratification systems.

Due to lack of uniform availability of the data for analysis that led to proposal of ISS system, there are a number of factors which still may have a significant role in risk stratification as individual factors, e.g. LDH was not available for all patients and was not considered in developing ISS. However, in the limited patients that had this information LDH was found to have significant influence in identifying risk.

Some of the features considered significant as individual factors are LDH, IgA,
extramedullary disease, renal failure, high serum free light chain and serum free kappa/serum free lambda ratio, plasmablastic disease, and plasma cell leukemia. These features are useful under some circumstances but, their general applicability is unknown. Also it is very much a constellation of features that often determine high risk, rather than a single factor which may make it intermediate risk. Unlike FISH/cytogenetics which may suggest a change in therapeutic approach to more aggressive treatment, no change in treatment approach is currently indicated based on such single higher risk features.

Consensus for evaluation of genomic changes

As described above, there is a consensus that both cytogenetics and FISH play an important and independent role in risk stratification. Both have some favorable and some adverse risk features. Both highlight different disease parameters, and both need to be performed to have better understanding of the behavior and biology of the disease.

FISH data should be reported specifically for clonal plasma cells determined by surface marker or cytoplasmic immunoglobulin light chain expression, and not all cells. The positivity is to be determined by % positive cells that are above the individual laboratories' standard.

No specific global cut off should be applied. It is unclear whether number of positive cells carries any different risk. For example, if a patient has 7% versus
57% cells positive for a specific FISH abnormality, the relative risk for both patients is considered same at the present time. This is not true for del(17p). In a report del(17p) is prognostic only if present in at least in 60% of the PCs. ²

There is consensus that 1) detection of t(4;14), t (14;16), or17p by FISH suggests higher risk disease; 2) del13 or 13q- detected only by FISH independently in the absence of other abnormality does not carry significant higher risk, while t(11;14) does not predict superior outcome; 3) There are some reports that 1q+, Del 1p may have clinical significance as a poor risk feature; however, the consensus is that the data is not yet adequate to suggest routine use of these FISH markers to predict prognosis.

ISS system

ISS system incorporating serum albumin and β2M is applicable as a prognostic system in majority of the settings. ISS is validated for conventional treatments as well as high-dose therapy. ¹⁷ However, its validity with combination novel agent therapy still needs to be confirmed.

Method used for measurement should be standard. ISS system, although extremely convenient for everyone to use, requires incorporation of additional myeloma-specific features to make it more robust or more applicable using the newer generation of drugs and studies.

The ISS system is a baseline lowest common denominator, to be supplemented
and not necessarily supplanted. There is a clear need and consensus to add cytogenetics/FISH or other markers to ISS.

**Durie Salmon classification system**

A clinical staging system at diagnosis using standard laboratory measurement, developed by Durie and Salmon, was predictive of clinical outcome after standard-dose chemotherapy. However, with the use of high-dose therapy and novel agents, the Durie-Salmon (DS) system is less predictive of outcome.\(^{17,23}\) This may be explained by the fact that the DS system is predominantly focused on tumor burden, and as these newer therapies are able to better reduce tumor burden, its significance has changed. There is increasing importance of tumor biology-related factors. DS system is still considered a means to measure tumor mass.

There is general agreement that DS system can supplement the diagnostic criteria for myeloma such as CRAB (hypercalcemia, renal dysfunction, anemia and bone disease); however, if a patient has already been diagnosed as having symptomatic myeloma based on current criteria, then there is no need to use the DS system in regular practice for diagnosis. As only patients with symptomatic disease should be placed in clinical trials, reporting of DS system is not considered essential. As Stage I represents early stage of disease, description of patients in clinical studies by DS staging system is encouraged. However, it routine clinical use is unclear.
**Incorporation of imaging results**

Number of bone lytic lesions, per DS system, is not considered of any prognostic significance.

Although there are small single institution studies that have indicated that achieving MRI-directed CR has prognostic significance, this observation requires further studies to include imaging parameters in risk stratification or response definition. Similarly, a recent study has pointed to presence of more than 3 FDG-avid focal lesions as the leading independent parameter associated with inferior overall and event-free survival. However, these results require further independent confirmation before they are widely applied.

None of the imaging studies or results are currently recommended for inclusion in risk stratification.

**Inclusion of expression/genomic profile**

Expression profile data generated by a number of groups have been very helpful in identifying an expression signature that may identify a poor risk group. Shaughnessy et al investigated expression profile of myeloma cells in 532 newly-diagnosed myeloma patients treated on 2 protocols incorporating tandem autologous transplantation. Using log-rank tests of expression quartiles, 70
genes linked to shorter durations of complete remission, event-free survival, and overall survival were identified. The ratio of mean expression levels of up-regulated to down-regulated genes defined a high-risk score which was an independent predictor of outcome endpoints in multivariate analysis ($P < .001$) that included the ISS and high-risk translocations. A subset of patients with high risk score had a 3-year continuous complete remission rate of only 20%, as opposed to a 5-year continuous complete remission rate of 60% in absence of a high-risk score. Interestingly, multivariate discriminant analysis identified a 17-gene subset that performed as well as the 70-gene model.  

A second large study published by Decaux et al from the IFM group studied gene expression profiles of myeloma cells obtained at diagnosis in 182 patients and identified the 15 strongest genes to calculate a risk score associated with the length of survival. This analysis divided patients into high-risk group, characterized by the overexpression of genes involved in cell cycle progression and its surveillance, and low-risk patients, with hyperdiploid signature and heterogeneous gene expression. The results were confirmed in a test set, as well as independent cohorts comprising 853 patients with multiple myeloma. Overall survival at 3 years in a low risk or high risk groups was 91% and 47%, respectively. These results were independent of traditional prognostic factors.  

It is interesting to note that although both these studies have included patients undergoing high-dose therapy, the 15 and 17 gene models do not share common genes. This highlights the complexity of biological behavior of the tumor and the fact that ultimate utilization of such expression data will require significantly more
work. Functional commonality or functional association between these various genes needs to be considered in developing a more composite model. It also highlights the molecular redundancy in tumor cells driving their clinical behavior.

The factors that require standardization are method used to assess expression profile, the data analysis technique, consensus and validation of genes to be considered important for risk stratification, and standardization of method to apply this definition to expression profile for a single patient.

A more robust and comprehensive analysis is needed to analyze significance of stratification using CGH/ SNP array.

In the future, a specific polymorphism may help identify patients with differential response profile and/or higher risk of toxicity. However, currently there is lack of data to propose any specific SNPs that can be utilized for such decision.

Consideration of risk factors in special therapeutic scenario.

There is emerging data that allogeneic transplant may have beneficial outcome in high-risk patients defined by cytogenetics/FISH. This data is limited and requires further confirmation. However, the group feels that allogeneic transplant could be considered in this group of patients.

Current level of evidence does not provide direction in deciding if patients with a
specific risk group will benefit from maintenance therapy.

Acknowledgements

The authors thank the following colleagues for their participation on the Consensus Panel: Drs. Shaji Kumar (Rochester, MN), David S. Siegel (Hackensack, NJ), Philippe Moreau (Nantes, France), Hermann Einsele, Seema Singhal (Chicago, IL), Irene M. Ghobrial (Boston, MA), Gösta Gharton (Stockholm, Sweden), James R. Berenson (West Hollywood, CA), Jayesh Mehta (Chicago, IL), Angela Dispenzieri (Rochester, NY), Sagar Lonial (Atlanta, GA), Ruben Niesvizky (New York, NY), Robert Schlossman (Boston, MA), David H. Vesole (New York, NY), Asher Chanan-Khan (Buffalo, NY), Jeffrey Wolf (San Francisco, CA), Michael Kuehl (Bethesda, MD), Johannes Drach (Vienna, Austria), Rik Schots (Brussels, Belgium), Giampaolo Merlini (Pavia, Italy), Maurizio Zangari (Salt Lake City, UT), Jeffrey Zonder (Detroit, MI), L. Thompson Heffner Jr. (Atlanta, GA), Ramon Garcia Sanz (Valencia, Spain), Philip R. Greipp (Rochester, MN), and Vinod Raina (New Delhi, India).

Authorship Contribution Statement: N.C.M., K.C.A., L.B., J.S., A.P., B.D., R.F., K.S., J.LH., M.D., S.J., R.H., O.S., R.K., P.S., M.C., V.R., J.SM., J.C. and H.A.L. developed the consensus, provided critical review and edits to the manuscript, and gave approval to the final manuscript. All authors significantly participated in the development of the consensus and the writing of the manuscript. This work was developed as part of the 12th International Myeloma Workshop in Washington, D.C.

Disclosure of Conflicts of Interest: N.C.M. is consultant/advisory board member of Millenium, Celegene, Novartis, and Onyx. K.C.A. is consultant/advisory board member of Millenium, Celegene, Novartis, Merck, BMS, Signalgenetics and Onyx and co-founded and owns stock in Acytelon. R.F.
has received a patent for the prognostication of MM based on genetic categorization of the disease, received consulting fees from Medtronic, Otsuka, Celgene, Genzyme, BMS and Amgen and research support from Cyline and Onyx. J.S. has been awarded patents, or has submitted patent applications on molecular diagnostics and therapeutics in cancer medicine; receives royalties related to patent licenses from Genzyme, Novartis, and Signal Genetics, LLC.; co-founded and owns stock in Signal Genetics, LLC, a molecular diagnostics company; consulted/advised Signal Genetics, LLC, Array Biopharma, Celgene, Genzyme, Millennium, Centocor Ortho Biotech and Novartis; and received honoraria from Array Biopharma, Celgene, Centocor Ortho Biotech, Genzyme, Millennium, and Novartis. A.K.S is an advisor for Signalgenetics. O.S. has received honoraria from Amgen, Celgene, Janssen, and Novartis. M.C. has received honoraria from Janssen-Cilag, Millennium Pharmac, and Celgene. The remaining authors declare no competing financial interests.

Correspondence: Nikhil Munshi, Dana-Farber Cancer Institute, 44 Binney St., Boston, MA 02115; email: nikhil_munshi@dfci.harvard.edu. and Herve Avet-Loiseau, 42, Rue Jules SIMON, 44000 Nantes, France e-mail: herve.avetloiseau@chu-nantes.fr.

References


Table 1: Investigation for risk stratification

Investigation recommended for risk stratification:

- Serum Albumin and b-2-microglobulin to determine ISS stage
- bone marrow examination for t(4;14); t(14;16) and del(17p) on identified PC by FISH
- LDH
- Immunoglobulin type – IgA
- Histology - plasmablastic disease

Additional investigation for risk stratification:

- Cytogenetics
- Gene expression profiling
- Labeling index
- MRI/PET Scan
- DNA copy number alteration by CGH/SNP array
Guidelines for risk stratification in multiple myeloma: report of the International Myeloma Workshop Consensus Panel 2

Nikhil C. Munshi, Kenneth C. Anderson, P. Leif Bergsagel, John Shaughnessy, Antonio Palumbo, Brian Durie, Rafael Fonseca, A. Keith Stewart, Jean-Luc Harousseau, Meletios Dimopoulos, Sundar Jagannath, Roman Hajek, Orhan Sezer, Robert Kyle, Pieter Sonneveld, Michele Cavo, S. Vincent Rajkumar, Jesus San Miguel, John Crowley and Hervé Avet-Loiseau