Response Assessment in Myeloma: Is Achieving Response Deeper than CR/sCR Important?

Ola Landgren, MD, PhD
Professor of Medicine
Chief Attending Physician
Myeloma Service
Memorial Sloan Kettering Cancer Center
New York, New York

Hello, welcome to Managing Myeloma. I am Ola Landgren. I am the Chief of the Myeloma Service at Memorial Sloan Kettering Cancer Center in New York City. Today, I am going to talk about response assessment in myeloma: is achieving response deeper than complete response (CR) and stringent complete response (sCR) important? I will address this as four topics; the first one is clonal diversity and selection, and clinical impact on minimal residual disease detection.

Multiple myeloma has been found to be massively genetically heterogeneous only at diagnosis.

Myeloma Massively Heterogenous at Diagnosis; Limited Set of Genes Mutated

Although there are many genes that have been recognized that may be mutated in multiple myeloma, we also know that there are a limited number of genes that are the more frequently mutated ones.

Beyond having mutations, if you look across the tumor cells that you can capture in a single myeloma patient, the diversity of the mutational spectrum will vary across the cells in that patient. Around 50% of the mutations that you can find in a patient are actually only present in a quarter of that patient’s tumor cells.
Currently, we still use diagnostic criteria based on the microscope, but if you are going to apply more sophisticated technologies based on what I’ve just showed you, you’ll be able to see that each and every patient has many subclones. Some papers indicate that there could be up to 10, or more, myeloma subclones at the same time in the same patient. On a clinical note, it is important to keep this in mind, because we know that different subclones seem to respond differently to different drugs.

If you follow a patient over time, and you use these more sophisticated molecular tools to follow the disease, you will be able to capture evidence of clonal tides. In other words, if you see relapses over time, you will see the increase and decrease of the subclones, as they come and go as you treat the patient.
Response Assessment in Myeloma: Is Achieving Response Deeper than CR/sCR Important?

Presented by Ola Landgren, MD, PhD

To further illustrate the issue of clonal tides, let’s look at this image. This comes from work by Johnathan Keats and colleagues in Arizona. In the very beginning, on the left, you see that the patient has a dominant clone illustrated in red, but there are subclones illustrated in orange, green, and gray. And you may be able to see there is a small little blue sliver of disease. Each of these colors represents a sub-clone of multiple myeloma cells that together make up the total tumor burden. After diagnosis, the patient is treated, in this case, with lenalidomide/dexamethasone. Unfortunately, the patient has a recurrence later, and again, as you can see, the red proportion of the tumor burden is the major clone that comes back at the relapse, and orange is the second largest. Again, the patient is treated with lenalidomide/dexamethasone. The next time the disease comes back, on a clinical note, the genetic distribution shows that this time the green subclone is the dominant one, the red one is now the second largest, and the orange is the third largest. And, as you see the disease coming and going when the patient is being exposed to different types of therapy, in the very end, unfortunately, the patient develops plasma cell leukemia and passes away. You can see that, at this point, almost the entire circle is now blue. In the beginning, there was only a very small proportion of the blue subclone, but this is the one that comes and goes throughout the course of the disease. Because the patient was given different therapies that did not deliver a cure, this caused clonal selection, and unfortunately, the most aggressive clone is the one that develops into plasma cell leukemia and leads to the patients’ passing away.
The clinical implications and treatment responses of MRD are included in the International Myeloma Working Group’s guidelines for response re-evaluation. As you probably know, complete response is based on the negative immunofixation of the serum and the urine, as well as on the disappearance of any soft tissue plasma cystoma and on a bone marrow biopsy showing less than 5% plasma cells in the bone marrow. Beyond complete response (CR), we use the terminology stringent complete response (sCR) but sCR is the same as complete response, as well as a normalization of the free-light chain ratio. The minimal residual disease, or MRD, criteria in the current guidelines imply that you can use flow cytometry or molecular technologies to rule out MRD. The International Myeloma Working Group is currently working on developing new defined response criteria for minimal residual disease; we anticipate that the new criteria will be published in 2016.
To illustrate the importance of using different drugs in the setting of MRD, I’ve developed two hypothetical scenarios. In scenario A, we treat 10 patients with drug A, and as we can see, all 10 achieve partial response, or PR, and they have a 50% reduction of the M spike. We could conclude that there is a 100% overall response rate, because they all achieved PR or better. You can also see that five of these patients passed the complete response, or CR, line so we can conclude in this scenario that drug A delivers a 50% complete response rate. Among the five complete responders, 80% or four are MRD 10⁻³.

In the second scenario, patients are treated with drug B. All 10 patients again achieved PR, so this drug also delivers 100% overall response rate. If you look for patients that have better than PR, you see that 6/10 actually achieved CR as well as MRD 10⁻⁵. They are all 10⁻⁵ negative, and four out of the six, or 67%, were 10⁻⁶ negative.
Now, we’re not supposed to compare studies like this side-by-side, because that could be biased across the studies. But, if you did put these two studies side-by-side, just as a comparison, one way of thinking about it is that the two studies are very similar: as you can see on the right, both studies deliver a 100% overall response rate, and when it comes to complete response, one is 50% and the other one is 60%. You could argue those response rates are very similar, and if you can just use the term MRD negativity in a loose way, you could say that one of them delivered some type of MRD negativity in 80% while the other one was 100% among the CR patients. However, as I just illustrated, you saw how the depth of response really differed in study B, with drug B much better in terms of delivering a deep response. So, I would like to caution you that when you use MRD, you have to be very specific and say how MRD was achieved. If you look across the board at the newer therapies that have been developed, we now have many therapies that can deliver 100% overall response rate.

If you look at the detail of the depth of response, we now have newer therapies that are reaching almost 100% for VGPR (very good partial...
Response Assessment in Myeloma: Is Achieving Response Deeper than CR/sCR Important?  
Presented by Ola Landgren, MD, PhD

The best studies without transplant that we have seen so far deliver more than 60% complete response rates, and some studies have indicated that, if you add transplant, response rates could go even higher. The new monoclonal antibodies may also help in facilitating response rates closer to 100% CR.

**MRD Status Predicts Progression-free Survival and Overall Survival**

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment arms</th>
<th>Test method</th>
<th>Outcomes (MRD-negative versus MRD-positive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paiva et al. (2008)</td>
<td>6 alternating cycles of VBMP + and VBAD, followed by HDT-ASCT (n=577)</td>
<td>4-colour flow cytometry</td>
<td>Median PFS 7.1 months vs 37 months (P&lt;0.001); Median OS not reached vs 89 months (P=0.02)</td>
</tr>
<tr>
<td>Paiva et al. (2011)</td>
<td>6 cycles of VBMP or VBAD (n=102)</td>
<td>4-colour flow cytometry</td>
<td>Median PFS not reached vs 35 months (P=0.02); Median OS not significantly different</td>
</tr>
<tr>
<td>Korthals et al. (2012)</td>
<td>Idarubicin or dexamethasone plus HDT-ASCT (n=53)</td>
<td>ASO-PCR</td>
<td>Median EFS 30 months vs 20 months (P=0.001); Median OS 70 months vs 45 months (P=0.04)</td>
</tr>
<tr>
<td>Rezvani et al. (2013)</td>
<td>CD48 or CD3 plus HDT-ASCT (n=372)</td>
<td>6-colour flow cytometry</td>
<td>Median PFS 28.6 months vs 15.5 months (P&lt;0.001); Median OS 80.6 months vs 59 months (P=0.018)</td>
</tr>
<tr>
<td>Puig et al. (2014)</td>
<td>VBMP or VBAD induction therapy plus HDT-ASCT or 6 cycles of VBMP or VBAD (n=410)</td>
<td>ASO-PCR</td>
<td>Median PFS 54 months vs 27 months (P=0.003); OS not significantly different; 6 cycles of VBMP or VBAD: median PFS not reached vs 31 months (P=0.029); OS not significantly different</td>
</tr>
<tr>
<td>Martinez-Lopez et al. (2014)</td>
<td>VBMP or VBAD induction therapy plus HDT-ASCT or 6 cycles of VBMP or VBAD (n=133)</td>
<td>Next generation VDJ sequencing</td>
<td>Median time to progression 80 months vs 31 months (P&lt;0.0001); Median OS not reached vs 81 months (P=0.02)</td>
</tr>
</tbody>
</table>

Here is a slide from the Spanish study group, and they have done a significant amount of work using flow cytometry to determine MRD. This study looked at patients who were MRD $10^{-5}$ negative, or, in other words, who came in under a cutoff of 1 in 100 cells. As you can see in the blue curve, these patients remained negative, and they had the best progression-free survival. The red line method, shows that MRD is highly correlated with both progression free and overall survival.

Recently, *Nature Reviews* published a review paper looking at all of the studies that have been published comparing the relationship between MRD status and progression free and overall survival. Importantly, in this paper, every study that has been used to determine MRD status, independent of
Response Assessment in Myeloma: Is Achieving Response Deeper than CR/sCR Important? 
Presented by Ola Landgren, MD, PhD

represents patients who were positive, from $10^{-3}$ to $10^{-5}$, ranging between 1 in 1,000 and 1 in 100,000; and then you can see in the black curve, patients who were more than $10^{-3}$ MRD positive, beyond 1 to 1,000 positive. So, clearly, MRD is not black and white; it is a continuous variable.

![MRC Myeloma IX Study: Significantly Longer OS for Each Log MRD Level](image)

<table>
<thead>
<tr>
<th>MRD Level</th>
<th>Median PFS, y</th>
<th>Median OS, y</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-4}$</td>
<td>3.1</td>
<td>Not reached</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>2.7</td>
<td>6.8</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>1.9</td>
<td>5.9</td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>1.7</td>
<td>4</td>
</tr>
</tbody>
</table>

MRD assessed by 6-color flow cytometry

~1 year median OS benefit per MRD level


Similarly, this data is from the British group, in Leeds in the UK. In their MRC Myeloma IX study, they showed that, for each improved level of MRD detection from $10^{-1}$, $10^{-2}$, $10^{-3}$, and $10^{-4}$, there is longer progression-free survival and overall survival. Based on this data, they claim that, for each improved log of MRD detection, you gain a 1-year median overall survival benefit; this means that, if you go from 1 in 1,000 to 1 in 10,000 negative without degree of sensitivity, you gain a year or from 10,000 to 100,000, etc.

We’ve used technologies in our studies. We’ve used flow cytometry that can deliver $10^{-5}$, or 1 in 100,000, and we’ve also used next generation sequencing where we’ve sequenced for VDJ. This has the highest sensitivity and higher reproducibility, and here we show that, if you distinguish patients as negative or positive for progression-free survival with these two technologies, you

![PFS in Newly Diagnosed Myeloma Patients Treated with CRd (N=45)](image)

MRD $10^{-5}$ by flow cytometry

MRD $10^{-6}$ by VDJ sequencing

see how they separate. As you can see on the top, in the patient group that is negative $10^5$, there is unfortunately evidence of progression. When you use VDJ sequencing in the same patient, that patient would move to the positive curve on the lower part of the slide. It is simply a matter of sensitivity. If I haven’t convinced you yet, I’ll say it again: MRD is not arbitrary, it’s a continuous marker. We don’t know if $10^5$ or $10^6$ or higher is what we need to achieve in order to reach a cure. So, for right now, although MRD stands for minimal residual disease, I think we, as clinicians, need to remember that we’re talking about measurable residual disease. I don’t say that in order to try to change the acronym, but I just want you to keep that concept in mind.

Now, I am going to talk about incorporating MRD testing into clinical practice. There are several different technologies that are available for MRD testing. I have already outlined flow cytometry and VDJ sequencing. This data is from the 2015 *Nature Reviews* paper by Dr. Mailankody. The four platforms that

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ASO-PCR</th>
<th>VDJ sequencing</th>
<th>Exome or genome sequencing</th>
<th>Flow cytometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal assay</td>
<td>No (patient-specific primers)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sensitivity*</td>
<td>$1 \times 10^6$</td>
<td>$1 \times 10^5$</td>
<td>$1 \times 10^5$</td>
<td>$1 \times 10^5$ to $10^6$</td>
</tr>
<tr>
<td>Sample source</td>
<td>Bone marrow aspirate</td>
<td>Bone marrow aspirate or peripheral blood</td>
<td>Bone marrow aspirate or peripheral blood</td>
<td>Bone marrow aspirate</td>
</tr>
<tr>
<td>Sample quality assessment</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Sampling error</td>
<td>Likely</td>
<td>Can be overcome by using peripheral blood</td>
<td>Can be overcome by using peripheral blood</td>
<td>Likely</td>
</tr>
<tr>
<td>Clonal evolution</td>
<td>Not detected</td>
<td>Limited detection</td>
<td>Detectable</td>
<td>Not detected</td>
</tr>
<tr>
<td>Turnaround time</td>
<td>Days</td>
<td>1 week</td>
<td>Days to weeks</td>
<td>Hours</td>
</tr>
<tr>
<td>Interobserver variation</td>
<td>Likely</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Substantial</td>
</tr>
<tr>
<td>Clinical benefit associated with MRD-negative status</td>
<td>Improvements in PFS and OS</td>
<td>Improvements in PPS and OS</td>
<td>Improvements in PFS and OS</td>
<td>Improvements in PFS and OS</td>
</tr>
</tbody>
</table>

*Expressed as the minimum cell sample size required for detection of up to 10% is higher than four-colour or dual-colour flow cytometry (1 in 10^5). Abbreviations: ASO, allele-specific oligonucleotide; MRD, minimal residual disease; OS, overall survival; PFS, progression-free survival; VDJ, variable diversity junction.
have been used around the world are listed here from left to right: ASO PCR, VDJ sequencing, exome or genome sequencing, or flow cytometry. I think it is fair to say that the VDJ sequencing and flow cytometry are the two leading platforms. ASO PCR is very tedious and has multiple limitations. As a result, I do not believe ASO PCR will be a significant method of testing for MRD in the future. Exome or genome sequencing has some future potential value in testing, but much more work needs to be done for this type of testing to be commonly used. The strength of VDJ sequencing is that it is more sensitive, and it has very low interobserver variation. The current weakness with VDJ sequencing is that it is not widely available. I do think that new platforms will begin to be available in the coming 12 months, and I think this will probably have a significant impact on the field. The major strength of flow cytometry is that it is commonly available. The downside is that there is significant interobserver variation and antibodies and gating strategy both vary between different hospitals. So, for the moment, I think flow cytometry is the leading platform, but I do foresee that VDJ sequencing has a high likelihood of becoming the standard in the future.

I would like to emphasize that, when we send samples for VDJ sequencing or for flow cytometry, we need to think about the sample itself. When we talk about sensitivity, the sample is obviously one important aspect, but there is a very practical aspect that is perhaps even more important than sensitivity; it is the quality of the sample. We did a survey several years ago in which we asked 30 leading myeloma centers in the United States if they conducted MRD testing for myeloma. As we can see on the right, 11 of those 30 centers answered yes. What was surprising, however, is how the 11 institutions responded when we asked them which specific bone marrow aspirate is submitted for flow cytometry: the first, the second, or the third pool, or if they didn’t know. As you can see, only 3 institutions used the first pool from the aspirate.
Hemodilution in Relation to Order of Bone Marrow Aspirate

![Graph showing hemodilution impact.]

Why is this important? Well, we know that as you keep on going in more pools, you will have more blood coming into the syringe, and hemodilution becomes a major problem. This is a study showing that the first pool has much less hemodilution compared to the second pool, and that subsequent pools have moderate-to-significant hemodilution impact.

So, we need to send samples from the first pool. This is very, very important if you want to have an accurate readout from your laboratory.

**Bone Marrow Aspirates and MRD**

- Sample quality of bone marrow aspirate is a major problem worldwide
- Send sample from **first** pull to the lab for MRD testing
I would like to give you my perspective on future directions in MRD testing in multiple myeloma. I think, in the future, we will have an increased focus on maintained MRD negativity. I believe that we will use modern combination therapy that will be more than three drugs. This could be our current 3-drug combinations, with a biologic drug added. We will reach very deep responses, and I think using MRD $10^{-6}$ will be very reasonable to use as the benchmark. We already have data showing that this level of MRD is associated with both improved progression-free and overall survival across the studies that have used this tool. As a result, this is where I would set the bar.

I’ve highlighted another box in orange following the modern combination therapy, and this box represents potentially transplant, potentially not. I don’t know if we need transplant if we reach $10^{-6}$; this is a question we need to investigate further.

After the orange therapy has been delivered, extended therapy or maintenance therapy will occur for some time, and we don’t really know what the optimal time for this is. So, I think it’s very reasonable to measure MRD status over time, and to ensure that the patient maintains MRD $10^{-6}$ during maintenance, as well as after, if we stop the continued therapy.

So, what I am outlining is a paradigm that uses very effective therapy upfront, either with or without autologous transplant, and with some extended therapy. The goal is to attempt to deliver a curative strategy. The new focus I foresee in the very near future is to deliver therapy and ensure maintenance of MRD $10^{-6}$ negativity. For those patients who do not reach MRD negativity, we need to direct research to dissect mechanisms of MRD positivity and develop both targets and intervention strategies. Also, for those patients that are $10^{-6}$ negative, if we follow these patients over time, some may become MRD positive again. For these patients, I think it will be increasingly important for us to develop new treatment studies to see if we can intervene in the disease before it becomes too dramatic. I think this will be another area for future drug development.
either become MRD $10^6$ negative or $10^6$ positive. Ideally, of course, we’re hoping for $10^6$ negative. For those patients who are $10^6$ negative, we feel that we’re almost at the juncture at which we can probably tell patients that, by combining the additional melphalan and high-dose therapy with stem cells, followed by maintenance, it is entirely reasonable to proceed immediately to maintenance and to collect and keep the cells in the freezer. We need more data to this back up, but I think we’re at the point where we need to address this question.

For patients that still are positive, we think it’s very reasonable to deliver high-dose melphalan with stem cells followed by maintenance, but it doesn’t seem completely unreasonable to do some additional modern therapy before you go there. Alternatively, if you give some additional modern therapy and you can convert the patient into negativity, the patient could then potentially be counseled the same way as those who reached MRD $10^6$ negativity immediately. I think this is a very exciting and provocative approach, and we’re trying to address this by building studies that are investigating these questions.

And as you know, the first paper sequencing myeloma patients came out in 2011 and identified several important genes that were mutated, including NRAS, KRAS, as well as BRAF and an unexpected gene that came up with BRAF, as well as many other genes that were mutated in a variable fashion.
The way I view the future of modern combination therapy, including MRD, is that, by using new drugs, we can induce rapid, deep, and sustained MRD negativity. As a consequence of that, I do think we have been able to increase our focus on quality of life. This is a photo of one of my patients who recently sent me this picture. He was diagnosed at the end of 2015, and after only a few months, he was doing very well and was out skiing. This patient with multiple myeloma had 50% plasma cell infiltration at diagnosis.

To summarize, today, MRD negativity is consistently associated with longer progression-free and overall survival. The two tools that are available are flow cytometry-based MRD assays which are commonly available, and VDJ-based sequencing MRD-based assays, which, while they aren’t as commonly available today, are likely to be in our clinics very soon. In my opinion, we need to develop non-bone marrow biopsy-based MRD $10^6$ technologies. I think they are

**Conclusion**

- Revisions in risk are based in part on genetics
- Newer technologies are needed to more accurately define risk subsets of patients
- As technology evolves, need broad trials to ensure risk is not treatment based but rather disease specific
- R-ISS is the best current and validated tool that incorporates genetics with standard criteria
urgently needed to match the newer therapies that are coming soon to clinic. As we will be able to treat patients with better therapies delivering better responses, we need to have improved tools so that we don’t have to poke the patient in the skeleton and cause pain in order to make sure that we’ve maintained MRD negativity. This is something that we, and a number of other groups, are working on.

I think it’s fascinating to see how this disease has changed in a very short time. A few years ago, overall survival for multiple myeloma was only 2 to 3 years. Now many of our patients will survive for 10, 15, or even 20 years after diagnosis with all of the new treatments, and we’re talking about the potential for developing a curative strategy. It is truly amazing.

Thank you very much for your attention.