Thank you very much. I wish to thank the 16 organizers, particularly John Senior, for inviting me and giving me the honor to speak after this distinguished scientist, after Professor Karmen and before Professor Dufour
My talk today will introduce you to what is normal for transaminase. I think that this has some relevance because, as you know from the FDA definition of DILI, the upper normal limit of liver enzymes, not only for ALT, but also for other liver enzymes, is still important for the definition of disease.
And the upper limit of normal still is used in the even more recent definition of the disease, like these definitions coming from Spain in which they try to renew the composite algorithm of the Hy's Law to predict the outcome of acute liver failure in patients with drug-induced liver injury. This paper appeared last year.
So, the definition of normal still has relevance. In this talk today I would like to introduce some concepts about normality and some considerations about the definition of normal values. And then, next, I will speak to you more in detail about the cutoffs for liver transaminase.

When we talk about normal individuality, we think that, on one side, normal is taken as diagnosed free from a given disease, but on the other side we define diseases as those that are different from normal individuals. And this is an example of circular reasoning.
I would like to draw your attention to the difficulties of defining normality. I would like to start with an idea of Sigmund Freud that, for the first time, more than a century ago, challenged the idea of normality. These are words from his late work, Analysis Terminable and Interminable. He said that every normal person, in fact, is only normal on the average, and this is a statistical concept, and that his ego approximates to that of the psychotic at some point or other, and to a greater or lesser extent. And this means that any definition requires individualization.
Also, a seminal paper about the concept of normality came from E. A. Murphy, published in 1972, and shows you some different definitions of normality from different disciplines. You can see here that in clinical medicine we identify normality as healthiness, but we need to define it from a probabilistic approach. And so, we need statistics.

<table>
<thead>
<tr>
<th>Definition</th>
<th>Discipline</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Probabilistic</td>
<td>Statistics</td>
<td>Gaussian (or other)</td>
</tr>
<tr>
<td>2. Most common in its category</td>
<td>Descriptive sciences</td>
<td>Modal</td>
</tr>
<tr>
<td>3. Common in its category</td>
<td>Descriptive sciences</td>
<td>Habitual</td>
</tr>
<tr>
<td>4. Most suitable to survive</td>
<td>Genetics</td>
<td>Best adapted</td>
</tr>
<tr>
<td>5. Without any disease</td>
<td>Clinical medicine</td>
<td>Healthy</td>
</tr>
<tr>
<td>6. Best in its category</td>
<td>Moral philosophy</td>
<td>Ideal</td>
</tr>
</tbody>
</table>

Mod. From Murphy EA. Perspect Biol Med 1972
So, we use the Gaussian distribution or other distributions. The Gaussian distributions, I remind you, mean that we define as normal all the values that fall within the Gaussian curve.
And to deal with more complex distributions like this one that is, as you can see, skewed on the right. So, it means that there is some contribution from subjects with subclinical disease. It means that we use the 95-percent percentile of the normal population.
So, in any of these definitions, anyway, we have some drawbacks. The most important drawbacks were underlined by Sackett and Haynes several years ago saying that the probability of an individual to be called normal depends on the number of tests we require. So, if we order one test, the probability of testing to be normal is 95 percent, but if we order 20 tests, the probability of testing normal is only 35 percent. It is only 0.6 percent when you order 100 tests. So, this is due to the statistical definition of normality.
The concept of reference values was, then introduced in place of the ambiguous term "normal values." Reference values would apply to any type of reference individuals whether healthy or unhealthy, and assuming that the reference values were properly qualified.


The concept of reference values was then introduced in place of the ambiguous term of "normal values," and the term of "reference values" would apply to any type of reference individuals, whether healthy or unhealthy, assuming that the reference values were properly qualified.
And so, the idea was to change terminology and not define anymore the normal range, but talk, rather, about reference ranges. When these ranges have been calculated in an apparently healthy population, we should use the healthy-related reference ranges, or healthy ranges.
And this also because having just a normal value can be quite a simplistic approach. As you can see here, we can see that the normal limit is the same throughout all the natural history of the disease, while probably the next table of a reference population in different phases of the disease and in different steps, like prevention, screening, and diagnosis, and that affect the survey. And so, we could use different populations like healthy individuals, those with recent complications, and those who have responded or been cured from the disease.
In this regard, I would like to draw your attention to this paper published by The Cochrane Hepato-Biliary Group last year. The senior author was a colleague from my group. As you can see here, probably the slide is not very clear, but we can divide the diagnostic process into phases. The architecture of diagnosis goes in phases, phases that are very similar to the clinical trial phases.
And so, starting from Phase Zero to Phase 4, we have different phases in which we define threshold among individuals at the beginning, but later on, also, defining with ratios on the basis of the outcomes. This process, any diagnostic process, deals with certain cutoffs or uses cutoffs for comparing them to predict different outcomes.
So, the possible approaches are the healthy ranges, the diagnostic thresholds in subjects with the disease. The thresholds with complications, to confirm complications, outcome studies. These are morbidity and mortality among both the healthy subjects and among the subjects with some complications, but, also, the idea of finding individual reference ranges, which is a very complicated task, meaning that the reference subject is the subject himself when he/she was healthy or uncomplicated. These are, of course, very complicated approaches.
So, in any case, if we deal with a continuous variable and we want to predict the risk of the disease, it is better to speak about the different levels of specificity and sensitivity that drive us to define the different thresholds, in other words, challenging the idea itself of normality.
For any distribution, moving the threshold from one level to another means that we decrease the number of false-negatives but we increase the number of false-positives. This can be done when we are approaching a disease. For example, it has particular characteristics, as being easily treatable with a drug, with a new dose of drug, for example, that has very low side effects. So, in this case we can probably tolerate a higher number of false-positive results while, before that, we could not tolerate it.
So, for liver injury and the normal range, we can apply different approaches, as these approaches have been all used to define the cutoffs for alanine aminotransferase. The low-risk population is what I showed you in the distribution among the normal individuals. The mortality definition, predicting mortality for chronic liver disease, differs from the values observed in individuals cured from liver disease.
The first step, the one from low-risk populations, means that we have to exclude the patients with a high risk of liver disease, including those with hepatitis B, hepatitis C, alcoholic liver disease, and fatty liver disease. So, even in this case, having different ratios for different clinical outcomes, it would probably be more appropriate.
When we started work in the early 2000s, we had a laboratory marker, serum ALT, that was widely used, but had very poor value for diagnosing liver disease. Technical standardization was lacking. It was improperly used as a marker of fibrosis, and advanced liver disease. And the reference values were not updated for a long time.

**Different approaches to define new thresholds for alanine aminotransferase**

1. **“Low risk population”**: values observed in subjects at low risk for liver disease (exclude those at high risk including HBV, HCV, alcoholic, NAFLD, etc)

2. **“Mortality definition”**: the values predicting mortality for chronic liver disease

3. **“Therapeutic definition”**: values observed in individuals cured from liver disease (eg, SVR for hepatitis C)
So, we found that reference populations proposed by laboratories as "normal" contain a substantial proportion of people with subclinical, undiagnosed liver disease. Our idea was to take a symptom-free population, a blood donor in this case, to exclude viral hepatitis carriers and to exclude also those at risk for steatosis, those with high lipids, high glucose levels, overweight individuals, and heavy drinkers, and calculate their reference interval.
In this way, we were able to show that we could lower the upper limit of "normal" from 40 units for males and 30 units for females that were average at that time, to 30 units for males and 19 units for females. This work was published, as I said, in the Annals of Internal Medicine.
Another approach that followed this was proposed by a Korean group that used receiver operating curves for ALT to identify people at risk from death of liver disease, a mortality approach. And this approach, surprisingly, gave similar results. So, as you can see here, 30 units per liter was identified on the basis of survival data.
We also tried to apply the definition, calculating the distributions in people that had been cured from liver disease from hepatitis C. Again, in different labs we had a 25-percent reduction, meaning that from 40 we went to 30. So, this summarizes these three approaches published between 2002 and 2006.
More recently, another group calculated from living donors, living liver donors, the distribution of ALT, taking into account not only the metabolic parameters, but also the histological data from these donors. They found that among the 600 individuals selected in the material published in our Annals of Internal Medicine paper, the healthy values were 33 units for men and 25 units for women.

<table>
<thead>
<tr>
<th>Approach</th>
<th>New thresholds (U/L)</th>
<th>Reduction from to historical thresholds</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk population</td>
<td>30 in males, 19 in females</td>
<td>25-35%</td>
<td>Prati, Ann Int Med 2002</td>
</tr>
<tr>
<td>Mortality</td>
<td>30 in males</td>
<td>25%</td>
<td>Kim, BMJ 2004</td>
</tr>
<tr>
<td>Therapeutic efficacy</td>
<td>-- (Expressed as xULN)</td>
<td>27%</td>
<td>Prati, J Hepatol 2006</td>
</tr>
</tbody>
</table>
And this is a slide from a recent review that shows you that almost every paper published so far shows higher levels in males, lower in females, higher levels in those whose metabolic features are not considered and in a group of those who have been cleaned up by the presence of patients at risk for liver disease. Also, this work has been done in children, almost five years ago. This is an American study showing that the 95 percentile in children with healthy weight, was 25 for boys and 22 for girls. If you wanted to review in depth all this data, you can read it. It is hard because it has been published by another Italian group and shows in very great detail almost all the studies that have been done during the last decade.

“Among the 665 individuals (346 men and 319 women) selected according Prati criteria, healthy ALT values were **33 IU/L for men** and **25 IU/L** for women.”
So, is everything definitely set now? I would say no, because the upper reference limits in different labs are still variable, and expressing the results that are done as upper normal limits does not solve the problem. We do not know what might be the impact of the introducing standardized information on chemistry based on the reference ranges. And there has been no clinical validation of these ranges so far. And I would also say that most clinicians don’t even know that this has been changed over time. So, I think that this would be the topic of Professor Dufour after this presentation.

In addition, we almost thought up until now about ALT, but the upper limit cutoffs for SD phosphatase and bilirubin, among all the others, have not been so far very well-identified, not with the same analytical attention. The cutoff varies for drug-induced liver injury. They are not based on prospective studies. This is a major problem because this implies that it would be difficult to define upper thresholds based on outcomes. And so, of course, we need new studies.
Just to summarize the message that I want to leave you with today, several studies conducted during the last decade suggest that the healthy ranges for ALT can provisionally be set around 30 U/L in males, and 20-25 U/L in females.

However, the simplistic approach of a universal cutoff value valid for all clinical situations should be abandoned. Results should be interpreted flexibly, taking into account the scope of the test and the patient’s characteristics (gender, age, clinical history, specific risk factors, symptoms, signs, etc).

Setting appropriate cutoffs needs close cooperation between hepatologists, clinical pathologists, statisticians, regulatory institutions.

Much work is still needed.

Messages

- Several studies conducted during the last decade suggest that the healthy ranges for ALT can be provisionally set around 30 U/L in males, and 20-25 U/L in females.
- However, the simplistic approach of a universal cutoff value valid for all clinical situations should be abandoned. Results should be interpreted flexibly, taking into account the scope of the test and the patient’s characteristics (gender, age, clinical history, specific risk factors, symptoms, signs, etc).
- Setting appropriate cutoffs needs close cooperation between hepatologists, clinical pathologists, statisticians, regulatory institutions.
- Much work is still needed.

But we have also to think that a simplistic approach of using a universal cutoff for all clinical situations should be abandoned. Results should be interpreted more flexibly, taking into account the scope of the tests and especially the patients' characteristics in terms of gender, age, history, and also other risk factors.

In addition, the definition of the ALT ranges, as we have seen, is only the first step in evidence-based diagnostic research, and a cutoff should be ideally identified on the basis of prospective studies based on clinical outcomes.

Setting the appropriate cutoffs needs close cooperation between the pathologists, clinical pathologists, statisticians, and the regulatory institutions. This is very important because it has not always been the case. So, sometimes hepatologists create their own thresholds and clinical pathologists may establish another. It is important to have multidisciplinary work. Much work is still needed.

To finish with an idea that comes from a very important but non-medical book published almost a century ago, let me quote from a book that is called The Confessions of Zeno, by Italo Svevo. The book focused on the relationship between
health and disease. Using as an example thyroid disease, that fits well also for liver disease, it says "All organisms extend along a line. At one end is Basedow's disease, which implies the generous, mad consumption of vital force at a precipitous pace, the pounding of an uncurbed heart. At the other end are the organisms depressed through organic avarice, destined to die of a disease that would appear to be exhaustion, but which is, on the contrary, sloth. The golden mean between the two diseases is found in the center and is improperly defined as health, which is only a way station. In the middle are those who have either incipient goiter or incipient myxedema, and along the entire line, in all mankind, absolute health is missing."

So, thank you very much. (Applause.)