Episode 35: A Comprehensive Discussion of Reference Intervals

Andrea Ott-Vasconi:

Hello, I'm Andrea Ott-Vasconi and welcome to Quidel Ortho Science BYTES. In this episode, we'll be talking about reference intervals, a fundamental tool used by clinicians to interpret laboratory reports. The establishment of reference intervals is complex and there are many factors that need to be considered when using reference intervals as a benchmark for comparing lab results. With me today to discuss this important topic is Dr. Lindy Crimmins, Manager of Medical Affairs at Quidel Ortho. Dr. Crimmins holds a bachelor degree in chemistry from the University of Wisconsin-Madison and a medical degree from the University of Illinois. Her clinical experience spanned from emergency medicine to critical care to primary care. Her time in industry focused on clinical chemistry and point of care technical support prior to her current role in medical affairs. Thank you, Dr. Crimmins for joining me today.

Lindy Crimmins:

Thank you, Andrea. It's good to be here.

Andrea Ott-Vasconi:

Let's start with some definitions. What are reference intervals? How are they defined and determined per CLSI guidelines?

Lindy Crimmins:

A reference interval is a powerful and useful tool that takes a single test result and makes it meaningful, determining whether it's too low, too high, or as expected. A reference interval is the range of values which are typical in a healthy population for a specific test. This healthy population is also called a healthy reference population, hence the term reference interval. Reference intervals are also called expected values, which may be a more intuitive term. They can be thought of as what do I expect as a result for this particular patient if she is healthy?

Statistically and classically, a reference interval is calculated to include 95 percent of apparently healthy individuals. Say we measure 200 apparently healthy people. To simplify the underlying statistics, imagine the values from this healthy population are plotted from lowest to highest on a number line. The 5 percent of healthy individuals not included in the reference interval are equally divided above and below the reference interval. For 200 people, this means 10 are excluded. The lower end of the reference interval range will exclude the five lowest values. The upper end of the reference interval range will exclude the five highest values, leaving 95 percent of people or 190 of our 200 apparently healthy included to define the reference interval. The Clinical and Laboratory Standards Institute, or CLSI, in the 2010 EP-28 document, defining, establishing and verifying reference intervals in the clinical laboratory, states the best means to establish a reference interval is to collect 120 samples from qualified reference individuals for analysis. I used 200 in the preceding example to make the math easier to follow.

Andrea Ott-Vasconi:

How have reference intervals evolved over time and when are partitioned reference intervals used?

Lindy Crimmins:

This is a test-specific topic. Reference intervals get established for every specific test, so let's use creatinine for example. Studies establishing a reference interval extend back to the 1940s for blood creatinine levels and even further back to 1905 for urine creatinine levels. These initial studies were done in males so that the creatinine reference interval in use up through the 1970s for adults was a single range applied to both male and female patients. Females began being evaluated in the literature for their own creatinine reference interval as early as 1953. Despite this early start, the single reference interval for creatinine based on males was still in use through the late 1970s. Throughout the late 1970s, there were several publications which incorporated sex as a factor in calculating renal function based on creatinine, leading up to today, where we have separate creatinine reference intervals for males and females.

Basically partitioned reference intervals get used when there's a significant difference between different groups of people. Some tests don't have partitioned reference intervals. Glucose, no matter how old you are, there's an expected reference interval for a glucose, but other values that are primarily dependent on differences between people like body weight will be different for adults and children, males and females, that sort of thing. A lot of different physiological factors influence what the reference interval is for specific tests.

Age has multiple components from pediatrics through puberty to adulthood and geriatrics. While children have their own separate reference intervals for adults for most nonelectrolyte tests, they show minimal differences between boys and girls. The sex-specific differences in tests become apparent at puberty. The differences for analytes, which are different between the sexes increases as puberty progresses to reach the final adult differences in sex-specific reference intervals. There are other factors that also impact reference intervals, such as hormonal status in pre-menopause, menopause, pregnancy, and even the time that a sample is collected. Many tests are impacted by fasting, meaning not eating for eight to 12 hours before the sample is obtained. Other tests are influenced by the time of sample collection due to variation with circadian rhythm. For example, cortisol reference intervals have one range before 10 a.m. and a different range after 5 p.m., but no real significant sex-specific differences.

Andrea Ott-Vasconi:

The concepts of reference intervals and clinical decision limits are sometimes confused. How do reference intervals differ from clinical decision limits?

Lindy Crimmins:

The International Federation of Clinical Chemistry’s committee on reference intervals and decision limits works on education around this, but basically I think of a reference interval as the range of values expected from a fully healthy person, while a clinical decision limit is the value beyond which disease

can be definitively diagnosed. I find this easiest to explain using hemoglobin A1C, the primary biomarker for diabetes. The reference interval for a person without diabetes is less than 5.7 percent. The clinical decision limit for using a hemoglobin A1C to diagnose diabetes is 6.5 percent. That leaves us with the range of 5.7 percent to 6.4 percent. This range is called prediabetes. Many assays have this sort of intermediate zone or area of uncertainty, also called a gray zone.

A value in the gray zone can be a warning sign that a person is on the pathway to disease or that that's where their body is even in a healthy state, but if they're on the pathway to disease, they have not yet arrived at the point where a pharmacologic treatment will make a consistent difference in their long-term clinical outcome. When results are in the gray zone, for many diseases, lifestyle changes are usually recommended, such as improving diet and improving physical activity. Additionally, a person with a result in the gray zone may get repeat testing more often than someone with a value within the reference zone. In research studies, clinical decision limits have been established as the point beyond which adverse clinical outcomes occur. For the A1C example, the studies to prove the association between elevated A1C and specific adverse outcomes related to diabetes took 20 years to finish because it takes that long for the disease to evolve and actually make significant changes in the body. These sorts of long-term studies have shown that values in the intermediate or gray zone were not consistently associated with adverse outcomes.

Andrea Ott-Vasconi:

You gave a great example of how hemoglobin A1C results are used. Let's talk more about the application of reference intervals. Why are they so important and how are they used in clinical practice?

Lindy Crimmins:

Yes, reference intervals are used to determine if nothing else needs to be done or if additional testing needs to be done, whether that be more blood testing, more testing of a different body's fluid type, any kind of radiological testing or even physical testing like pulmonary function tests or something where you test the actual function of the body. Values outside of a reference interval would also be used to determine if treatment needs to be continued or modified, or maybe a conversation about if the patient is able to comply with the treatment that's already been recommended, but because reference intervals influence clinical practice and changes recommended for specific patients, it's really important that the reference population be selected to match the population that's being tested.

There are two primary assumptions about a reference population. One, that they are healthy, and two, that they match and represent the population in which clinical testing will be performed. This importance of population matching leads to all regulatory bodies and diagnostics referring to the concept, it is recommended that each laboratory establish its own expected values for the population it serves. This concept appears in the documentation for all diagnostics.

Andrea Ott-Vasconi:

What factors need to be considered when interpreting lab test results and comparing the results to reference intervals.

Lindy Crimmins:

To review some and add some to what I said earlier, fasting, food intake, like the kind of food and the timing of that food, diurnal variation or changes throughout the day, seasonal variation, sex-specific variation, hormone replacement therapy, and what stage of life you're in. For fasting, the time when you last ate will impact your results for many tests. Glucose, for example, has no reference interval for someone who just ate. Glucose only has a reference interval for someone who has been fasting. However, glucose after eating is sometimes measured in a very controlled manner, such as an oral glucose tolerance test.

Oral glucose tolerance testing is done in mid-pregnancy. In this test, a specific amount of glucose is consumed. Curiously, it's a drink that tastes horrible even though it's full of sugar. Glucose measurements are then taken each hour for a total of three hours after drinking the glucose. Each timeframe has a unique specific reference interval to identify if the patient is at risk for or already has diabetes associated with pregnancy or gestational diabetes. For this specific oral glucose tolerance test, fasting is also required before beginning the test to ensure all patients are starting from the same baseline.

As for food intake, other tests impacted by diet are lipid panels or cholesterol tests. When you eat fat, it's absorbed fairly quickly after the meal and travels through the blood to reach the liver, and then the liver further processes that fat. If a blood sample is taken while this fat is traveling through the bloodstream soon after a meal, that sample will contain more fat in multiple forms. Ultimately this makes the sample look cloudy. You can actually see the fat in the sample after it gets spun down and the solid liquid parts are separated out. Therefore, a sample taken during this after meal phase won't reflect the average steady state concentration that all of our lipid tests reference intervals are based on, so it's really important that you either fast for tests where it's recommended or at least let your doctor know, oh, I wasn't fasting for that. I forgot, or you know what? I did grab a coffee on the way and it had milk in it. Does that count as not fasting?

Diurnal variation is variation throughout the day. This can be related to sleep wake cycles and just natural variations throughout the day. Then seasonal variation is sort of a newer concept. There is research looking into this, particularly focusing on populations who live very far north, where in the summer they have constant daylight, the sun never sets, and in the winter they have constant darkness, the sun never rises. Some of these studies have shown differences in thyroid hormone levels compared to the sunny summertime and the dark wintertime. Also related to light exposure is vitamin D, whose formation is dependent on sunlight, and this will vary between seasons and between places where you live, depending on how much time you're able to spend outdoors and how much skin is exposed when you're outdoors.

Additionally, sex-specific reference intervals and hormones or hormone replacement therapy all influence some tests. Not every test, but some tests. Because most of the sex-specific differences become apparent at puberty, it is assumed that hormones are responsible for these differences. This assumption is supported by the fact that female reference intervals change again after we go through menopause. Then the influence of hormones on reference intervals for some tests is also seen on patients taking hormones as therapy, whether for contraception, for transgender or for management of the symptoms of menopause.

Then finally, one of my favorite things from training as a physician is that children are not little adults. Their values for certain tests are very different. Children, for example, get to grow. There are two main growth spurts during childhood in the toddler phase between ages two and three and in the adolescent phase. In both these phases, bones grow rapidly. Therefore, the markers of bone growth are significantly different for children in these growth spurt ages than children not in growth spurts. The main markers that get influenced by this are alkaline phosphatase, calcium and phosphate. For alkaline phosphatase, the pediatric reference interval is higher than the adult interval, which is the opposite for how most age influenced reference intervals change from childhood into adulthood. Typically, adult values are higher than children's.

Andrea Ott-Vasconi:

Let's dive a little deeper into this. Tell us more about the physiological changes occurring during these life stages that have led to the partitioning of reference intervals.

Lindy Crimmins:

Yes, so there have been studies done where they've been able to take samples of children at every age from birth up through 18. When you plot these out for a large population, you find that the hemoglobin, creatinine, albumin, ferritin, some liver enzymes have equivalent reference intervals for boys and girls during childhood. Even when you plot them and try to look for a difference, there really isn't a difference between the boys' and girls' values when they're little, but when these kids reach puberty, the girls' values remain similar to childhood values, while boys' values increase starting around age 12 to reach adult male values. Conversely, for HDL cholesterol and total cholesterol, girls' values still remain similar, but boys' values actually drop around puberty. Then as far as hormonal changes over time, when women reach menopause and ovaries stop producing hormone and it's a slow fade away; the sex-specific differences for women at that age range start getting smaller. Women's values get closer to men's values as our hormones fade away.

Andrea Ott-Vasconi:

What are some routine lab tests that have different reference intervals for men and women?

Lindy Crimmins:

Creatinine, the difference in creatinine is primarily due to males having larger muscle mass on average than females. Related to this difference in body size is many of the markers of bone health men not only have larger muscle mass puts more tension on the bones, makes your bones grow larger, and men on average are taller, so that larger body mass means higher bone markers. Then also anemia. Values are the same for young children, but then adult male hemoglobin and iron values are higher and they raise during puberty. In comparison, adult female values for hemoglobin and iron are a little bit closer to childhood values.

Andrea Ott-Vasconi:

You mentioned bone lab tests. Let's get into the specifics of the reference interval partitions for a test

referred to as NTx, starting with what is NTx?

Lindy Crimmins:

So NTx stands for cross-linked N-telopeptides of type one collagen. This test is measured in human urine, so it's completely noninvasive. It is an indicator of human bone resorption. In a normal adult who is finished growing, bone is continuously being remodeled through a balance between the cells that resorb bone and the cells that form bone. The primary component of bone is type one collagen, and when resorption and formation are in balance, your N-telopeptide level will be lower. As a normal part of aging, bone resorption increases. When the bone resorption significantly outpaces bone formation, there is a risk for weakened bones or osteoporosis. In individuals who have osteoporosis, their N-telopeptide or NTx levels will be higher. The reference intervals for premenopausal females are closer to the reference intervals for males. Then once hormone production falls off, postmenopausal female NTx values get higher. Bone resorption increases when the hormones fade away.

So how the NTx gets used to monitor response to treatment for osteoporosis, it's an alternative to getting a DEXA scan. The major class of drugs for osteoporosis are bisphosphonates, which slow bone resorption. If bone resorption is decreasing, then your NTx value will decrease. Measuring NTx can identify poor response to treatment with bisphosphonates. NTx can also be used before starting treatment to predict which postmenopausal women will have a good skeletal response to hormone replacement therapy. This is just a simple urine test that's an alternative to a DEXA scan. Again, this test, if you take it at a certain time of day, when you do the follow-up test the next year and the year after, you want to keep that time of day consistent because how much NTx is in your urine will change throughout the day, so you want to keep it consistent so that the values can be comparable.

Andrea Ott-Vasconi:

Thank you for providing us with a great overview of the past and present state of reference intervals for lab testing. Let's conclude with discussing the future. What concepts are influencing how reference intervals are determined?

Lindy Crimmins:

One of the biggest new resources we have is data science or use of big data aggregating results from all the routine testing done. The traditional manner is a direct reference interval, which requires the potentially challenging recruitment of 120 people in each partition and then getting blood results from all of them, enough blood to run it on every single test in the lab. There's many, many challenges to calculating a direct reference interval and making it match the population in which clinical testing is done. Using big data, we can do instead an indirect reference interval. This will use existing data within the laboratory information system on patients whose full medical records are available to prove that the patients are healthy. This does require more complex statistics, but it enables a laboratory to more quickly and efficiently establish assay-specific and lab-specific and population specific reference intervals, making things a lot more aligned and matching the population that's being tested to the population being referred to.

Additional concepts influencing reference intervals are continuous, common and personalized.

Continuous is best applied to age partitions. Rather than assuming that a child's expected value will abruptly change on their 10th birthday, a continuous reference interval uses an equation to account for the true age of the child. A common reference interval requires a reference method that all assays are traceable to and a multi-center reference interval study has been done to identify the reference interval. This is a significant undertaking to get every test from every manufacturer on every available platform to have matching results is quite an accomplishment, but it has been done for hemoglobin A1c, meaning that a 6 percent from California will be the same and comparable to a 6 percent in New Jersey, a 6 percent in New York. The International Federation of Clinical Chemistry has committees working on this to try to standardize every analyte once the technology for testing enables us to do so, and this is a test-by-test undertaking.

Finally, a personalized reference interval is akin to getting a bespoke suit tailored to yourself, rather than buying something off the rack. A personalized reference interval is specific to one person and will be a much smaller range than the reference interval shared by a whole partition or the whole entire population. The idea of a personalized reference interval is to enable earlier detection of potential disease by identifying more subtle changes that have clinical meaning. Then finally, we have the impact of hormonal therapies. There is research being done to establish reference intervals for individuals who are transgender and who are on stable hormone therapy. This is not something that was being looked at back in the ‘50s and ‘70s when differences between simple sex -specific hormone values were being established. This is a burgeoning area of new research.

Andrea Ott-Vasconi:

What are some take home messages you'd like to give our listeners?

Lindy Crimmins:

I think the most important thing to do is to get your labs drawn every year, whatever your primary care physician recommends, and then look at your results. I think most electronic medical records have some sort of online patient portal or a smartphone app or a way for the patient to see what their numbers are. If you have access to your numbers, you can see if they've changed year over year. If you see something that you're not sure about or something you have a question on, then discuss it with your physician.

A lot of times when we're younger or when we don't have any medications or diseases, we go in and we don't have questions, so it's something to start a dialogue with and to learn more about if you're curious. You don't have any control over the reference intervals used by the lab or if the lab where your test gets run, what reference intervals they're using, but if you get your labs taken every year, that's one step in the direction of trying to establish your own personalized reference interval. What was normal for you when you're 20, it can help follow how things change for you as your life moves on.

Andrea Ott-Vasconi:

Thank you, Dr. Crimmins for this very insightful and informative discussion today on the important topic of reference intervals. I hope everyone enjoyed this podcast episode. Make sure to review sections within the podcast description with suggested reading materials and links to learn more. Based on

today's podcast, I'll leave you with our pop quiz. Which lab tests are influenced by physiological changes that occur across our life journey? You can go back and listen again if you'd like some more details. Thank you so much for listening today. Please subscribe to QuidelOrtho Science Bytes, our monthly podcast, brought to you by QuidelOrtho Corporation, where we are transforming the power of diagnostics into a healthier future for all. Take care.

References and Reading Materials:

1. Bjornsson TD. Use of serum creatinine concentrations to determine renal function. *Clin Pharmacokinet*. 1979;4(3):200-222. doi:10.2165/00003088-197904030-00003

2. Sikaris KA. Physiology and its importance for reference intervals. *Clin Biochem Rev*. 2014;35(1):3-14.

3. Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory, 3rd Edition. Accessed May 3, 2023. https://clsi.org/standards/products/method-evaluation/documents/ep28/

4. Ozarda Y, Sikaris K, Streichert T, Macri J; IFCC Committee on Reference intervals and Decision Limits (C-RIDL). Distinguishing reference intervals and clinical decision limits - A review by the IFCC Committee on Reference Intervals and Decision Limits. *Crit Rev Clin Lab Sci*. 2018;55(6):420-431. doi:10.1080/10408363.2018.1482256

5. Shoback D, Rosen CJ, Black DM, Cheung AM, Murad MH, Eastell R. Pharmacological Management of Osteoporosis in Postmenopausal Women: An Endocrine Society Guideline Update. *J Clin Endocrinol Metab*. 2020;105(3):dgaa048. doi:10.1210/clinem/dgaa048

6. Bunch, D. Modern Reference Intervals When developing reference intervals, clinical laboratories must consider what data sources and statistical methods to use, among other factors. *Clinical Laboratory News*. July 1, 2021. Accessed May 4, 2023. https://www.aacc.org/cln/articles/2021/july/modern-reference-intervals