



**Universidad**  
Zaragoza

Trabajo Fin de Grado

**LOS INTERVALOS DE REFERENCIA BIOLÓGICOS.**

**BIOLOGICAL REFERENCE INTERVALS.**

**Autor:**

**Antonio Muñoz Anadón.**

**Tutor:**

**Dr. Enrique Sánchez Óriz.**

Facultad de Medicina de Zaragoza.  
Curso académico 2016-2017.

## RESUMEN:

Los intervalos de referencia biológicos son la herramienta más utilizada para la interpretación de los resultados de laboratorio y pueden condicionar decisiones clínicas. Es preciso conocer los principios en los que se basan para hacer un uso adecuado de ellos en la práctica clínica habitual.

Se calculan a partir individuos sanos seleccionados procedentes de una población concreta, con distintos criterios. Es importante que se expliquen con el detalle necesario estos criterios, así como con todos los pasos llevados a cabo para obtener los intervalos de referencia.

En este trabajo hemos pretendido realizar una revisión de las distintas etapas necesarias para la obtención de unos intervalos de referencia, incluyendo la selección de los individuos sanos, la preparación preanalítica, distintos controles de calidad analítica, los procedimientos estadísticos aplicados o la presentación de los datos.

En conjunto es un proceso costoso y que requiere tiempo, por lo que muchos laboratorios no pueden permitírselo. Una posible solución es la de compartir datos para determinar unos intervalos de referencia comunes.

Algunos autores proponen la armonización del proceso global de una prueba y la estandarización de la medida como respuestas a la variabilidad que existe entre distintos laboratorios. Suponen una oportunidad para lograr hacer comparables resultados obtenidos en centros diferentes.

Por último se exponen la situación actual y las perspectivas futuras de los intervalos de referencia.

Palabras clave: Laboratorio, interpretación, fase preanalítica, ayuno, armonización, estandarización, intervalos de referencia comunes.

Biological reference intervals are the most widely used tool in the interpretation of laboratory results and may determine clinical decisions. A basic knowledge of the principles in which they are based is necessary to achieve a proper use in daily clinical practice.

They are calculated from healthy individuals belonging to a specific population, with different health criteria. It is important to provide enough information about the decisions taken along the whole process of determining the reference intervals.

In this essay we have aimed to perform a review of the different phases of this process, including healthy individual selection, preanalytical considerations, quality controls, statistic procedures applied or data presentation.

All in all, determining RI is an expensive and highly time-demanding task. Therefore many laboratories can not afford to follow strictly the recommendations. A possible solution can be to share information between laboratories to determine common reference intervals via multicenter collaborative studies.

Some authors propose the harmonization of the total testing process in laboratory and the standardization of measurements as the answer for reducing interlaboratory variation. Those initiatives are a chance to make results more comparable and consequently easier to share.

Finally, we deal with the present situation and the future perspectives of reference intervals.

Keywords: laboratory, interpretation, preanalytical phase, fasting, harmonization, standardization, common reference intervals.

## ÍNDICE.

<b>1.- INTRODUCCIÓN.</b>	<b>5</b>
1.1.- Evolución de la teoría de los valores de referencia.....	6
<b>2.- JUSTIFICACIÓN Y OBJETIVOS.</b>	<b>7</b>
<b>3.- CONCEPTOS BÁSICOS.</b>	<b>8</b>
3.1.- Intervalos de referencia.....	8
3.2.- Significado de los intervalos de referencia.....	9
3.3.- Fuentes de variación en el laboratorio.....	10
3.4.- Tipos de variación biológica.....	10
3.5.- Índice de individualidad.....	11
3.6.- Límites de decisión clínica.....	11
<b>4.- SELECCIÓN DE LA MUESTRA DE REFERENCIA.</b>	<b>13</b>
4.1.- Definición de salud. Criterios de inclusión y exclusión.....	13
4.2.- Otros aspectos a tener en consideración.....	14
4.3.- Ejemplo de nuestro entorno.....	15
4.4.- El supuesto estudio Miss Manhattan.....	15
4.5.- La selección con criterios <i>a priori</i> y <i>a posteriori</i> .....	16
4.6.- Valores de referencia indirectos. El estudio REALAB.....	16
<b>5.- CONSIDERACIONES PREANALÍTICAS: PREPARACIÓN DE LOS INDIVIDUOS Y EJECUCIÓN DE LA RECOGIDA DE SANGRE.</b>	<b>18</b>
5.1.- Preparación del sujeto.....	19
5.2.- Recogida de la muestra.....	21
5.3.- Procesamiento de la muestra.....	23
<b>6.- ASPECTOS ANALÍTICOS: CONTROL DE LA VARIACIÓN ANALÍTICA.</b>	<b>24</b>
6.1.- Definiciones.....	24
6.2.- Control interno.....	24
6.3.- Control externo.....	25
<b>7.- PROCEDIMIENTOS ESTADÍSTICOS.</b>	<b>26</b>
7.1.- Definiciones teóricas.....	26
7.2.- Métodos estadísticos.....	26
7.3.- Formas de estimar los percentiles con muestras más pequeñas.....	28
7.4.- Tamaño de la muestra y error estándar de los límites de referencia.....	28
7.5.- Criterios de estratificación o partición.....	29

7.6.- Detección y exclusión de valores anormales.....	31
<b>8.- PRESENTACIÓN DE LOS VALORES OBSERVADOS EN RELACIÓN A LOS VALORES DE REFERENCIA.</b> .....	<b>32</b>
8.1.- Valores de referencia.....	32
8.2.- Transformación de los valores.....	32
<b>9.-ARMONIZACIÓN Y ESTANDARIZACIÓN. PROCESO GLOBAL DE UNA PRUEBA.</b> .....	<b>33</b>
9.1.- Proceso global de una prueba.....	33
9.2.- Armonización.....	33
9.3.- Avances en la armonización.....	35
<b>10.-INTERVALOS DE REFERENCIA COMUNES.</b> .....	<b>36</b>
10.1.- Procedimiento para establecer intervalos de referencia comunes.....	37
10.2.- Adopción de intervalos de referencia comunes.....	37
<b>11.- SITUACIÓN ACTUAL Y PERSPECTIVAS DE LOS VALORES DE REFERENCIA.</b> .....	<b>39</b>
11.1.- Normativa y situación actual.....	39
11.2.- ¿De qué forma puede determinar sus intervalos de referencia un laboratorio?..	41
11.3.- Retos y margen de mejora en el futuro.....	42
<b>12.- CONCLUSIONES.</b> .....	<b>44</b>
<b>13.- AGRADECIMIENTOS.</b> .....	<b>45</b>
<b>14.- BIBLIOGRAFÍA.</b> .....	<b>46</b>

## 1.- INTRODUCCIÓN.

Los intervalos de referencia son un importante factor en la toma de decisiones clínicas, su desarrollo es paralelo al de los test de laboratorio. Influyen en la interpretación de los resultados. En un principio se denominaban valores normales.

Cuando se ha hecho un análisis de sangre a un paciente, la primera pregunta que se plantea es: '¿Es normal el resultado?' [1]

Es difícil definir de forma práctica la normalidad, por la cantidad de significados que tiene este término. Estadísticamente puede ser el valor que con más frecuencia se presenta en una clase o el más representativo. Pero esto en medicina tiene un valor limitado. Médicamente interesa definir la normalidad como el valor más adecuado para la supervivencia, o por otra parte que no produce daño. Llega incluso a considerarse como un ideal, una ausencia de factores de riesgo que no se produce nunca. [2]

En todo caso, lo que interesa al paciente que se ha hecho la prueba es si está sano, enfermo, o en situación de riesgo de alguna enfermedad. En algunos casos el criterio del médico y del laboratorio no concordarán, ya que manejan informaciones diferentes. Muchas veces el laboratorio no conoce detalladamente la razón que ha motivado la analítica o la situación clínica del paciente.

La analítica puede haber sido pedida por varias razones:

- Chequeo de salud o screening.
- Valoración de un factor de riesgo (probabilidad de tener una enfermedad en el futuro).
- Diagnóstico de una enfermedad.
- Manejo de una patología, tratamiento por objetivos.
- Seguimiento.

Para valorar los resultados, habrá que compararlos. Según cuál sea el objetivo, se establecerá la comparación con referencias distintas. Pueden usarse:

- Los resultados de un grupo de individuos supuestamente sanos para confrontarlos con los obtenidos. En eso consisten los intervalos de referencia, objeto de este trabajo.
- También se podrá recurrir a los límites de decisión, valores establecidos a partir de nueva información científica y por tanto cambiantes sobre en qué momento es necesario tomar una determinada decisión diagnóstica o terapéutica.
- Por último, resultados previos del propio paciente (valor de referencia del cambio) para el seguimiento.

Cada una de estas herramientas tendrá valor en función de la finalidad de la analítica. Por ejemplo, en el control de una enfermedad crónica establecida, los intervalos de referencia no nos serán útiles, ya que es muy probable que los valores del paciente excedan los límites de referencia, que se establecen con individuos sanos. Probablemente podremos recurrir a resultados previos de otros análisis del mismo paciente para confrontarlos.

## 1.1.- EVOLUCIÓN DE LA TEORÍA DE LOS VALORES DE REFERENCIA.

Entre los años sesenta y ochenta se asistió a una automatización triunfalista en la medicina de laboratorio. Conseguián producir una gran cantidad de datos con gran nivel de calidad. Consideraban la variabilidad biológica y analítica para evaluar sus posibles implicaciones médicas y fisiológicas. Se evidenció que era importante dominar y entender la variabilidad biológica de todos los constituyentes del plasma en personas sanas. Ya en 1960 un pediatra se planteaba cuestiones sobre cómo definir los valores normales y cuál podía ser su importancia en distinguir salud y enfermedad. [3]

El concepto de intervalos de referencia fue desarrollado por Gräsbeck y Saris a finales de los años sesenta y presentado en un congreso en Helsinki en 1969. El interés de Gräsbeck comenzó porque desde pequeño había aprendido que las variables biológicas se distribuían siguiendo la campana de Gauss, pero al participar en un estudio poblacional sobre la vitamina B<sub>12</sub> en suero esto no sucedía así. Era necesario tomar el logaritmo de los valores para que estos se ajustasen a una distribución normal. Cuando preguntó al estadístico por qué, solo le pudo responder que era lo que se hacía habitualmente. Hasta entonces se hablaba de valores normales, sin una clara definición, quizá haciéndolos análogos a los controles de las ciencias experimentales. El tema despertó gran interés. [4]

Con el tiempo se eligió a un panel de expertos para que trabajara en la teoría de los valores de referencia. El trabajo que desarrollaron fue lento, lo que les mereció críticas. Fundamentalmente esto era debido a desacuerdos en los significados precisos de valor de referencia y salud, pero también en detalles a la hora de recoger las muestras. Algunos eran partidarios de criterios rígidos, mientras que otros abogaban por ser más laxos. Finalmente se llegó a una recomendación preliminar en 1978, lo que supuso un hito importante por ser el primer documento en el que llegaron a un acuerdo común. [5]

En 1986 la Federación Internacional de Química Clínica (IFCC) a través de su panel de expertos en la teoría de los valores de referencia emitió una serie de recomendaciones que dividían en seis partes: [6-11]

- Parte 1. El concepto de los valores de referencia.
- Parte 2. Selección de los individuos para la producción de los valores de referencia.
- Parte 3. Preparación de los individuos y ejecución de la recogida de sangre para la producción de los valores de referencia (y observados).
- Parte 4. Control de la variación analítica en la producción y uso de los valores de referencia.
- Parte 5. Tratamiento estadístico de los valores de referencia recogidos. Determinación de los límites de referencia.
- Parte 6. Presentación de los valores observados en relación a los valores de referencia.

Esta estructura es la que va a seguirse en el desarrollo de este trabajo, añadiendo después otros apartados que pueden resultar de interés como son los esfuerzos de armonización en el laboratorio, la situación actual sobre el terreno y las soluciones que se plantean para su determinación.

## 2.- JUSTIFICACIÓN Y OBJETIVO.

Este trabajo surge de un interés en saber qué pasos teóricos se siguen a la hora de establecer esos valores que acompañan a los resultados en las analíticas y nos permiten afirmar si están demasiado elevados, normales o demasiado descendidos.

Los intervalos de referencia son la herramienta más utilizada en la interpretación de los resultados del laboratorio. Pueden llegar a tener una importancia significativa en la toma de decisiones clínicas, de ahí la conveniencia de determinarlos de la forma más objetiva y rigurosa que sea posible.

Debido a que se refieren a poblaciones concretas y ya que no hay dos poblaciones iguales, los intervalos de referencia podrán ser distintos en distintos laboratorios. Puede suceder que un mismo resultado sea considerado 'normal' en un laboratorio y 'anormal' en otro, incluso aunque estén usando la misma técnica.

La teoría al respecto está bien definida, pero existe un agujero en su aplicación práctica, que en muchos casos dista de ser óptima. [12]

El objetivo del trabajo es realizar una revisión de bibliografía relacionada con este tema. Se explican distintas alternativas y se habla de las ventajas e inconvenientes de cada uno de los métodos. Se pretende dar una visión global de la teoría enunciada, la situación actual, las posibilidades de mejora y las limitaciones de los intervalos de referencia y si fuera posible despertar un interés sobre el origen de estos valores, que en ocasiones damos por hecho que son exactos e infalibles, pero no siempre ha de ser así.

Hemos encontrado en algunos casos la dificultad de no poder acceder a artículos que trataban el tema. Es especialmente importante en el caso del documento C28-A3, la guía internacional para la definición, establecimiento y verificación de los intervalos de referencia en el laboratorio clínico.

### 3.- CONCEPTOS BÁSICOS.

#### 3.1.- INTERVALOS DE REFERENCIA.

Los intervalos de referencia (IR) permiten comparar el valor observado en una medición con unos valores de referencia, obtenidos en una población de individuos bien definida. Se llaman de referencia porque se refieren a una población concreta. La finalidad de esta comparación es facilitar la valoración de los resultados.

Habitualmente se determinan con población general saludable. En ese caso se llamarán intervalos de referencia biológicos o fisiológicos.

También pueden determinarse IR en condiciones fisiológicas como el embarazo o en grupos específicos de población como puedan ser atletas profesionales o personas con estados patológicos como por ejemplo enfermedad pulmonar obstructiva crónica.

Representan a una población específica y por tanto, su valor depende de la elección de los sujetos a partir de los cuales se han determinado. [12]

Sin información sobre la población de referencia y cómo se han obtenido, es imposible saber su validez y su aplicabilidad en un contexto diferente. Esto debe hacernos ser críticos con valores 'normales' procedentes de la literatura que pretendan poder ser aplicados de forma universal.

El esquema original de su descripción fue propuesto por Siest y Wilding en 1976.

Unos INDIVIDUOS DE REFERENCIA

Constituyen una

POBLACIÓN DE REFERENCIA

De la que se selecciona una

MUESTRA DE REFERENCIA

En la que se determinan unos

VALORES DE REFERENCIA

Que presentan una

DISTRIBUCIÓN DE REFERENCIA

De la que se calculan unos

LÍMITES DE REFERENCIA

Que podrían definir

INTERVALOS DE REFERENCIA.

O a la inversa: un intervalo de referencia es el intervalo comprendido entre, y que incluye, a dos límites de referencia; que son valores derivados de la distribución de resultados obtenida a partir de una muestra de la población de referencia.

La población de referencia es una entidad hipotética que incluye un número indeterminado de individuos de referencia.

Un individuo de referencia es una persona elegida para ser analizada a través de unos criterios bien definidos. El intervalo que obtengamos se usará para compararlo con los valores observados en un único individuo. [12]

Por convención, para calcular los límites se establece que entre ellos deben estar comprendidos el 95 % de los valores centrales medidos. Esto hace que queden fuera los valores inferiores al 2'5 % y superiores al 97'5 %.

### 3.2.- SIGNIFICADO DE LOS INTERVALOS DE REFERENCIA.

Un valor que esté fuera del IR no siempre se asocia con patología. Es cierto que un resultado dentro del rango de referencia se asocia a menor probabilidad de estar enfermo y uno que esté fuera con una mayor probabilidad de ser patológico. Pero hay que tener en cuenta que por el modo que tienen de ser calculados, un 5 % de los individuos teóricamente sanos sobre los que se determinan habrá quedado fuera de esos límites. Lo mismo sucederá con una persona sana que se haga un análisis, la probabilidad de que su resultado quede fuera del IR es de 0.05.

#### SOLAPAMIENTO DE VALORES ANALÍTICOS LÍMITE ENTRE EL GRUPO NORMAL Y EL PATOLÓGICO.

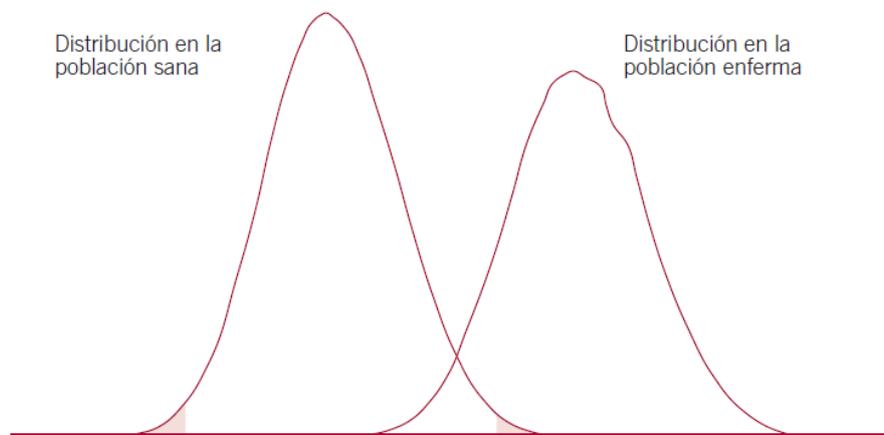


Figura 1. Tomada de [13].

Así pues, unos valores que no estén dentro de los intervalos de referencia no tienen significado patológico por sí mismos. De ahí que en algunos casos se haya criticado el sistema de aviso con asteriscos que identifica los resultados fuera de los límites. Los asteriscos dirigen el interés solo hacia los valores que están alterados, desviando la atención de los valores que están próximos al límite y que es posible que sean igualmente importantes. Además, alarman al paciente, que no tiene por qué ser consciente de cómo se determinan estos límites y se preocupará por el resultado extraño. [1]

### 3.3.- FUENTES DE VARIACIÓN EN EL LABORATORIO [tomado de 14].

*Todas las magnitudes biológicas de tipo cuantitativo presentan una variación en el resultado que es consecuencia de la variación preanalítica, la analítica y la biológica, propia de cada magnitud.*

- *La variación preanalítica está relacionada con la preparación del paciente, el proceso de obtención de la muestra, la postura y tiempo de aplicación del torniquete en el momento de la extracción, el transporte, las interferencias producidas por la medicación, etc. El laboratorio intenta minimizar los errores que se puedan producir estandarizando todos estos procesos y haciendo las debidas recomendaciones a todos los profesionales que intervienen en ellos; de la correcta aplicación de estas medidas dependerá la verosimilitud del resultado.*
- *Por otra parte, el laboratorio tiene establecidos sistemas muy rigurosos de control de todo el proceso analítico con el objetivo de minimizar las fuentes de variación que se puedan producir durante todo este proceso.*
- *La última fuente de variación es la llamada variación biológica, que no se puede minimizar ya que es una característica intrínseca a cada magnitud biológica.*

*Es importante conocer todas estas fuentes de variación a la hora de interpretar un resultado. Para aquellas magnitudes con una variación biológica pequeña y poco afectadas por las fuentes de variación preanalítica, el resultado es muy fiable, ya que el laboratorio minimiza todas las fuentes de variación analítica. Además, en este caso, una única medida con una sola muestra del paciente es una buena estimación de su verdadero valor. En cambio, para las magnitudes de variación biológica alta, el resultado hay que interpretarlo con un poco más de cautela.*

### 3.4.- TIPOS DE VARIACIÓN BIOLÓGICA. [tomado de 14]

*La variación preanalítica y analítica se analizarán más adelante. Existen dos tipos de variación biológica: la intraindividual y la interindividual.*

- *La variación intraindividual consiste en la fluctuación aleatoria del resultado de una magnitud alrededor del punto de equilibrio homeostático de un individuo. Es constante para cada magnitud y similar en todos los individuos sanos. Se obtiene a través de la ponderación de los valores de un conjunto de individuos. Se representa como CVI (coeficiente de variación intraindividual).*
- *La variación interindividual es la diferencia en el punto de equilibrio entre distintas personas, la variación de sus tendencias centrales. Se representa como CVG (coeficiente de variación interindividual o grupal).*

*Podría resumirse en que la variación intraindividual son distintos valores en un mismo individuo, mientras que la variación interindividual son distintos valores en distintos individuos. Se presentan conjuntamente.*

*Hay magnitudes como el colesterol o la creatinina que tienen una variación biológica pequeña, pero hay otras como los triglicéridos o el hierro cuya variación biológica es más alta. Las cifras de variación de la mayoría de magnitudes del laboratorio son conocidas y están recogidas en bases de datos que se pueden consultar fácilmente.*

### 3.5.- ÍNDICE DE INDIVIDUALIDAD. [tomado de 14]

El índice de individualidad (II) es habitualmente el cociente entre el valor de la variación biológica intraindividual y la interindividual para cada magnitud. Si este cociente es bajo significará que la magnitud tiene un alto grado de individualidad, y si el cociente es alto tendrá un bajo grado de individualidad. En algunos casos se introduce en el numerador también la variabilidad analítica.

$$II = \frac{CV_I}{CV_G}$$

Una magnitud con poca variabilidad producirá un intervalo de referencia estrecho. El uso de un IR será más adecuado cuanto mayor sea el índice de individualidad de una magnitud, ya que cualquier cambio significativo en un individuo se detectará porque probablemente caerá fuera de dicho intervalo.

#### DATOS DE VARIACIÓN BIOLÓGICA INTRA E INTERINDIVIDUAL E ÍNDICE DE INDIVIDUALIDAD PARA ALGUNAS MAGNITUDES EN SUERO Y SANGRE VENOSA.

	Variación biológica intraindividual	Variación biológica interindividual	Índice de individualidad*
<b>Suero</b>			
Alfa amilasa	8,7	28,3	0,31
Alanina amino transferasa	24,3	41,6	0,58
Albumina	3,1	4,2	0,74
Antígeno CA125	24,3	54,6	0,44
Antígeno específico de la próstata	18,1	72,4	0,25
Aspartato amino transferasa	11,9	17,9	0,66
Calcio	1,7	1,9	0,89
Colesterol	5,4	15,2	0,36
Creatinina	5,3	14,2	0,37
Ferritina	14,2	15	0,95
Hierro	26,5	23,2	1,14
Potasio	4,8	5,6	0,86
Sodio	0,7	1	0,7
Triglicéridos	20,9	37,2	0,56
<b>Sangre venosa</b>			
Hematies	3,2	6,1	0,52
Hemoglobina	2,8	6,6	0,42
Leucocitos	10,9	19,6	0,56
Plaquetas	9,1	21,9	0,42
*La mayoría de magnitudes tienen un alto grado de individualidad (Índice de individualidad < 0,5).			

Figura 2. Tomada de [14]

### 3.6.- LÍMITES DE DECISIÓN CLÍNICA.

Debe evitarse su confusión con los IR. Esta confusión deriva de que en algunas situaciones se pueden usar los IR como límites de decisión clínica (LDC). Por ejemplo, al realizar una actividad de screening podría decidirse que todos los individuos que tengan un valor mayor al límite superior de referencia se sometan a un examen instrumental de confirmación o a algún tipo de profilaxis preventiva. Aun así, son conceptualmente diferentes.

Los IR describen las características biológicas de una población y dependen de ella, es decir son los propios de esa población. Si se han establecido de manera cuidada, se mantendrán estables mientras lo haga la población a la que se refieran.

Los límites de decisión se establecen a través de un proceso de consenso y dependen de la decisión que se ha de tomar. Pueden cambiar a medida que se dispone de nueva información científica.

Por ejemplo, el límite de decisión de los triglicéridos pasó de 250 mg/dL a 200 mg/dL, y más tarde a 150 mg/dL en tres documentos consecutivos del National Cholesterol Education Program (años 1988, 1993 y 2001).

Se definen como la mejor línea para separar lo ‘normal’ y lo ‘enfermo’, o aquello sobre lo que no es necesario actuar y lo que sí.

Identifican un umbral de riesgo (valores de los lípidos o de la hemoglobina glucosilada) o unos límites para clasificar una patología. Aunque se determinan por consenso, la elección se basa en la comparación entre los valores obtenidos en la población sana (valores de referencia) y aquellos obtenidos en una población afectada por la patología con la que se relacionan.

Así, con una superposición de distribuciones podríamos saber la proporción de falsos positivos y negativos que habría con un valor dado, y la sensibilidad y especificidad de cada uno de los valores que se eligieran. Esto se representa a través de las curvas ROC (Receiver Operating Characteristic). Ponderando estas cifras con las consecuencias clínicas que podría tener una clasificación errónea se llega a un valor por consenso.

En muchos casos, en la elección precisa de un valor se tiene en cuenta la facilidad para recordarlo y se redondea. Por ejemplo en Europa el límite de decisión para el colesterol total es de 200 mg/dL (5.17 mmol/L), mientras que en EEUU es de 5 mmol/L (190 mg/dL). En otros se establece como límite de decisión un valor que se corresponda con x veces el límite superior de referencia.

**DIFERENCIAS ENTRE INTERVALOS DE REFERENCIA Y LÍMITES DE DECISIÓN CLÍNICA.**

	Intervalos de referencia.	Límites de decisión clínica.
Depende de	<ul style="list-style-type: none"> <li>• Tipo de población.</li> <li>• Rango de edad.</li> <li>• Género.</li> </ul>	<ul style="list-style-type: none"> <li>• Problema clínico.</li> <li>• Categoría del paciente.</li> </ul>
Información obtenida.	Valor incluido o no en el 95% central de la distribución de la población de referencia.	Elegibilidad del paciente para un procedimiento determinado.
Estadística.	Distribución de la población de referencia.	Ninguna (por consenso) Curvas ROC y valores predictivos.
Datos proporcionados.	Límite inferior y límite superior.	Uno o más, en función de la probabilidad de la situación clínica o de las diferentes cuestiones clínicas.

*Tabla 1. Tomada de [15].*

## 4.- SELECCIÓN DE LA MUESTRA DE REFERENCIA.

Pasamos ahora a describir las etapas necesarias para establecer unos IR. La selección de la muestra es el punto inicial y uno de los más importantes. Se necesitan sujetos sanos. Con el tiempo se ha visto que los criterios deben adaptarse al tipo de estudio que se va a realizar, no serán necesarios los mismos en un gran estudio multicéntrico que en un estudio a pequeña escala. Sin embargo, habrá que proporcionar siempre una información adecuada sobre los criterios aplicados, con el fin de permitir a otros investigadores saber cómo se han obtenido y si les son válidos. La selección de individuos puede realizarse en distintos momentos en relación con la toma de la muestra (a priori o de forma indirecta).

Para determinar un intervalo de referencia para la hemoglobina, por ejemplo, sería razonable excluir a sujetos con déficit de hierro, deficiencia pronunciada de vitamina B12 o ácido fólico, inflamación o enfermedad respiratoria crónica, tumores, anomalías genéticas en la síntesis de hemoglobina. Todos estos factores tienen influencia en su síntesis. Además, en el caso de la hemoglobina convendría registrar la altitud y el hábito tabáquico. Más adelante podrían ser útiles como criterios de estratificación. Para determinar los criterios de exclusión es importante el conocimiento de la fisiopatología del elemento que se analiza.

### 4.1.- DEFINICIÓN DE SALUD. CRITERIOS DE INCLUSIÓN Y DE EXCLUSIÓN.

La OMS define la salud como 'un estado de completo bienestar físico, mental y social y no únicamente la ausencia de afección o enfermedad'. [16] Pero este no es un punto de partida realista. El concepto de salud puede ser diferente según las culturas y los países.

En 1975 el Comité Escandinavo de Valores de Referencia definió una lista de condiciones que debían excluirse para considerar a alguien sano. Sirvieron como modelo a seguir. [17]

- Entre las condiciones patológicas se incluían la insuficiencia renal, la insuficiencia cardiaca congestiva, enfermedades respiratorias crónicas, patologías hepáticas, síndromes de malabsorción o anemias.
- Respecto a la toma de sustancias, el documento de la IFCC abogaba por excluir a los individuos que estuviesen tomando tratamientos para enfermedades, bien como suplemento o reemplazo, o a los consumidores de drogas. La lista incluía alcohol, tabaco o anticonceptivos orales. En el caso del alcohol se debía excluir a los que tuvieran una ingesta excesiva y como criterio preanalítico, se exigía a los individuos que no consumieran alcohol las 24 horas antes de realizarse el análisis.
- Hay estados fisiológicos alterados que debían excluirse. Entre ellos están el embarazo o el ejercicio físico anómalo. El embarazo da lugar a cambios hormonales y metabólicos. El ejercicio puede tener efectos a corto plazo como deshidratación y daño en los tejidos, o más a largo plazo como retraso de la pubertad.
- Patologías mentales o psicológicas como estrés o depresión pueden dar lugar a desregulaciones hormonales y metabólicas.

La aplicación estricta era poco práctica, en particular en sujetos de más edad. En personas mayores es habitual la toma de medicaciones de forma crónica, por lo que es más difícil que cumplan de forma estricta todos estos criterios.

Se requiere una aproximación más pragmática, entendiendo salud como ausencia de signos de enfermedad o condiciones que se puedan relacionar específicamente con los

elementos a medir. El individuo deberá cumplir unos determinados criterios de inclusión y no deberá tener otros determinados criterios de exclusión, que dependerán del parámetro a determinar y de los recursos de los que se dispone. Para saber si el individuo cumple los requisitos, se puede emplear un cuestionario, la realización de una exploración física u otras investigaciones más detalladas (pruebas de imagen o de laboratorio).

En el protocolo para la realización de grandes estudios multicéntricos, se recomienda para el reclutamiento de voluntarios: [18]

#### **4.1.1.-Criterios de inclusión.**

- Los participantes deben sentirse subjetivamente bien.
- Deben ser mayores de 18 años. Al menos el 80 % debe estar entre 18 y 65 años, con igual mezcla de sexo y distribución de edades, salvo en los mayores de 65 años.
- Idealmente los participantes no deberían estar tomando ninguna medicación. Se deberá dejar por escrito si algún sujeto está tomando medicamentos o suplementos dietéticos (nombre, dosis y frecuencia). De este modo se podría realizar una exclusión secundaria después de la medición si fuera necesario. Los anticonceptivos orales, los estrógenos o las hormonas tiroideas se permiten si el sujeto está compensado, pero se debe anotar en la historia. Por ejemplo, alguien que tome hormonas tiroideas se considerará compensado si su valor de TSH es inferior al límite superior de la normalidad.

#### **4.1.2.-Criterios de exclusión.**

- Diabetes conocida en tratamiento con antidiabéticos orales o insulina. Si el tratamiento es solo dietético, puede aceptarse su inclusión.
- Historia de enfermedad crónica hepática o renal.
- Resultados de análisis de sangre que apuntan claramente a enfermedad severa.
- Historia de hospitalización o enfermedad seria en las cuatro semanas previas.
- Donación de sangre en los 3 meses previos.
- Portador conocido de VHB, VHC o VIH.
- Mujeres participantes que estén embarazadas, en periodo de lactancia o en el primer año tras haber dado a luz.
- Cualquier otra enfermedad significativa o trastorno que en opinión del investigador pudiera poner en riesgo al paciente por participar en el estudio o podría influir en los resultados del estudio.
- Participación en algún estudio de investigación en los últimos tres meses.

Nótese que el componente subjetivo sigue siendo importante en algunos de los puntos.

## 4.2.- OTROS ASPECTOS A TENER EN CONSIDERACIÓN.

Respecto al alcohol y el tabaco, pueden ser criterios de exclusión en estudios pequeños. En estudios con gran número de participantes se debe registrar la cantidad consumida y evaluar estadísticamente sus efectos por análisis de regresión múltiple.

Las diferencias étnicas, incluyendo las fenotípicas y las culturales, no pueden ser desatendidas. Así lo demostraron Johnson et al. donde caucásicos y asiáticos de la India que vivían en la misma comunidad en Reino Unido mostraron diferentes concentraciones de proteínas en suero. En un estudio limitado por usar una muestra pequeña, Ichihara et al. describieron grandes diferencias en elementos analizados entre seis ciudades asiáticas diferentes. Como ejemplo clásico, en el cálculo de la tasa de filtrado glomerular a partir de la creatinina se tiene en cuenta la etnia.

El Índice de Masa Corporal (IMC) o factores como la dieta también pueden influir. [19]

La muestra empleada deberá ser un fiel reflejo de toda la población en la que se vayan a aplicar los intervalos de referencia.

## 4.3.- EJEMPLO DE NUESTRO ENTORNO.

Se adjuntan a continuación, a título de ejemplo, los criterios de inclusión y exclusión y el cuestionario realizado en un estudio sobre valores de referencia en plasma de electrolitos en una muestra de 146 donantes de sangre en Toledo. [20] Puede apreciarse que las recomendaciones no requieren un cumplimiento estricto, sino que cada investigador puede adaptarlas según considere necesario en su estudio.

### CUESTIONARIO REALIZADO Y CRITERIOS DE EXCLUSIÓN EN ESTUDIO DE IR EN TOLEDO [20].

Encuesta de salud de la muestra de donantes	
¿En general Vd. diría que su salud es buena?	(S) (N)
¿Realiza alguna actividad física regularmente?	(S) (N)
¿Con qué frecuencia?	(horas/semana)
Tipo de actividad	(moderada) (fuerte)
¿En su puesto de trabajo está expuesto a sustancias tóxicas o volátiles?	(S) (N)
¿Ha padecido alguna enfermedad recientemente?	(S) (N)
Si la ha padecido diga cuál	
¿Está tomando algún medicamento?	(S) (N)
¿Es fumador? Si lo es indique los cigarrillos/día	(S) (N)
¿Sigue alguna dieta especial?	(S) (N). Describala
¿Toma bebidas alcohólicas?	(S) (N). Indique cuales y con qué frecuencia.

Criterios de exclusión
Enfermedad reciente
Hospitalización reciente
Cirugía reciente
Hipertensión
Enfermedades genéticas
Embarazo
Lactancia
Consumo de alcohol (> 100 g etanol/semana) o drogas de abuso
Uso crónico de fármacos
Tabaquismo
Obesidad
Actividad física intensa
Dietas especiales

Figuras 3 y 4. Tomadas de [20]

#### 4.4.- EL SUPUESTO ESTUDIO MISS MANHATTAN.

En un monográfico sobre valores de referencia, Solberg y Gräsbeck describen el supuesto estudio Miss Manhattan con tono humorístico, como respuesta a las quejas por la dificultad de hacer un estudio propio sobre IR. [21]

*“Hay una concepción equivocada de que la producción de valores de referencia es sumamente complicada. Por ejemplo, se ha dicho que los valores de referencia “pueden usarse en muy raras ocasiones porque han de cumplirse todos los requisitos estadísticos y todos los factores fisiológicos deben tenerse en cuenta. Una perfección difícil de conseguir en este mundo imperfecto”. Es una interpretación incorrecta. Lo que se requiere en esencia es que se dé una cantidad razonable de información sobre:*

*-Las características de la población, especialmente cómo se seleccionó a los sujetos y cómo se valoró su salud. No se requiere que estos criterios sean rígidos, sino que se mencionen. Por tanto sería perfectamente suficiente decir que los valores de un estudio se obtuvieron a partir de las cien primeras mujeres atractivas que el doctor King vio en Broadway. Sus criterios de salud fueron que él las considerara atractivas y que pudieran andar.*

*-El estado fisiológico de los sujetos, su preparación y la ejecución de la toma de muestras también se deben referir. Sería compatible con la teoría de los valores de referencia decir que no se les preguntó si habían comido, que las venopunciones se realizaron en la esquina de la calle mientras estaban de pie, después de haber usado un torniquete durante 30 minutos, con agujas no estériles y recogida de las muestras en latas de cerveza vacías.*

*-La forma de manejar las muestras obtenidas y el procedimiento analítico llevado a cabo se deben citar. De nuevo, sería compatible referir que se almacenaron las latas en el tejado de un edificio desde el 15 de Julio hasta el 22 de Julio de 1981 para evaluar su cantidad de oxígeno de acuerdo con la técnica de Lavoisier (1790).*

*(Por supuesto sospechamos que los resultados de este estudio no terminarían en la revista Química Clínica).”*

#### 4.5.- SELECCIÓN CON CRITERIOS A PRIORI Y A POSTERIORI.

La selección a priori consiste en decidir antes de recoger la muestra a qué individuos seleccionar y cómo realizar la estratificación. En la selección a posteriori se recoge un número importante de sujetos y se les analiza. Más tarde, en función de los resultados, se decide si se les mantiene dentro de la población de referencia y si hay necesidad de estratificarlos. En caso de que la biología del elemento a analizar sea conocida y se tengan los recursos necesarios, se recomienda optar por la selección con criterios a priori.

Ambos constituyen métodos directos de selección, ya que en los dos casos los datos son recogidos con la expresa finalidad de determinar los intervalos de referencia, con control sobre la forma en que se realiza la determinación al paciente y la información que se recoge de él.

#### **4.6.- VALORES DE REFERENCIA INDIRECTOS. EL ESTUDIO REALAB.**

En esta aproximación se establecen los intervalos de referencia a partir de datos que existían previamente en el laboratorio, usando bases de datos con historias de pacientes reales. Esto es lo que se hizo por ejemplo en el estudio REALAB italiano. [22]

En él se analizaron los resultados de analíticas realizadas en un centro de diagnóstico más especializado en tareas preventivas que en el diagnóstico y seguimiento de la enfermedad. Se recopilaron los resultados de todos los análisis realizados entre los años 1997 y 1999, un total de 197.350. Eligieron a aquellos que se habían hecho al menos una de las 23 determinaciones básicas que iban a analizar. Además excluyeron a quienes se habían realizado más de una analítica en ese periodo, considerando que tenían mayor probabilidad de padecer enfermedad. (n= 93.649)

Luego ejecutaron un algoritmo para detectar correlación significativa entre cada cantidad y todas las demás, usando el coeficiente de correlación de Spearman porque algunas de las variables tenían una distribución no gaussiana. La disponibilidad de la información les permitió analizar otras 13 cantidades más.

Después, para cada resultado de cada individuo se analizaban las variables correlacionadas. Si el valor de esas variables se encontraba dentro de los intervalos de referencia previos del laboratorio, el resultado se consideraba elegible como valor de referencia. El número de muestras que se emplearon en total fue de 61.246.

Para 23 de las determinaciones los nuevos IR fueron más estrechos, mientras que en los 13 restantes fueron más anchos. Lo más importante es la originalidad del método, pero es quizá poco aplicable a laboratorios con un perfil asistencial distinto.

##### **4.6.1.- Limitaciones de la selección indirecta.**

La selección indirecta tiene limitaciones importantes. No hay una definición cuidadosa de la población de referencia. No se conoce bien a la población, se confía principalmente en métodos estadísticos para excluir a los sujetos que no están sanos. No hay ningún control en las variables preanalíticas. Es muy difícil además demostrar la trazabilidad metrológica de los resultados obtenidos (ausencia de variabilidad analítica), con lo que es posible que los resultados obtenidos solo sean aplicables en el laboratorio que los produjo. [12]

En la mayoría de los casos representa una forma de confirmar y validar los hallazgos obtenidos con una selección a priori, más precisa y científica. También pueden ser útiles en situaciones locales (pequeños laboratorios) o en grupos difíciles como pueden ser neonatos, niños o ancianos.

Sin embargo, algunos autores defienden esta aproximación debido a que los resultados son clínicamente relevantes para el laboratorio que la realiza y es mucho más sencilla su obtención.

## 5.- CONSIDERACIONES PREANALÍTICAS: PREPARACIÓN DE LOS INDIVIDUOS Y EJECUCIÓN DE LA RECOGIDA DE SANGRE.

La fase preanalítica incluye una serie de procesos que empiezan en el momento en el que el clínico realiza la petición analítica y terminan cuando la muestra está en condiciones de ser analizada. Es en esta fase donde se produce el mayor número de errores en el laboratorio clínico, hasta un 70% del total. La gran mayoría de ellos surgen a partir de las actividades manuales como son la recogida, manejo, transporte, preparación y almacenamiento de la muestra. Además, la preparación del sujeto tiene lugar fuera de la supervisión médica o del personal del laboratorio. [23]

### EL ICEBERG DE LOS ERRORES DE LABORATORIO.

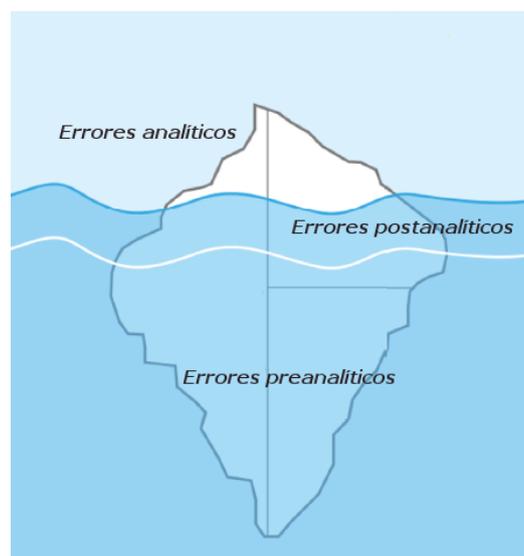


Figura 5. Tomada de [24]

Es importante definir y describir adecuadamente las condiciones preanalíticas para que puedan ser reproducidas. Los estudios que se encuentran sobre la forma en que estas condiciones influyen los resultados son de escaso tamaño muestral, por lo general realizados con los trabajadores de un laboratorio. En una revisión sobre los estudios al respecto (36 estudios cumplieron sus requisitos), se encontró que se tomaban muestras o de voluntarios (N=20, rango intercuartílico (RIQ): 17.5 - 30.0) o de pacientes ambulatorios (N=88, RIQ: 54.5-220.5) [25]

Serían necesarios estudios con muestras mayores para evaluar las repercusiones de esta fase en el resultado final.

Podríamos dividir la fase preanalítica en tres periodos diferentes: preparación del sujeto, recogida de la muestra y procesamiento de la muestra.

## 5.1.- PREPARACIÓN DEL SUJETO.

### 5.1.1.- El ayuno.

Este es un concepto conocido pero poco delimitado que se refiere a que el paciente debe abstenerse de consumir una serie de elementos como comida, alcohol, café, tabaco o quizá incluso medicación. Desafortunadamente estos ítems no están bien descritos ni armonizados. La duración necesaria tampoco es inequívoca. Se ha demostrado por ejemplo que los triglicéridos aumentan a partir de un cierto tiempo de ayuno por el inicio del catabolismo graso. Además, muchos clínicos tienden a pensar que solo se afectan unos pocos elementos, pero el ayuno afecta de forma significativa a distintos parámetros hematológicos, hemostáticos y bioquímicos. [26]

Con la comida se producen marcadas respuestas metabólicas y hormonales a la absorción de sustancias. El periodo inmediatamente posterior se llama postprandial y se caracteriza por estrés oxidativo, inflamación y disfunción endotelial. La mayoría de cambios tiene lugar hasta 4 horas después de la ingesta. Afecta a triglicéridos, albúmina, ALT, AST, calcio, sodio, magnesio, potasio, PCR, ácido úrico o bilirrubina total. También tiene influencia significativa en valores del hemograma. Por ello se recomienda el ayuno.

Pero además, existen otros factores que condicionan la respuesta postprandial. Se ha observado un incremento de la cantidad de lipoproteínas ricas en triglicéridos después de fumar un solo cigarrillo. El café también se ha visto que aumenta los resultados de la glucemia en ayunas. Podrían armonizarse las sustancias que se permite o no consumir durante este periodo.

El grupo de trabajo de la fase preanalítica de la Federación Europea de Química Clínica y Medicina de Laboratorio (EFLM) recomienda un ayuno de 12 horas, durante el cual se permita el consumo de agua. El alcohol debe ser evitado durante las 24 horas previas a la toma de sangre. En la mañana del análisis los pacientes deberán evitar fumar o tomar bebidas con cafeína. Recomiendan también una cierta severidad con el cumplimiento de estos criterios en análisis de rutina. No tomar una muestra es mejor que una mala muestra. [24]

Es preciso recordar que el cumplimiento por parte del paciente de las recomendaciones no será total. Pero sería interesante conocer en qué medida esta falta de adherencia es debida a falta de información. Con un criterio consensuado y uniforme quizá se podrían reducir las diferencias entre los laboratorios.

### 5.1.2.- Ejercicio físico.

Puede modificar los valores de parámetros hematológicos. Se distinguen dos conjuntos de situaciones: adaptaciones a corto plazo ante ejercicio extenuante puntual o adaptaciones a largo plazo por la realización de ejercicio regularmente. Con el ejercicio tiene lugar una hemoconcentración por la deshidratación que supone. En deportistas de fondo se ha observado una mayor volemia, por expansión del plasma y del volumen eritrocitario.

Además, el ejercicio aumenta el metabolismo basal y en caso de que sea extenuante hay un cierto daño celular por estrés oxidativo. En atletas entrenados se ha constatado que muchos valores de la analítica no se corresponden con los valores de referencia, no por patología subyacente sino por su adaptación al ejercicio regular. Se propone considerar la abstención de ejercicio físico 48 horas antes del análisis de sangre. [27]

En mi opinión, esto debería recomendarse en caso de que fuera un ejercicio anormalmente intenso. Por otra parte, considero que sería recomendable registrar en el análisis de los individuos de referencia la cantidad de ejercicio que realizan habitualmente, tratando este aspecto de igual modo que el consumo de alcohol y tabaco, dado que parece que también puede influir sobre los resultados.

### **5.1.3.- Otros factores.**

En caso de que la muestra no sea sanguínea, por ejemplo, líquido cefalorraquídeo, es posible que haya que tener en cuenta otros factores diferentes. A su vez, en algunos elementos a analizar habrá que tener en cuenta otras variables. Por ejemplo, en las determinaciones hormonales influye el nivel de estrés emocional. La vitamina D por su relación con la exposición al sol se verá condicionada por la época del año en que se haga el análisis. Habrá determinaciones que dependan de la fase del ciclo menstrual.

## 5.2.- RECOGIDA DE LA MUESTRA.

La toma de sangre venosa es un procedimiento invasivo indispensable para la medicina de laboratorio. Permite obtener la matriz biológica que se ha de analizar.

Es una técnica que se emplea desde la antigüedad. En la época de Hipócrates el cuidado médico se basaba en la estabilidad de los cuatro humores: sangre, flema, bilis amarilla y bilis negra. Por ello una de las prácticas que se realizaban era la obtención de sangre con una pequeña aguja. Entre los siglos XVIII y XIX se difundió la práctica de las sangrías con finalidad terapéutica. Las flebotomías eran la técnica más común. En ellas se practicaba una incisión en las venas externas del antebrazo o del cuello. Muchas veces el procedimiento no terminaba hasta que el paciente no mostraba síntomas sincopales. Ya a finales del siglo XIX se definió como charlatanería la terapia con sangrados, pero la utilidad diagnóstica se mantuvo. La introducción de agujas y jeringuillas tuvo lugar a principios del siglo XX. Otra innovación significativa fue la introducción por parte de la marca comercial Becton Dickinson de sistemas de vacío tipo Vacutainer en torno a los años 50. [28]

La Organización Mundial de la Salud (OMS) publicó una guía sobre las mejores prácticas en la técnica de extracción de sangre venosa. [29]

### 5.2.1.- Hora del día.

Es habitual recoger las muestras de sangre por las mañanas, de 7 am a 10 am. Esto responde a una cuestión de comodidad a la hora de realizar el ayuno. Por ello en las determinaciones con variabilidad circadiana, los límites de referencia o los cambios en el valor de referencia se calculan en este momento del día.

### 5.2.2.- Dispositivos para la extracción sanguínea.

Se recomienda usar dispositivos que integren las agujas monouso, sistemas de soporte tipo holder y probetas con sistema de vacío. Las jeringas convencionales son una alternativa en situaciones de emergencia en las que no se disponga del sistema anterior o en caso de situaciones anatómicas particulares. El uso de agujas de diámetro inferior a 23 gauge puede producir hemólisis, por lo que se aconseja emplear agujas de 20 ó 21 G.

### 5.2.3.- Colocación del torniquete.

El torniquete resulta en una estasis venosa que tiene influencia en los parámetros medidos. Para comprobarlo, Lippi et al. analizaron la determinación en 30 médicos voluntarios sin ninguna compresión, después de un minuto de compresión a una presión de 60 mmHg y después de 3 minutos con esa presión. Las realizaron secuencialmente, cambiando de brazo y esperando diez minutos para realizar la última determinación. Demostraron cambios con significación estadística. Tal y como esperaban, hubo aumentos en la serie roja y disminuciones en la blanca.

En otros estudios se ha visto un aumento en la respuesta procoagulante, alteraciones de la función plaquetaria o cambios en el pH y el equilibrio electrolítico. [30]

Se recomienda usar el torniquete solo si las venas no son visibles o palpables, y durante el menor tiempo posible. Esto se aleja bastante de la práctica clínica habitual.

#### **5.2.4.- Posición previa.**

En un estudio similar al anterior, los mismos investigadores se preguntaron si la posición que había adaptado el sujeto podía influir sobre los valores de la determinación. Determinaron que era importante la posición corporal que había tenido durante los minutos antes de la extracción. La presión venosa en las partes declives del cuerpo aumenta tras estar de pie de forma prolongada, generando un aumento de la presión capilar que da lugar a mayor ultrafiltración del plasma al espacio intersticial. [31]

#### **5.2.5.- Procedimiento.**

El paciente se colocará sentado o tumbado, cómodamente, en posición supina. Se prefieren las venas centrales del antebrazo (cubital y cefálica), como alternativa se puede usar la vena basílica o las del dorso del brazo. Se deben usar un par de guantes no estériles, con previa higiene de manos. Para la desinfección de la piel se prefiere el alcohol porque se ha visto que la povidona yodada puede aumentar falsamente el potasio, el fósforo o el ácido úrico. Debe permitirse que se seque completamente. Se debe sostener el brazo del paciente y palpar con el pulgar por debajo del sitio de punción. Se entra delicadamente con la aguja con un ángulo de 30 o menos grados y se sigue introduciendo la aguja. El torniquete se ha de retirar antes de extraer la sangre y no se puede mantener en ningún caso más de dos minutos. Una vez que se retira la aguja hay que colocar una gasa o trozo de algodón y aplicar presión sobre ella. El paciente debe evitar doblar el brazo para que no se le forme hematoma. [29]

#### **5.2.6.- Orden de los tubos de recogida.**

La OMS recomienda seguir un orden determinado en las extracciones. Primero los tubos de hemocultivo o estériles, luego aquellos lisos o de gel y por último aquellos que contengan aditivos. Así se evita la contaminación con aditivos de tubos usados antes. La recomendación se basa en un caso clínico presentado en 1977 por Sun y un estudio de seguimiento realizado por Calam y Cooper en 1982 que observó hiperpotasemia e hipocalcemia, marcadores de contaminación in vitro de EDTA (un anticoagulante para muestras), si se seguía un orden incorrecto. Los autores reconocieron que la presencia de contaminación no se observaba en condiciones de flebotomía ideales, sino solo en venopunciones difíciles. En otro estudio con sistemas de asa cerrada, por el contrario, se midió la cantidad de EDTA tras la extracción y se vio que no se producía contaminación. Un estudio reciente sobre muestras con hiperpotasemia prueba que la contaminación con EDTA y en menor medida con citrato sódico, es todavía relativamente frecuente y difícil de identificar [32].

### **5.3.- PROCESAMIENTO DE LA MUESTRA.**

El modelo actual del laboratorio clínico hace que a veces sea necesario transportar las muestras obtenidas en centros periféricos hasta un centro en el que se realicen los análisis. El contacto prolongado entre plasma o suero (el plasma es el suero con los factores de la coagulación) con las células es una causa frecuente de resultados alterados. Deberían ser separados lo antes posible para evitar el transporte pasivo o activo de sustancias. Algunas proteínas del plasma son muy lábiles.

Así, serán importantes tanto el tiempo que transcurre hasta la centrifugación y el posterior análisis como la temperatura a la que se mantienen las muestras. La estabilidad es la capacidad de una muestra de mantener la propiedad inicial de un elemento medido por un periodo de tiempo dentro de unos límites determinados, siempre que la muestra sea almacenada en unas determinadas condiciones. [33]

Idealmente las muestras deberían ser transportadas en caso de que fuera necesario a temperatura ambiente (15-22 °C) lo antes posible. Los exámenes de coagulación deben haberse realizado en las primeras 4 horas desde su recolección. Deben evitarse temperaturas extremas.

#### **5.3.1.- Conservación.**

En caso de que el procesamiento de la muestra no estuviera disponible, se debería proceder al centrifugado y separación del suero o el plasma, seguido de refrigeración o congelación. A -20°C la muestra debería procesarse en las siguientes 2-4 semanas, mientras que a -80°C pueden mantenerse meses o incluso años.

#### **5.3.2.- Centrifugación.**

El objetivo de la centrifugación es separar los componentes de una muestra en función de su densidad, para asegurar que la determinación de magnitudes o células de interés puede ser evaluada. Una centrifugación inadecuada puede requerir más centrifugación o dar lugar a resultados inapropiados. Existen distintos equipos de centrifugado (de ángulo fijo, típicamente a 45° respecto al eje de rotación o basculantes, que se mueven cuando el mecanismo centrifugador gira) y distintas condiciones de centrifugado (velocidad, tiempo, temperatura, aceleración, deceleración y radio).

Por ser la etapa limitante en el laboratorio, muchos fabricantes proveen de indicaciones para maximizar su rendimiento. Sin embargo, ocurre que para distintos tipos de muestras se dan distintas recomendaciones, por lo que la estandarización es difícil.

## 6.- ASPECTOS ANALÍTICOS: CONTROL DE LA VARIACIÓN ANALÍTICA.

La precisión a la hora de realizar una medición nunca puede ser total, todas las determinaciones vienen acompañadas de incertidumbre. En los laboratorios se llevan a cabo una serie de controles para minimizar las repercusiones que pueda tener esta inexactitud en la toma de decisiones.

### 6.1.- DEFINICIONES.

El control de la calidad analítica es un conjunto de procedimientos llevados a cabo para monitorizar la calidad de los resultados de las pruebas, detectando problemas antes de la entrega de resultados, de manera que se asegure la prestación necesaria para cumplir con los requisitos clínicos.

- La exactitud o error total es la diferencia entre el resultado de una medida y el valor verdadero.
- La precisión es el grado de concordancia entre los resultados independientes de mediciones realizadas en condiciones estipuladas.
- El error total es la suma del error aleatorio (imprecisión) y el error sistemático (sesgo).
- El error aleatorio es la diferencia entre un resultado individual de una magnitud y la media de un número determinado de mediciones de la misma, en condiciones de repetibilidad.
- El error sistemático es la diferencia entre la media de un número definido de medidas de una misma magnitud en una muestra y el valor verdadero

### 6.2.- CONTROL INTERNO.

Se realiza dentro del propio laboratorio. Permite aceptar o rechazar las series analíticas. Una serie analítica es el intervalo (en tiempo o en número de muestras) dentro del cual el procedimiento de medida es estable. El control interno de calidad habitualmente utiliza controles estables a uno, dos o tres niveles de concentración, y siempre el mismo lote hasta su caducidad. [34]

Para cada magnitud y procedimiento analítico el laboratorio debe conocer los dos componentes de su prestación analítica, el error aleatorio y el sistemático.

#### 6.2.1.- Imprecisión o error aleatorio.

La imprecisión inherente al procedimiento analítico se puede estimar con un mínimo de 20 resultados control, obtenidos en condiciones óptimas (instrumento bien ajustado, reactivos recientes, personal entrenado) en un tiempo relativamente corto (por ejemplo, 4 controles por día durante 5 días, 2 resultados de control en 10 días).

Se puede expresar en términos de DE o de CV, mediante la siguiente fórmula:

$$CV = \frac{100 \times DE}{Media}$$

Donde: DE = desviación estándar interserial de los 20 valores.

Esta estimación inicial no incluye la variabilidad debida a factores como lotes diferentes de reactivos y calibradores, variaciones a largo plazo, recalibraciones, cambios de personal técnico, etc. Por ello, se recomienda reevaluar en un periodo de tiempo más amplio, donde se haya observado que la imprecisión se mantiene constante (por ejemplo, 6 meses), para tener en cuenta todos estos factores de variabilidad.

Existen 2 variedades, el modelo de gestión interna, donde el tratamiento estadístico de los resultados control se realiza únicamente con los datos obtenidos por el propio laboratorio, y el control interno con gestión externa, donde el procesamiento estadístico se realiza con los datos obtenidos por el propio laboratorio y por otros laboratorios. [35]

Ambos sirven para calcular la imprecisión analítica pero no son adecuados para evaluar el error total ni el sesgo. Para determinar el sesgo es necesario conocer el valor verdadero convencional o el mejor estimado posible del material control.

### **6.3.- CONTROL EXTERNO.**

El control externo de la calidad es la determinación del desempeño de cada laboratorio mediante la comparación con otros laboratorios, a través de una organización externa. Se le conoce como programa de intercomparación.

En este programa la muestra de control debe ser ciega, es decir, de valor desconocido para el laboratorio en el momento del análisis. Como desventaja práctica ocurre que se realiza habitualmente solo con una determinación, de forma que no evalúa la actividad continuada, sino una sola determinación.

Existen 3 modelos, los 2 primeros: evaluación externa de la calidad y ensayo de aptitud son muy similares, y se centran en las prestaciones analíticas, mientras que el tercero, denominado garantía externa de la calidad, tiene en cuenta todas las fases del laboratorio, incluyendo la interpretación de las pruebas y el asesoramiento a los clínicos.

Con el control externo de la calidad, en cualquiera de sus modelos, se mide el error total de cada mensurando, porque la muestra de control, que es ciega para el participante, se analiza una única vez. A largo plazo, cuando se dispone de todos los resultados del programa, se puede medir el error sistemático o sesgo.

## 7.- PROCEDIMIENTOS ESTADÍSTICOS.

Este es el aspecto que más ha centrado la atención de quienes han estudiado los valores de referencia. Los problemas principales son:

- Encontrar el método estadístico más eficiente para extrapolar los resultados obtenidos en la población muestral al conjunto de la población.
- La estratificación de los resultados en distintos grupos.
- La detección y descarte de valores anormales.

### 7.1.- DEFINICIONES TEÓRICAS.

La estadística inferencial es la rama de la estadística que busca inducir las características de una población a partir de la observación de una parte de ella, la muestra, seleccionada habitualmente de forma casual.

Los métodos estadísticos paramétricos son aquellos que asumen la hipótesis de que trabajan con una distribución normal o gaussiana.

Una transformación es una operación matemática que se aplica a una distribución para aproximarla a la distribución normal. La de Box-Cox fue la que inicialmente se recomendó por el panel de expertos de la IFCC. Fue posteriormente modificada por Ichicara y Kawai. [36]

Los métodos no paramétricos no asumen que trabajen sobre un tipo concreto de distribución, para calcular los límites de referencia estiman el percentil 2'5 y el 97'5.

### 7.2.- MÉTODOS ESTADÍSTICOS.

Harris y Boyd citan la aproximación de Wootton et al. que en 1951 aplicaron métodos paramétricos por primera vez para calcular valores normales. Se dieron cuenta de que este modelo estadístico solo era aplicable en una pequeña parte de las situaciones y dos años más tarde propusieron una transformación logarítmica de los datos para conseguir una distribución gaussiana.

En 1987 se publicó la recomendación de la Federación Internacional de Química Clínica (IFCC) que aún ahora mantiene su validez. Promovía la elección arbitraria de determinar los intervalos de referencia con el 95 % central de la distribución. Esto generalmente ha sido aceptado. Una excepción es la propuesta de Jørgensen et al. que propusieron ampliar el intervalo hasta incluir un 99.8 % de los valores centrales, basándose en que su principal utilidad actual es servir de elemento de comparación en pruebas de screening [37]. Así evitarían falsos positivos a los que habría que hacer grandes baterías de pruebas. Esta corriente no ha sido muy seguida.

La IFCC también propugnaba que los límites de referencia se debían establecer con intervalos de confianza (IC) del 90 %. La anchura del IC se reduce con el aumento del número de sujetos evaluados y es un buen indicador de la incertidumbre de los límites de referencia.

Se recomienda el cálculo no paramétrico. En teoría los cálculos paramétricos son más precisos, sobre todo si la muestra de sujetos es pequeña. Sin embargo, la falta de certeza de que la distribución original (o su transformación) sea normal aumenta la incertidumbre de la estimación final.

Se propuso un método para calcular los intervalos de referencia, basado en calcular los fractiles 0.025 y 0.975.

El percentil  $\alpha$  sólo se puede calcular si  $\alpha \geq 1/n$ , donde  $n$  es el tamaño de la muestra.

$$n \geq 1/\alpha \rightarrow n \geq 1/0.025 \rightarrow n \geq 40.$$

Aunque 40 sea el número mínimo necesario de sujetos para calcular estos fractiles, se corresponderían con el valor mínimo y el máximo de la muestra, y sería por tanto imposible estimar sus intervalos de confianza. Para calcular la incertidumbre de los límites se requieren al menos 120 individuos. Así  $\alpha = 0.0083$ . Cuando los datos se disponen de menor a mayor, el tercer valor se corresponde con el límite inferior de referencia (percentil 2.5) y el 118º valor será el límite superior de referencia (percentil 97.5). El IC del 90 % del límite inferior abarcará desde el primer valor hasta el séptimo y el del límite superior desde el 114º valor hasta el 120º. [12]

Para un número  $n$  de individuos, se puede aplicar tres métodos [38]:

-Primero habrá que hallar el orden  $r$  correspondiente al percentil 2'5, lo que se puede hacer con tres fórmulas diferentes:

$$r = 2.5 \times \frac{n}{100} + 0.5$$

$$r = 2.5 \times \frac{n+1}{100}$$

$$r = 2.5 \times \frac{n-1}{100} + 1$$

Si el orden obtenido es un número entero, entonces ese valor se corresponderá con el percentil 2'5. En caso contrario habrá que obtenerlo por interpolación lineal. Se redondeará por defecto el valor del orden. Se obtendrán los valores de percentiles  $p_1$  y  $p_2$ , correspondientes a los valores de  $x_{[r]}$  y  $x_{[r+1]}$ .

$$p_1 = \frac{r-0.5}{n} \times 100$$

$$p_2 = \frac{r+1-0.5}{n} \times 100$$

$$p_1 = \frac{r}{n+1} \times 100$$

$$p_2 = \frac{r+1}{n+1} \times 100$$

$$p_1 = \frac{r-1}{n-1} \times 100$$

$$p_2 = \frac{r+1-1}{n-1} \times 100$$

La estimación será pues:

$$\text{Percentil 2.5} = x_{[r]} + \frac{2.5 + p_1}{p_2 - p_1} \times (x_{[r+1]} - x_{[r]})$$

En teoría es un proceso automático, simple de ejecutar, pero en la práctica se complica. El número relativamente alto de individuos de referencia requeridos crea dificultades en poblaciones pediátricas, pruebas caras o en elementos que varían en función de sexo y edad en los que hay que estratificar en varias clases.

### 7.3.- FORMAS DE ESTIMAR LOS PERCENTILES CON MUESTRAS MÁS PEQUEÑAS.

Como alternativa, Horn et al. propusieron un 'método robusto' basado en la transformación de los datos originales según Box y Cox, seguida de un algoritmo relativamente complejo. Este algoritmo da diferentes pesos a los valores en función de su distancia de la media. El uso de este método permitiría calcular los límites de referencia con tamaños muestrales de tan solo 20 sujetos, pero sin IC.

Para calcular los IC sería posible usar el método de 'bootstrap' o remuestreo (el término hace alusión a la tira de tela de la parte posterior del calzado, sería como tirar de ahí para levantarse a uno mismo). Consiste en que las observaciones se reelaboran sustituyendo los datos entre distintas muestras, de forma que se crean 'pseudomuestras' de las que se derivan los intervalos de referencia. Se repite este proceso muchas veces (1000-2000) y se obtiene una distribución de límites de referencia derivados. En esa distribución los percentiles 5 y 95 pueden usarse para determinar los límites superior e inferior del IC 90 %. El problema es que serán muy amplios si la muestra es pequeña, por lo que se requerirán al menos 80 individuos para obtener IC de tamaño razonable. [39]

### 7.4.- TAMAÑO DE LA MUESTRA Y ERROR ESTÁNDAR DE LOS LÍMITES DE REFERENCIA.

La elección del tamaño de la muestra se realiza en función de cuánta es la precisión con la que se pretende estimar los límites de referencia. La recomendación es que se deben calcular con al menos 120 individuos. En un estudio se simuló una distribución normal con una media de 0.0 y una desviación estándar de 1.0 y se calcularon los intervalos de referencia con muestras generadas aleatoriamente desde 40 a 2000 individuos. Para el cálculo se emplearon tres aproximaciones. La primera (1) era la paramétrica, que los calculaba como la media  $\pm 1.96$  Desviaciones estándar. La segunda (2) era paramétrica previa transformación potencial con la fórmula modificada Box-Cox (que se usaría para acercar los datos a una distribución normal). Por último (3) se empleó la opción recomendada por la IFCC de usar métodos no paramétricos, calculando el rango entre los percentiles 2'5 y 97'5. [38]

#### INFLUENCIA DEL TAMAÑO DE LA MUESTRA EN LA AMPLITUD DE LOS INTERVALOS DE CONFIANZA DE LOS IIRR.

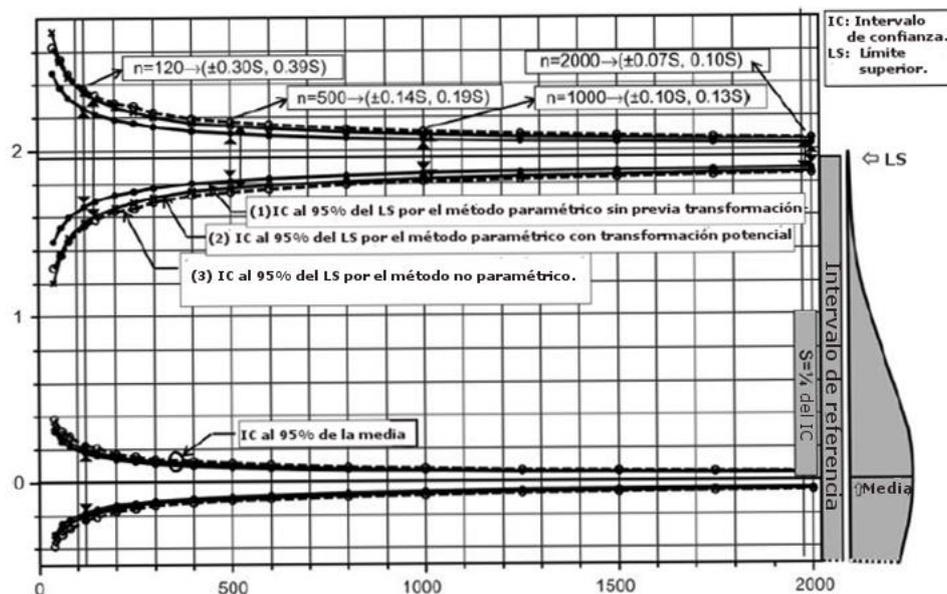


Figura 6. Tomada de [36]

Como se esperaba, el primer método dio IC más estrechos en cualquier tamaño de la muestra. Debe tenerse en cuenta que en la práctica suele ser necesaria la transformación de los datos, que ya conlleva un aumento de la incertidumbre casi equiparable al del método no paramétrico. Por eso se recomienda este último. Una posible ventaja del método paramétrico sería su capacidad para identificar y excluir valores extremos basándose en la media y la desviación estándar de la distribución durante el análisis.

De esta simulación podemos deducir también que la muestra de 120 individuos, descrita como la mínima necesaria para un estudio de intervalos de referencia, produce un intervalo de confianza cuya mitad es igual a 0'39 veces la desviación estándar, mientras que si se emplean 400 individuos la mitad del intervalo de confianza se reduce a 0'2 veces la desviación estándar (un 5% del total del intervalo de referencia). Esta cantidad sería más recomendable en grandes estudios colaborativos realizados conjuntamente por varios laboratorios.

### 7.5.- CRITERIOS DE ESTRATIFICACIÓN O PARTICIÓN.

La estratificación o partición es la creación de dos o más subgrupos dentro de una misma variable por la existencia de un factor causal que los condiciona.

En un estudio multicéntrico sobre proteínas séricas se obtuvieron valores significativamente más bajos en la concentración de IgG en fumadores, tal y como era ya conocido. Sin embargo había también asociación entre el género masculino y unas concentraciones menores. [38]

#### ASOCIACIÓN POR SEPARADO ENTRE HÁBITO TABÁQUICO Y GÉNERO Y CONCENTRACIÓN DE IGG.

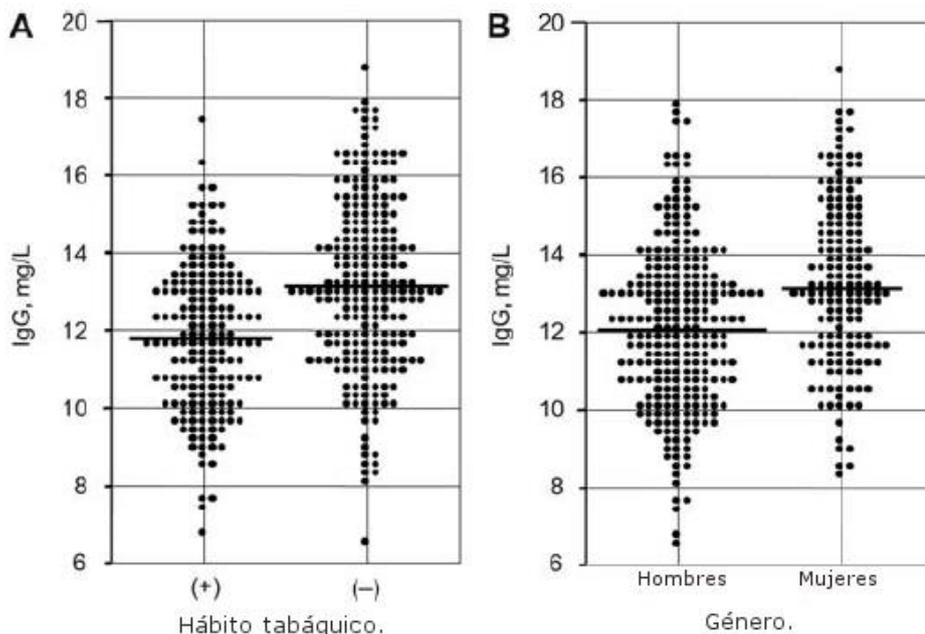


Figura 7. Tomada de [38]

El hábito tabáquico actuaba como factor de confusión, distorsionando el grado de asociación entre género y concentración en suero de IgG. Al haber una mayor proporción de fumadores entre los hombres que entre las mujeres, parecía que las concentraciones de IgG de estos fuesen menores. Estratificando por hábito tabáquico se aprecia que no hay diferencia entre sexos.

ASOCIACIÓN ENTRE CONCENTRACIÓN DE IGG Y SEXO DESPUÉS DE ESTRATIFICAR POR HÁBITO TABÁQUICO.

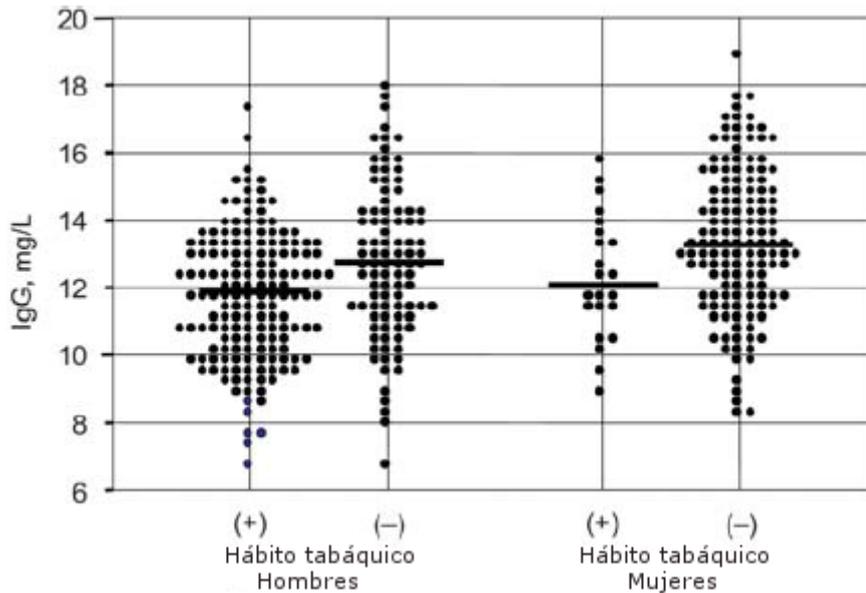


Figura 8. Tomada de [38]

De forma intuitiva, se podría plantear la necesidad o no de estratificación buscando significación estadística en la diferencia entre las dos medias de las subclases. Pero este planteamiento lleva a separar en exceso en diferentes subclases. Identifica diferencias estadísticamente significativas, pero clínicamente irrelevantes; en particular cuando el número de sujetos de las subclases es grande.

El método de Harris y Boyd es aplicable cuando hay dos subgrupos. Se basa en métodos para identificar diferencias entre dos conjuntos de datos, pero introduce adaptaciones en función del tamaño de muestra. Asume que el número de individuos y las desviaciones estándar de los subgrupos son iguales. Al final, la diferencia crítica es de 0.38 veces la desviación estándar o lo que es lo mismo, un cuarto del intervalo de referencia obtenido tras haber separado los grupos.

Una vez establecido este criterio, se centraron en los grupos de individuos fuera de los intervalos de referencia y sugirieron que el problema se reducía a saber si ese límite común de ambos posibles subgrupos se aproximaba razonablemente al que tendrían cada uno por separado. Propusieron que si el percentil de las subclases que se correspondía con el valor del límite de referencia era mayor que 4 % o menor de 1 % era necesario definir límites de referencia diferentes. Este método en principio solo era válido para distribuciones normales, pero Lahti et al. han propuesto un método basado en conceptos similares que pueda funcionar en cualquier situación. [12]

## 7.6.- DETECCIÓN Y EXCLUSIÓN DE VALORES ANORMALES.

Llamamos valores anormales o *'outliers'* a las observaciones cuya discordancia con respecto a la mayoría de la muestra es excesiva en relación al modelo de distribución que se asume, lo cual lleva a la sospecha de que son valores que no se han generado por ese modelo. Su presencia puede tener una gran influencia en los límites. [40]

Un método simple y eficaz para su detección es la inspección visual de la distribución de los datos. Si se detecta un valor anormal y no hay motivos claros para descartarlo como condiciones del sujeto, problemas analíticos, de cálculo o de transcripción; es útil aplicar criterios estadísticos para justificar su exclusión.

El método más intuitivo es el que propuso Dixon en 1953, que se basa en el cociente D/R. D es el valor absoluto de la diferencia entre el valor extremo y el siguiente (si es mínimo) o el precedente (si es máximo). R representa el rango entero de la distribución. La IFCC propone 1/3 como límite máximo de esta ratio, a partir de este valor se debe realizar el descarte. El problema es que puede haber más de un valor anormal. Estos valores menos extremos pueden enmascarar al primero. Se aconseja comprobar si es anormal el valor menos extremo y si lo es, descartar a la vez el más extremo.

Horn et al. propusieron un método algo más sofisticado. Usaban un algoritmo de dos pasos. Primero transformaban los datos por el método de Box y Cox a una distribución normal y luego identificaban valores anormales por la aproximación robusta de Tukey. Esto consiste en identificar los extremos usando solo el 50 % central de la distribución, lo que elimina la influencia de otros valores anormales. El percentil 25 es el primer cuartil ( $Q_1$ ) y el percentil 75 es el tercer cuartil ( $Q_3$ ). El rango intercuartílico (IQR) es la diferencia entre los dos. Los límites inferior y superior para descartar los valores que los excedan se establecen como  $Q_1 - 1.5 \times IQR$  y  $Q_3 + 1.5 \times IQR$ . Todo valor que los exceda es descartado.

No siempre es una buena opción descartar un valor solo por su posición en la distribución. Una alternativa es el método de exclusión del valor anormal latente (EVAL). Primero deriva unos intervalos de referencia provisionales de cada una de las variables teniendo en cuenta todos los valores obtenidos. Luego utiliza de forma repetida métodos para detectar si hay correlación entre cada variable y todas las demás. Hecho esto, excluirá para el cálculo de los intervalos de referencia definitivos los valores obtenidos en muestras que hayan salido fuera de los límites provisionales en las variables correlacionadas.

La ventaja de este método es que no se produce un corte en la distribución de los datos. Sin embargo, si no hay correlación entre distintas variables o esta es débil, este algoritmo reducirá el número de datos disponibles. [12]

## 8.- PRESENTACIÓN DE LOS VALORES OBSERVADOS EN RELACIÓN A LOS VALORES DE REFERENCIA.

### 8.1.- VALORES DE REFERENCIA.

A la hora de presentar unos valores de referencia para un parámetro dado en un laboratorio, se nos presentan dos opciones iniciales.

- Si hubiera pocos valores de referencia disponibles, una opción podría ser incluir todos los valores disponibles. Por lo general esto no es muy práctico y casi nunca se hace.
- La opción más habitual es acotar los valores de referencia deseables (el 95 % central) dentro de un intervalo de referencia. Si la dispersión de los valores es grande el IR tenderá a ser más amplio, si los valores son homogéneos será más estrecho.

### 8.2.-TRANSFORMACIÓN DE LOS VALORES.

Por otro lado, en algunos casos puede ser práctico transformar el valor observado para mostrar su relación con los valores de referencia.

- Intervalos de referencia. Dado que contamos con dos límites, podemos dividir los resultados en tres categorías:
  - Valor por debajo del límite inferior o inusualmente bajo.
  - Valor habitual.
  - Valor por encima del límite superior o inusualmente alto.

Una posibilidad es destacar un valor inusual con un signo determinado, generalmente un asterisco. Esta caracterización en tres clases deriva en pérdida de información. Debe recordarse que un valor cualquiera que estuviera dentro del IR pudiera también ser un valor raro en su distribución patológica debido a la variación intraindividual o al contrario. Ha de tenerse en cuenta que los valores predictivos de los IR no son totales. Un valor mucho más elevado que el límite superior será más indicativo de patología que uno que esté un poco elevado. Es por comodidad y sencillez lo que se realiza actualmente.

- Valor observado menos la media aritmética, dividido para la desviación estándar. En una distribución normal, un IR del 95% quedaría limitado con este indicador entre  $-1.96$  y  $+1.96$ . No se puede recomendar de forma genérica esta transformación porque daría lugar a confusión si la distribución de los resultados no fuera normal.
- Fractil. Consiste en indicar el orden que ocuparía el valor observado en la distribución de valores de referencia. Si es inferior al valor mínimo o superior al máximo del total de la distribución de referencia, no se podrá calcular. [11]

## **9.- ARMONIZACIÓN Y ESTANDARIZACIÓN. PROCESO GLOBAL DE UNA PRUEBA.**

Los intervalos de referencia son una herramienta para la interpretación de los resultados del laboratorio. A la par, en su determinación dependen del funcionamiento del laboratorio donde se hayan obtenido, su forma de realizar las pruebas. Los intervalos se verán condicionados por las mismas variables que puedan influir sobre cualquier determinación que allí se realice.

Surge, por tanto, la necesidad de armonizar los procedimientos y estandarizar los resultados, para permitir hacer los resultados de distintos laboratorios más comparables entre sí. La actividad del laboratorio debe ser llevada a cabo de forma global.

### **9.1.- PROCESO GLOBAL DE UNA PRUEBA.**

El proceso global de una prueba (total testing process, TTP) tiene distintas etapas diferenciadas que exceden a la determinación numérica de un valor dentro de un sistema de medida, que es la fase analítica.

Según describió Lundberg en un editorial en 1981, en una prueba de laboratorio tiene lugar un bucle de cerebro a cerebro, de la persona que decide que es necesario el análisis a quien lo debe interpretar y toma una decisión en consecuencia.

El proceso comienza con la petición de la prueba por parte de un clínico. Se sigue la recolección de la muestra, su transporte y su separación (o preparación). Estas son las fases preanalíticas. Después de realizar el análisis habrá una elaboración de informes y una interpretación de los resultados por parte del médico solicitante que dará lugar a una acción. Esto es la fase postanalítica. En distintas etapas deberá identificarse el paciente, la muestra o el informe con el resultado.

El objetivo final es que la acción tomada por parte del médico basándose en la información proporcionada por el laboratorio sea adecuada, oportuna en el tiempo y segura. [41]

### **9.2.- ARMONIZACIÓN.**

El diccionario de Oxford define armonizar como hacer consistente o compatible. La armonización en el laboratorio es el proceso de reconocer, comprender y explicar las diferencias en todas las fases de la realización de una prueba (TTP). Aunque su objetivo inicial era solo la fase analítica, el objetivo de la armonización incluye también aspectos como terminología y unidades, formas de proporcionar la información, intervalos de referencia y límites de decisión. A esto se le llama el enfoque global de la armonización.

No deja de ser una forma de sistematizar. Su finalidad es establecer equivalencias entre valores obtenidos a través de procedimientos de medida diferentes sobre el mismo elemento a medir. Busca, por tanto, conseguir la uniformidad de la información que proporciona el laboratorio, y como mínimo, es un medio de identificar y divulgar políticas y procedimientos comunes con el fin de que grupos de investigación distintos puedan usar los datos procedentes de laboratorios diferentes de forma intercambiable. Es una oportunidad para compartir resultados. [42]

La armonización de una medida puede conseguirse a través de dos métodos: la estandarización y la armonización empírica.

### 9.2.1 - La estandarización.

La estandarización es la realización de la definición de una cantidad dada, con un valor de cantidad y una incertidumbre en la medida asociada, que se usan como referencia. El término cantidad se refiere a una propiedad medible de una sustancia o material.

En el Sistema Internacional se emplean unidades básicas como el Kg para la masa o los metros para la longitud; o derivadas en caso de que se obtengan a partir de ecuaciones con unidades básicas, por ejemplo  $m^3$  para el volumen.

La realización de una definición se refiere al establecimiento por convención de un material o fenómeno físico, o una constante, que por definición o acuerdo es el estándar último.

El prototipo del Kg es un cilindro de Iridio-Platino que se encuentra en Sèvres, Francia. Los estándares de menor orden serán réplicas de ese estándar último.

Calibrar consiste en comparar los valores obtenidos por un instrumento de medición con la medida correspondiente de un patrón de referencia (o estándar). [43]

La trazabilidad metrológica es la propiedad de un resultado que le permite relacionarse con una referencia o estándar a través de una cadena ininterrumpida de calibraciones, cada una de las cuales contribuye a la incertidumbre de la medida. Para que un elemento sea trazable metrológicamente, se deberá asegurar la validez de los sistemas de medida con una clara definición de las relaciones entre el procedimiento de medida de rutina y su calibración con el estándar de referencia más alto.

**EL SISTEMA DE MEDIDA DE REFERENCIA. NIVELES DE TRAZABILIDAD.**

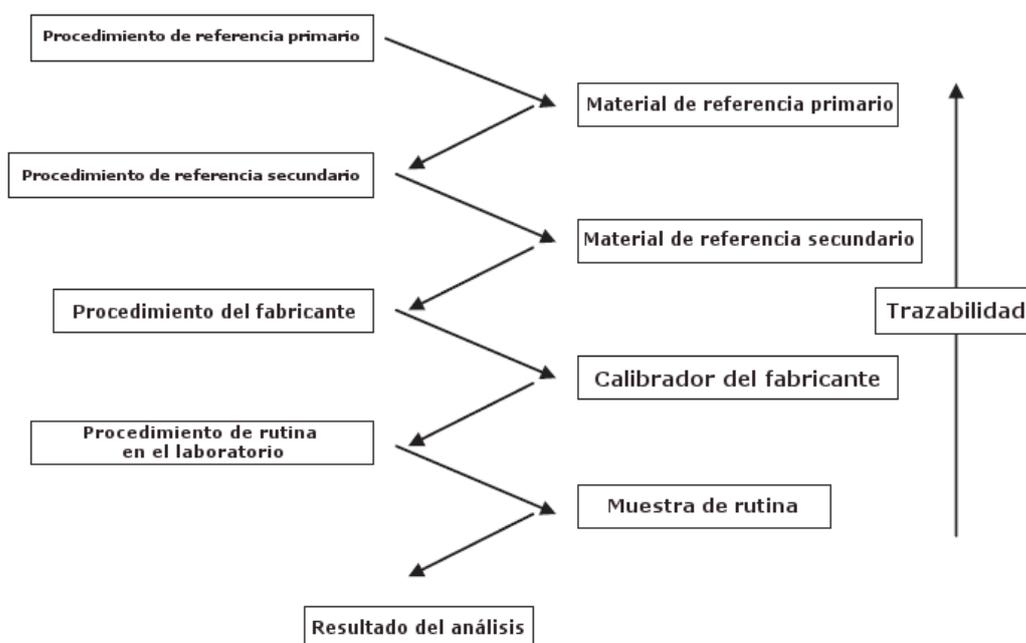


Figura 9. Tomada de [44]

Si el nivel de estandarización es bueno, solo las diferencias entre las poblaciones justifican diferentes intervalos de referencia. [45]

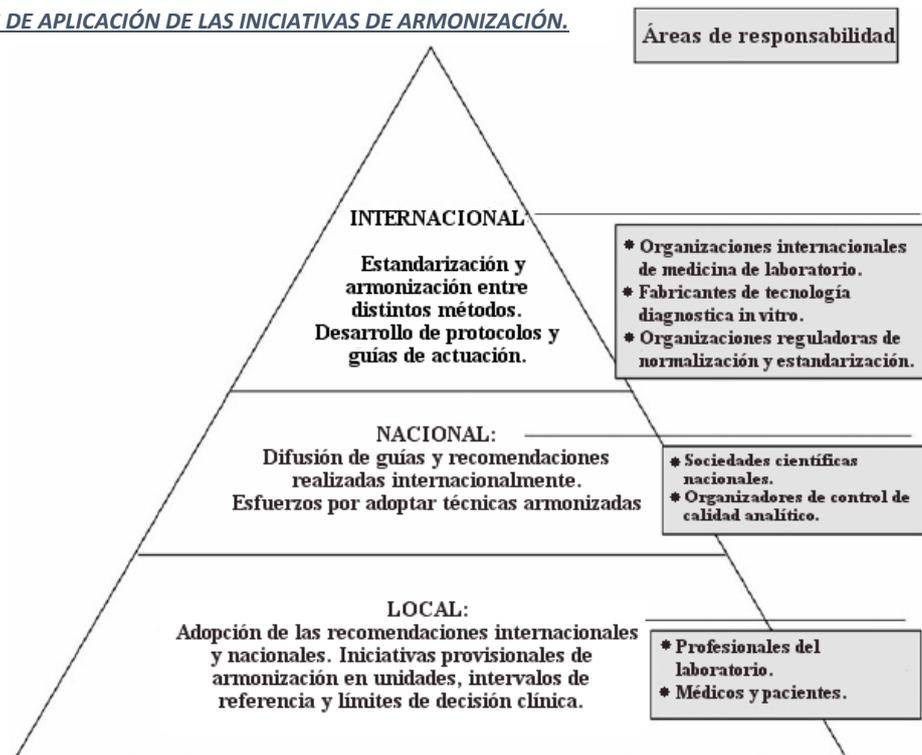
### 9.2.2.- Armonización empírica.

Este segundo método, se basa en la aplicación de un factor corrector u otra estrategia de manipulación de los datos obtenidos, para así eliminar su disparidad. Se lleva a cabo la comparación de dos procedimientos de medida que se encuentran en la parte más baja de la jerarquía metrológica, comparando paneles compartidos de muestras. Esta armonización empírica ha sido empleada para la TSH o la insulina.

### 9.3.- AVANCES EN LA ARMONIZACIÓN.

La armonización en el proceso global de la prueba es un proceso que dependerá mucho del éxito de implantación de las recomendaciones que se den. Y este, a su vez, dependerá de que la armonización sea percibida como necesaria no solo por parte de los técnicos de laboratorio, sino también de médicos y pacientes. Las iniciativas de armonización se basarán en tres niveles de aplicación: [42]

#### NIVELES DE APLICACIÓN DE LAS INICIATIVAS DE ARMONIZACIÓN.



La armonización en la medicina de laboratorio: niveles y áreas de responsabilidad.

Figura 10. Tomada de [42]

Un ejemplo de iniciativa de armonización a nivel local sería el documento que se adjunta en el Anexo 1, en el que una médico de atención primaria preparaba una serie de instrucciones para uniformar la preparación de sus pacientes para los análisis sanguíneos.

## 10.- INTERVALOS DE REFERENCIA COMUNES.

Consiste en el establecimiento y adopción de unos IR compartidos por diferentes laboratorios. Si el método analítico es el mismo o proporciona resultados comparables por estar correctamente estandarizado y la población tiene características similares o se sabe que una magnitud específica no depende significativamente del ambiente o la etnicidad, entonces se pueden usar IR comunes. Para su correcta aplicación en la práctica se deben cumplir una serie de requisitos. [46]

### **PRERREQUISITOS PARA EL USO DE INTERVALOS DE REFERENCIA COMUNES.**

Tipo de requisito.	Requisito.	Responsabilidad.
Analítico	Existencia de un sistema de medida de referencia.	IFCC, JCTLM, Institutos nacionales e internacionales de metrología.
	Existencia de métodos de rutina trazables.	Fabricantes.
	Implementación correcta de estos en los laboratorios clínicos.	Laboratorios clínicos.
	Control de calidad de los métodos de rutina para mantener los niveles de incertidumbre dentro de unos límites.	Laboratorios clínicos. Organizadores de controles de calidad (EQAS)
Clínico.	Definición adecuada de los intervalos de referencia, dando información sobre la influencia de factores biológicos y ambientales.	Esfuerzo conjunto – IFCC, fabricantes y laboratorios clínicos.
	Fase preanalítica comparable.	Laboratorios clínicos.
	Validación de la aplicabilidad de los intervalos de referencia comunes a la población particular de cada laboratorio.	Laboratorios clínicos.
	Adopción de los intervalos de referencia comunes.	Laboratorios clínicos.

IFCC es la Federación Internacional de Química Clínica.

JCTLM es la Junta-Comité para la Trazabilidad en la Medicina de Laboratorio.

EQAS son Esquemas de Valoración Externa de la Calidad.

*Tabla 2. Tomada de [46].*

La transferencia de un IR es el proceso por el cual se adapta un IR previamente establecido a otro procedimiento de medida o a una población diferente. No es necesario recoger muestras de individuos de referencia y, simplemente, se pueden usar muestras de individuos de los que no se disponga información sobre su estado de salud.

La revisión de un IR es el proceso por el cual se confirma que un IR puede ser transferido, con una razonable confianza, usando un número relativamente pequeño de individuos de referencia. [47]

## **10.1.- PROCEDIMIENTO PARA ESTABLECER INTERVALOS DE REFERENCIA COMUNES.**

Asumiendo un grupo de laboratorios con similares condiciones preanalíticas y analíticas, la tarea más exigente será la de definir una muestra adecuada de referencia. Deberá incluir sujetos con características distintas representativas de la población que se estudia. Una buena opción es realizar un estudio multicéntrico que implique laboratorios clínicos en diferentes regiones o países.

El número de individuos requeridos para cada grupo deberá ser de unos 500, con un mínimo de 120 para el cálculo no paramétrico. Debe haber una definición clara de toda la fase preanalítica. Idealmente los análisis se deberían hacer sobre muestras frescas para reproducir las condiciones habituales de los laboratorios. Sin embargo, para reducir la variabilidad analítica puede ser útil congelar las muestras y analizarlas de forma centralizada en un laboratorio de referencia, siempre que se haya demostrado que la congelación no afecta al resultado.

Es necesario usar métodos que den resultados trazables al sistema de medida y que haya una alta estandarización interlaboratorial. La comparabilidad entre laboratorios debe asegurarse a través de materiales de control de calidad comunes. Con un sistema de control de calidad interno se definirán a priori criterios para la aceptación o el rechazo de cada serie analítica.

Respecto al análisis de los datos, se han de comparar los datos de los distintos centros para identificar posibles sesgos analíticos o distribuciones atípicas. Después deben descartarse los valores anormales.

## **10.2.- ADOPCIÓN DE INTERVALOS DE REFERENCIA COMUNES.**

Se recomienda que los laboratorios clínicos validen los intervalos de referencia que van a adoptar con un pequeño grupo de muestra procedente de su propia población. Así se revisa si unos valores de referencia son adecuados, es decir similares en cuanto a sus condiciones preanalíticas, el método analítico usado y las características de la población a la que se da servicio.

Para adecuar las condiciones preanalíticas es suficiente con adoptar procedimientos similares en caso de que sea necesario (si no se ha demostrado que cambios en estas condiciones modifican los resultados).

El método analítico debe proporcionar resultados trazables a un sistema de medida de referencia para la determinación requerida. En países europeos, si el reactivo tiene el certificado de 'CE' (marca de conformidad europea) y se usa siguiendo las especificaciones del fabricante, esto debería ser así.

Para evaluar si las características de las poblaciones son similares se aconseja examinar a 20 individuos que cumplan los criterios de selección de individuos sanos. Después de descartar valores anormales, si no más de dos valores quedan fuera del intervalo de referencia que se pretende adoptar, puede hacerse. Si tres o más valores salen fuera del intervalo de referencia propuesto, se repetirá la prueba con otros 20 sujetos. Si dos o menos valores quedan fuera, los

intervalos se pueden adoptar, si no es así, probablemente querrá decir que las poblaciones son diferentes para ese parámetro y que se necesita un intervalo de referencia específico.

Esto es útil si la distribución es normal, pero si no está centrada deberían llevarse a cabo análisis estadísticos más potentes como el de Kolmogorov-Smirnov, que compara la batería de datos de los individuos analizados para determinar los límites de referencia con los valores de los 20 sujetos usando un algoritmo robusto.

Una alternativa puede ser la selección de análisis históricos por métodos indirectos para compararlos con los límites de referencia que se proponen. Así se podrá juzgar su aplicabilidad. [46]

## 11.- SITUACIÓN ACTUAL Y PERSPECTIVAS FUTURAS DE LOS INTERVALOS DE REFERENCIA.

### 11.1.- NORMATIVA Y SITUACIÓN ACTUAL.

La Organización Internacional por la Estandarización (ISO) determina que “los intervalos de referencia biológicos deberán ser periódicamente revisados”, en particular cuando se verifiquen variaciones de técnica analítica o preanalítica o si se sospecha que el intervalo ha dejado de ser adecuado para la población. [49]

Esto es difícil de satisfacer por el gran número de análisis que se desarrollan y su variación casi cotidiana. La directiva europea 98/79 sobre los diagnósticos in vitro obliga a los fabricantes a establecer en el manual de instrucciones “los intervalos de referencia de las cantidades que se determinan, con inclusión de una descripción de la población de referencia que deba considerarse”. Para mejorar la comparabilidad entre distintos métodos establece que “la correlación de los valores asignados a los calibradores o a los materiales de control se garantizará mediante procedimientos de medida de referencia disponibles o materiales de referencia disponibles de grado superior.” [50]

En las recomendaciones originales se aconsejaba que cada laboratorio determinara sus propios intervalos de referencia a partir de una muestra de al menos 120 individuos sanos, seleccionados previamente. [51] Este es un método costoso y que consume mucho tiempo. Muchos laboratorios no pueden llevarlo a cabo. Más aún si se debe repetir con cada modificación en la técnica analítica. Por ello se ha definido una serie de alternativas como la determinación por medio de estudios multicéntricos y su posterior validación en cada laboratorio o la selección indirecta de individuos sanos a posteriori. Aun así, hay que tener cuidado con la adopción acrítica de intervalos de referencia comunes. [52]

Pero interesa saber cuál es la situación sobre el terreno. En 2007 se publicó un estudio que evaluaba a través de un cuestionario el origen y el grado de concordancia de los IR de 163 laboratorios que proporcionaron sus datos en EEUU principalmente. Se les pidieron sus límites superior e inferior de referencia de siete valores: potasio, calcio total, magnesio, TSH, hemoglobina (Hb), conteo plaquetario y tiempo de tromboplastina parcial activada (aPTT); tanto para adultos como para pacientes pediátricos. En el caso de la hemoglobina se recogieron datos también para mujeres adultas y niños que precisaron ingreso hospitalario.

Además para cada elemento a medir se recogió la unidad de medida, el tipo de muestra (suero o plasma), instrumento analítico, fabricante, año en que se estableció originalmente el IR, año de la última validación del IR y año en que se puso en funcionamiento el instrumento [53]

Como resultados se obtuvo que en el 80% central de IR, la variación era ligera, pero algunos laboratorios fuera de este grupo tenían valores sorprendentemente bajos o altos. Llegó a ocurrir, como ejemplo extremo, que un laboratorio considerara una cifra de Hb de 13'8 g/dL como baja mientras que otro consideraba el mismo valor alto. Dicho de otro modo, no había ningún solapamiento entre esos IR tan extremadamente bajo uno y alto el otro.

**VARIACIÓN EN LOS INTERVALOS DE REFERENCIA EN 161 LABORATORIOS, ESTUDIO Q-PROBES.**

Determinación, población y límite.	n	Percentiles de todos los laboratorios						Máximo
		Mínimo	10	25	50 (Mediana)	75	90	
Potasio, mmol/L								
Adultos, bajo	162	3.0	3.3	3.5	3.5	3.6	3.6	4.1
Adultos, alto	162	4.5	5.0	5.0	5.1	5.2	5.3	5.7
Pediátrico, bajo	144	3.0	3.3	3.5	3.5	3.6	3.6	4.1
Pediátrico, alto	144	4.1	5.0	5.0	5.1	5.2	5.4	5.6
Calcio (total), mg/dL								
Adultos, bajo	162	7.6	8.3	8.4	8.5	8.5	8.8	9.0
Adultos, alto	161	9.6	10.1	10.1	10.2	10.5	10.5	10.8
Pediátrico, bajo	145	7.6	8.4	8.4	8.5	8.8	9.0	9.6
Pediátrico, alto	145	9.6	10.1	10.2	10.4	10.7	10.8	11.5
Magnesio, mg/dL								
Adultos, bajo	157	1.3	1.5	1.6	1.6	1.8	1.8	2.1
Adultos, alto	157	2.0	2.2	2.3	2.4	2.6	2.7	3.2
Pediátrico, bajo	141	1.2	1.5	1.6	1.6	1.8	1.8	2.2
Pediátrico, alto	141	2.0	2.1	2.3	2.4	2.5	2.7	3.2
Tirotropina (TSH), mIU/L								
Adultos, bajo	154	0.10	0.30	0.34	0.35	0.40	0.47	0.50
Adultos, alto	155	3.00	4.20	4.68	4.94	5.50	5.60	6.00
Pediátrico, bajo	133	0.10	0.30	0.34	0.35	0.46	0.50	0.70
Pediátrico, alto	136	4.00	4.40	4.75	5.00	5.50	5.90	8.00
Hemoglobina (hombres), g/dL								
Adultos, bajo	161	11.5	13.0	13.0	13.6	14.0	14.0	14.8
Adultos, alto	161	15.0	16.5	17.0	17.5	18.0	18.0	18.1
Pediátrico, bajo	148	8.5	10.5	11.0	11.5	11.7	12.5	14.0
Pediátrico, alto	148	12.7	13.7	14.5	15.5	15.5	16.3	20.0
Hemoglobina (mujeres), g/dL								
Adultos, bajo	159	10.4	11.5	11.7	12.0	12.0	12.1	14.0
Adultos, alto	159	13.5	15.0	15.5	16.0	16.0	16.0	18.0
Pediátrico, bajo	148	9.0	10.5	11.0	11.5	11.5	12.0	13.0
Pediátrico, alto	148	12.7	13.7	14.5	15.0	15.5	16.0	18.0
Plaquetas, ×10 <sup>9</sup> /L								
Adultos, bajo	162	100	130	130	142	150	150	174
Adultos, alto	161	316	400	400	400	440	450	500
Pediátrico, bajo	146	100	130	140	150	150	184	242
Pediátrico, alto	143	325	392	400	400	450	450	539
Tiempo de tromboplastina parcial activada, s								
Adultos, bajo	156	17	21	22	23	25	26	32
Adultos, alto	159	28	32	33	35	37	38	43
Pediátrico, bajo	135	17	21	22	23	25	26	29
Pediátrico, alto	138	28	32	33	35	37	38	43

\* 'Bajo' indica el límite inferior del intervalo de referencia, 'alto' indica el límite superior del intervalo de referencia.

Figura 11. Tomada de [53]

Aproximadamente la mitad de los participantes dijo llevar a cabo estudios internos en individuos sanos para establecer IR químicos y hematológicos en adultos, mientras que la otra mitad confiaba en fuentes externas (indicaciones de fabricantes, literatura publicada, intervalos de otros laboratorios o recomendaciones médicas). Un 84 % refería recibir algún tipo de asistencia por parte de los fabricantes para establecer los IR, bien en forma de apoyo con la estadística, asesoramiento en las distintas etapas o incluso recibiendo las muestras de individuos sanos ya seleccionados.

No había una diferencia apreciable entre los IR de laboratorios que tomaban directamente los valores proporcionados por los fabricantes con los que hacían análisis sobre sujetos sanos para validar los IR proporcionados. Tampoco existían diferencias con los que calculaban sus propios IR.

**ORIGEN DE LOS DISTINTOS INTERVALOS DE REFERENCIA, ESTUDIO Q-PROBES.**

	Laboratorios, Nº, (%). n= 162							
	Potasio	Calcio	Magnesio	TSH	Hgb (Hombres)	Hgb (Mujeres)	Plaquetas	aPTT
<b>Adultos.</b>								
Estudio interno de individuos sanos.	70 (44.3)	70 (43.8)	70 (44.6)	75 (48.4)	85 (53.5)	83 (52.5)	81 (51.3)	130 (82.3)
Recomendaciones del fabricante.	55 (34.8)	56 (35.0)	57 (36.3)	46 (29.7)	17 (10.7)	18 (11.4)	19 (12.0)	12 (7.6)
Literatura publicada/libros de texto.	19 (12.0)	17 (10.6)	16 (10.2)	12 (7.7)	43 (27.0)	43 (27.2)	44 (27.8)	6 (3.8)
Otros laboratorios (con val. interna)	8 (5.1)	9 (5.6)	8 (5.1)	10 (6.5)	8 (5.0)	8 (5.1)	8 (5.1)	5 (3.2)
Recomendaciones médicas.	2 (1.3)	4 (2.5)	2 (1.3)	4 (2.6)	3 (1.9)	3 (1.9)	3 (1.9)	2 (1.3)
Otros laboratorios (sin validación)	0 (0)	0 (0)	0 (0)	2 (1.3)	1 (0.6)	1 (0.6)	1 (0.6)	2 (1.3)
Otros métodos.	4 (2.5)	4 (2.5)	4 (2.5)	6 (3.9)	2 (1.3)	2 (1.3)	2 (1.3)	1 (0.6)
<b>Pacientes pediátricos.</b>								
Estudio interno de individuos sanos.	35 (25.2)	34 (24.1)	36 (25.9)	21 (15.7)	35 (23.8)	34 (23.0)	34 (23.8)	85 (63.0)
Recomendaciones del fabricante.	45 (32.4)	42 (29.8)	45 (32.4)	59 (44.0)	13 (8.8)	14 (9.5)	15 (10.5)	11 (8.1)
Literatura publicada/libros de texto.	42 (30.2)	44 (31.2)	43 (30.9)	33 (24.6)	75 (51.0)	75 (50.7)	68 (47.6)	24 (17.8)
Otros laboratorios (con val. interna)	7 (5.0)	8 (5.7)	6 (4.3)	7 (5.2)	7 (4.8)	8 (5.4)	8 (5.6)	5 (3.7)
Recomendaciones médicas.	3 (2.2)	5 (3.5)	3 (2.2)	5 (3.7)	4 (2.7)	4 (2.7)	4 (2.8)	2 (1.5)
Otros laboratorios (sin val. interna)	1 (0.7)	2 (1.4)	1 (0.7)	4 (3.0)	10 (6.8)	10 (6.8)	9 (6.3)	4 (3.0)
Otros.	6 (4.3)	6 (4.3)	5 (3.6)	5 (3.7)	3 (2.0)	3 (2.0)	5 (3.5)	4 (3.0)

\* TSH indica tirotropina u hormona estimulante de tiroides. aPTT, tiempo de tromboplastina parcial activado.

Figura 12. Tomada de [53]

## 11.2.- ¿DE QUÉ FORMA PUEDE DETERMINAR SUS INTERVALOS DE REFERENCIA UN LABORATORIO?

En base a la norma ISO 15189, todos los laboratorios que quieran acreditarse deberán demostrar que los IR comunicados a todos los usuarios del servicio son apropiados para la población de referencia y para los pacientes a los que se atiende, al igual que adecuados a los sistemas de medida empleados. Esta demostración se realizará periódicamente.

Esta demostración se puede llevar a cabo de distintas maneras, en función del origen de los IR biológicos, la transferibilidad de estos valores y la información metrológica y preanalítica de que se dispone. Se ordenan las opciones de mayor a menor rigor en el cumplimiento de las recomendaciones.

- Primera opción: Producción propia de valores de referencia biológicos y estimación del intervalo de referencia biológico, según la IFCC, con estimación de la imprecisión interdiaria y el sesgo del sistema de medida.
- Segunda opción: Producción de valores de referencia biológicos por el mismo laboratorio clínico conjuntamente con otros laboratorios que usan el mismo sistema de medida, siguiendo un diseño multicéntrico y estimación del IR biológico común según la IFCC, con estimación de la imprecisión interdiaria y el sesgo del sistema de medida.
- Tercera opción: Adopción, tras su validación, de un IR biológico estimado, según las recomendaciones de la IFCC, por otro laboratorio clínico, o por varios laboratorios siguiendo un diseño multicéntrico.

- Cuarta opción: Adopción, sin validación previa, de un IR biológico estimado por otro laboratorio clínico según las recomendaciones de la IFCC, siempre y cuando las poblaciones biológicas sean muy parecidas y las propiedades metrológicas de ambos sistemas de medida sean aproximadamente las mismas.
- Quinta opción (que debería evitarse): Adopción, sin validación previa de un IR biológico hallado en una publicación sin ninguna descripción de la población de referencia ni de las propiedades metrológicas del sistema de medida. Esta opción es la que se da en muchos laboratorios clínicos, al adoptar límites de referencia biológicos de un libro o de un folleto comercial. Esta opción podría adoptarse temporalmente, durante el periodo inicial de la acreditación.

Respecto a la frecuencia con la que los IR deberían revisarse, no existe ninguna recomendación al respecto. El autor que desarrolla esta lista de alternativas opina que puede hacerse anualmente. [48]

Se adjuntan como Anexo 2 diversos resultados de analíticas llevadas a cabo en laboratorios de la comunidad autónoma de Aragón.

### **11.3.- RETOS Y MARGEN DE MEJORA EN EL FUTURO.**

Los intervalos de referencia puedan parecer una teoría ya terminada y sin repercusiones en el futuro de la medicina. Es cierto que tienden a ser sustituidos por límites de decisión clínica. Sin embargo, serán siempre útiles para estudiar cómo va cambiando la sociedad y así se podrán adecuar a ellos los límites de decisión. Además son un reflejo de cómo funciona un laboratorio, estarán condicionados por la población a la que el laboratorio presta servicios, las técnicas que emplea y la forma que tiene de ponerlas en práctica. Todo esto por no hablar de que son la herramienta más utilizada hoy en día en la interpretación de datos de laboratorio.

En el futuro, si realmente queremos llegar a una medicina de laboratorio personalizada, en la que se pueda comparar al paciente con otras personas con sus mismas circunstancias (o muy similares) a través de enormes bases de datos; primero deberemos poner en práctica las recomendaciones que nos permitan obtener unos resultados del laboratorio 'universales' y comparables entre sí. Para ello habrá que adoptar en la práctica diaria unos procedimientos comunes en relación a aspectos que podrían condicionar los resultados, a lo largo de lo que se conoce como el proceso global de la prueba. Sin esa comparabilidad entre distintos resultados, no se podrá ir más allá de los actuales intervalos de referencia determinados de forma poco uniforme en cada laboratorio.

En ancianos o niños los IR son más difíciles de calcular. En el primer caso por la elevada comorbilidad que dificulta encontrar individuos estrictamente sanos. En los grupos pediátricos influye no solo la cambiante fisiología a lo largo del desarrollo, sino además el problema ético de tener sacar sangre a un recién nacido o a un niño sano. [54]

Otro posible desarrollo que puede implementarse en el futuro puede ser la estratificación de los IR usando información genética. Es decir, integrar los resultados de la secuenciación de genes con la clínica y los resultados de laboratorio obtenidos para cada individuo, permitiendo calcular unos IR de la población que tenga una mutación o alteración genética dada.

Como ejemplo en un estudio sobre este tema, evalúan la posibilidad de estratificar los valores de bilirrubina en individuos con la mutación propia del síndrome de Gilbert (enzima UDP-glucuronosiltransferasa, codificada por el gen UGT1A1, y frecuentemente asociada con el alelo UGT1A1\*28, un 5% de cuyos individuos homocigotos desarrollan hiperbilirrubinemia asintomática). [55]

## 12.- CONCLUSIONES.

Los intervalos de referencia son la herramienta más utilizada para la interpretación de resultados de laboratorio y pueden condicionar decisiones clínicas. Es preciso conocer los principios en los que se basan para hacer un uso adecuado de ellos. Deben distinguirse de otras herramientas de interpretación como los límites de decisión clínica o del cambio en el valor de referencia.

Los intervalos de referencia se originan con individuos seleccionados procedentes de una población concreta, con distintos criterios. Es importante que se expliquen con el detalle necesario estos criterios, así como los pasos llevados a cabo para obtener los intervalos de referencia.

La variabilidad en los resultados del laboratorio es inevitable en cierta medida, pero sí se puede reducir. El desarrollo de más estudios y por tanto un mayor conocimiento de la influencia que tienen las variables preanalíticas en los resultados finales podría evidenciar su repercusión y, de haberla, acelerar los esfuerzos en la estandarización de esta fase.

No nos parece disparatado exigir un rigor similar al que se tiene en las determinaciones analíticas, con continuos controles de calidad a distintos niveles, a la hora de calcular los intervalos de referencia.

Para la mayoría de los laboratorios es imposible seguir las recomendaciones iniciales de la IFCC. Se proponen por tanto alternativas de distinto rigor teórico y empeño práctico. En nuestra profana opinión sería interesante hacer esfuerzos por comprender cuál es la situación actual de los intervalos de referencia en nuestro entorno, buscando diferencias entre intervalos obtenidos a través de métodos distintos y, posteriormente, ver si estas diferencias pudieran tener repercusión sobre los comportamientos clínicos que motivan.

No tenemos recursos para hacer lo anterior. Nos hemos limitado a intentar suscitar un interés. La mayor parte de la bibliografía reciente sobre el tema procede de otros países, lo que contrasta con la intensa actividad que se llevaba a cabo en España sobre el tema en los años ochenta. Es llamativa esta aparente interrupción (siempre con excepciones), ya que es un campo en el que aún se trabaja, se avanza y en el que se siguen tomando decisiones importantes. La amplitud del tema nos ha impedido profundizar con el detalle que hubiéramos querido en cada aspecto tratado. Es una revisión limitada, por tiempo y espacio.

Se proponen como soluciones a la variabilidad entre laboratorios la armonización y la estandarización. La automatización del laboratorio ha traído consigo una sensación de seguridad, de confianza ciega en los resultados. Es importante el componente humano en la toma de decisiones del laboratorio, que debe ser objeto de estudio, análisis y debate para garantizar su mejora continua. La comparabilidad de resultados entre distintos laboratorios sería deseable.

### **13.- AGRADECIMIENTOS.**

A mi tutor, el Dr. Enrique Sánchez Óriz, por su gran ayuda en la sistematización, dirección y por el continuo apoyo en el desarrollo del trabajo.

A los Drs. Pedro Latorre Marcellán y Marta Vicente Aznar por las incontables explicaciones sobre el funcionamiento del laboratorio de hematimetría en el hospital Miguel Servet.

A la Dra Francisca González Rubio, médico de familia del centro Delicias Sur, por poner en común conmigo una propuesta de preparación preanalítica que elaboró para sus pacientes.

A la Dra. Ana Matarredona Pareja, de la Dirección General de Asistencia Sanitaria, y al Dr. José Puzo Focillas, del servicio de Bioquímica Clínica del hospital Miguel Servet, por el tiempo dedicado a recibirme y dar su opinión sobre el tema del trabajo.

## 14.- BIBLIOGRAFÍA.

1. Ceriotti F, Henny J. "Are my laboratory results normal?" Considerations to be made concerning reference intervals and decision limits. *European Journal of the IFCC*. 2008 Oct; 19(2): 106–114.
2. Wellman M. The concept of normal in medicine. *Canad MAJ*, 1958, 79: 43-44.
3. Schneider AJ. Some thoughts on normal, or standard, values in clinical medicine. *Pediatrics*. 1960, 26.6: 973-984.
4. Gräsbeck R. The evolution of the reference value concept. *Clinical chemistry and laboratory medicine*, 2004, 42.7: 692-697.
5. Siest G, Henny J, Gräsbeck R, Wilding P, Petitclerc C, Queraltó JM, Petersen PH. The theory of reference values: an unfinished symphony. *Clinical chemistry and laboratory medicine*. 2013, 51(1), 47-64.
6. Solberg HE. Approved recommendation on the theory of reference values. Part 1. The concept of reference values. *J Clin Chem Clin Biochem*. 1986; 25, 337-42.
7. PetitClerc C, Solberg HE. Approved recommendation on the theory of reference values. Part 2. Selection of individuals for the production of reference values. *Clinica Chimica Acta*, 170(2-3), S1-S11.
8. Solberg HE, PetitClerc C. Approved recommendation on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. *Clinica Chimica Acta*. 1988; 177(3), S3-S11.
9. Solberg HE, Stamm D. Approved recommendation on the theory of reference values, Part 4: Control of analytical variation in the production transfer and application of reference values. *European Journal of Clinical Chemistry and Clinical Biochemistry*. 1991; 29, 531-535.
10. Solberg, HE. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *Clinica Chimica Acta*. 1987; 170(2-3), S13-S32.
11. Dybkoer, R. Approved recommendation on the theory of reference values. Part 6. Presentation of observed values related to reference values. *Clinica Chimica Acta*. 1987; 170(2-3), S33-S41.
12. **Ceriotti F, Hinzmann R, Panteghini M. Reference Intervals: the way forward. *Annals of Clinical Biochemistry*. 2009; 46: 8-17.**
13. Gomís-Castellví M, Miró-Balagué J. Decisiones clínicas basadas en valores críticos o discriminantes de los análisis clínicos. *FMC*. 2006; 13(2):68-72.
14. Álvarez Funes V, Güell Miró R. Interpretación de los datos de laboratorio. *FMC*. 2011; 18 (6):360-5.
15. Ceriotti F. Gli intervalli di riferimento nel nuovo millennio. *Biochimica Clinica*. 2007, 31(4): 254-266.
16. WHO. Constitution of the World Health Organization. 2006. [www.who.int/governance/eb/who\\_constitution\\_en.pdf](http://www.who.int/governance/eb/who_constitution_en.pdf)
17. Alström T, Gräsbeck R, Hjelm M et al. Committee on Reference Values. Scandinavian Society for clinical chemistry and clinical physiology. Recommendations concerning the collection of reference values in clinical chemistry and activity report. *Scand J Clin Lab Invest* 1975; 35 (Suppl. 144), 1-74.
18. Ozarda, Y, Ichihara K, Barth JH, Klee G. Protocol and standard operating procedures for common use in a worldwide multicenter study on reference values. *Clinical chemistry and laboratory medicine*, 2013, 51.5: 1027-1040.
19. Koerbin G, et al. Evidence-based approach to harmonised reference intervals. *Clinica Chimica Acta*, 2014, 432: 99-107.
20. Martín Calderón JL, et al. Valores de referencia en plasma de osmolalidad, electrolitos, calcio, magnesio y urea en una población del centro de España. *Revista del Laboratorio Clínico*, 2014, 7.2: 49-54.
21. Solberg HE, Gräsbeck R. Reference Values. En: Spiegel HE. *Advances in Clinical Chemistry*. San Diego: Academic Press, INC; 1989. p. 2-69.
22. Grossi E, Colombo R, Cavuto S, Franzini C. The REALAB Project: A new method for the formulation of reference intervals based on current data. *Clinical Chemistry*. 2005; 51:7, 1232-1240.
23. Lippi G, et al. Preanalytical quality improvement: from dream to reality. *Clin Chem Lab Med* 2011; 49(7): 1113-26.
24. Lippi G, et al. Preanalytical quality improvement. In pursuit of harmony, on behalf of European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Working group for Preanalytical Phase (WG-PRE). *Clinical Chemistry and Laboratory Medicine (CCLM)*, 2015, 53.3: 357-370.
25. Ialongo C, Bernardini S. Preanalytical investigations of phlebotomy: methodological aspects, pitfalls and recommendations. *Biochimica Medica* 2017; 27(1):177-191.
26. Lippi G, Lima-Oliveira G, Salvagno GL, et al. Influence of a light meal on routine haematological tests. *Blood Transfus* 2010;8:94-99.
27. Sanchis-Gomar F, Lippi G. Physical activity-an important preanalytical variable. *Biochimica medica*, 2014, 24.1: 68-79.
28. Lippi G. Raccomandazioni per il prelievo di sangue venoso. *Biochimica clinica*. 2008, 32,6 :569-577.

29. World Health Organization. WHO guidelines on drawing blood: best practices in phlebotomy. WHO Press: Geneva, 2010.
30. Lippi G, Salvagno GL., Montagnana M et al. Venous stasis and routine hematologic testing. *Clinical & Laboratory Haematology*, 28(5), 332-337.
31. Lippi G, et al. Influence of posture on routine hemostasis testing. *Blood Coagulation & Fibrinolysis*, 2015, 26.6: 716-719.
32. Cornes MP, Ford C, Gamma R. Spurious hyperkalaemia due to EDTA contamination: common and not always easy to identify. *Annals of clinical biochemistry*, 2008, 45.6: 601-603.
33. Alcaraz Quiles J, Rico Santana N, Bedini Chesa JL. Estabilidad de 27 magnitudes bioquímicas en muestras de suero conservadas en refrigeración. *Rev Lab Clín*. 2014; 7(1):9-16.
34. Perich Alsina C, Álvarez Ríos AI, Blázquez R, et al. Aplicación práctica del control interno de la calidad en los procedimientos de medida cuantitativos. *Revista del Laboratorio Clínico*, 2014, 7.1: 25-32.
35. Prada E, Blázquez R, Gutiérrez Bassini G, et al. Control interno de la calidad vs control externo de la calidad. *Revista del Laboratorio Clínico*, 2016, 9.2: 54-59.
36. Ichihara K, Kawai T. Determination of reference intervals for 13 plasma proteins based on IFCC international reference preparation (CRM470) and NCCLS proposed guideline (C28-P,1992): a strategy for partitioning reference individuals with validation based on multivariate analysis. *J Clin Lab Anal* 1997; 11:117-24.
37. Jørgensen GM, Brandslund I, Petersen PH. Should we maintain the 95 percent reference intervals in the era of wellness testing? A concept paper. *Clin Chem Lab Med*, 2004, 42.7: 747-751.
38. Ichihara K, Boyd JC. An appraisal of statistical procedures used in derivation of reference intervals. *Clin Chem Lab Med*, 2010; 48(11): 1537-1551.
39. **Ozarda Y. Reference Intervals: current status, recent developments and future considerations. *Biochimica Medica* 2016; 26 (1):5-16.**
40. **Horn PS, Pesce A. Reference intervals: an update. *Clinica chimica acta*, 2003, 334: 5-23.**
41. Plebani M, Panteghini M. Armonizzazione in laboratorio: verso una visione globale. *Biochimica clinica*. 2015; 39,1.
42. Plebani M. Harmonization in laboratory medicine: Request, samples, measurements and reports.
43. Greenberg N. Update on current concepts and meanings in laboratory medicine – Standardization, traceability and harmonization.
44. Panteghini M. Traceability as a unique tool to improve standardization in laboratory medicine. *Clin Biochem* 2009; 42:236–40.
45. Panteghini M, Ceriotti F. Obtaining reference intervals traceable to reference measurement systems: is it possible, who is responsible, what is the strategy? *Clin Chem Lab Med*. 2012; 50(5): 813-817.
46. Ceriotti F. Prerequisites for use of common reference intervals. *Clin Biochem Rev*, 2007, 28: 115-121.
47. Sociedad Española de Bioquímica Clínica y Patología Molecular. González de la Presa F, Canalias Reverter S, Esteve Poblador S, et al. Procedimiento para la transferencia y revisión de intervalos de referencia biológicos. *Revista del Laboratorio Clínico*. 2017.
48. Fuentes-Arderiu X. Biological reference intervals and ISO 15189. *Clinica Chimica Acta*, 2006. 364: 365–366.
49. International Organization for Standardization. Medical Laboratories – Particular Requirements for Quality and Competence ISO 15189. Geneva: ISO, 2007.
50. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. *Offic J Eur Commun* 7. December 1998, L331/1-L331/37.
51. **CLSI and IFCC. C28-A3 document: Defining, establishing and verifying reference intervals in the clinical laboratory: approved guideline-third edition, 2008;28:1-76. \*CONSULTA NO DISPONIBLE.**
52. Boyd JC. Cautions in the adoption of common reference intervals. *Clinical Chemistry*. 2008; 54(2):238-239.
53. **Friedberg RC, Souers R, Wagar E, et al. The origin of reference intervals. A College of American Pathologists Q-Probes study of “normal ranges” used in 161 clinical laboratories. *Arch Pathol Lab Med*. 2007; 131: 348-357.**
54. Ceriotti F. Establishing pediatric reference intervals: a challenging task. *Clinical chemistry*, 2012; 58(5): 808-810.
55. Shirts BH, Wilson AR, Jackson BR. Partitioning reference intervals by use of genetic information. *Clinical Chemistry*, 2011; 57(3): 475-481.