



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS  
APLICADAS Y EN SISTEMAS**

**MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA  
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA**

**“Propuesta metodológica para verificación de precisión y  
desempeño en pruebas de ELISA espectrofotométricas en  
enfermedad de Chagas”**

*Tesina*

*Que para obtener el:*

*el título de especialista*

*en:*

**“ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA”**

Presenta:

**Ing. Aldo Pasten Mayen**

Asesor:

**Dra. Silvia Ruiz-Velasco Acosta**

**Ciudad Universitaria, CD. MX. 2022**





Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



# Agradecimientos

- A los Estados Unidos Mexicanos por mantener y fortalecer las instituciones para el desarrollo de la nación.
- A la Universidad Nacional Autónoma de México por formar capital humano con perspectiva y alcance para el desarrollo de la nación y de nuevo capital humano.
- Al Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas por tener un fuerte compromiso por formar individuos con un alto conocimiento técnico y metodológico.
- A mis padres Elvira y Sergio por siempre brindarme su cariño y apoyo.
- A mi hermano Sergio por siempre brindarme su cariño y apoyo.
- A la doctora Silvia Ruiz por aceptarme como su asesorado.
- Al jurado conformado por; la doctora Silvia Ruiz; la doctora Lizbeth Naranjo; la maestra Patricia Romero; la maestra Leticia Gracia-Medrano; el maestro Benjamín Alvares. Sus comentarios enriquecieron y complementaron el trabajo, alineándolo con los objetivos de la especialidad. Se los agradezco mucho.
- Al laboratorio de la enfermedad de Chagas del InDRE-SSA por apoyarme con la información para esta tesina.



# Resumen

Esta tesina propone una metodología, la cual pueda servir de guía para la planificación, ejecución y análisis de resultados obtenidos en los procedimientos de verificación en pruebas basadas en ensayos inmunoenzimáticos de interpretación cualitativa en enfermedad de Chagas que se realizan en los laboratorios de diagnóstico clínico.



# Índice de contenido

Índice de contenido	VII
Índice de tablas	XIII
Índice de figuras	XV
Objetivo	XVII
Introducción	XIX
<b>Parte I Aspectos Biológicos</b>	<b>1</b>
<b>Capítulo 1 Aspectos biológicos</b>	<b>3</b>
1.1 Enfermedad de <b>Chagas</b> . . . . .	3
1.2 Historia natural de la enfermedad . . . . .	4
1.2.1 Fase aguda . . . . .	4
1.2.2 Fase crónica . . . . .	4
1.2.3 Fase crónica indeterminada(asintomática) . . . . .	5
1.2.4 Diagnóstico . . . . .	5
1.3 Inmunodiagnóstico . . . . .	6

<b>Parte II Metodología</b>	<b>9</b>
<b>Capítulo 2 Alcance de la verificación</b>	<b>11</b>
2.1 Verificación . . . . .	13
2.2 Tipo de pruebas . . . . .	15
2.2.1 Pruebas cuantitativas . . . . .	16
2.2.2 Pruebas cualitativas . . . . .	16
2.3 Verificación de la precisión . . . . .	18
2.4 Verificación de desempeño . . . . .	20
2.5 Recursos de una verificación . . . . .	21
2.6 Procedimientos operativos(Rutinarios) . . . . .	21
2.6.1 Prácticas de laboratorio . . . . .	22
2.6.1.1 Recursos humanos . . . . .	22
2.6.1.2 Recursos materiales . . . . .	22
2.7 Recursos biológicos . . . . .	23
2.7.1 Material de referencia . . . . .	23
2.7.2 Criterio de precisión diagnóstico . . . . .	23
2.7.2.1 Construcción de un criterio de precisión diagnóstico(estándar de referencia) cuando no existe y es posible construirlo . . . . .	24
<b>Capítulo 3 Aspectos teóricos y diseño de la muestra</b>	<b>27</b>
3.1 Aspectos teóricos . . . . .	28
3.1.1 Definición estadística de Población . . . . .	28
3.1.1.1 Población objetivo . . . . .	29
3.1.2 Definición operativa de espectro de la enfermedad . . . . .	30
3.1.3 Intervalo de respuesta no reactivo . . . . .	31
3.1.4 Intervalo de respuesta reactivo . . . . .	31
3.1.5 Intervalo de respuesta de la enfermedad de <b>Chagas</b> . . . . .	32
3.2 Selección de población de referencia y muestra de precisión . . . . .	32
3.3 Muestra para precisión . . . . .	33

3.3.1	Paneles de desempeño . . . . .	33
3.3.2	Tamaño de muestra . . . . .	34
3.4	Muestra para validez externa . . . . .	35
3.4.1	Muestra para desempeño ó población de referencia . . . . .	35
3.4.2	Tamaño de muestra . . . . .	36
3.5	Criterios de Inclusión . . . . .	38
3.5.1	Criterios de inclusión para muestra de precisión . . . . .	39
3.5.2	Criterios de inclusión para población de referencia . . . . .	39
3.5.2.1	Criterios de inclusión epidemiológicos . . . . .	39
3.5.2.2	Criterios de inclusión clínicos . . . . .	39
3.5.2.3	Criterios de inclusión de laboratorio . . . . .	40
<b>Capítulo 4</b>	<b>Ensayos experimentales</b>	<b>41</b>
4.1	Ensayos para precisión . . . . .	42
4.1.1	Asignación de especímenes para verificación de precisión . . . . .	43
4.2	Diseño experimental para desempeño . . . . .	45
4.2.1	Estratificación de material biológico . . . . .	45
4.2.2	Algoritmo de selección para población de referencia . . . . .	46
<b>Capítulo 5</b>	<b>Análisis de datos</b>	<b>49</b>
5.1	Análisis del experimento de precisión . . . . .	49
5.1.1	Análisis de precisión intracorrida . . . . .	49
5.1.2	Análisis de precisión del laboratorio . . . . .	50
5.1.3	Comparación de la Repetitividad Estimada con la definida por el fabricante . . . . .	51
5.1.4	Comparación de la Precisión Estimada Intralaboratorio con la definida por el fabricante . . . . .	51
5.2	Análisis del ensayo de desempeño . . . . .	52
5.2.1	Sensibilidad y Especificidad . . . . .	52
5.2.2	Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo . . . . .	53

5.2.3	Interpretación de los parámetros Sensibilidad, Especificidad, Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo . . .	54
5.3	Intervalos de confianza para Sensibilidad, Especificidad, Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo . . . . .	55
5.3.1	I.C para Sensibilidad y Especificidad [1] . . . . .	55
5.3.2	I.C para Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo [1] . . . . .	56
<b>Parte III Ejemplo de Aplicación</b>		<b>57</b>
<b>Capítulo 6 Ejemplo de Aplicación</b>		<b>59</b>
6.1	Definición del Problema . . . . .	59
6.2	Verificación de precisión . . . . .	60
6.2.1	Muestra de precisión . . . . .	60
6.2.1.1	Criterios de inclusión de especímenes en la muestra . . .	60
6.2.2	Selección de especímenes . . . . .	60
6.2.3	Asignación de los especímenes a la microplaca . . . . .	61
6.2.4	Cálculo de precisión intracorrída y entrecorrída . . . . .	61
6.3	Verificación de desempeño . . . . .	61
6.3.1	Población objetivo . . . . .	62
6.3.1.1	Espectro de la enfermedad . . . . .	63
6.3.1.2	Intervalo de respuesta . . . . .	63
6.3.2	Referencia . . . . .	63
6.3.2.1	Material de referencia . . . . .	63
6.3.2.2	Caso confirmado . . . . .	64
6.3.2.3	Caso confirmado . . . . .	64
6.3.2.4	Caracterización de material biológico . . . . .	64
6.3.3	Muestra para desempeño(Población de referencia) . . . . .	65
6.3.3.1	Criterios de inclusión de la muestra . . . . .	65

6.3.4	Tamaño de la muestra . . . . .	67
6.3.4.1	Tamaño de muestra para subpoblación reactiva . . . . .	67
6.3.4.2	Tamaño de muestra para subpoblación no reactiva . . . . .	68
6.3.5	Estratificación del material biológico . . . . .	69
6.3.5.1	Definición de límites superiores e inferiores . . . . .	69
6.3.5.2	Definición de los límites de cada nivel . . . . .	69
6.3.5.3	Niveles para reactividad positiva . . . . .	70
6.3.5.4	Niveles para reactividad negativa . . . . .	70
6.3.6	Selección de especímenes y conformación de la población de referencia	71
6.3.6.1	Selección de especímenes en subpoblación reactiva y conformación de subpoblación de referencia positiva . . . . .	71
6.3.6.2	Selección de especímenes en subpoblación no reactiva y conformación de subpoblación de referencia negativa . . . . .	72
6.3.7	Ejecución del algoritmo . . . . .	73
6.3.7.1	Selección de especímenes de la subpoblación reactiva . . . . .	73
6.3.7.2	Gráfico de densidades del banco de material biológico(población objetivo) . . . . .	77
6.3.7.3	Gráfico de densidades de la muestra(población de referencia)	78
6.3.8	Asignación de los especímenes a la microplaca(con arreglo 8 filas x 12 columnas) . . . . .	78
6.3.9	Cálculo de parámetros de desempeño . . . . .	79
6.3.9.1	Alcance tipo de parámetros . . . . .	79
6.3.10	Cálculo de intervalos de confianza . . . . .	80

**Parte IV Conclusiones** **81**

Discusión **83**

Recomendaciones **87**

Conclusiones	<b>91</b>
<b>Referencias</b>	<b>93</b>
Referencias	93
<b>Anexos</b>	<b>97</b>
Anexo A: Definiciones operativas.	<b>99</b>
Anexo B: Técnicas estadísticas.	<b>103</b>
B.1 Muestreo Aleatorio Simple[2] . . . . .	103
B.1.1 Muestreo Aleatorio Simple con reemplazo . . . . .	103
B.1.2 Muestreo Aleatorio Simple sin reemplazo . . . . .	104
B.2 Muestreo Estratificado[3] . . . . .	104
B.3 Tablas de contingencia[4] . . . . .	105
B.3.1 Sensibilidad y Especificidad[4] . . . . .	106
Anexo C: Desempeño de algoritmos	<b>109</b>
C.1 Desempeño algoritmo para selección de la población de referencia . . . . .	109

# Índice de tablas

5.1	Tabla de contingencia para evaluar sensibilidad y especificidad. . . . .	52
5.2	Tabla de contingencia para evaluar porcentajes de acuerdo positivo y negativo. . . . .	53
B.1	Tabla de contingencia para evaluar una prueba diagnóstica. . . . .	106
C.1	Tamaño de población para selección de población de referencia. . . . .	111
C.2	Tamaño de población para selección de población de referencia. . . . .	111



# Índice de figuras

2.1	Valor de corte que clasifica a las poblaciones reactiva y no reactiva. . . . .	18
6.1	Histograma subpoblación reactiva. . . . .	74
6.2	Histograma subpoblación no reactiva. . . . .	75
6.3	Histograma población de referencia. . . . .	76
6.4	Gráfico densidades de población de referencia. . . . .	77
6.5	Gráfico densidades banco de material. . . . .	78



# Objetivo

Desarrollar una metodología con el fin de verificar pruebas basadas en ensayos inmunoenzimáticos espectrofotométricos en enfermedad de **Chagas**. Articulando la normatividad vigente, protocolos consensuados por grupos de expertos, conceptos estadísticos, metodológicos y el conocimiento documentado de la enfermedad de **Chagas**, que además pueda servir para el desarrollo de protocolos que pretendan verificar pruebas cualitativas para otras enfermedades.



# Introducción

La **verificación** es un requisito que debe ser cumplido de manera que, la prueba pueda ser utilizada como instrumento de detección, satisfaciendo los requisitos conforme se establece en la normatividad[5]. Hay que hacer énfasis en que; dependiendo del tipo de prueba que se quiera verificar, los parámetros a evaluar son diferentes, siendo de tipo cualitativo y cuantitativo. Para el caso de la enfermedad de **Chagas** son pruebas cualitativas.

Los instrumentos existen y se utilizan para realizar verificaciones, sin embargo, lo que no existe es una metodología que los conjunte para el desarrollo de los protocolos para realizarla. La normatividad sólo dice “¿Qué se debe hacer?”, los protocolos responden a la pregunta “¿Cómo se debe de hacer?”, mientras que los conceptos estadísticos y el conocimiento del fenómeno, responde “¿Con que se hace?”. Al no existir una metodología que conjunte estos elementos, se aumenta la probabilidad de emitir un veredicto sesgado.

En esta tesina se desarrolló una metodología que sirva como guía para el planteamiento de protocolos que verifiquen pruebas de detección de anticuerpos para la enfermedad de **Chagas**. Siendo posible que esta metodología pueda ser usada para otros padecimientos o enfermedades, con sus pertinentes modificaciones.



**Parte I**  
**Aspectos Biológicos**



# CAPÍTULO 1

## Aspectos biológicos

### Definiciones operativas

<b>Acuerdo</b>	Porcentaje total de casos en el cual dos métodos tienen el mismo resultado.
<b>Historia natural de la enfermedad</b>	Se refiere a la progresión de un proceso de la enfermedad en un individuo con el tiempo, en ausencia de tratamiento.

### 1.1. Enfermedad de Chagas [6]

La enfermedad de **Chagas** es autóctona del continente americano. Es causada por el protozooario hemoflagelado *Trypanosoma cruzi*, se transmite al humano y otros mamíferos por insectos vectores hemípteros de la subfamilia Triatominae, chupadores de sangre. Aunque también puede transmitirse a través de transfusiones de sangre o trasplante de órganos, verticalmente de la madre infectada al hijo durante el embarazo, y por ingestión accidental de alimentos contaminados. Se ha estimado que en los primeros años de la década de los 90, de 16 a 18 millones de individuos estaban infectados con *T. cruzi*. [Organización Mundial de la Salud (OMS), 1991] sin embargo, gracias al éxito de las intervenciones para el control vectorial y el tamizaje en donadores de sangre las cifras han disminuido entre 6 a 8 millones en el 2010 (OMS 2014). Se estima, que 120 millones de individuos continúan en riesgo de

adquirir la infección.

## 1.2. Historia natural de la enfermedad

La enfermedad de **Chagas** tiene una evolución natural y se divide dos fases, una aguda y una crónica. La fase crónica se divide a la vez en dos tipos, asintomática y sintomática. A la fase crónica asintomática también se le conoce como indeterminada, mientras que a la fase crónica sintomática tiene asociados alteraciones gastrointestinales, cardiológicas y neurológicas. Utilizando criterios de diagnóstico y terapéuticos distintos en cada fase.

### 1.2.1. Fase aguda[7]

Las fase aguda de la enfermedad puede presentarse a cualquier edad. Sin embargo, la mayoría de los casos son detectados antes de los 15 años, siendo entre 1 a 5 años la de más alta frecuencia. La fase aguda de la enfermedad empieza cuando el *Trypanosoma cruzi* entra al cuerpo. Una reacción local en la zona de entrada es seguida por un malestar general. Todas las manifestaciones clínicas desaparecen después de 4-8 semanas. En esta fase las pruebas parasitológicas son utilizadas para demostrar parásitos circulantes en la mayoría de los individuos infectados.

### 1.2.2. Fase crónica[7]

La fase crónica de la enfermedad de **Chagas** comienza con la disminución de la parasitemia a niveles indetectables y manifestación clínica de miocarditis aguda o meningoencefalitis. Forma sintomática y asintomática, respectivamente. Esos cambios parasitológicos y clínicos pueden suceder de 4-8 semanas después de la infección.

En individuos inmunocompetentes no tratados, se piensa que el nivel de parasitemia disminuye a consecuencia de un equilibrio alcanzado entre el parásito y la respuesta inmunológica del individuo infectado. Este equilibrio puede durar por el resto de la vida del individuo, y anticuerpos de la clase IgG contra *T. cruzi* pueden ser detectados. Los métodos de diagnósti-

### 1.2.3 Fase crónica indeterminada(asintomática)

---

co tales como, xenodiagnóstico o hemocultivo, pueden demostrar parásitos circulantes en al menos la mitad de todos los individuos infectados varios años después de que la infección se originó.

### 1.2.3. Fase crónica indeterminada(asintomática)[7]

La fase aguda es seguida por la forma crónica indeterminada de la enfermedad de **Chagas**. Aproximadamente 50-70 % de los individuos infectados permanecerán bajo esta condición por el resto de su vida. Aquellos con la forma indeterminada asintomática cuentan, por tanto, con infección con *T. cruzi*. En áreas endémicas donde la transmisión vectorial aún ocurre, el humano actúa como reservorio natural del *T. cruzi* y contribuyendo al mantenimiento del ciclo de vida del parásito. La mayoría de los pacientes con la forma indeterminada de la enfermedad se encuentran entre 20-50 años(i.e son económicamente productivos). Pueden ser identificados por encuestas epidemiológicas, o por exámenes serológicos tales como aquellos que donan sangre.

### 1.2.4. Diagnóstico[6]

El diagnóstico etiológico de la enfermedad de **Chagas** se basa en la valoración clínica, epidemiología y pruebas de laboratorio. Para el diagnóstico de laboratorio, los exámenes adecuados dependen de la fase clínica del paciente. En la fase aguda los estudios se centran en la búsqueda y reconocimiento del *T. cruzi* en sangre (metodología: técnica directa), porque en las etapas iniciales de la enfermedad se encuentran parasitemias importantes y a medida que transcurre la infección van disminuyendo hasta hacerse mínimas. En la fase crónica (asintomática o sintomática) las parasitemias son transitorias y por ello el diagnóstico se realiza fundamentalmente mediante el hallazgo de anticuerpos circulantes contra el *T. cruzi* (metodología: técnica indirectas).

### 1.3. Inmunodiagnóstico[6]

El inmunodiagnóstico se basa en la detección de anticuerpos específicos contra *T. cruzi* sin embargo, se debe entender que lo que se busca no es el parásito y por tanto sus resultados nunca proporcionarán certeza diagnóstica y deben evaluarse en términos de probabilidad. Para que la probabilidad se acerque a la certeza, es importante seleccionar adecuadamente el método a emplear, el momento de la toma de muestra y la interpretación apropiada de los resultados.

El diagnóstico por laboratorio se realiza en una muestra de sangre o suero del paciente, el tipo de muestra y la técnica a realizar se determina con base en la fase en la que se encuentra la enfermedad.

Para inmunodiagnóstico se utiliza suero o plasma, para técnicas de cultivo o inoculación en modelo animal es sangre completa y laminillas con extendido y gota gruesa para técnicas parasitológicas en fase aguda.

El inmunodiagnóstico de la infección por *T. cruzi* involucra una mezcla de anticuerpos dirigidos contra un gran número de antígenos parasitarios, con diferente concentración, afinidad y tiempo de aparición. Puede haber expresión de diferentes mosaicos antigénicos entre las distintas cepas y reactividad cruzada con microorganismos relacionados filogenéticamente que comparten zonas geográficas.

Los distintos métodos y estuches de reactivos para el diagnóstico pueden comportarse de manera diferente frente a una misma muestra y la evaluación de desempeño de diferentes lotes no son siempre consistentes con estudios previos. La especificidad del inmunodiagnóstico para la tripanosomiasis americana es buena sin embargo, la sensibilidad es mayor hacia la fase crónica sintomática. La fuente de antígenos en los estuches comerciales o reactivos de preparación local son obtenidos a partir de diferentes cepas y fases del parásito, van desde el parásito íntegro, extractos crudos, extractos parcialmente purificados, antígenos recombinantes de fase aguda o crónica, péptidos sintéticos, hasta antígenos quiméricos.

Los estuches de diagnóstico son muy variables en su composición antigénica y ninguno de ellos alcanzan por sí solo el 100 % de certeza diagnóstica (WHO, 2010), es por esto que, los grupos de expertos internacionales de Tropical Disease Research (TDR)/World Health Organization (WHO) (TDR-WHO) recomiendan para tamiz, una técnica altamente sensible y para diagnóstico al menos dos pruebas en paralelo, de diferente formato. Con este diseño, el diagnóstico, puede alcanzar un rango de sensibilidad del 98-99.5 %. Es decir, para obtener resultados más seguros y concluyentes, se debe realizar más de una técnica diagnóstica. Sin embargo, una limitante es la presencia de resultados indeterminados (resultados discrepantes), la frecuencia de estos obedece a las características del desempeño de las pruebas utilizadas.



**Parte II**  
**Metodología**



# CAPÍTULO 2

## Alcance de la verificación

### Definiciones operativas

<b>Acuerdo</b>	Porcentaje total de casos en el cual dos métodos tienen el mismo resultado.
<b>Criterio de precisión diagnóstico</b>	El mejor criterio disponible para establecer la presencia o ausencia de la condición, evento o característica de interés utilizando un solo método o combinación de métodos que incluyen pruebas de laboratorio, pruebas por imágenes, patología e información clínica, incluido el seguimiento.[1]
<b>Especificidad</b>	Porcentaje de individuos sin la condición blanco (como fue determinado por el criterio de precisión diagnóstico) los cuales tienen valores negativos en la prueba.[1]
<b>Inserto</b>	Manual que provee el desarrollador en el cual se especifican la ejecución de la prueba, el estudio de precisión, la sensibilidad y la especificidad.
<b>Precisión intra-laboratorio</b>	precisión en un tiempo definido y operadores, dentro de la misma instalación y con el mismo equipo. La calibración y reactivos pueden variar.[8]
<b>Repetitividad</b>	el acuerdo más cercano entre los resultados de mediciones sucesivas del mismo mesurando llevado a cabo bajo las mismas condiciones de medición.[8]
<b>Sensibilidad</b>	Porcentaje de individuos con la condición de blanco (como fue determinado por el criterio de precisión diagnóstico) los cuales tienen valores positivos en la prueba.[1]
<b>Valor de corte</b>	para una prueba cualitativa, el umbral por encima del cual el resultado se reporta como positivo y por debajo del cual el resultado se reporta como negativo.[1]

<b>Validez Externa</b>	Representatividad de los elementos estudiados de la población objetivo, siendo las características relevantes del proceso estudiado representativas de toda la población.
<b>Validez Interna</b>	Comparabilidad de grupos que no difieren en características relevantes a la asociación entre la variable X e Y.
<b>Verificación</b>	Confirmación mediante la aportación de evidencia objetiva de que se han cumplido los requisitos especificados.[5]
<b>Zona de baja veracidad</b>	Región en donde la interacción de las poblaciones reactivas y no reactivas afectan la exactitud de la prueba, disminuyendo la probabilidad de clasificación del espécimen, con respecto a los parámetros especificados de la prueba. A los resultados obtenidos en esta región se les clasifica como <b>falso positivo</b> , <b>falso negativo</b> o <b>indeterminado</b> .

Cuando se realiza cualquier procedimiento siempre es importante comprender cuál es el fin y consecuentemente su importancia, considerando siempre que cualquier acción forma parte de un proceso, el cual, se verá fuertemente influido por los resultados obtenidos de cada una de las acciones que lo componen. Por lo anterior, es importante que en cada acción se consideren todos los aspectos teóricos y empíricos que se conocen para obtener evidencia objetiva que sea lo más precisa con el problema que se quiere resolver(realidad).

En síntesis y concretamente, una **verificación** consiste en demostrar que las características declaradas por el fabricante/desarrollador se cumplan. Desde un punto de vista técnico es relativamente sencillo demostrar que se cumplan las características, no obstante se debe de tener en cuenta que la prueba debe de ser de utilidad en la población en donde se pretende emplear.

Recordando que cada resultado obtenido con la prueba/método está asociado a un paciente, ya sea que esté o no infectado. Siendo el resultado de la prueba la evidencia para la gestión clínica de tratar o no a un individuo.

Si en una verificación no se articulan los aspectos necesarios(teóricos/prácticos) para obtener conclusiones objetivas y la prueba/método no tiene el alcance especificado por su fabricante/desarrollador en la población objetivo(ver sección 3.1.1.1 Población objetivo), y si esta es

implementada como herramienta de detección o confirmación, se podrían reportar estados de salud que conlleven a decisiones en políticas de salud o médicas erróneas, dejando libres potenciales fuentes de infección o atentando contra la calidad y/o vida de un individuo, como la salud de la población.

Hasta este punto sólo se ha definido a la verificación como un proceso aislado, faltando describir su lugar en el proceso rutinario de un laboratorio clínico o de salud pública.

Una verificación no es procedimiento que se realice en un laboratorio clínico en su operación rutinaria, es un servicio que generalmente ofrece el laboratorio de referencia que cuenta con la autoridad y suficiencia técnica(estabilidad) para evaluar de manera experimental que determinada prueba/método cumple con los requisitos especificados/declarados por su desarrollador. Sin embargo, la normatividad mexicana vigente establece que los laboratorios clínicos antes de implementar un nuevo método en su operación rutinaria deben verificarlo, lo que implica el desarrollo de un protocolo en donde se articulen los procedimientos operativos(rutinarios) y recursos biológicos del laboratorio, además de la prueba/método que provee el fabricante/desarrollador para verificar si este instrumento es adecuado para el fin perseguido y sobre todo en la población que se quiere aplicar.

## 2.1. Verificación

Para comprender el contexto de una verificación, es importante comprender primeramente la definición operacional de verificación. A continuación se presentan definiciones consensuadas por grupos internacionales;

VIM (Vocabulario internacional de Metrología) *Aportación de evidencia objetiva de que un elemento dado satisface los requisitos especificados;*[9]

NMX-CC-9000-IMNC-2015 *Confirmación mediante la aportación de evidencia objetiva de que se han cumplido los requisitos especificados para un método;*[10]

NMX-CH-152-IMNC-2005 *La verificación consiste en evaluar el desempeño del método para demostrar que cumple con los requisitos para el uso previsto, que fueron especificados como resultado de su validación.*[11]

Las definiciones presentadas tienen su contexto en:

- Metrología;
- Sistemas de gestión de calidad;
- Química.

Estas definiciones se pueden sistematizar para generar normas dirigidas a contextos más específicos. Una de ellas es la norma mexicana NMX-EC-15189-IMNC-2015 (“Laboratorios clínicos - Requisitos de la calidad y competencia”). Esta norma es la que usualmente se emplea para verificar pruebas/métodos. En el apartado referente a la verificación de “procedimientos de exámenes” específicamente el apartado 5 hace referencia a;

*“Los procedimientos de examen validados, utilizados sin modificación, se deben someter a una verificación independiente por el laboratorio antes de iniciar su uso en la rutina del laboratorio.*

*El laboratorio debe obtener información de los métodos desarrollados por el fabricante/desarrollador del método, para confirmar las características del desempeño del procedimiento.*

*La verificación independiente por parte del laboratorio, debe confirmar, a través de la obtención de evidencia objetiva (en forma de características del desempeño), que se han cumplido las especificaciones declaradas para el procedimiento de examen. Las características de desempeño declaradas para el procedimiento de examen conformadas durante el proceso de verificación deben ser aquellas relevantes para el uso previsto de los resultados de los exámenes.*

*El laboratorio debe documentar el procedimiento utilizado para la verificación y registrar los resultados obtenidos”*

Por lo general la variable que se mide al individuo con determinada prueba/método en una muestra biológica es la respuesta cualitativa y/o cuantitativa. La respuesta, en pruebas de interpretación cualitativa, que está expresada en escala nominal, puede ser de dos tipos; positiva p.ej. (presencia de anticuerpos específicos o, sustancia de interés, etc.) o negativa (ausencia de anticuerpos específicos o sustancia de interés, etc.).

Como se menciona en el punto 5, los exámenes de verificación se reducen a comprobar las especificaciones declaradas por el fabricante en el inserto de la prueba/método. Por lo general el inserto de una prueba/método especifica:

- Principio de su funcionamiento;
- Procedimiento de ejecución, validación e interpretación;
- Valores de precisión y desempeño.

Dependiendo del tipo de prueba (cuantitativa o cualitativa), los parámetros que se evalúan varían.

## 2.2. Tipo de pruebas

Por su uso y desempeño, las pruebas se dividen principalmente en 3 grupos que son;

**Tamizaje** Estas pruebas se caracterizan por tener alta sensibilidad para garantizar que los verdaderos positivos sean detectados. Por lo general las pruebas de tamizaje arrojan una mayor cantidad de resultados falsos positivos que las pruebas con fin diagnóstico o confirmatorio. Esta baja especificidad puede ser permitida si existe una buena prueba confirmatoria, y si las consecuencias social/económico de resultados falso positivo no son demasiado severas.[1]

**Diagnóstico** Clínicamente, se requiere establecer una gestión apropiado. El diagnóstico por laboratorio debe tener excelente sensibilidad y especificidad. Si una prueba confirmatoria sigue después de las prueba diagnóstica de alto desempeño y precisión, el requerimiento de las especificidad puede ser de alguna manera menor.[1]

**Confirmatoria** Se utilizan para ratificar los resultados obtenidos por pruebas de tamizaje o diagnóstico (presuntivas). La confirmación de resultados previos permite al médico establecer un diagnóstico. Las pruebas confirmatorias están diseñadas para ser específicas (sacrificando sensibilidad, si es necesario) teniendo un valor predictivo positivo alto.[1]

El proceso de confirmación de una enfermedad requiere una prueba cuya especificidad sea alta. Cuando dos o más pruebas están disponibles para este propósito, aquellas con la más alta especificidad son normalmente preferidas.

Cuando una prueba es utilizada tanto para tamizaje ó para excluir una posibilidad diagnóstica, debe de ser sensible. Cuando dos o más pruebas están disponibles, aquellas con la más alta sensibilidad son normalmente preferidas.[12]

### **2.2.1. Pruebas cuantitativas**

Este tipo de pruebas se caracteriza por poder medir una concentración específica de determinada sustancia en un volumen de muestra p.ej(mg/dL, mg/Kg, etc.).

### **2.2.2. Pruebas cualitativas**

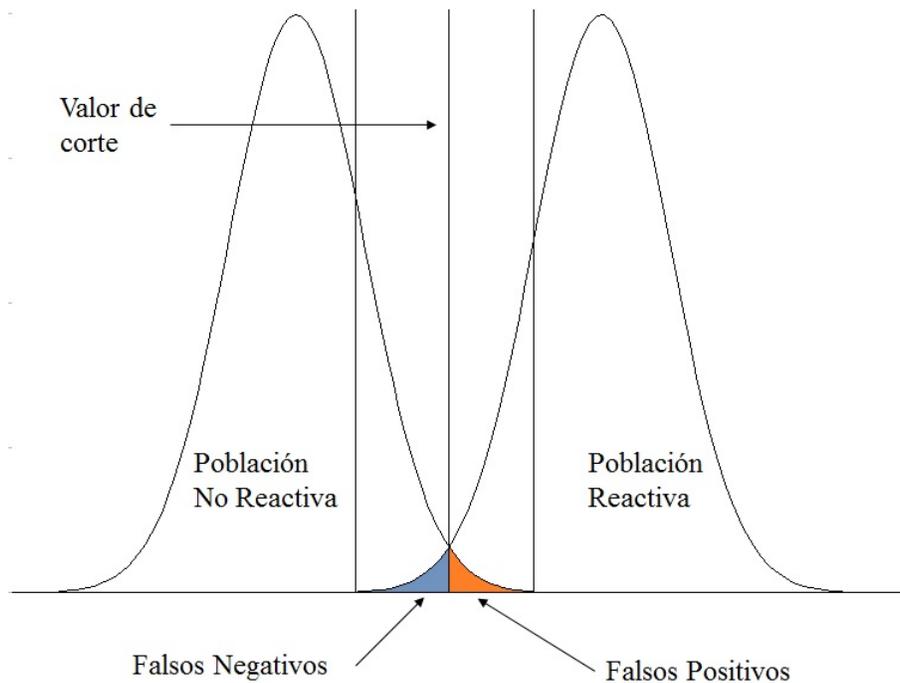
Las pruebas cualitativas solo pueden clasificar la presencia/ausencia de determinada sustancia en una muestra, lo que se traduce en clasificar reactivos y no reactivos. Esta clasificación se hace con base en un valor crítico que se conoce como “valor de corte”. El valor de corte clasifica a los especímenes ya sea como reactivo(positivo) o no reactivo(negativo) . Una muestra es positiva si el resultado obtenido es mayor al valor de corte y es negativa en caso contrario. El valor de corte es un parámetro de cada prueba y es especificado por el fabricante/desarrollador. Es importante considerar que en las pruebas diagnósticas la sensibilidad y especificidad no varían con la prevalencia, pero sí lo hacen con el valor de corte. Los índices como los valores predictivos y la eficiencia, pueden variar por la prevalencia y por el valor de corte.

Los parámetros que se evalúan en las pruebas/métodos, por su enfoque se pueden clasificar en parámetros de estabilidad y parámetros de desempeño. Los parámetros de estabilidad (condiciones de repetitividad o reproducibilidad) son aquellas que demuestran la congruencia de resultados de una prueba/método en diferentes lugares en condiciones similares, en esta clasificación se encuentran la repetitividad intra corrida y repetitividad entre corrida o también llamada precisión de laboratorio. Estos parámetros los poseen todas las pruebas.

Análogamente, los parámetros de desempeño son aquellos que demuestran la efectividad de detección del método/prueba en una población con el padecimiento de estudio, estos parámetros son;

- Verdaderos positivos (Sensibilidad);
- Verdaderos negativos (Especificidad);
- Falsos positivos;
- Falsos negativos.

La tasa de falsos positivos, falsos negativos, si no están declarados en el inserto como parámetro, se pueden calcular utilizando una tabla de contingencia 2x2 una vez obtenidos los resultados de las series de trabajo. En la Figura 2.1 se observa de manera gráfica lo que se espera suceda en la práctica, en el [capítulo 6](#) se puede observar una gráfica mostrando lo que en la realidad sucede. Se encuentran tres zonas: una para los individuos no reactivos, una intermedia en donde los casos son dudosos, para esta tesina se definió como **zona de baja veracidad** clasificándose como; falso positivo o falso negativo, y la tercera en donde los resultados indican la presencia de la infección y/o enfermedad (reactivos).



**Figura 2.1:** Valor de corte que clasifica a las poblaciones reactiva y no reactiva.

Es muy importante que se identifiquen, en los ensayos a realizar, las características específicas de desempeño que se deben evaluar a una prueba/método para su verificación. Estos ensayos son;

- Verificación de la precisión;
- Verificación de desempeño.

Para realizar una verificación se siguen las recomendaciones de las directrices EP15-A2 y EP12-A2; para precisión y desempeño de pruebas de interpretación cualitativa. Metodológicamente equivalen a verificar **validez interna** y **validez externa**.

### 2.3. Verificación de la precisión

Esta etapa lo que busca es comprobar la variación que puede existir en diferentes laboratorios, y que, en condiciones de operación similares, se espera que esta variabilidad fluctúe en

el rango que fue ajustado al momento de diseñar la prueba/método, las condiciones especificadas son condiciones de repetitividad y condiciones de precisión intermedia. Teniéndose en cuenta que el método se probará para detectar correctamente niveles de reactividad, solo se verifica la repetitividad de la prueba, enfocándose en verificar la prueba como herramienta de detección y no de clasificación.

Para verificar la precisión, se requiere de utilizar material biológico el cual cuente con reactividades iguales o similares a las que especifica el fabricante, a este conjunto de material se le denomina “**muestra de precisión**”. Para lo que se pueden emplear dos tipos de materiales

- Paneles (material altamente caracterizado);
- Utilizar especímenes de la población objetivo;

Dado que únicamente se busca comparar la variabilidad obtenida especificando una reactividad, no es necesario emplear especímenes de alguna población objetivo en donde se pretende emplear la prueba, por lo que utilizar un panel es una opción adecuada para medir este error, siempre y cuando el panel contenga las reactividades adecuadas o sea el mismo que fue empleado para el desarrollo de la prueba. La otra opción es utilizar especímenes de la población objetivo con reactividades similares a las declaradas en el inserto de la prueba.

Habiendo seleccionado el material biológico, ahora se debe de planificar la manera de realizar las series de trabajo por día. Para obtener resultados lo más confiable posible, es necesario controlar factores que pueden conllevar a correlacionar las respuestas entre ellos, para lo que es recomendable;

- Aleatorizar y bloquear; los especímenes seleccionados, se asignan al azar en cada uno de los días, formando bloques en donde se incluyen muestras con los niveles de respuesta. Estos niveles (estratificación), se hacen con base en las reactividades que se especifican en el inserto;
- Homogeneización; Se deben seleccionar especímenes, los cuales tengan los niveles de respuesta más parecidos a los especificados en el inserto;

- Métodos estadísticos; se utiliza cuando los valores obtenidos están “muy cercanos” a los esperados y se quiere determinar si se encuentran dentro del rango del error.

## 2.4. Verificación de desempeño

En esta parte de la verificación lo que se busca es determinar si la prueba cumple con lo declarado por el fabricante. Para este ensayo la recomendación del uso de especímenes varia dependiendo de los parámetros de desempeño que se quieran verificar. Estos parámetros pueden ser;

- porcentaje de acuerdo positivo y negativo;
- sensibilidad y especificidad;

Para la verificación del rendimiento de precisión diagnóstica, es decir; sensibilidad y especificidad, lo que recomienda el protocolo EP12-A2 es lo siguiente;

*“Para los estudios que evalúan el rendimiento de precisión diagnóstica, las muestras de pacientes utilizadas para la comparación del método, deben incluir una población representativa de los estados clínicos esperados en la práctica clínica, criterio de precisión diagnóstica que es más que un simple método analítico que sustituye el termino estándar de referencia”*

En el caso de que los parámetros declarados sean de acuerdo (porcentaje de acuerdo positivo y negativo), el protocolo EP12-A2 recomienda lo siguiente;

*“Para estudios que evalúan acuerdo, muestras de pacientes, paneles de referencia y muestras para prueba de competencia se puede utilizar para estudiar el desempeño de las pruebas.”*

De las recomendaciones por parte del protocolo EP12-A2, y teniendo en cuenta que en la definición de verificación, sólo menciona el hecho de “demostrar con evidencia objetiva que se cumplen las especificaciones declaradas para el procedimiento del examen”, lo que se puede cumplir utilizando especímenes de diferentes niveles de reactividad. Una verificación además de demostrar que se cumplan las especificaciones declaradas, debe de demostrar utilidad en la población en la que se pretende emplear para diagnosticar individuos potencialmente in-

fectados/enfermos. Por lo que los especímenes, criterio de precisión diagnóstica, con los que se retará a la prueba/método deben seleccionarse articulando conceptos metodológicos, clínicos y estadísticos obteniéndose conclusiones aplicables a la población de interés.

## 2.5. Recursos de una verificación

Ya se sabe qué es una verificación, su alcance y lo que se evalúa, ahora falta definir los recursos que se necesitan para llevar a cabo una verificación. Se mencionó que para llevar a cabo una verificación, el protocolo de verificación es la guía conceptual hace uso de los procedimientos operativos(rutinarios) y de recursos biológicos del laboratorio para realizarla.

## 2.6. Procedimientos operativos(Rutinarios)

Como se mencionó al principio del capítulo 2, todas las acciones forman parte de un proceso.

Las buenas prácticas de laboratorio (BPL) aseguran la calidad e integridad de los datos obtenidos en estudios o investigaciones. Las BPL establecen las condiciones bajo las cuales se planifican, realizan, controlan, registran, archivan e informan los estudios realizados por un laboratorio, asegurando la veracidad de los resultados. Estas prácticas conforman acciones del proceso de diagnóstico que; si no son acordes con los procedimientos internos establecidos se pueden convertir en factores potenciales que afectarán el desempeño de las pruebas/métodos y en consecuencia los resultados obtenidos no serán válidos. Las buenas prácticas de laboratorio se pueden dividir en 4 principios, que manifiestan estabilidad operativa.

1. Instalaciones adecuadas: El laboratorio debe cumplir con todas las normas de seguridad que apliquen para el trabajo que ahí se realiza
2. Personal calificado: Se debe proporcionar capacitación continua para garantizar que el personal conoce la técnica y sabe utilizar el equipo o material empleado
3. Equipo adecuado y calibrado: Se debe dar mantenimiento continuo a los equipos para

garantizar su correcto funcionamiento y calibrarlos de forma regular

4. Procedimientos estándares de operación (SOPs): Procedimientos escritos, suficientemente claros para que cualquier técnico que trabaja en el laboratorio pueda seguirlos al pie de la letra. De esta forma se garantiza que todos los técnicos trabajan bajo las mismas directrices

## **2.6.1. Prácticas de laboratorio**

### **2.6.1.1. Recursos humanos**

Hace referencia a las habilidades desarrolladas por los técnicos del laboratorio para utilizar de manera correcta el equipo y reactivos. Siendo importante la supervisión y evaluación periódica para asegurar que el error asociado sea el mínimo. Es importante resaltar que, para que existan buenas prácticas de laboratorio es recomendable que se lean las especificaciones del equipo/reactivo a emplear con el fin de hacer la selección y el correcto uso de ellos para desempeñar la actividad, algunas de estas prácticas se enlistan:

- Uso/selección correcto de la pipeta;
- Selección de la punta adecuada;
- Capacitación(entrenamiento) en los procesos automatizados;
- Manejo adecuado de los reactivos para evitar contaminarlos;
- Manejo adecuado de las muestras

### **2.6.1.2. Recursos materiales**

Hace referencia al mantenimiento del equipo, que puede ser preventivo ó correctivo. Haciéndolo el técnico ó proveedor respectivamente. Otra parte del mantenimiento del material es la verificación, que consiste en calibrarlo en caso de que sea necesario para asegurar que las mediciones que realiza sean consistentes con el patrón de referencia.

Además de los aspectos mencionados también se tienen los ambientales, que se refieren a las condiciones en el lugar de trabajo, teniendo a la temperatura y humedad relativa como los principales que afectan al equipo y al reactivo empleado. Siendo conveniente tener control de temperatura y humedad en los espacios de trabajo.

Todos los aspectos anteriores son factores potenciales que pueden influir en el resultado (negativa o positivamente) si no son evaluados periódicamente (Control de calidad). Además de los aspectos técnicos se tienen los aspectos teóricos, los cuales también se pueden considerar factores potenciales que pueden influir en los resultados si no se hacen las reflexiones y razonamientos adecuados al problema planteado.

## 2.7. Recursos biológicos

### 2.7.1. Material de referencia

Son especímenes caracterizados y documentados que pueden ser empleados como material de referencia. Este material de referencia puede ser empleado para verificación de pruebas, así como para control de calidad. Por lo general, los especímenes se caracterizan utilizando un criterio de precisión diagnóstico el cual debe de ser definido en cada contexto.

### 2.7.2. Criterio de precisión diagnóstico [13]

Se le denomina criterio de precisión diagnóstico al mejor método disponible para establecer la presencia o ausencia del padecimiento estudiado. En un estudio de precisión puede haber una o más pruebas de referencia[14]. Un criterio de precisión diagnóstica puede estar conformado por una prueba o método, combinación de métodos y técnicas en donde se puede incluir el seguimiento. Si el criterio es una combinación de métodos, el algoritmo debe especificar cómo los diferentes resultados son combinados para hacer la clasificación final positivo/negativo, pudiendo incluir la selección y orden de los métodos[13].

Para la enfermedad de **Chagas** existe un criterio de precisión diagnóstico que ayuda a con-

firmar infección por *T. cruzi*. Sin embargo, éste sólo es útil en la fase aguda de la infección, uno de los métodos de confirmación parasitológicos consiste en la observación del parásito en sangre como es el examen de gota gruesa(prueba confirmatoria).

Por otro lado, para la fase crónica asintomática, la única manera de confirmar infección es buscando anticuerpos de respuesta específica a la infección, sin embargo, para esta fase no se ha establecido un criterio de precisión diagnóstico(no hay prueba confirmatoria), siendo necesario construirlo para poder caracterizar las muestras que conforman el banco de material biológico que además puedan ser empleadas para una verificación.

#### **2.7.2.1. Construcción de un criterio de precisión diagnóstico(estándar de referencia) cuando no existe y es posible construirlo**

Al construir el criterio de precisión diagnóstico es importante definir el padecimiento y su comportamiento en determinada(s) población(es) con el fin de no introducir un error sistemático. Evitando sobreestimar o subestimar el criterio definido. Este tipo de errores pueden surgir por problemas en el diseño, que es principalmente un problema de validez interna. Mientras que seleccionar elementos de la población no adecuados, puede ocasionar un problema con la validez externa[15]. Esto sucede cuando se pretende aplicar a más de una población objetivo.

Para un criterio de precisión diagnóstico de referencia los estimadores son la sensibilidad y especificidad, y estos se verán afectados si no se consideran todos los aspectos teóricos necesarios para evitar introducir algún error generando estimadores imprecisos[13].

Al construirse un criterio de precisión diagnóstico, es posible determinar sensibilidad y especificidad.

Algunas veces no existe un criterio de precisión diagnóstico, ni tampoco es posible construirlo. Para estos casos se define un estándar que no es de criterio de precisión diagnóstico. Este criterio sólo puede estimar concordancia expresada como porcentaje de acuerdo entre resultados, con lo que se especifica que no se está estimado precisión diagnóstica, sino acuerdo

### 2.7.2 Criterio de precisión diagnóstico

---

del método bajo estudio con el estándar que no es de precisión diagnóstica. Al utilizar este tipo de estándar se tiene la desventaja de que los porcentajes de acuerdo no son una medida de aserción, esto porque la prueba bajo estudio y el estándar pueden concordar en resultados, pero ambos estar errados.



# CAPÍTULO 3

## Aspectos teóricos y diseño de la muestra

### Definiciones operativas

<b>Espectro de la enfermedad</b>	Es el intervalo de la historia natural de la enfermedad en donde un individuo ya está infectado y puede ser diagnosticado, siendo probable o no la aparición de síntomas asociados a la enfermedad.
<b>Panel</b>	Material caracterizado que no contiene una población objetivo en específico.

En las pruebas de interpretación cualitativa basadas en detección instrumental (e.g. colorimetría, quimioluminiscencia, etc.), la respuesta que se obtiene se debe transformar en una respuesta de tipo binario presencia/ausencia, reactivo/no reactivo. Las pruebas que utilizan un espectrofotómetro como instrumento de detección, utilizan el principio de colorimetría. La respuesta obtenida por este principio se llama **absorbancia**, esta absorbancia se transforma a respuesta binaria, en este caso reactivo/no reactivo a partir de un valor seleccionado en el intervalo de valores de la respuesta instrumental.

La clasificación está condicionada por un valor umbral que discrimina entre reactivo/no reactivo, a este valor se le conoce como **valor de corte** (ver sección 2.2.2).

Para verificar los parámetros de las pruebas, que consiste en la correcta clasificación de los in-

dividuos dado el resultado de la prueba y su condición actual, se requiere de una muestra que contenga especímenes de las subpoblaciones reactivas(positivas) y no reactivas(negativas), a esta muestra se le denominará **población de referencia**.

### **3.1. Aspectos teóricos**

Hasta este punto se sabe que una verificación tiene como fin evaluar la precisión y los parámetros de desempeño de una prueba/método que se pretende emplear de manera rutinaria para clasificar a individuos potencialmente infectados con algún padecimiento, en este caso con enfermedad de **Chagas**, el resultado tiene un impacto directo sobre la decisión médica, sin embargo, la existencia de factores asociados reales o potenciales pueden modificar los resultados obteniéndose información inválida. Lo anterior son aspectos necesarios e importantes para la verificación.

No obstante, lo que guiará acerca del desempeño de una prueba/método en una verificación son los aspectos teóricos, específicamente el concepto de **población**, además de definir operativamente el concepto de **espectro de la enfermedad** como variable. Si estos conceptos no se comprenden correctamente, no importa que todos los aspectos técnicos sean correctos, ya que las conclusiones obtenidas serán de alcance limitado o en el peor de los casos equivocadas.

#### **3.1.1. Definición estadística de Población**

La población desde el punto de vista estadístico se define como; colección (finita o infinita) de elementos o unidades de estudio  $U_i$ , que tienen ciertos atributos, características o factores comunes a todos ellos A, B, C y D. Como los atributos posibles que puede tener cualquier elemento o unidad son conceptualmente infinitos, al definir una población especificando de que unidades se trata y qué atributos tienen en común, se dejan sin definir un número infinito de factores, E, F, G, . . . , estos factores pueden variar de un elemento a otro.[16]

Se dice que una población es finita cuando pueden enumerarse todos sus elementos, mien-

### 3.1.1 Definición estadística de Población

---

tras que cuando los elementos resultan de un proceso generador, se dice que la población es infinita[16] . El proceso generador en este caso son la caracterización de las muestras biológicas provenientes de la población objetivo.

Sin embargo, ¿por qué es tan importante el concepto de población? Recordando que una **Verificación**(ver sección 2.1) consiste en demostrar que las características declaradas por el desarrollador/fabricante se cumplan, en este caso nos referimos a métodos/pruebas basadas en ensayos inmunoenzimáticos y, al ser un fenómeno biológico la variabilidad en la respuesta es intrínseca dada su naturaleza, ya que depende del polimorfismo genético en las poblaciones el cual se expresa particularmente en función de la raza, región geográfica, condiciones nutricionales, ambiente epidemiológico y aún hasta de cada individuo.[17]

Además de la variabilidad en la respuesta inmunológica de las diferentes poblaciones, las pruebas/métodos son validadas utilizando especímenes de algunas poblaciones(no todas) seleccionadas por el desarrollador, lo que metodológicamente limita la validez externa de la misma. Por esta disminución en la validez externa lo metodológicamente adecuado sería verificar que el método/prueba sea funcional en cada una de las poblaciones en la que se pretende implementar como instrumento clasificación, siendo necesario definir a cada población.

Descrito el concepto de población y de la presencia de variabilidad biológica asociada a la raza, ubicación geográfica, etc., hay que definir la población sobre la que se hará la verificación del método/prueba, refiriéndose a esta población como **población objetivo**.

#### 3.1.1.1. Población objetivo

La población objetivo considera las variables; edad, raza, región geográfica, sexo, presunción de algún padecimiento o enfermedad. Las variables se definen al momento de seleccionar la población en donde se hará la verificación teniendo en cuenta el comportamiento de la enfermedad (ver ejemplo sección 6.3). Dado que el objetivo es verificar una prueba de interpretación cualitativa, cuyos únicos resultados son reactivo y no reactivo. La población objetivo se divide en 2 subpoblaciones; **población objetivo reactiva** y **población objetivo**

**no reactiva.**

Para conocer la magnitud de la presencia de una enfermedad en una población, se determina la prevalencia. La prevalencia con base en su definición, es una variable que cambia con el tiempo, pudiendo disminuir o aumentar, dificultando conocer con exactitud la cantidad de individuos infectados y enfermos.

Por lo general es muy difícil conocer con exactitud la cantidad de individuos infectados con alguna enfermedad, de tal manera que el intervalo de respuesta reactivo contribuye a definir a la población objetivo como infinita al no conocerse el intervalo completamente.

Las subpoblaciones se definen de la siguiente manera;

**Población objetivo reactiva:** En esta población se encuentran individuos con respuesta positiva a alguna enfermedad. Se cuantifica con el **intervalo de respuesta reactivo**.

**Población objetivo no reactiva:** En esta población se encuentran individuos con respuesta negativa a alguna enfermedad. La población no reactiva está cuantificada por el **intervalo de respuesta no reactivo**.

### **3.1.2. Definición operativa de espectro de la enfermedad**

Como se describe en el Anexo A, el espectro de la enfermedad representa un intervalo de la historia natural de la enfermedad. Al diagnosticar un padecimiento en una población siempre existe un espectro de enfermedad. Para documentar y caracterizar este espectro de enfermedad es necesario contar con algún método/metodología (ej. criterio de precisión diagnóstico, pruebas serológicas, métodos confirmatorios, etc.). La documentación y caracterización del espectro de la enfermedad es importante siendo que se puede generar alguna variable o conjuntos de variables que permitan definir la amplitud del espectro de la enfermedad. La(s) variable(s) generada(s) a partir de esta documentación por lo general se deben de expresar en términos de los métodos (clínicos, de laboratorio, etc.) diseñados para medir esta(s) variable(s) (ej. absorbancia, luminiscencia, cantidad de réplicas, etc.). Esta documentación es útil para poder diseñar verificaciones/evaluaciones de nuevos métodos de diagnóstico, programas

### 3.1.3 Intervalo de respuesta no reactivo

---

de control de calidad, etc.

Para trabajar el concepto de espectro de la enfermedad se define el concepto de intervalo de respuesta, dividiéndose el intervalo de respuesta en dos; **reactivo** y **no reactivo**.

Aunque se conoce el espectro de la enfermedad, no se conocen los valores máximos del intervalo de respuesta que puede desarrollar un individuo, ya que, la respuesta está asociada fuertemente a las variables;

- Respuesta específica de cada individuo;
- Mosaico antigénico asociado a cada cepa;

Además, al estado clínico del paciente, teniéndose un conjunto infinito de valores.

En resumen, el espectro de la enfermedad representa las diferentes etapas de evolución y gradiente de la enfermedad, la inclusión de individuos sin la infección o enfermedad en individuos con diagnósticos diferencial que permite determinar los falsos positivos.

### **3.1.3. Intervalo de respuesta no reactivo**

Consiste en todos los valores del intervalo de respuesta que están asociados a individuos que no reaccionan positivamente a la prueba. Con una respuesta mínima de cero como límite inferior y una respuesta máxima menor ó igual al valor de corte menos un porcentaje de seguridad, pudiendo ser 10, 20 hasta 30 %. Este porcentaje es establecido dependiendo del uso y de la precisión aceptada.[1]

### **3.1.4. Intervalo de respuesta reactivo**

Consiste en todos los valores del intervalo de respuesta que están asociados a individuos que reaccionan positivamente a la prueba. Con una respuesta mínima igual o mayor al valor de corte mas un porcentaje de seguridad, pudiendo ser 10, 20 hasta 30 %. Este porcentaje es establecido dependiendo del uso y de la precisión aceptada.[1]

### **3.1.5. Intervalo de respuesta de la enfermedad de Chagas**

El intervalo de respuesta de alguna enfermedad o padecimiento está compuesto por el intervalo de respuesta no reactivo y el intervalo de respuesta reactivo.

## **3.2. Selección de población de referencia y muestra de precisión**

Para evitar sesgar las conclusiones del estudio, la teoría indica que la muestra debe ser seleccionada al azar. En las técnicas de muestreo siempre se hace referencia a una muestra representativa. Una muestra representativa se define como; muestra que refleje las características de la población de donde fue seleccionada[16] (ver sección 3.1.1.1).

En la sección 3.1.1.1 se definió a la población objetivo como infinita, no siendo posible generar un marco de muestreo, y en consecuencia no se puede seleccionar una muestra aleatoria con igual probabilidad de selección para todos los elementos; por lo que se recurre a las muestras denominadas “disponibles o a conveniencia”. Seleccionando como muestra a un conjunto de elementos disponibles (accesibles) al investigador, y se conceptualiza una población infinita de elementos para la cual la muestra disponible es representativa. En este caso se deben de tener precauciones para que la muestra resulte, en la medida de lo posible, representativa o lo más cercano a esto[16]. Siendo que, en este caso, la población esta conceptualizada por el material biológico del que dispone el laboratorio el cual, es un conjunto de especímenes (muestras biológicas) mucho mayor que la muestra que se selecciona para efectuar la verificación.

Para un estudio de verificación se tienen dos tipos de muestras, asociadas al experimento que se vaya a realizar estas son;

- Muestra de precisión(validez interna);
- Muestra de desempeño(validez externa).

### 3.3. Muestra para precisión

Para el experimento que verifica la (precisión) validez interna, se tienen respuestas especificadas en el inserto del método/prueba con su respectivo coeficiente de variación. Estas respuestas fueron con las que el desarrollador validó la repetitividad de la prueba.

Con base en esas respuestas se seleccionan especímenes con una respuesta igual o cercana a las especificadas en el inserto cuidando que se encuentren dentro del intervalo de respuesta, en donde los límites superior e inferior están acotados por el coeficiente de variación.

Además, existen materiales de desarrolladores que tienen utilidad para realizar estos experimentos, a estos materiales se les conoce comúnmente como control positivo de anticuerpos *anti-Trypanosoma cruzi* independientes, que consisten en material que no tienen valores asignados. Los niveles de reactividad variarán entre los ensayos de diferentes fabricantes, diferentes procedimientos, diferentes números de lote y aún en diferentes laboratorios.

#### 3.3.1. Paneles de desempeño

Los paneles de desempeño consisten en material caracterizado y datos complementarios para realizar análisis comparativos, provistos para fabricantes de reactivos, investigadores y laboratorios clínicos. Proporciona los resultados obtenidos utilizando las pruebas comerciales de tamizaje, de seguimiento y confirmatorias.

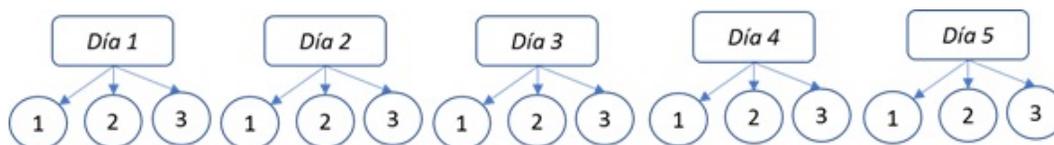
Hay diferentes tipos de paneles, por lo general los que se utilizan son los paneles de tercera opinión, que son desarrollados de manera independiente al laboratorio en donde se pretende utilizar y al desarrollador de la prueba/método. Para el caso de una verificación, se deben de utilizar aquellos que están enfocados a verificar el **desempeño**. Un panel consiste en 14 o 21 especímenes de diferentes poblaciones. Estos especímenes están caracterizadas por pruebas comerciales, en donde se reporta el rango de reactividad desde negativo hasta alto positivo. Por lo que es muy importante leer la hoja técnica para obtener detalles de este tipo de materiales.

**Verificación de precisión** En la precisión lo que se busca es que los especímenes se encuentren dentro los límites declarados por el desarrollador, independientemente de la población de la que provengan las muestras. Para esta evaluación se pueden utilizar; ya sean paneles, sueros control o muestras de banco dependiendo de la disponibilidad del material.

**Verificación de desempeño** La desventaja de utilizar paneles es que, éstos pueden no contener el espectro de interés o no representar a la población objetivo.[1] Por lo que, para la verificación de desempeño siempre se deben utilizar muestras de la población objetivo en donde se aplicará la prueba/método.

### 3.3.2. Tamaño de muestra

Para evaluar precisión, el protocolo EP15-A2 (sección 8.2) recomienda que se seleccione una muestra para cada repetición y para cada concentración, el protocolo utiliza dos (2) concentraciones, teniéndose quince(15) muestras biológicas por nivel de concentración para todo el experimento de 5 días considerándose 3 repeticiones por día. En total, se tienen 30 muestras biológicas para todo el experimento de precisión. Sin embargo, el número de niveles de concentración se deben seleccionar dependiendo de lo que especifica el fabricante. En esta tesina se consideraron 3 niveles de concentración (bajo, medio y alto). Además de que, se puede reducir el número de muestras a solo 5 por nivel de respuesta para todo el experimento, teniéndose un total de 15 muestras de diferente concentración. Pero, para seguir la recomendación de tres (3) repeticiones por día, por cinco (5) días. Cada muestra se fracciona para obtener tres fracciones de igual reactividad teniéndose así las 15 muestras por nivel de concentración y por consiguiente las 45 muestras que se recomiendan para el experimento de precisión.



## 3.4. Muestra para validez externa

### 3.4.1. Muestra para desempeño ó población de referencia

Como se mencionó al principio del Capítulo 3, se denomina población de referencia a un grupo de especímenes que incluyen a la subpoblación reactiva y no reactiva. Esta población de referencia debe de representar a la población objetivo de la mejor manera.

Al interior de cada una de las subpoblaciones, se deben generar grupos con diferente nivel de respuesta que representen a la subpoblación lo más fidedignamente posible. La cantidad de grupos a generar queda a criterio del experto. Sin embargo, los grupos que siempre deben de estar presentes son;

- Alto no reactivo;
- Bajo reactivo.

Los especímenes de los grupos mencionados se les denomina **especímenes desafiantes** al estar “muy cercanos” a los límites en donde la probabilidad de ser clasificado de manera errónea aumenta;

Además, el espécimen (muestra biológica) del grupo bajo reactivo que siempre debe estar incluido es;

- Especimen de menor reactividad del grupo bajo reactivo;

Para el grupo alto negativo, no es necesario que se incluya el valor más negativo, aunque sería lo preferente, siendo aceptable la selección de valores cercanos al límite de detección de no reactividad.

Para la población objetivo reactiva, representada por el intervalo de respuesta reactivo se conforman grupos de respuesta(bajo, medio y alto o los que se consideren necesarios).

Para la población objetivo no reactiva, representada por el intervalo de respuesta no reactivo, al igual que el intervalo reactivo se forman grupos(bajo,medio y alto o los que se conside-

ren). Para este caso el grupo en donde se tiene interés es en el límite superior del intervalo no reactivo, constituido por las reactividades en el grupo alto no reactivo, esto al tener mayor probabilidad de ser clasificados erróneamente ya sea como **falso positivo** o **falso negativo**. Para esta población no es muy crítico agregar los valores de mínima respuesta, ya que conforme disminuye la respuesta, la probabilidad de ser clasificado como no reactivo (negativo) es un evento seguro, sin embargo, metodológicamente tiene mayor sustento el agregar todo el intervalo de respuesta del que se disponga.

A la población objetivo reactiva se le presta especial atención al grupo bajo ya que, se ha observado que la respuesta conforme más cercana se encuentre al límite de detección, mayor probabilidad tienen de ser clasificados erróneamente ya sea como **falso positivo** o **falso negativo**. De forma similar que en la población objetivo no reactiva, el límite superior de la zona de baja veracidad está representado por los valores de mínima respuesta de la población reactiva. Conforme la respuesta aumenta, la probabilidad de que sea clasificado correctamente tiende a uno, es decir; los grupos de mediana y alta respuesta se podrían considerar como un “**evento seguro**”, por lo que, se puede incluir un menor número de especímenes para completar el intervalo de respuesta, aunque metodológicamente tiene mayor sustento y representatividad el agregar todo el intervalo de respuesta disponible, siendo conveniente incluirlo.

Además de los especímenes de las poblaciones objetivo, también se consideran especímenes para verificar que no existan reacciones cruzadas o inespecíficas.

Simplificando, para minimizar el sesgo de espectro es necesario idear el diseño de muestreo adecuado.

### **3.4.2. Tamaño de muestra**

Si el propósito del estudio es conocer si la reactividad o no reactividad de la prueba a verificar se asocian con la clasificación por el estándar de referencia, enfermo/sano, un diseño adecuado sería de casos (enfermos) y controles (sanos), la razón de los enfermos comparada

con los sanos es la parte del diseño que el investigador debe determinar.

Para el tamaño de muestra se debe considerar;

- Los parámetros operativos de la prueba(especificidad, sensibilidad o Porcentaje de acuerdo positivo, porcentaje de acuerdo negativo);
- La disponibilidad del material;
- El error que se obtendrá con determinado tamaño de muestra.

Se puede realizar un ejercicio numérico para estimar el tamaño de la población de referencia, que consiste en seleccionar los especímenes (muestras biológicas) de las poblaciones reactiva y no reactiva en función de los parámetros de desempeño(sensibilidad, especificidad, etc.) de la prueba que se quiere verificar, que en conjunto determinaran el tamaño mínimo de la población de referencia.

Para lo que se utiliza la expresión para calcular el tamaño de muestra, para un IC95 %, de una proporción;

$$n = \frac{Z_{\alpha/2}^2 pq}{d^2} \quad (3.1)$$

Para tener una estimación se necesita conocer;

- Probabilidad de éxito(p). El cuál esta representado por la especificidad, sensibilidad, porcentaje de acuerdo positivo, porcentaje de acuerdo negativo;
- Probabilidad de fracaso(q). Esta representado por los valores falsos negativos, falsos positivos ó  $(1 - p)$  ;
- Precisión(d);
- Nivel de confianza(z), por lo regular al 95 %

Una vez que se conoce el tamaño de la población de referencia, se selecciona teniendo en

cuenta las características que debe cumplir. Recordando que el propósito de este conjunto de especímenes es demostrar que se cumplan las características de la prueba.

Se debe tener en cuenta que, en los estudios de verificación, sólo se hacen estimaciones basadas en la población de referencia evaluadas con el criterio de precisión diagnóstico. Estos criterios son susceptibles a sesgos, si no se consideran los factores potenciales que lo pueden generar. Para minimizar este efecto, se pueden seleccionar los sujetos (muestras biológicas, unidades de estudio, etc.) adecuadas, modificar la manera de llevar a cabo el estudio, o utilizando métodos de análisis que ayuden a remover o quitar el sesgo.

La selección por conveniencia, muchas veces es la mejor manera de seleccionar los elementos que se incluyen en la muestra, ya que, el investigador (experto) conoce a fondo la problemática del fenómeno[18]. También se puede utilizar alguna técnica de muestreo.

Siempre teniendo en mente que un amplio conocimiento del fenómeno para la selección es fundamental.

Un aspecto importante, en la población de referencia, es que no se quiere estimar media o varianza en las unidades de estudio, sólo se utiliza el método de selección para poder sistematizar la selección.

### **3.5. Criterios de Inclusión**

Los conceptos mencionados en este capítulo (Población de referencia, estratos de reactividades, espectro de la enfermedad), son los aspectos metodológicos que deben tener la **muestra de precisión y población de referencia** al momento de seleccionar e incluir los especímenes a cada conjunto.

Los aspectos metodológicos varían dependiendo del ensayo que se vaya a realizar.

#### **3.5.1. Criterios de inclusión para muestra de precisión**

Retomado lo mencionado, la muestra para el ensayo de precisión no requiere que los especímenes provengan de la población objetivo en donde se pretende emplear la prueba/método. Por lo que el único criterio que deben de cumplir los especímenes es:

- No es necesario que provengan de la población objetivo en donde se desee implementar la prueba;
- Las reactividades deben de ser iguales o parecidas a las declaradas por el fabricante.

#### **3.5.2. Criterios de inclusión para población de referencia**

La población de referencia es la muestra que se utiliza para verificar los parámetros de desempeño de una prueba; sensibilidad y especificidad.

Para la población de referencia se tienen tres categorías de criterios de inclusión para los especímenes, los poblacionales, metodológicos y los de laboratorio.

##### **3.5.2.1. Criterios de inclusión epidemiológicos**

Son las características de la población objetivo.

- Raza;
- Edad;
- Sexo;
- Región;

##### **3.5.2.2. Criterios de inclusión clínicos**

Son referentes al espectro de la enfermedad.

- Especímenes de las subpoblaciones reactiva y no reactiva,

- Representatividad del espectro de la enfermedad. Incluir especímenes de diferente reactividad(i.e. baja, media, etc.) de ambas subpoblaciones,
- Incluir especímenes desafiantes(i.e. cercanas a los límites de detección reactivo y no reactivo). Estos especímenes se encuentran en los grupos **bajo reactivo** y **alto negativo**.

### **3.5.2.3. Criterios de inclusión de laboratorio**

Especímenes caracterizados con la referencia. Estos especímenes deben ser adecuados para realizar estudios de verificación.

# CAPÍTULO 4

## Ensayos experimentales

### Definiciones operativas

<b>Microplaca</b>	Arreglo rectangular que contiene los pocitos en donde se asigna la sustancia a analizar. Esta compuesta de 8 filas y 12 columnas, teniendo un total de 96 pocitos.
<b>Fración</b>	Parte representativa de una muestra biológica.

Los ensayos realizados durante una verificación, son lo que ayudarán a concluir el desempeño de la prueba. Los ensayos se basan en los protocolos/guías EP(Evaluation Protocol) por sus siglas en Inglés, desarrollados por la CLSI(Clinical & Laboratory Standards Institute), con el fin de proveer estándares y guías para los profesionales dedicados a la salud. Para la sección de ensayos experimentales se utilizarán los protocolos;

- EP12-A2 (Protocolo de usuario para la evaluación del desempeño de la prueba cualitativa) Este protocolo se utiliza para evaluar los parámetros de sensibilidad-especificidad o porcentajes de acuerdo positivo y negativo.
- EP15-A2 (Verificación del Desempeño de la Precisión y Veracidad por el usuario) Se emplea para medir la repetitividad en el laboratorio y entre cada uno de los ensayos (error entre corrida) y el error entre cada uno de los especímenes(error Intracorrida).

## 4.1. Ensayos para precisión

El planteamiento del ensayo se hace con base en el protocolo EP15-A2 de la sección (Experimento para Evaluar la Precisión (sección 6.2)).

*“El experimento para evaluar la precisión proporciona al usuario procedimientos guía para demostrar el desempeño de precisión. Usualmente, el fabricante realiza dos tipos de señalamientos de precisión-repetitividad (precisión intracorrída) ( $\sigma_r$ ) y precisión del laboratorio ( $\sigma_l$ ). Esta sección proporciona métodos estadísticos para identificar desviaciones gruesas para ambos tipos de señalamientos.”*

Para el ensayo primero se necesitan identificar los factores que influyen en la respuesta. Para lo que se hace referencia a la sección “Procedimiento específicos” (sección 8.2) del protocolo EP15-A2 (ver sección 5.1) en la que establece;

*“Analice una corrida por día con tres réplicas para cada una de dos concentraciones, diariamente por cinco días.”*

Con la frase anterior se identifican los siguientes factores;

- cantidad de días(cinco);
- número de réplicas por día(tres);
- cantidad de concentraciones(dos);

El objetivo del experimento es comparar la precisión obtenida en el laboratorio con lo declarado en el inserto por el desarrollador. Se verifican la precisión; **intracorrída (repetitividad)** y **entrecorrída (Precisión del Laboratorio)**. Para calcular la precisión se utilizan las expresiones del protocolo EP15 sección (8.4 y 8.5 con todos sus incisos o sólo los aplicables).

Hay que hacer énfasis en la cantidad (dos) de concentraciones o niveles de respuesta especificadas en la EP15-A2, pudiendo variar entre desarrollos aumentando o disminuyendo.

La definición de los niveles de respuesta (o concentraciones) bajo, medio, alto, etc., son conceptos de utilidad técnica. Los niveles se emplean en ambos ensayos (precisión y des-

#### 4.1.1 Asignación de especímenes para verificación de precisión

---

empeño). En el caso de los ensayos de precisión, los niveles se seleccionan acorde a lo especificado en el inserto de la prueba, siendo que, lo que se busca es demostrar que la precisión(repetitividad) de la prueba se encuentre en el rango especificado por el fabricante.

Para esta tesina se utilizarán los niveles bajo, medio y alto para ambos ensayos. Para el ensayo de desempeño, los límites de cada nivel están definidos en las tablas de las secciones 6.3.5.3 y 6.3.5.4. Para el ensayo de precisión solo se utilizan los niveles (bajo, medio y alto) sin especificar algún valor.

Algunas debilidades estadísticas identificadas en el ensayo para la precisión son que, el protocolo EP15-A2 no hace recomendaciones de:

- cómo seleccionar los especímenes;
- cómo asignarlos a la microplaca.

Para complementar esta debilidad, se propuso un algoritmo que seleccione y asigne los especímenes a la microplaca. En la sección 4.1.1 se propone un algoritmo para la asignación de los especímenes de manera aleatoria en la microplaca.

#### **4.1.1. Asignación de especímenes para verificación de precisión**

Como se menciona en la sección 3.3.2, en esta tesina se utilizan 5 especímenes de cada concentración (bajo, medio y alto), quince (15) en total. Cada espécimen de cada nivel, se fracciona en 3 fracciones (ver [Definiciones operativas](#)) para así obtener las 15 muestras (fracciones) recomendadas por nivel de concentración. El fraccionamiento de los especímenes se realiza para cada uno de ellos, para así obtener las 45 muestras(fracciones) para el ensayo de precisión.

Como el ensayo se ejecuta en cinco días, esto implica que se asignarán nueve fracciones por día a cada microplaca, tres (3) de cada nivel de respuesta. Una microplaca se puede representar como una matriz de 8 filas x 12 columnas, en donde cada entrada representa un pocito de reacción en la microplaca. A continuación, se propone un algoritmo que considera

a los tres niveles de respuesta utilizados.

Una vez caracterizados los grupos, las fracciones se deben asignar de manera aleatoria a los pocitos en la microplaca, con el fin de no introducir errores sistemáticos.

### **Algoritmo de asignación de especímenes para ensayo de precisión**

- 1) Con los grupos caracterizados se generan conjuntos, que representan a cada población de nivel de respuesta. Para lo que se generan los conjuntos B, M y A, en donde;
  - B es el conjunto de fracciones con nivel de respuesta bajo;
  - M es el conjunto de fracciones con nivel de respuesta media;
  - A es el conjunto de fracciones con nivel de respuesta alta;

Con los conjuntos poblacionales generados, se generan los conjuntos de fracciones que se evaluarán por día, a los que se les denominó: **“bloque”**. Para generar los bloques se genera la secuencia de números  $dia = \{1, 2, 3, 4, 5\}$  en donde cada elemento representa un día de ensayo. El fin de generar los bloques es conjuntarlos en una matriz a la que se le denominó como: Población de precisión. De los elementos de la secuencia “día”, se selecciona uno al azar sin reemplazo para generar los “bloques” hasta tener todos los bloques (cinco).

- 2) Del número al azar seleccionado en el punto uno, se genera el grupo de fracciones(bloque) con las que se hará el ensayo en cada día. Estos grupos se generan seleccionando una muestra al azar sin reemplazo de cada grupo (B, M y A).
- 3) Hay que recordar de la sección 4.1.1 que, para los cinco días, solo se requieren quince fracciones en total, cinco de cada nivel. Esto implica que en cada bloque solo hay una muestra de cada nivel (B, M y A) diferente. Por lo que, cada vez que se selecciona un elemento de los grupos B, M y A, se produce un proceso iterativo de tres ciclos, esto es, para un elemento seleccionado de B, M o A se ejecuta lo siguiente;
  - Iteración 1: Se selecciona un elemento de un grupo (B, M o A) al azar sin reem-

plazo y se selecciona en un pocito de la microplaca al azar;

- Iteración 2: Se asigna el mismo elemento de la iteración uno a otro pocito de la microplaca, seleccionado al azar y sin reemplazo;
  - Iteración 3: Se asigna el mismo elemento de la iteración uno a otro pocito de la microplaca, seleccionado al azar y sin reemplazo. Termina la asignación del elemento de alguno de los grupos.
- 4) El proceso del punto tres es iterativo de tres ciclos, por lo que, en cada bloque se tienen nueve(9) asignaciones en los pocitos de la microplaca (tres de cada grupo B, M y A). Terminando la generación del bloque.
- 5) El procedimiento del punto cuatro es iterativo de cinco ciclos. Cuando se generan los bloques se conforma la “Población de precisión”, mencionada en el punto uno. Teniéndose así las 45 observaciones independientes que se utilizaran en todo el ensayo de precisión.

## **4.2. Diseño experimental para desempeño**

En la sección 3.4.1 se mencionaron las características que debe de contener el grupo de especímenes para el experimento de desempeño, este experimento se ejecuta una sola vez y los resultados obtenidos con el método se utilizan para calcular la sensibilidad y especificidad de la prueba. Para lo que es necesario estratificar la población (banco de material biológico), para posteriormente seleccionar la población de referencia.

### **4.2.1. Estratificación de material biológico**

Se tiene el banco de material biológico con las subpoblaciones reactivas y no reactivas, la generación de los estratos son determinados por el experto técnico. Para generar los estratos;

- Se definen los límites inferiores y superiores, partiendo del valor de corte con su error.  
A partir de este valor, se generan los grupos de baja reactividad positiva y de alta

reactividad negativa. El valor de corte puede variar entre pruebas;

- Posteriormente se definen el resto de los límites con los que se generan los estratos faltantes de las subpoblaciones. Los de media y alta reactividad en el caso de la subpoblación reactiva y mediana y baja reactividad para la subpoblación no reactiva.

#### **4.2.2. Algoritmo de selección para población de referencia**

La muestra de desempeño tiene como objetivo evaluar la clasificación de las subpoblaciones utilizando una prueba/método, requiriéndose para dicho objetivo una muestra que represente de la mejor manera el intervalo de respuesta (espectro de la enfermedad) del que se disponga, a la cual se le denomina: población de referencia. Para seleccionar a esta población de referencia, se generan tres estratos, bajo, medio y alto, en cada subpoblación (reactiva y no reactiva), en donde cada estrato representa las diferentes reactividades. Con estos estratos, se seleccionan al azar especímenes de cada uno. El procedimiento para generar los estratos, se describe en el siguiente algoritmo.

##### **Descripción del algoritmo de selección y asignación de la población de referencia**

- Con el material biológico estratificado (sección 4.2.1), se seleccionan especímenes aleatoriamente de cada estrato, hasta tener el número especímenes de cada nivel de reactividad de cada subpoblación;
- Del estrato bajo reactivo de la subpoblación reactiva, es importante que, al momento de la selección, ésta incluya los especímenes desafiantes.
- Con los especímenes de las subpoblaciones reactiva y no reactiva, se conforma la población de referencia. Con la que se realizará la verificación de desempeño (validez externa).
- Una vez seleccionadas los especímenes se asignan de manera aleatoria a la microplaca.

En la discusión se habla acerca de la muestra (población de referencia).

#### 4.2.2 Algoritmo de selección para población de referencia

---

En el capítulo 6 se desarrolla un ejemplo aplicando la metodología descrita.



# CAPÍTULO 5

## Análisis de datos

En la sección 2.2.2 se mencionaron los parámetros que se pueden evaluar en una verificación.

Siendo el análisis y la ejecución de los ensayos experimentales los mismos para todos los parámetros independientemente de los parámetros que se quieran verificar (de acuerdo o de precisión diagnóstico).

Aunque la ejecución y análisis de los parámetros se realice de la misma manera, tienen diferente utilidad.

### **5.1. Análisis del experimento de precisión**

Para analizar la precisión se utilizan las expresiones del protocolo EP15-A2 (sección 8 con todos los subincisos)

#### **5.1.1. Análisis de precisión intracorrida**

Para calcular la precisión intracorrida se utiliza la siguiente expresión;

$$s_r = \sqrt{\frac{\sum_{d=1}^D \sum_{i=1}^n (x_{di} - \bar{x}_d)^2}{D(n-1)}} \quad (5.1)$$

En donde:

$s_r$  = Varianza intracorrída,

D = número de días(cinco),

n = número total de replicas por día (tres),

$x_{di}$  = resultado de la réplica i por día d, y,

$\bar{x}_d$  = promedio de todos los resultados por día d.

### 5.1.2. Análisis de precisión del laboratorio

Se calcula el término de varianza (Varianza de la media diaria),  $s_b^2$ , para la media diaria

$$s_b^2 = \frac{\sum_{d=1}^D (\bar{x}_d - \bar{\bar{x}})^2}{D-1} \quad (5.1)$$

En donde:

$\bar{x}_d$  = promedio de todos los resultados por día d ( $\bar{x}_1$  es el promedio por día 1), y

$\bar{\bar{x}}$  = promedio de todos los resultados.

Después se calcula  $s_1$  (Precisión del laboratorio) de la siguiente manera:

$$s_1 = \sqrt{\frac{n-1}{n} \cdot s_r^2 + s_b^2} \quad (5.2)$$

En donde n = número de réplicas por corrida (tres).

### **5.1.3. Comparación de la Repetitividad Estimada con la definida por el fabricante**

Para realizar la comparación se utiliza el valor calculado en la sección 5.1.1 y se compara con la declarada por el fabricante/desarrollador. Si el fabricante la reporta en términos de coeficiente de variación, se transforma a la desviación estándar de la concentración promedio de todos los resultados de los especímenes analizadas, teniéndose;

$$\sigma_r = CV \%_r \cdot \bar{x} \quad (5.1)$$

En donde:

$CV \%_r$  Es la repetitibilidad definida por el fabricante

Si la desviación estándar de la repetitibilidad estimada es menor a la definida por el fabricante, entonces se ha demostrado que la precisión es consistente con lo declarado.

### **5.1.4. Comparación de la Precisión Estimada Intralaboratorio con la definida por el fabricante**

Para realizar la comparación se utiliza el valor calculado en la sección 5.1.2 y se compara con la declarada por el fabricante/desarrollador. Si el fabricante la reporta en términos de coeficiente de variación, se transforma a la desviación estándar de la concentración promedio de todos los resultados de los especímenes analizados, teniéndose;

$$\sigma_1 = CV \%_{01} \cdot \bar{x} \quad (5.1)$$

En donde:

$CV \%_{01}$  Es el coeficiente de variación intralaboratorio declarado por el fabricante

Si la desviación estándar estimada intralaboratorio es menor a la definida por el fabricante, entonces se ha demostrado que la precisión es consistente con lo declarado.

## 5.2. Análisis del ensayo de desempeño

Para estimar la sensibilidad, especificidad, porcentaje de acuerdo positivo y negativo se utilizan las tabla de contingencia 2 X 2 como la tabla 5.1.

		Criterio de precisión diagnóstico		
		Positivo (+)	Negativo (-)	Total
Método evaluado	Positivo (+)	Verdaderos Positivos	Falso Positivo	VP + FP
	Negativo (-)	Falso Negativo	Verdaderos Negativos	VN + FN
	Total	VP + FN	FP + VN	N

**Tabla 5.1:** Tabla de contingencia para evaluar sensibilidad y especificidad.

### 5.2.1. Sensibilidad y Especificidad

Para la sensibilidad y especificidad, se utilizan las siguientes expresiones;

$$\text{Sensibilidad} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Negativos}}$$

## 5.2.2 Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo

---

$$\text{Especificidad} = \frac{\text{Verdaderos Negativos}}{\text{Verdaderos Negativos} + \text{Falsos Positivos}}$$

$$\text{Tasa de Falsos Negativos} = \frac{\text{Falsos Negativos}}{\text{Verdaderos Positivos} + \text{Falsos Negativos}}$$

$$\text{Tasa de Falsos Positivos} = \frac{\text{Falsos Positivos}}{\text{Verdaderos Negativos} + \text{Falsos Positivos}}$$

### 5.2.2. Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo

Al igual que la sensibilidad y especificidad, los porcentajes de acuerdo positivo y negativo, se calculan utilizando una tabla similar a la tabla 5.1. Con expresiones similares a las de la sección 5.2.1.

Ahora la tabla que se utiliza adopta la siguiente forma;

		Prueba de referencia		
		Positivo (+)	Negativo (-)	Total
Método evaluado	Positivo (+)	a	b	a + b
	Negativo (-)	c	d	c + d
	Total	a + c	b + d	n

**Tabla 5.2:** Tabla de contingencia para evaluar porcentajes de acuerdo positivo y negativo.

Las expresiones con las que se calculan los porcentajes de acuerdo son;

$$\text{Porcentaje de Acuerdo Negativo} = 100 \times \left( \frac{d}{b + d} \right)$$

$$\text{Porcentaje de Acuerdo Positivo} = 100 \times \left( \frac{a}{a + c} \right)$$

$$\text{Porcentaje de Acuerdo de la prueba} = 100 \times \left( \frac{a + d}{n} \right)$$

### 5.2.3. Interpretación de los parámetros Sensibilidad, Especificidad, Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo

Es importante diferenciar los parámetros, ya que, aunque aritméticamente se obtienen de la misma manera su confiabilidad no es la misma. Estas diferencias residen en qué;

- La tabla 5.2, no representa la condición de interés, proveyendo la frecuencia de concordancia con la prueba de referencia, contrario a la tabla 5.1 que provee información de la frecuencia en que el método evaluado identifica correctamente (asignando una probabilidad) la condición de interés;
- La mayor desventaja de los Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo es que, no son una medida de aserción, ya que dos pruebas pueden concordar y ambas ser incorrectas, ó dos pruebas pueden concordar pero ambas tener un Porcentaje de Acuerdo Positivo y Porcentaje de Acuerdo Negativo deficiente. Sin embargo, cuando dos pruebas no concuerdan no implica que el método evaluado clasifique incorrectamente y que el de referencia lo haga correctamente.[13]

### 5.3. Intervalos de confianza para Sensibilidad, Especificidad, Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo

Los intervalos de confianza son la precisión que se obtuvo en el experimento realizado, el cual está fuertemente asociado al tamaño de muestra. Para obtener los intervalos de confianza se utilizan las siguientes expresiones para Sensibilidad, Especificidad, Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo, los valores cambian según las tablas de contingencia 5.1 y 5.2.

Los intervalos de confianza consideran un 95 % de confianza para los parámetros de las tablas de contingencia. La manera de obtener los intervalos es;

#### 5.3.1. I.C para Sensibilidad y Especificidad [1]

$$\left[ 100 \times \frac{(Q_{1,se} - Q_{2,se})}{Q_{3,se}}, 100 \times \frac{(Q_{1,se} + Q_{2,se})}{Q_{3,se}} \right] \quad (5.1)$$

En donde para sensibilidad;

$$Q_{1,se} = 2 \times VP + 1.96^2 \quad (5.2)$$

$$Q_{2,se} = 1.96 \sqrt{1.96^2 + 4 \times VP \times \left( \frac{FN}{VP + FN} \right)} \quad (5.3)$$

$$Q_{3,se} = 2 \left( VP + FN + 1.96^2 \right) \quad (5.4)$$

En donde para especificidad;

$$Q_{1,esp} = 2 \times VN + 1.96^2 \quad (5.5)$$

$$Q_{2,esp} = 1.96 \sqrt{1.96^2 + 4 \times FP \times \left( \frac{VN}{FP + VN} \right)} \quad (5.6)$$

$$Q_{3,esp} = 2 \left( FP + VN + 1.96^2 \right) \quad (5.7)$$

### 5.3.2. I.C para Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo [1]

Para el caso de Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo se utilizan las expresiones 5.2, 5.3 y 5.4, cambiando las variables por las de la tabla 5.2, con lo que se obtiene;

para Porcentaje de Acuerdo Positivo;

$$Q_{1,PAP} = 2a + 1.96^2 \quad (5.1)$$

$$Q_{2,PAP} = 1.96 \sqrt{1.96^2 + \left( \frac{4ac}{a + c} \right)} \quad (5.2)$$

$$Q_{3,PAP} = 2 \left( a + c + 1.96^2 \right) \quad (5.3)$$

para Porcentaje de Acuerdo Negativo;

$$Q_{1,PAN} = 2d + 1.96^2 \quad (5.4)$$

$$Q_{2,PAN} = 1.96 \sqrt{1.96^2 + \left( \frac{4bd}{b + d} \right)} \quad (5.5)$$

$$Q_{3,PAN} = 2 \left( b + d + 1.96^2 \right) \quad (5.6)$$

# **Parte III**

## **Ejemplo de Aplicación**



# CAPÍTULO 6

## Ejemplo de Aplicación

Para poner en práctica los conceptos expuestos en la tesina, se desarrolló un ejemplo en donde se emplean para la verificación de una prueba. El ejemplo solo contempla el aspecto de la muestra, esto por tener de manera innata conceptos metodológicos y estadísticos, además la verificación de desempeño al involucrar aspectos metodológicos y estadísticos que no están en las recomendaciones de los protocolos EP12-A2 y EP15-A2 .

Para la verificación se utilizaron las definiciones operacionales del Laboratorio Nacional de Referencia de la enfermedad de Chagas (InDRE-SSA), así como información del material de referencia.

Como comentario; cada laboratorio debe de definir sus propias definiciones operaciones y material de referencia que representen sus condiciones de operación.

### **6.1. Definición del Problema**

Se desea verificar una prueba serológica para diagnosticar la enfermedad de **Chagas**, de interpretación cualitativa cuyo método de detección es instrumental basado en el principio de colorimetría.

Los parámetros de desempeño son;

- sensibilidad 99 %
- especificidad 99 %
- valor de corte 0.317

Se verificarán la precisión y desempeño de la prueba utilizando las recomendaciones de los protocolos EP12-A2 (desempeño) y EP15-A2 (precisión).

## **6.2. Verificación de precisión**

El objetivo de la verificación de la precisión es demostrar que las variaciones de las series de trabajo en las condiciones del laboratorio sean aceptables dentro de los límites especificados por el desarrollador.

### **6.2.1. Muestra de precisión**

Para la verificación de la precisión se debe de seleccionar un conjunto de especímenes(muestra) conformada por niveles de reactividad cercanas o iguales a las declaradas por el desarrollador.

#### **6.2.1.1. Criterios de inclusión de especímenes en la muestra**

La población de donde provienen estos especímenes no es un criterio de inclusión necesario para que forme parte de la muestra, siendo que el objetivo de la verificación de la precisión es demostrar la variación entre las series de repeticiones(intracorrida y entrecorrida) en las condiciones del laboratorio en donde se pretende implementar el método, para después comparar que la variación observada sea consistente con la declarada por el desarrollador.

#### **6.2.2. Selección de especímenes**

Para la selección de estos especímenes se generan estratos con las diferentes reactividades(baja, media, alta, etc.) con base en lo que esta declarado por el desarrollador y se selec-

cionan el número de especímenes (ver sección 3.3.2) hasta tener el número que recomienda el protocolo EP15-A2.

### **6.2.3. Asignación de los especímenes a la microplaca**

Para la asignación se puede hacer utilizando un algoritmo que emplee técnicas de muestreo (ver secciones 4.1, 4.1.1 y Algoritmo de asignación de especímenes para ensayo de precisión)

### **6.2.4. Cálculo de precisión intracorrida y entrecorrida**

Al terminar las series de trabajo de la precisión entrecorrida se hacen los cálculos correspondientes al ensayo (ver sección 5.1.1).

De igual manera se hacen los cálculos correspondientes al ensayo para la precisión entrecorrida (precisión de laboratorio ver sección 5.1.2)

Finalmente se calculan los coeficientes de variación intracorrida (ver sección 5.1.3) y entrecorrida (laboratorio ver sección 5.1.4) y se comparan con lo declarado. Obteniéndose la conclusión de la precisión de la prueba en las condiciones en donde se desea implementar.

## **6.3. Verificación de desempeño**

La verificación de desempeño se enfoca en evaluar la correcta clasificación de un conjunto de especímenes caracterizados (población de referencia) para calcular los parámetros de desempeño, estos parámetros son; sensibilidad, especificidad, tasa de falsos positivos y tasa de falsos negativos con base en lo observado en el ensayo. Finalmente se comparan los parámetros calculados con los declarados, de esta comparación se espera que lo observado sea al menos igual a lo declarado.

Para poder comparar la sensibilidad y especificidad de una prueba, es necesario compararla contra un criterio de precisión diagnóstico (ver secciones 2.7.2 y 2.7.2.1).

Para la verificación del desempeño, se utiliza como guía el protocolo EP12-A2.

### 6.3.1. Población objetivo

La población objetivo de manera general se define de la siguiente manera; Individuos mestizos mexicanos mayores a diez meses, de sexo indistinto, de cualquier región geográfica de la República Mexicana con presunta infección de **Chagas**, dividiéndose en 2 subpoblaciones; **población objetivo reactiva** y **población objetivo no reactiva**.

Las subpoblaciones se definen de la siguiente manera;

**Población objetivo reactiva:** En esta población se distinguen dos grupos, los crónicos sintomáticos y los crónicos asintomáticos la cual está cuantificada por el **intervalo de respuesta reactivo**;

**Crónicos sintomáticos:** Individuos mestizos mexicanos mayores a diez meses de edad, sexo indistinto, infectados con *Trypanosoma cruzi*, agente causal de la enfermedad de **Chagas** con sintomatología asociada a la enfermedad (afecciones cardíacas, afecciones gastrointestinales), de cualquier región geográfica de la República Mexicana;

**Crónicos asintomáticos:** Individuos mestizos mexicanos mayores a diez meses de edad, sexo indistinto, infectados con la enfermedad de **Chagas** sin sintomatología asociada a la enfermedad, de cualquier región geográfica de la República Mexicana;

**Población objetivo no reactiva:** Individuos mestizos mexicanos mayores a diez meses de edad, de sexo indistinto, no infectados con *T. cruzi*, agente causal de la enfermedad de Chagas, de cualquier región geográfica de la República Mexicana sin resultado positivo a enfermedad de Chagas. La población no reactiva está cuantificada por el intervalo de respuesta no reactivo. Los especímenes caracterizados para esta población corresponden a casos descartados o confirmados como no reactivos.

### 6.3.1.1. Espectro de la enfermedad

Para la enfermedad de **Chagas** el espectro de la enfermedad se refiere a la fase crónica (sintomática y asintomática), conociéndose la sintomatología que se presenta cuando se desarrolla el estado de enfermedad.

### 6.3.1.2. Intervalo de respuesta

El intervalo de respuesta a la enfermedad de **Chagas** está compuesto por el intervalo de respuesta no reactivo y el intervalo de respuesta reactivo

## 6.3.2. Referencia

### 6.3.2.1. Material de referencia

El Laboratorio Nacional de Referencia (LNR-InDRE SSA) cuenta con un banco de material biológico caracterizado. El objetivo del banco es obtener material de control y referencia con una calidad adecuada que permita calibrar reactivos diagnósticos de elaboración propia, evaluar el desempeño de equipos de diagnóstico comerciales y ser utilizado en protocolos de investigación[6].

La selección de las muestras biológicas que conforman el banco deben de cumplir características:

- epidemiológicas;
- clínicas;
- laboratorio.

Si cumplen estas características, se confirman si son reactivas a la enfermedad de **Chagas**.

El tipo de especímenes que conforman el banco de material biológico son de dos tipos:

- Caso confirmado;

- Caso descartado.

#### **6.3.2.2. Caso confirmado**

Un caso confirmado se define como: “*Todo caso probable en quien se demuestre, por técnicas directas o indirectas (mayor a 10 meses de edad) reconocidas por el InDRE, la presencia de T. cruzi.*” [6]

#### **6.3.2.3. Caso confirmado**

Un caso confirmado se define como: “*Caso probable en el que no se encuentra evidencia, por técnicas directas o indirectas reconocidas por el InDRE, de T. cruzi*” [6]

#### **6.3.2.4. Caracterización de material biológico**

Antes de seleccionar los especímenes que conformarán la población de referencia, estas se tienen que caracterizar con base en un método, si satisfacen este método, se agregan al banco de material biológico, que es de donde se selecciona la población de referencia. A este método se le denomina criterio de precisión diagnóstica.

El laboratorio de la enfermedad de Chagas caracteriza especímenes provenientes de diferentes estados de la República Mexicana. Para que estos especímenes puedan ser consideradas para para tal fin, deben de contar con las características de punto 6.3.2.1. Si los especímenes cumplen con esos requisitos, el laboratorio verifica la congruencia utilizando 4 pruebas diagnósticas de diferente formato y que, en caso de ser cierto, se envían al banco de material biológico. De lo contrario no se considera para esta aplicación.

Se utilizan 4 pruebas diagnósticas para caracterizar con base en las recomendaciones de la OMS(WHO) que recomienda al menos 3. Disminuyéndose la probabilidad de tener una clasificación errónea.

### **6.3.3. Muestra para desempeño(Población de referencia)**

Se selecciona un conjunto de especímenes, del banco de material biológico(ver sección 3.4.1).

La recomendación que sugiere el protocolo EP12-A2 en cuanto a los especímenes hace referencia únicamente a “la población blanco”. Especificando aquellos especímenes que presenten el padecimiento.

Los grupos de las subpoblaciones seleccionados tienen las siguientes características;

Para la población objetiva reactiva, los grupos de baja respuesta corresponden por lo general al grupo crónico asintomático, mientras que los de respuesta media y alta al grupo crónico sintomático. Haciendo nuevamente énfasis en los especímenes desafiantes seleccionados en el grupo bajo reactivo.

Para la población no reactiva, la única consideración es la representatividad del intervalo de respuesta no reactivo, de igual manera haciendo énfasis en el grupo alto no reactivo.

Algunos padecimientos comparten características biológicas, en el caso de la enfermedad de Chagas estos padecimientos son la **Leishmaniasis** y **Toxoplasmosis**. Por lo que además de los especímenes asociados al padecimiento(Chagas) también se deben incluir especímenes para verificar que no existan reacciones cruzadas o inespecíficas con los padecimientos con los que comparte características.

Para la inclusión de los especímenes, en este caso se utilizan tres criterios.

#### **6.3.3.1. Criterios de inclusión de la muestra**

Los criterios de inclusión son epidemiológicos, clínicos y de laboratorio

##### **Epidemiológicos**

- Raza: Individuos mestizos mexicanos,

- Edad: mayores a 10 meses,
- Sexo: Femenino o Masculino,
- Región: Cualquier estado de la República Mexicana

### **Clínicos**

Para este ejemplo se utilizaron tres niveles de reactividad para cada subpoblación, los niveles fueron alto, medio y bajo.

La **subpoblación reactiva** está conformada por especímenes crónicos sintomáticos y asintomáticos. El nivel de reactividad bajo lo conforman los crónicos asintomáticos(infectados). Este grupo no tiene manifestaciones clínicas, teniendo un diagnóstico clínico negativo, siendo la respuesta positiva a pruebas serológicas la manera de confirmar la presencia de anticuerpos anti *Trypanosoma Cruzi*.

Los niveles medio y alto de la subpoblación reactiva lo conforman los crónicos sintomáticos (infectados y enfermos). Este grupo por lo regular cuentan con un diagnóstico clínico positivo siendo que presentan manifestaciones clínicas (alteraciones gastrointestinales, cardiológicas y neurológicas). Además de contar con pruebas serológicas que confirman la presencia de anticuerpos anti *Trypanosoma Cruzi*.

La **subpoblación no reactiva** fueron especímenes que cuentan con diagnóstico clínico negativo y además cuentan con resultados de pruebas serológicas negativas a anticuerpos anti *Trypanosoma Cruzi*. Los especímenes desafiantes se encuentran en los grupos bajo reactivo y alto negativo.

### **Laboratorio**

Todos los especímenes tienen 4 pruebas de diferente formato con resultado concordante.

### 6.3.4. Tamaño de la muestra

Por lo regular el tamaño de la muestra se calcula utilizando la expresión de una proporción (expresión 3.1), que tiene la siguiente forma;

$$n = \frac{Z_{\alpha/2}^2 pq}{d^2} \quad (6.1)$$

En donde:

- Probabilidad de éxito(p),
- Probabilidad de fracaso(q),
- Precisión(d);
- Nivel de confianza(z)

Los tamaños de muestra obtenidos son los **mínimos**. Se puede aumentar el tamaño de la muestra si se dispone de material.

#### 6.3.4.1. Tamaño de muestra para subpoblación reactiva

Sustituyendo los valores de la prueba (sensibilidad/falsos negativos o especificidad/falsos positivos) y el nivel de confianza se tiene que el tamaño de muestra es;

- Probabilidad de éxito (sensibilidad) = 99 % ó 0.99,
- Probabilidad de fracaso (falso negativo) = 0.01 % ó 0.01,
- Precisión(d) = 5 % ó 0.05;
- Nivel de confianza(z) = 95 % (1.96)

$$n = \frac{(1.96^2)(0.99)(0.01)}{(0.05)^2} \quad (6.2)$$

$$n = 15.21 \Rightarrow 15 \quad (6.3)$$

Teniéndose un tamaño de muestra de 15. Sin embargo, por la disponibilidad de material se decide incrementar a 30 especímenes.

#### 6.3.4.2. Tamaño de muestra para subpoblación no reactiva

Sustituyendo los valores de la prueba y el nivel de confianza se tiene que el tamaño de muestra es;

- Probabilidad de éxito(especificidad) = 99 % ó 0.99,
- Probabilidad de fracaso(falso positivo) = 0.01 % ó 0.01,
- Precisión(d) = 5 % ó 0.05;
- Nivel de confianza(z) = 95 %(1.96)

$$n = \frac{(1.96^2)(0.99)(0.01)}{(0.05)^2} \quad (6.4)$$

$$n = 15.21 \Rightarrow 15 \quad (6.5)$$

Teniéndose un tamaño de muestra de 15. Sin embargo, por la disponibilidad de material se decide incrementar a 30 especímenes.

Con el cálculo del tamaño de muestra de cada subpoblación se tiene que la población de referencia contiene 60 especímenes. No olvidando agregar los especímenes para verificar reactividad cruzada. Estos especímenes se seleccionan a criterio del experto.

Para seleccionar los especímenes es necesario primero estratificar el material biológico(banco de material) para su posterior selección.

Los algoritmos secciones 4.2.1 y 4.2.2 se ejecutan a continuación.

## **6.3.5. Estratificación del material biológico**

### **6.3.5.1. Definición de límites superiores e inferiores**

Se tiene un total de 1590 especímenes caracterizadas, de las cuales 1035 son de reactividad positiva y 555 de reactividad negativa. Se generaron las subpoblaciones de referencia reactiva y no reactiva a partir de especímenes seleccionadas de cada uno de los estratos (bajo, medio y alto).

Para definir los límites inferiores y superiores utilizó un **valor de corte de 0.317** con un error del **10 %**(ver sección 3.1.4).

Para la reactividad positiva se obtuvo el límite inferior de **0.348**, con lo que los especímenes para reactividad positiva deben de ser seleccionadas con una reactividad mayor o igual al límite inferior. La menor reactividad que se encuentra en el grupo positivo es **0.35**, mientras que la reactividad máxima es de **2.03**. Entre este rango se deben de definir los niveles que se quiera estudiar.

En la reactividad negativa se obtuvo el límite superior de **0.285**, siendo este límite la reactividad máxima que el grupo negativo puede contener. En este grupo la reactividad máxima que se tiene es **0.282** y la mínima reactividad es de **0.033**. Siendo el intervalo de selección.

### **6.3.5.2. Definición de los límites de cada nivel**

Los niveles generados fueron tres, bajo, medio y alto para la reactividad de cada grupo; recordando que esta reactividad se mide en absorbancia (ver inicio Capítulo 3).

### 6.3.5.3. Niveles para reactividad positiva

Para las reactividades de la subpoblación reactiva se establecieron los siguientes niveles A  
CRITERIO;

---

Límites de selección reactividad positiva	
Reactividad baja	$0.348 \leq m \leq 0.450$
Reactividad media	$0.450 < m \leq 1.400$
Reactividad alta	$m > 1.400$

---

### 6.3.5.4. Niveles para reactividad negativa

En el caso de las reactividades de la subpoblación no reactiva se establecieron los niveles;

---

Límites de selección reactividad negativa	
Reactividad baja	$m < 0.100$
Reactividad media	$0.100 \leq m < 0.200$
Reactividad alta	$0.200 \leq m \leq 0.282$

---

En donde **m** representa la reactividad del espécimen seleccionado.

Con los límites establecidos se obtuvieron los siguientes estratos.

Para la subpoblación reactiva, los estratos tienen los siguientes tamaños;

- Estrato bajo: 79;
- Estrato medio: 861;
- Estrato Alto: 95.

La subpoblación no reactiva se conformó de la misma manera que la reactiva, los estratos tienen el siguiente tamaño:

- Estrato bajo: 207;
- Estrato medio: 256;
- Estrato Alto: 92.

Con los estratos generados en las subpoblaciones se procedió a generar la población de referencia.

#### **6.3.6. Selección de especímenes y conformación de la población de referencia**

Para la selección de los especímenes el protocolo EP12-A2 no recomienda alguna forma/técnica. Por lo que se propuso un algoritmo (ver sección 4.2.2) que emplea estratos de reactividades para la selección de los especímenes de la población de referencia.

##### **6.3.6.1. Selección de especímenes en subpoblación reactiva y conformación de subpoblación de referencia positiva**

Los especímenes se seleccionaron al azar y sin reemplazo en cada estrato (ver sección 4.2.2). Recordando que el estrato de mayor importancia en esta subpoblación es el bajo reactivo y que debe de incluir especímenes con reactividades **desafiantes**. Esto es, con valores iguales o cercanos al límite inferior de la reactividad positiva.

**Comentario sobre el tamaño la población de referencia** Originalmente se propuso un tamaño de 30 para la población de referencia, sin embargo, al momento de seleccionar los especímenes en el estrato bajo reactivo se observó que de cada 10 poblaciones de referencia seleccionadas solo 3 poblaciones de referencia seleccionaban al espécimen desafiante. Por lo que se decidió aumentar el tamaño de muestra a 50 especímenes, con un tamaño de 35 para el estrato bajo reactivo. Con este tamaño se observó que, de cada 10 poblaciones de referencia seleccionadas, 7 poblaciones de referencia incluían al espécimen desafiante lo que es “bueno” dada la variabilidad del fenómeno. En el Anexo D (ver sección C.1), se puede

ver el desempeño de la selección de las muestras con diferentes tamaño de estrato.

Finalmente, la población de referencia tuvo los estratos con los siguientes tamaños;

- Estrato bajo está compuesta en este caso por el 44 % de los especímenes en el estrato (79 espécimenes en total en el estrato), el tamaño de muestra fue 35;
- Estrato medio: 10;
- Estrato alto: 5.

En total la subpoblación de referencia reactiva tiene 50 especímenes.

#### **6.3.6.2. Selección de especímenes en subpoblación no reactiva y conformación de subpoblación de referencia negativa**

Al igual que para la subpoblación de referencia reactiva, los especímenes de la subpoblación no reactiva se seleccionaron al azar y sin remplazo.

Para este caso el estrato de mayor importancia fue el alto negativo, con reactividades iguales o muy cercanas al límite alto negativo.

Los tamaños de muestra fueron;

- Estrato bajo 6;
- Estrato medio: 8;
- Estrato alto: 16.

Teniéndose una población de referencia negativa de tamaño 30.

Conjuntando ambas subpoblaciones (reactiva y no reactiva), se tiene que la población de referencia contiene 80 observaciones (especímenes).

El algoritmo se ejecutó una sola vez, implementado en R. Los resultados se muestran a continuación.

## 6.3.7. Ejecución del algoritmo

### 6.3.7.1. Selección de especímenes de la subpoblación reactiva

La muestra del estrato bajo reactivo cuenta con las siguientes reactividades;

```
[1] 0.350 0.359 0.362 0.362 0.365 0.367 0.380 0.383 0.385 0.387
[11] 0.387 0.389 0.396 0.402 0.402 0.407 0.409 0.414 0.416 0.419
[21] 0.424 0.425 0.426 0.426 0.428 0.428 0.428 0.429 0.429 0.432
[31] 0.435 0.435 0.436 0.436 0.442
```

La muestra del estrato medio reactivo cuenta con las siguientes reactividades;

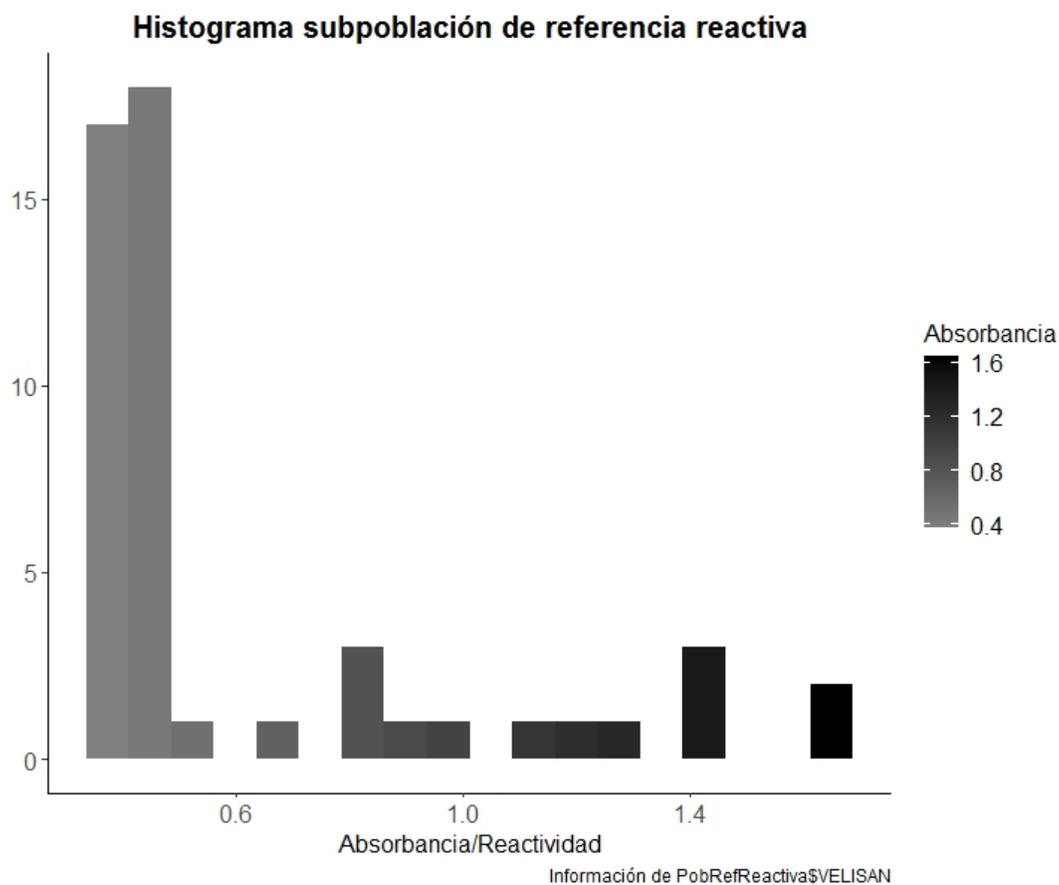
```
[1] 0.523 0.658 0.795 0.801 0.835 0.914 1.012 1.141 1.222 1.307
```

La muestra del estrato alto reactivo cuenta con las siguientes reactividades;

```
[1] 1.427 1.448 1.449 1.615 1.621
```

Con lo que la subpoblación reactiva se compone de 50 especímenes.

El histograma de la subpoblación reactiva fue el siguiente;



**Figura 6.1:** Histograma subpoblación reactiva.

La subpoblación no reactiva se obtuvo de la misma manera que la reactiva, obteniéndose los siguientes tamaños de muestra en los estratos;

Estrato bajo;

[1] 0.039 0.057 0.061 0.079 0.096 0.098

Estrato medio;

[1] 0.110 0.147 0.150 0.156 0.177 0.182 0.190 0.196

Estrato alto;

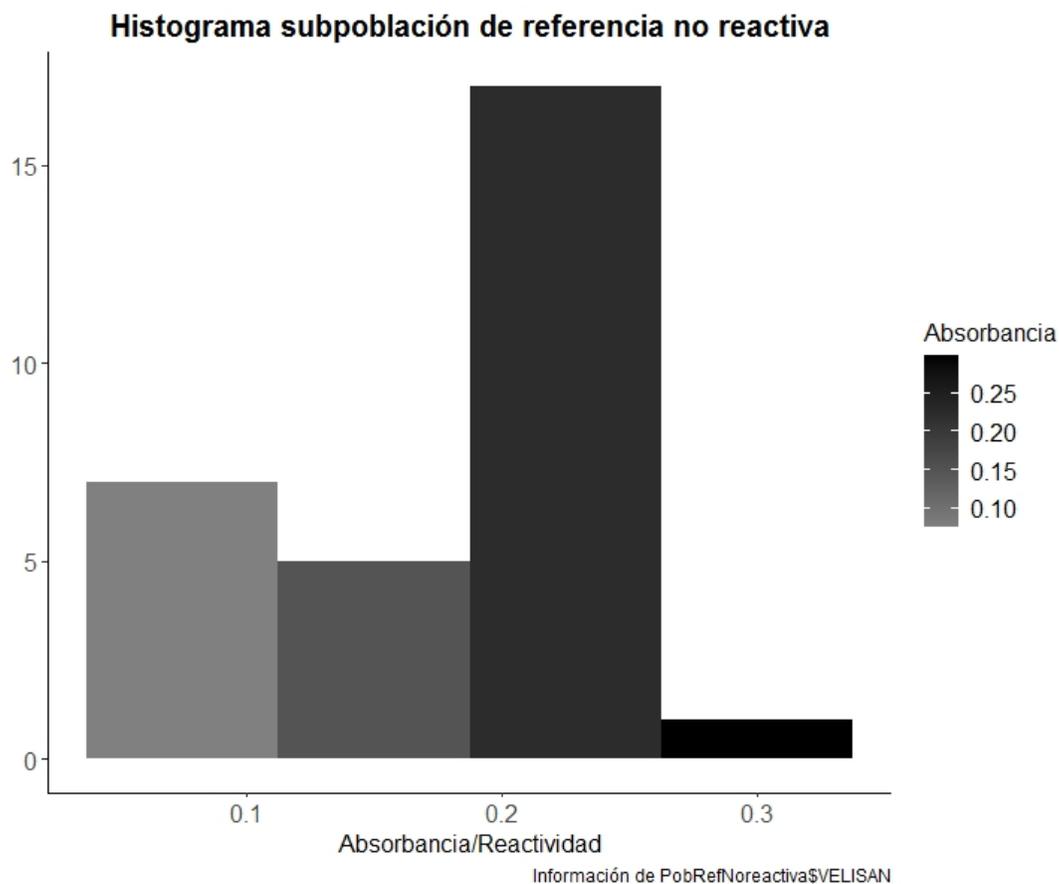
### 6.3.7 Ejecución del algoritmo

---

```
[1] 0.202 0.206 0.213 0.215 0.218 0.219 0.219 0.225 0.227 0.229 0.234  
[12] 0.236 0.244 0.246 0.254 0.274
```

Con lo que la subpoblación no reactiva se compone de 30 especímenes.

Su histograma se muestra a continuación;



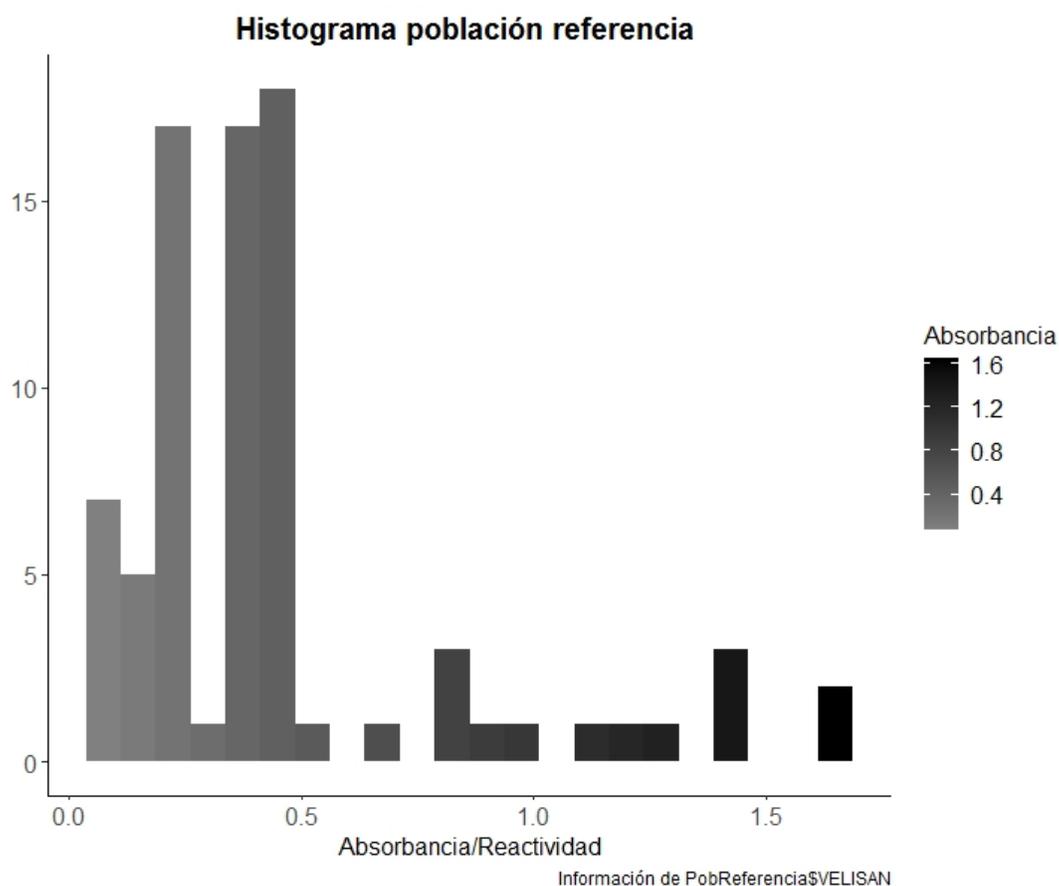
**Figura 6.2:** Histograma subpoblación no reactiva.

Con ambas subpoblaciones obtenidas se conformó la siguiente población de referencia;

```
[1] 0.039 0.057 0.061 0.079 0.096 0.098 0.110 0.147 0.150  
[10] 0.156 0.177 0.182 0.190 0.196 0.202 0.206 0.213 0.215  
[19] 0.218 0.219 0.219 0.225 0.227 0.229 0.234 0.236 0.244  
[28] 0.246 0.254 0.274 0.350 0.357 0.362 0.363 0.363 0.365  
[37] 0.367 0.368 0.371 0.373 0.374 0.375 0.380 0.383 0.385
```

```
[46] 0.386 0.392 0.399 0.401 0.409 0.410 0.410 0.412 0.414
[55] 0.414 0.417 0.418 0.418 0.421 0.428 0.428 0.432 0.435
[64] 0.435 0.442 0.591 0.592 0.766 0.795 0.801 0.811 0.916
[73] 0.973 1.081 1.202 1.403 1.425 1.443 1.745 1.820
```

El histograma se muestra continuación

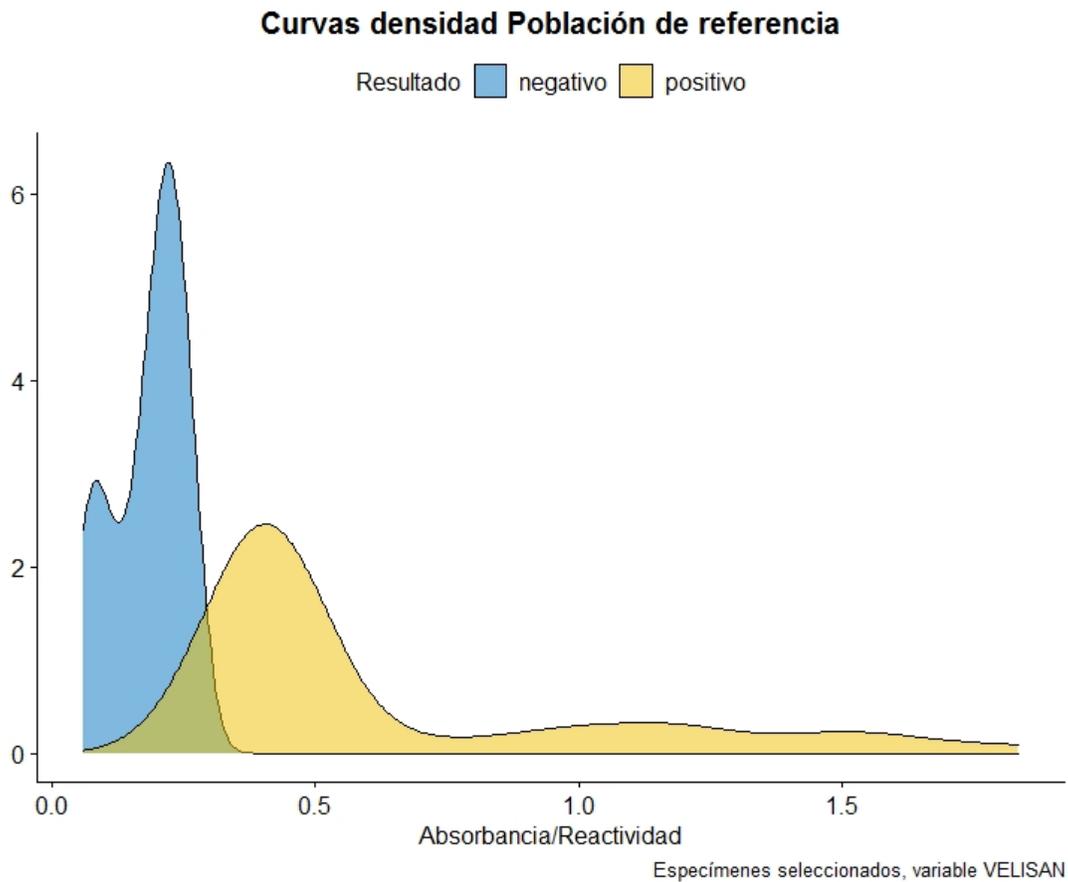


**Figura 6.3:** Histograma población de referencia.

De los histogramas se observa que en la supoblación reactiva falta información, sin embargo, todos los especímenes satisfacen los criterios de inclusión para la población de referencia, teniendo utilidad la selección. En la población negativa, no se observa falta de información.

**6.3.7.2. Gráfico de densidades del banco de material biológico(población objetivo)**

En la sección anterior se mostraron los histogramas, sin embargo, no se aprecia en donde se superponen las subpoblaciones reactiva y no reactiva.

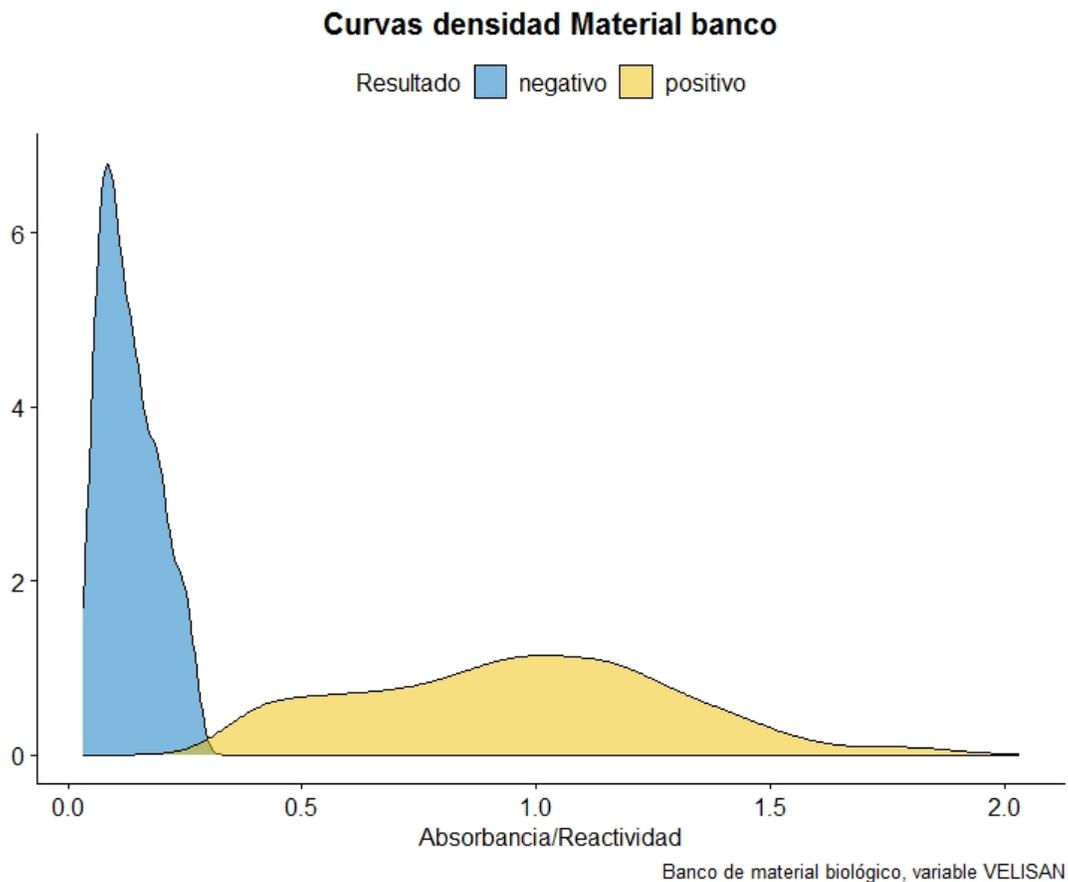


**Figura 6.4:** Gráfico densidades de población de referencia.

En la gráfica 6.4 se observa que las subpoblaciones no siguen una distribución normal, aunque la selección tiene utilidad para la verificación.

### 6.3.7.3. Gráfico de densidades de la muestra(población de referencia)

La gráfica 2.1 es lo que de manera teórica sucede en donde se observan poblaciones con distribución normal, sin embargo, en el gráfico 6.5 se puede observar que el comportamiento en la realidad no se acerca a la normal.



**Figura 6.5:** Gráfico densidades banco de material.

### 6.3.8. Asignación de los especímenes a la microplaca(con arreglo 8 filas x 12 columnas)

Para la asignación se puede hacer utilizando un algoritmo que emplee técnicas de muestreo(ver sección 4.2.2)

#### **6.3.9. Cálculo de parámetros de desempeño**

Al terminar el experimento, se calculan los parámetros de desempeño.

Los parámetros de desempeño se calculan contando la frecuencia de las diferentes salidas de los resultados.

Las posibles salidas son positivo(reactivo), no reactivo(negativo), falso positivo(resultado positivo cuando no tiene el padecimiento), falso negativo(resultado negativo cuando tiene el padecimiento).

Para calcular la sensibilidad y la especificidad, la frecuencia de los resultados del ensayo se insertan en una tabla de contingencia de 2 x 2(ver sección 5.2).

Retomando que los parámetros pueden ser de(ver sección 2.4);

- Acuerdo(Porcentaje de Acuerdo Positivo y Negativo) ver sección 5.2.2,
- Precisión diagnóstica(sensibilidad y especificidad)

Cuando se calculan se comparan con los declarados. Lo que se espera es que sean al menos iguales.

##### **6.3.9.1. Alcance tipo de parámetros**

Independientemente del tipo de parámetros que se declaren, ambos se calculan de la misma manera, con la tabla de contingencia de 2 x 2. Aunque numéricamente se calculen de la misma forma, su alcance para toma de decisión no es el mismo.

Los parámetros declarados de acuerdo indican si el método que se esta verificando tiene el suficiente acuerdo con respecto al método con el que se está comparando.[1]

Para la sensibilidad y especificidad del método que se esta verificando se compara contra un criterio de precisión diagnóstico. La precisión diagnóstica de una prueba candidata se refiere al grado de acuerdo entre la información de la prueba bajo evaluación y el criterio de

precisión diagnóstica. Además, con la sensibilidad y especificidad es posible calcular Valores Predictivos Positivos y Negativos, no siendo posible con los parámetros de acuerdo.[1]

Ver sección 5.2.3 Interpretación de los parámetros Sensibilidad, Especificidad, Porcentaje de Acuerdo Negativo y Porcentaje de Acuerdo Positivo

### **6.3.10. Cálculo de intervalos de confianza**

Con los parámetros de desempeño calculados se pueden obtener intervalos de confianza. Los intervalos de confianza muy rara vez se declaran en las especificaciones de la prueba, sin embargo, resultan útiles a los usuarios para registrar el comportamiento en el tiempo de la prueba. Muestra la precisión de los estimadores de sensibilidad y especificidad, tamaño del intervalo, así como los límites superior e inferior.

Para calcular los intervalos de confianza el protocolo EP12-A2 recomienda las expresiones de la secciones 5.3.1 y 5.3.2. Dependiendo de los parámetros verificados.

**Parte IV**  
**Conclusiones**



# Discusión

Durante el desarrollo de la metodología, se identificaron en los protocolos algunas debilidades estadísticas y metodológicas, proponiendo recomendaciones (ver sección [Recomendaciones](#)) las cuales ayudarían a desarrollar protocolos adecuados para el fin, que es verificar una prueba/método para clasificación de individuos con alguna enfermedad.

Definir operativamente el concepto con el cual se definirán los niveles de estudio requiere de un amplio conocimiento del fenómeno, en este caso el concepto intervalo de respuesta reactivo y no reactivo, que representa la cuantificación del concepto de espectro de enfermedad a partir del cual se pueden definir los niveles del estudio. Esto es importante saberlo siendo aplicable a cualquier fenómeno con conocimiento similar que se desee estudiar.

En los experimentos de precisión, ya se establecía previamente en la metodología la selección de las muestras a diferente nivel de respuesta (estratificación), a lo que se complementó la metodología agregando la selección de la muestra de precisión y la asignación, ambas de manera aleatoria en la microplaca.

Para la selección y asignación se propuso un algoritmo que emplea las técnicas de muestreo como se revisan en la literatura (i.e. de una manera “rígida”). Esto es posible porque, la población que consiste en mediciones de reactividades, no se considera como criterio de inclusión el origen (étnico, socioeconómico, etc.) de estas reactividades, es decir; se pueden utilizar cualesquier especímenes siempre y cuando cumpla con los requisitos especificados (reactividad) para el experimento.

La principal duda y pregunta al momento de desarrollar algún estudio, yace en el diseño de muestra con la que se pretende extrapolar las conclusiones a la población objetivo. Este diseño involucra la forma de selección de elementos y el tamaño que esta debe tener. Aunque desde el punto de vista estadístico, los especímenes de la muestra se deben seleccionar

con alguna técnica de muestreo para evitar una posible dependencia, en algunos estudios de carácter biológico, tales como los serológicos, puede ser contraproducente, esto al no tener la misma probabilidad de selección la variable que se desea medir en los especímenes, lo cual puede sesgar la muestra, obteniéndose una muestra no útil. Para atenuar este efecto, se suele estratificar la población en función de lo que se quiera estudiar, por lo que, es imperante que se tenga un vasto conocimiento del fenómeno. Para la población de referencia se propuso un algoritmo para la selección de las muestras en cada uno de los estratos definidos.

Al seleccionar una muestra, una de las principales preguntas recae en el tamaño de la misma y, aunque es importante para obtener estimaciones con la precisión establecida, en algunos estudios, tales como en una verificación de desempeño no se busca inferir algún parámetro a partir de la muestra. Lo que se busca es tener una muestra que sea representativa del espectro de la enfermedad en una población, por lo que la obtención de una muestra “buena” depende en gran medida del conocimiento del experto al momento de estratificar la población y al momento de seleccionarla (población de referencia). Sin embargo, la “reactividad” que es la variable que se mide, no tiene la misma frecuencia en la población (i.e. la gran mayoría de los casos se presentan en las reactividades medias a altas) siendo las reactividades bajas las de menor frecuencia. Como se mencionó en la sección 3.4.1 hay ciertos especímenes que siempre deben de estar incluidos en la población de referencia, siendo esta una restricción importante al momento de emplear técnicas de muestreo. Esto porque se debe de buscar la manera de que siempre se incluyan, para lo que se puede hacer de dos maneras; agregándolas de manera consciente (a conveniencia) o aumentar el tamaño de muestra cómo se realizó en la sección 6.3.6, en este caso se consideró un tamaño de muestra del 44 % del total del estrato bajo reactivo, seleccionando 35 de los 79 especímenes que tiene el estrato bajo reactivo, con este tamaño, prácticamente se “garantiza” que los especímenes de interés se encontraran en la muestra (población de referencia), siendo además que como se explicó en la sección 3.4.1 conforme aumenta la reactividad o disminuye, según sea el caso tiende a un evento seguro pudiendo disminuir el número de especímenes de esas reactividades y aumentando en donde es más crítico detectar posibles clasificaciones erróneas.

Los resultados de los ensayos experimentales con ambas muestras (muestra de precisión y población de referencia), se comparan con los declarados en el inserto. Los resultados declarados en el inserto se presentan como valores puntuales. Sin embargo, desde el punto de vista de un fenómeno con gran variabilidad, no es del todo acertado reportarlos de esta manera. En estos casos es mejor representarlos como intervalo de confianza, con un rango que se considere plausible para la población[19].

Aunque en el protocolo EP12-A2 los menciona e indica la manera de obtenerlos, muchas veces no se pueden comparar con los del método que se está comparando al no estar reportados, sin embargo, son de gran utilidad para los técnicos que aplican el método. Se recomienda obtenerlos si el fabricante no los declara con los datos reportados en el inserto. En el algoritmo para seleccionar las muestras biológicas para la población de referencia, se le asignó un porcentaje de error al valor de corte, a partir del cual se seleccionaron los especímenes. Al no estar reportado el error en el inserto se puede sobreestimar o subestimar el parámetro, lo que eventualmente conlleva a una posible selección de muestras en la población de referencia que no se puedan clasificar correctamente por la prueba bajo estudio, afectando los resultados del estudio. Representando una debilidad metodológica en el diseño de la prueba.

Al reportar los resultados de una prueba verificada se debe tener precaución, esto porque; aunque se utilice un criterio de precisión diagnóstico con el cual se puede comparar cualquier prueba, no implica que se deba de reportar como sensibilidad y especificidad que son los parámetros operativos del criterio de precisión diagnóstico, sino que se debe de reportar de acuerdo a los parámetros especificados en el inserto, ya sea sensibilidad, especificidad o acuerdo.

Al reportar pruebas cuyos parámetros son porcentajes de acuerdo; el porcentaje de acuerdo total puede ser engañoso, ocurriendo que el porcentaje de acuerdo total es bueno en comparación cuando el porcentaje de acuerdo positivo y negativo son muy bajos, siendo mejor utilizar los porcentajes de acuerdo positivo y negativo como medida para caracterizar el desempeño de la prueba.

La generalidad de los elementos que articulan a la metodología en una verificación se determina según el objetivo que se establezca. Elementos tales como, criterios de comparación, población(es), muestra(s), etc.

El criterio de precisión diagnóstico es el instrumento con el que se compara alguna prueba, por lo que, es importante identificar el grado de generalidad que éste representa dado el tipo de especímenes de determinada población con el que está sustentado, siendo aplicable únicamente a esa población.

Metodológicamente, el criterio de precisión diagnóstico utilizado por el laboratorio de **Chagas** tiene validez interna y muy poca validez externa, esto sucede al estar ajustado únicamente con especímenes provenientes de los diferentes estados de la República Mexicana. Si se pretendiera utilizar este criterio de precisión diagnóstico en otras poblaciones ajenas a la mexicana, el error esperado estaría alejado del observado.

La verificación del desempeño implica verificar la validez interna y externa de la prueba. Esto sucede porque, como se mencionó, la prueba que se está evaluando fue ajustada con poblaciones ajenas a la mexicana o en el mejor de los casos, también se utilizó una muestra de la población mexicana para su ajuste, no necesariamente representativa (sesgada). Esto se hace de esta manera siendo, que, lo que se busca con estas pruebas es poder tener mayor poder de clasificación del padecimiento en múltiples poblaciones, por lo que tiene mucha validez interna para las poblaciones con la que fue ajustada y limitada validez externa para otras. Siendo necesario efectuar la verificación de desempeño para poder determinar si el grado de validez externa que posee es útil para ser empleado como método de clasificación en una población diferente, en este caso la mexicana.

# Recomendaciones

## **Recomendaciones metodológicas muestra de precisión**

### **Criterios de inclusión**

Como se explicó el objetivo de la verificación de la precisión es la comparación de variaciones entre series de trabajo. No teniendo mucha importancia el origen de los especímenes, sino solo su reactividad. Considerándose como único criterio de inclusión especímenes con niveles adecuados para comparar las variaciones observadas con las declaradas.

### **Estratificación del material**

En el protocolo EP15-A2, se recomiendan 2 niveles(estratos) de reactividad, sin embargo, el número de niveles debe de ser acorde a los especificado en el inserto/manual de la prueba, estos por lo general son 3 niveles.

## **Recomendaciones estadísticas muestra de precisión**

### **Selección de los especímenes**

Es posible utilizar técnicas de muestreo sin riesgo de sesgar la muestra. Esto porque los niveles de reactividad ya están especificados. Además complementarías al protocolo EP15-A2, no haciendo mención/recomendación alguna de la selección de los especímenes.

### **Asignación de los especímenes**

Es recomendable emplear técnicas de muestreo para la asignación de los especímenes a la microplaca, evitando introducir algún error sistemático. Complementando de igual manera al protocolo EP15-A2, no haciendo mención o recomendación de los posibles errores que se pueden introducir al momento de asignar los especímenes.

## **Recomendaciones metodológicas población de referencia**

### **Recomendaciones epidemiológicas**

#### **Origen de los especímenes**

Para realizar el ensayo de los parámetros de desempeño se utiliza el protocolo EP12-A2 como guía. Sin embargo, el protocolo hace algunas diferenciaciones en cuanto al origen de los especímenes.

Esta diferenciación consiste en el material biológico dependiendo de cómo se reporten los parámetros de desempeño, algunas pruebas reportan parámetros de acuerdo positivo y negativo que no es lo mismo que sensibilidad y especificidad. La diferencia está en el alcance que tiene un resultado emitido con la prueba, es decir, con la sensibilidad y especificidad es posible calcular valores predictivos positivos y valores predictivos negativos, mientras que con los parámetros de acuerdo no es posible. Además, para verificar la sensibilidad y especificidad de una prueba es necesario emplear un criterio de precisión diagnóstico.

Independientemente de cómo se declaren los parámetros de desempeño es deseable que se utilicen especímenes provenientes de la población objetivo en donde se pretende emplear la prueba.

---

## **Recomendaciones de laboratorio y clínicos**

### **Caracterización de los especímenes**

Los especímenes utilizados para la verificación de desempeño se seleccionan del banco de material biológico. Este banco de material biológico tiene la utilidad de criterio de precisión diagnóstico, siendo que sus especímenes tienen criterios epidemiológicos, clínicos y de laboratorio. Por las características mencionadas, el criterio de precisión diagnóstico tiene “muchas” validez interna y limitada validez externa para poblaciones ajenas para la cual fue ajustado lo cual metodológicamente es adecuado para una verificación. Por todo lo expuesto, el objetivo de una verificación se podría complementar redefiniéndose como: **demostrar que el método/prueba verificado tenga utilidad para clasificar con el desempeño declarado individuos mestizos mexicanos con sospecha de enfermedad de Chagas.**

## **Recomendaciones estadísticas para población de referencia**

### **Estratificación del material**

La estratificación del material se genera dependiendo de su disponibilidad. Los niveles se definen por parte del experto. El estratificar complementa al protocolo EP12-A2, que es el que sirve de apoyo para la verificación de desempeño, no mencionando o recomendando criterios para generar niveles de reactividad. Por lo regular se utilizan 3 niveles para cada subpoblación de reactividad (reactiva y no reactiva); bajo, medio y alto. La cantidad de niveles puede variar dependiendo del padecimiento/enfermedad.

### **Selección de los especímenes**

Recordando que el objetivo es representar lo mejor posible el espectro de la enfermedad, seleccionando especímenes de los niveles establecidos, incluyendo además especímenes que desafíen a la prueba/método. Esto se puede lograr utilizando técnicas de muestreo, sin em-

bargo, la limitación(ver Anexo D, sección C.1) de las técnicas existe al intentar seleccionar los especímenes desafiantes ya que se tiene que aumentar el tamaño de la selección para “garantizar” su selección. Por lo que su uso se limitaría a los grupos medio y alto. Mientras que, para el grupo bajo reactivo y alto negativo, habría que analizar otras opciones para garantizar la inclusión de los especímenes desafiantes utilizando métodos de muestreo y/o estadísticos. Por lo regular y de manera empírica la población de referencia se selecciona a conveniencia.

Al igual que para la selección de la muestra de precisión, el protocolo EP12-A2, no hace mención/recomendación de cómo seleccionar a los especímenes complementándolo habiendo utilizado el método empleado.

### **Asignación de los especímenes**

Es recomendable emplear técnicas de muestreo para la asignación de los especímenes a la microplaca, evitando introducir algún error sistemático. Complementando de igual manera al protocolo EP12-A2, al hacer mención o recomendación de los posibles errores que se pueden introducir al momento de asignar los especímenes.

# Conclusiones

En esta tesina se expusieron conceptos que reforzaron y complementaron los protocolos vigentes (EP12-A2 , EP15-A2). Con lo expuesto se pretendió dar consistencia a este conocimiento, que, si bien no es ajeno para los técnicos de esta área, es a veces complicado de articular, con lo que se espera que este trabajo, pueda ayudar a explicar de manera sencilla y concreta los conceptos en los que se centra una verificación y los instrumentos para su ejecución.

De los conceptos expuestos, la población de referencia es el concepto sobre el que una verificación concluirá el cumplimiento y la utilidad de la prueba. Articulando conceptos teóricos de la enfermedad, metodológicos y estadísticos para su definición. Considerándose el concepto más importante de una verificación.

Finalmente, para obtener conclusiones aplicables a la población objetivo, es muy importante identificar la validez que aplica a los conceptos que articulan una verificación, por ejemplo, el criterio de precisión diagnóstico, el origen de los especímenes, confección de la prueba, etc.



## Referencias

- [1] Clinical and Laboratory Standards Institute, *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline-Second Edition*, Clinical and Laboratory Standards Institute Document CLSI EP12, Rev. A2, 2008.
- [2] Cochran G. William, *Técnicas de muestreo*, 1991.
- [3] Ignacio Méndez Ramírez, Guillermina Eslava Gómez y Romero Mares Patricia, *Conceptos básicos de muestreo*, 1<sup>ra</sup> ed., ser. Monografías, María Ochoa, Ed. UNAM-IIMAS, Mayo 2004, vol. 12, no. 27.
- [4] Agresti Alan, *Categorical Data Analysis*. New York: Wiley, 2012.
- [5] *Laboratorios clínicos - Requisitos de la calidad y competencia*, IMNC Norma NMX-E-15 189-IMNC-2015, 2015.
- [6] (2019) LINEAMIENTOS PARA LA VIGILANCIA POR LABORATORIO DE LA ENFERMEDAD DE CHAGAS (TRIPANOSOMIASIS AMERICANA). Instituto de Diagnóstico y Referencia Epidemiológicos “DR. MANUEL MARTÍNEZ BÁEZ”. [En línea]. Disponible: [https://www.gob.mx/cms/uploads/attachment/file/483705/Lineamientos\\_Chagas\\_\\_4T.pdf](https://www.gob.mx/cms/uploads/attachment/file/483705/Lineamientos_Chagas__4T.pdf)
- [7] World Health Organization, “Control of Chagas Disease: Second Report of the WHO Expert Committee,” World Health Organization, Geneva, WHO Technical Report Series 905, 2002.
- [8] Clinical and Laboratory Standards Institute, *Verificación del Desempeño de la Precisión y Veracidad por el Usuario; Directriz Aprobada-Segunda Edición*, Clinical and

Laboratory Standards Institute Documento CLSI EP15, Rev. A2, 2005.

- [9] *Vocabulario Internacional de Metrología (VIM)*, JCGM, 2012.
- [10] *Sistemas de gestión de calidad - Fundamentos y vocabulario*, IMNC Norma NMX-CC-9000-IMNC-2015, 2015.
- [11] *METROLOGIA EN QUIMICA-VOCABULARIO*, IMNC Norma NMX-CH-152-IMNC-2005, 2005.
- [12] Griner PF, Mayewski RJ, Mushlin AI and Greenland P, “Selection and interpretation of diagnostic tests and procedures. Principles and applications,” *Annals of internal medicine*, vol. 94, no. 4(Part 2), pp. 557–592, 1981.
- [13] *Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests*, FDA, 2007.
- [14] *STARD guidelines for reporting diagnostic accuracy studies: explanation and elaboration*, BMJ, 2015.
- [15] Julian PT Higgins and Sally Green, Eds., *Cochrane Handbook for Systematic Reviews of Interventions*, no. Version 5.1.0. The Cochrane Collaboration, [Actualizado Marzo 2011]. [En línea]. Disponible: [www.handbook.cochrane.org](http://www.handbook.cochrane.org)
- [16] Ignacio Méndez Ramírez, *Objetividad, probabilidad y población*. México: UNAM-IIMAS, 2013.
- [17] JM Rodríguez, *Respuesta inmune y mecanismos de autoinmunidad*. México, DF: Editores Noriega, 1996.
- [18] Silva AyÇaguer Luis Carlos, *Muestreo para la investigación en ciencias de la salud*. Madrid: Diaz de Santos, 1993.
- [19] Altman Douglas, Machin David, Bryant Trevor, and Gardner Martin, *Statistics with Confidence Confidence Intervals and Statistical Guidelines*. BMJ Books.
- [20] *Para la vigilancia epidemiológica.*, NOM Norma NOM-017-SSA2-2012, 2013.

## REFERENCIAS

---

- [21] Ignacio Méndez Ramírez, *Método Científico, aspectos epistemológicos y metodológicos para el uso de la estadística*. México: UNAM-IIMAS, 2010.
- [22] Shabir Banoo, David Bell, Patrick Bossuyt, Alan Herring, David Mabey, Freddie Poole, Peter G Smith, N. Sriram, Chansuda Wongsrichanalai, Ralf Linke, Rick O'Brien, Mark Perkins, Jane Cunningham, Precious Matsoso, Carl Michael Nathanson, Piero Olliaro, Rosanna W. Pelling and Andy Ramsay, "Evaluation of diagnostic tests for infectious diseases: general principles," *Nature Reviews Microbiology*, vol. 8, no. 12, pp. S16–S28, 2010.
- [23] Ivan Roit, *Essential Immunology*. Oxford, London: Blackwell Scientific Publication, 1994.
- [24] Köhler G and Milstein, "Continuous cultures of fused cell secreting antibody of predefined specificity," *Nature*, vol. 251, no. 12, pp. 495–497, 1975.
- [25] Ransohoff DF. and Feinstein AR., "Problems of spectrum and bias in evaluating the efficacy of diagnostic test," *N. Engl. J Med*, vol. 17, no. 229, pp. 926–930, 1978.
- [26] John R. Crowther, *The ELISA Guidebook*, ser. Methods in Molecular Biology. Spring Street, New York: Humana Press, 2009, no. 516.
- [27] E.Trullols, I.Ruisánchez, F.X.Rius, and J.Huguet, "Validation of qualitative methods of analysis that use control samples," *Trends in Analytical Chemistry*, vol. 24, no. 6, pp. 516–524, 2005.
- [28] —, "Validation of qualitative analytical methods," *Trends in Analytical Chemistry*, vol. 23, no. 2, pp. 137–145, 2004.
- [29] Stewart L. Fossee and Nathan A. Curtis, "Exploring Enzyme-Linked Immunosorbent Assay (ELISA) Data with the SAS<sup>®</sup> Analyst Application," no. 283.
- [30] Velasco-Castrejón O, Valdespino JL, Tapia-Conyer R, Salvatierra B, Guzmán-Bracho C, Magos C, Llausás A, Gutiérrez G, and Sepúlveda J, "SEROEPIDEMIOLOGÍA DE

LA ENFERMEDAD DE CHAGAS EN MÉXICO,” *Salud Pública de México*, vol. 34, no. 2, pp. 186–196, 1991.

- [31] Romero Mares Patricia, *Técnicas de Muestreo I*, IIMAS UNAM, 2014.
- [32] Ignacio Méndez Ramírez, *Método científico, aspectos epistemológicos y metodológicos para el uso de la estadística*, 1<sup>ra</sup> ed., ser. Monografías, María Ochoa, Ed. UNAM-IIMAS, Mayo 2010, vol. 9, no. 22.
- [33] Ignacio Méndez Ramírez, *El paradigma cuantitativo vs. el cualitativo en la investigación*, ser. Monografías. UNAM-IIMAS, Mayo 1993.
- [34] Centers for Disease Control and Prevention. (2012) Section 9: Natural History and Spectrum of Disease. [En línea]. Disponible: <http://www.cdc.gov/ophss/csels/dsepd/ss1978/lesson1/section9.html>
- [35] Yamane Taro, *Elementary Sampling Theory*. Prentice Hall.
- [36] Kenneth J. Rothman and Sander Greenland, *Modern Epidemiology*, 2nd ed. Lippincott Williams & Wilkins, 1998.

# **Anexos**



# Anexo A

## Definiciones Operativas

**Acuerdo** Porcentaje total de casos en el cual dos métodos tienen el mismo resultado.

**Chagas**(*Tripanosomiasis Americana*) La enfermedad de Chagas es autóctona del continente americano. Es causada por el protozooario hemoflagelado *Trypanosoma cruzi*, y se transmite a humanos principalmente por el insecto vector triatóminos hematófagos o por transfusión sanguínea.[6]

**Criterio de precisión diagnóstico** El mejor criterio disponible para establecer la presencia o ausencia de la condición, evento o característica de interés utilizando un solo método o combinación de métodos que incluyen pruebas de laboratorio, pruebas por imágenes, patología e información clínica, incluido el seguimiento.[1]

**ELISA** Acrónimo del inglés *Enzyme-Linked ImmunoSorbent Assay* ó Ensayo por inmunabsorción ligado a enzimas.

**Especificidad** Porcentaje de individuos sin la condición blanco (como fue determinado por el criterio de precisión diagnóstico) los cuales tienen valores negativos en la prueba.[1]

**Espectro de la enfermedad** Es el intervalo de la historia natural de la enfermedad en donde un individuo ya está infectado y puede ser diagnosticado, siendo probable o no la aparición de síntomas asociados a la enfermedad.

**Factor de confusión** Es un factor que puede variar en el interior de las poblaciones que interesa comparar, que provoca cambios importantes en los valores de Y, y que además en, en el proceso de obtención de muestras, se presenta con más intensidad o frecuencia en algunas que en otras.[16]

**Fracción** Parte representativa de una muestra biológica.

**Historia natural de la enfermedad** Se refiere a la progresión de un proceso de la enfermedad en un individuo con el tiempo, en ausencia de tratamiento.[34]

**Inserto** Manual que provee el desarrollador en el cual se especifican la ejecución de la prueba, el estudio de precisión, la sensibilidad y la especificidad.

**Microplaca** Arreglo rectangular que contiene los pocitos en donde se asigna la sustancia a analizar. Esta compuesta de 8 filas y 12 columnas, teniendo un total de 96 pocitos.

**Muestra biológica** Muestra de suero o plasma de algún paciente.

**Panel** Material caracterizado que no contiene una población objetivo en específico.

**Pocentaje de Acuerdo Negativo(PAN)** Pocentaje de casos que concuerdan cuando la referencia del método es negativa.

**Pocentaje de Acuerdo Positivo(PAP)** Pocentaje de casos que concuerdan cuando la referencia del método es positiva.

**Precisión(de medición)** la cercanía al acuerdo entre resultados de pruebas independientes obtenidas bajo condiciones estipuladas. La precisión no se representa típicamente como un valor numérico, pero se expresa cuantitativamente en términos de imprecisión - la desviación estándar (SD) o el coeficiente de variación (CV %) de los resultados en un grupo de mediciones replicadas.[8]

**Precisión intralaboratorio** precisión en un tiempo definido y operadores, dentro de la misma instalación y con el mismo equipo. La calibración y reactivos pueden variar.[8]

**Repetitividad** el acuerdo más cercano entre los resultados de mediciones sucesivas del

---

mismo mesurando llevado a cabo bajo las mismas condiciones de medición.[8]

**Sensibilidad** Porcentaje de individuos con la condición de blanco (como fue determinado por el criterio de precisión diagnóstico) los cuales tienen valores positivos en la prueba.[1]

**Valor de corte** para una prueba cualitativa, el umbral por encima del cual el resultado se reporta como positivo y por debajo del cual el resultado se reporta como negativo.[1]

**Validez Externa** Representatividad de los elementos estudiados de la población objetivo, siendo las características relevantes del proceso estudiado representativas de toda la población.

**Validez Interna** Comparabilidad de grupos que no difieren en características relevantes a la asociación entre la variable X e Y.

**Verificación** Confirmación mediante la aportación de evidencia objetiva de que se han cumplido los requisitos especificados.[5]

**Zona de baja veracidad** Región en donde la interacción de las poblaciones reactivas y no reactivas afectan la exactitud de la prueba, disminuyendo la probabilidad de clasificación del espécimen, con respecto a los parámetros especificados de la prueba. A los resultados obtenidos en esta región se les clasifica como **falso positivo, falso negativo o indeterminado**.



# Anexo B

## Técnicas estadísticas

### **B.1. Muestreo Aleatorio Simple[2]**

El muestreo aleatorio simple es un método de selección de  $n$  unidades en un conjunto de  $N$  de tal modo que cada una de las  ${}_N C_n$  muestras distintas tengan la misma probabilidad de ser elegidas.

#### **B.1.1. Muestreo Aleatorio Simple con reemplazo**

Se mide la unidad seleccionada y se regresa a la población. Si se hace esta operación  $n$  veces, se obtiene una muestra aleatoria simple seleccionada con reemplazo.

### B.1.2. Muestreo Aleatorio Simple sin reemplazo

Se mide la unidad seleccionada y ya no se regresa a la población. Se seleccionan las siguientes unidades con igual probabilidad de las unidades que quedan en la población. Si se hace esta operación  $n$  veces, se obtiene una muestra aleatoria simple seleccionada sin reemplazo.

Para el muestreo aleatorio simple, se tiene una población

$$\text{Población} = \{U_1, U_2, \dots, U_N\}$$

$$\text{muestra} = \{u_1, u_2, \dots, u_N\}$$

Por lo que en una muestra aleatoria cualquier elemento  $U_j = 1, \dots, N$  tiene una probabilidad  $1/N$  de ser seleccionado en cualquiera de las extracciones, siendo la probabilidad  $n/N$  de que el elemento este incluido en la muestra.

Los estimadores de media y varianza son;

$$\widehat{Y} = \sum_{i=1}^n \frac{y_i}{n} = \bar{y}$$

$$V(\bar{y}) = E(\bar{y} - \bar{Y})^2 = \left(1 - \frac{n}{N}\right) \frac{S^2}{n}$$

Para estimar el tamaño de muestra de una proporción, se utiliza la expresión 3.1 (página 37)

## B.2. Muestreo Estratificado[3]

Consiste en dividir a la población en  $L$  subconjuntos o estratos, y cada uno de ellos seleccionar una muestra probabilística, de manera independiente de un estrato a otro.

Las 3 principales razones por las que se estratifica son:

- Estadística: ocurre cuando la población está constituida por unidades heterogéneas y tenemos una idea previa de los grupos de unidades más homogéneas entre si, siendo conveniente formar estratos.
- Por marco: Sucede cuando si para alguna parte de la población se tiene un buen marco, este se utiliza para el muestreo de esa parte y la o las otras partes de la población se muestrean utilizando otros marcos más imprecisos y, posiblemente distintos esquemas(diseños) de muestra.
- Costo: Se presenta cuando levantar la información de las unidades es complicado, para lo que se construye un estrato con un tamaño de muestra pequeño.

Lo más frecuente es que los tres criterios para la formación de estratos coincidan, de modo que los estratos formen unidades homogéneas con un mismo tipo de marco y con costos de localización y captación de información semejantes.

### **B.3. Tablas de contingencia[4]**

Sean  $X$  e  $Y$  dos variables de respuesta categórica,  $X$  con  $I$  categorías y  $Y$  con  $J$  categorías. La clasificación de sujetos en ambas variables tiene  $IJ$  posibles combinaciones. La respuesta  $(X,Y)$  de un sujeto elegido aleatoriamente de alguna población tiene una distribución de probabilidad. Una tabla rectangular que tiene  $I$  filas para las categorías de  $X$  y  $J$  columnas para las categorías de  $Y$  muestra esta distribución. Las celdas de la tabla representan las  $IJ$  posibles salidas. Cuando las celdas contienen cuentas de frecuencia de salidas para una muestra, la tabla es llamada “tabla de contingencia”, un término introducido por Karl Pearson.

Se utiliza  $\pi_{ij}$  para denotar la probabilidad de que  $(X,Y)$  ocurra en una celda en la fila  $i$  y la columna  $j$ . La probabilidad de distribución  $\{\pi_{ij}\}$  es la *distribución conjunta* de  $X$  y de  $Y$ . Las *distribuciones marginales* son la fila y columna totales que resultan de sumar las probabilidades conjuntas. Se denotan por  $\{\pi_{i+}\}$  para la variable fila y  $\{\pi_{+j}\}$  para la variable columna, en donde es subíndice “+” denota la suma sobre el índice; esto es,

$$\pi_{i+} = \sum_j \pi_{ij} \quad y \quad \pi_{+j} = \sum_i \pi_{ij}$$

Estos satisfacen  $\pi_{i+} = \sum_j \pi_{ij} = \sum_i \sum_j \pi_{ij} = 1.0$ . Las distribuciones marginales proveen información de una sola variable.

En la mayoría de las tablas de contingencia, una variable, dígase Y, es una variable de respuesta y la otra (X) es una variable explicativa. Cuando X es fija en vez de aleatoria, la noción de distribución conjunta para X e Y ya no es significativa.

### B.3.1. Sensibilidad y Especificidad[4]

Se define a la sensibilidad como; Probabilidad condicional de que obtenga un resultado positivo, dado que realmente sea positivo.

Se define a la especificidad como; Probabilidad condicional de obtener un resultado negativo, dado que realmente sea negativo.

En una tabla de contingencia se representa de la siguiente manera

		Prueba de referencia	
		Positivo (+)	Negativo (-)
Prueba diagnóstica	Positivo (+)	Verdaderos Positivos (a)	Falso Positivo (b)
	Negativo (-)	Falso Negativo (d)	Verdaderos Negativos (d)

**Tabla B.1:** Tabla de contingencia para evaluar una prueba diagnóstica.

Para obtener la sensibilidad y especificidad se utilizan las siguientes expresiones;

$$\text{Sensibilidad} = \frac{VP}{VP + FN}$$

$$\text{Sensibilidad} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos} + \text{Falsos Negativos}}$$

$$\text{Especificidad} = \frac{VN}{VN + FP}$$

$$\text{Especificidad} = \frac{\text{Verdaderos Negativos}}{\text{Verdaderos Negativos} + \text{Falsos Positivos}}$$

Hay que recordar que se está evaluando una herramienta (prueba) de clasificación, para la cual se utiliza una población de referencia para su comparación y que está diseñada para ser utilizada en diversas poblaciones, existiendo variabilidad en la variable de respuesta (reactivo o no reactivo) al interior de la población (sección 3.2). Con ello en mente se espera que la prueba tenga resultados erróneos, que se expresan como falsos positivos y falsos negativos. La tasa de falsos negativos y falsos positivos se calculan con las siguientes expresiones;

$$\text{Tasa de Falsos Negativos} = \frac{FN}{VP + FN}$$

$$\text{Tasa de Falsos Negativos} = \frac{\text{Falsos Negativos}}{\text{Verdaderos Positivos} + \text{Falsos Negativos}}$$

$$\text{Tasa de Falsos Positivos} = \frac{FP}{VN + FP}$$

$$\text{Tasa de Falsos Positivos} = \frac{\text{Falsos Positivos}}{\text{Verdaderos Negativos} + \text{Falsos Positivos}}$$

Con las expresiones presentadas en esta sección se pueden determinar los parámetros de

desempeño de una prueba, los cuales se contrastarán con los declaradas en el inserto de la prueba que se está verificando. Si estos son al menos iguales o mayores, se concluye que el desempeño de la prueba en la población objetivo es el esperado. Concluyendo que la prueba puede ser utilizada para el diagnóstico de alguna enfermedad. De lo contrario se concluye que no tiene el desempeño esperado en la población objetivo.

# Anexo C

## Desempeño de algoritmos

### **C.1. Desempeño algoritmo para selección de la población de referencia**

En esta sección se mostrará el desempeño con diferentes tamaños de poblaciones de referencia. Esto ayudará a poder determinar si el tamaño de muestra es el adecuado, para satisfacer los criterios de inclusión de la población de referencia. El algoritmo es el siguiente;

```

absMinimaEstratoBajo <- min(DFDCAB$VELISAN)
contadorMinimaContenida <- 0
contadorMinimaNoContenida <- 0
for(i in 1:1000){
  EstratoBajoReactivoSim = MuestraPobRef(n_bajo,DFDCAB)
  sort(EstratoBajoReactivoSim$VELISAN)
  EstratoMedioReactivoSim = MuestraPobRef(n_medio,DFDCAM)
  EstratoAltoReactivoSim = MuestraPobRef(n_alto,DFDCSA)
  DFreactivoStrataSamplesSim = list(EstratoBajoReactivoSim,
  EstratoMedioReactivoSim, EstratoAltoReactivoSim)

  PobRefInfectadaSim = mergeDFStratasamples(DFreactivoStrataSamplesSim)

  if(is.null(EstratoBajoReactivoSim) == FALSE){
    if(length(EstratoBajoReactivoSim)>0){
      absMinimaMuestra = min(EstratoBajoReactivoSim$VELISAN)
      if(absMinimaMuestra == absMinimaEstratoBajo)
      {
        contadorMinimaContenida = contadorMinimaContenida + 1
      }
      else
      {
        contadorMinimaNoContenida = contadorMinimaNoContenida + 1
      }
    }
  }
}
}

```

En donde:

DFDCAB, DFDCAM, DFDCSA: Vectores que contienen las reactividades bajo, media y alta;

VELISAN: Es la reactividad(absorbancia) de los especímenes;

n\_bajo, n\_medio, n\_alto: Es el número de especímenes a seleccionar de los estratos;

contadorMinimaContenida: Número de veces que se seleccionó el espécimen desafiante;

contadorMinimaNoContenida: Número de veces que no se seleccionó el espécimen desafiante;

El algoritmo se ejecutó 1000 veces para los siguientes tamaños de población de referencia.

Tamaño de la población	Tamaño Estratos		
	Bajo	Medio	Alto
15	9	4	2
30	15	10	5
50	35	10	5

**Tabla C.1:** Tamaño de población para selección de población de referencia.

Los resultados de cada población fueron los siguientes:

Tamaño de la población	Conteo espécimen desafiante	
	Contado	No Contado
15	217	783
30	318	682
50	678	322

**Tabla C.2:** Tamaño de población para selección de población de referencia.

Como se observa y se espera, de la tabla C.2, conforme aumenta el tamaño de selección del estrato, aumenta el número de veces que se selecciona el espécimen desafiante. En este caso tiene una reactividad de **0.35**.

Aunque en la selección de la población de referencia, al ejecutar el algoritmo, el espécimen desafiante estaba incluido, la información de la tabla C.2 da indicios que para la selección de los especímenes bajo reactivo, no es muy conveniente emplear este tipo de muestreo (por estratos), para el resto de los estratos (medio y alto) no habría inconveniente en emplearlo.