



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

FACULTAD DE CIENCIAS

ORIGEN DE LOS DATOS CENSURADOS EN  
INMUNOENSAYOS LUMINISCENTES DEL TIPO  
MULTIPLEX Y SU INTEGRACIÓN A ANÁLISIS  
ESTADÍSTICOS

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

**Licenciado en Física**

PRESENTA:

**Ernesto Paas Oliveros**

TUTOR:

Dr. Felipe Vadillo-Ortega

Ciudad de México, 2019





Universidad Nacional  
Autónoma de México



**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.

# Reconocimientos

---

Agradezco a mi familia que me ha soportado por este largo proceso, siempre respetando mis gustos y aspiraciones. A mi madre, Mercedes Oliveros, por siempre estar al pendiente de que avance en mis proyectos y estar orgullosa de mí con cada paso que he dado. A mi padre, Dieter Paas, por apoyarme en la elaboración de las primeras tentativas de la estructura de la tesis, siempre mostrar interés y dar la mayor parte del apoyo económico. A mi hermana, Lya Katarina Paas, por siempre escuchar mis problemas, estar al tanto de mi avance en este proyecto de la manera mas comprensiva y quererme como pocas personas en este mundo. Al Dr. Felipe Vadillo, por permitirme trabajar en estos proyectos de importancia para el país y apreciar lo que puedo contribuir a ellos desde un punto de vista más matemático. A Carmen Alvarez-Buylla, por introducirme a este ambiente y siempre brindarme un apoyo incondicional en todos los aspectos, hasta el emocional. A la Dra. Silvia Ruiz por guiarme en los aspectos matemáticos. A todo la comunidad del laboratorio de vinculación UNAM-INMEGEN, por ser compañeros y amigos cercanos, que me hicieron sentir incluido en un ambiente de investigación, donde se valora la búsqueda por la verdad, la excelencia académica y el estar abierto a las críticas constructivas. Finalmente, a mis amigos Ana Hernández, Andrew Shaw, Daniel Gómez, Daniel Ritter, David Contreras, Lucía Irene, Tatiana Gonzalez, Tora Murakami y Tomás Boils por ser las personas mas cercanas desde mucho antes y durante todo este proceso y que siempre me aseguraron o apoyaron cuando tenía alguna paranoia, me acompañaron en muy buenos momentos y ocuparon el mayor rango de mi espacio emocional.

# Índice general

---

<b>Índice de figuras</b>	<b>vii</b>
<b>1. Introducción</b>	<b>1</b>
<b>2. Antecedentes</b>	<b>3</b>
2.1. Breve historia de los inmunoensayos <i>multiplex</i> . . . . .	3
2.2. Tecnologías actuales para detección de proteínas . . . . .	6
2.3. Inmunoensayo luminescente con perlas magnéticas en MAGPIX <sup>®</sup> . . . . .	10
2.3.1. Captura de las señales . . . . .	11
2.3.2. Curva de dosis-respuesta . . . . .	12
2.3.3. Ejecución del ensayo por software . . . . .	14
2.4. Modelo teórico de la sensibilidad en inmunoensayos . . . . .	16
2.4.1. Ley de acción de masas . . . . .	18
2.4.2. Error del instrumento óptico . . . . .	20
2.4.3. Estimación del LOQ y LOD y la influencia de parámetros en el LOD . . . . .	21
2.5. Citocinas . . . . .	28
<b>3. Objetivo e Hipótesis</b>	<b>31</b>
3.1. Planteamiento del problema . . . . .	31
3.2. Objetivo . . . . .	31
3.3. Objetivos particulares . . . . .	31
3.4. Hipótesis . . . . .	32
<b>4. Caso de estudio</b>	<b>33</b>
4.1. Criterios de inclusión . . . . .	33
4.2. Criterios de exclusión y eliminación . . . . .	33
4.3. Toma de las muestras . . . . .	34
4.4. Medición y captura de datos . . . . .	34

<b>5. Herramientas estadísticas</b>	<b>37</b>
5.1. Panorama . . . . .	37
5.2. Comparación de distribuciones con la prueba de Mann-Whitney . . . . .	38
5.2.1. El estadístico de Mann-Whitney . . . . .	39
5.2.2. Ajustes para empates . . . . .	42
5.2.3. Comparación con otros métodos . . . . .	43
5.2.4. Simulaciones . . . . .	43
5.3. Imputación con distribuciones truncadas . . . . .	45
5.4. Imputación con random forest . . . . .	47
<b>6. Resultados y Discusión</b>	<b>51</b>
6.1. Análisis exploratorio . . . . .	51
6.2. Imputación con Random Forest . . . . .	55
6.3. Imputación con distribuciones truncadas . . . . .	57
6.4. Comparación de distribuciones con Mann-Whitney . . . . .	61
6.5. Comparación . . . . .	63
<b>7. Conclusiones</b>	<b>69</b>
<b>A. Código/Manuales/Publicaciones</b>	<b>71</b>
A.1. Apéndice . . . . .	71
<b>Bibliografía</b>	<b>73</b>

# Índice de figuras

---

2.1. SP-PLA . . . . .	9
2.2. Inmunoensayo tipo <i>sandwich</i> . . . . .	11
2.3. Compartimiento de la cámara digital en el <i>MAGPIX</i> <sup>®</sup> . . . . .	12
2.4. Visualización de parámetros en la curva log4p . . . . .	13
2.5. Efecto de parámetros en curva log5p . . . . .	15
2.6. LOB, LOB, LOQ . . . . .	17
2.7. Límite de detección en molares, según un modelo en el que solo se toma en cuenta la ley de acción de masas, con la Ec.2.11. Cuando no se varían los valores para las parámetros, estos valen $[Ab1]_t = 210^{-12}M$ , $[Ab2]_t = 10^{-8}M$ , $K1_{eq} = K2_{eq} = 10^{12}\frac{1}{M}$ y $[Ab2]_b = 810^{-16}M$ . . . . .	24
2.8. Límite de detección en molares, según un modelo en el que se toma en cuenta la ley de acción de masas y el ligado no específico NSB con un valor de 0.1%. Los cortes abruptos en las gráficas <i>a</i> , <i>c</i> , <i>d</i> se deben a la singularidad que surge al dividir entre 0 y que se explican arriba. Cuando no se varían los valores para las parámetros, estos valen $[Ab1]_t = 210^{-12}M$ , $[Ab2]_t = 10^{-13}M$ , $K1_{eq} = K2_{eq} = 10^{12}\frac{1}{M}$ y $[Ab2]_b = 810^{-16}M$ . . . . .	25
2.9. Aglomerados . . . . .	27
2.10. Perfil de Precisión . . . . .	28
2.11. Diferenciación célula T . . . . .	29
2.12. Tansducción de señal . . . . .	30
4.1. Imágen del equipo <i>MAGPIX</i> <sup>®</sup> . . . . .	35
5.1. Aproximaciones a la normal . . . . .	41
5.2. Distribuciones normales truncadas . . . . .	45
5.3. Mapa de particiones en Random Forest . . . . .	49
6.1. Histogramas sin LOD . . . . .	52
6.2. Histogramas sin LOD . . . . .	54
6.3. Histogramas con LOD . . . . .	54
6.4. No. de variables vs. RMSE en RF . . . . .	56
6.5. R2 para predicciones en RF . . . . .	56
6.6. Comparación histogramas con y sin RF imputado . . . . .	57

## ÍNDICE DE FIGURAS

---

6.7. Histogramas con LOD y prevalencia . . . . .	58
6.8. Histogramas con optimización . . . . .	59
6.9. Histogramas con imputación . . . . .	60
6.10. Histograma IL4 con NA imputados . . . . .	61
6.11. Histogramas con imputación y RF . . . . .	62
6.12. Comparación de distribuciones sin LOD . . . . .	64
6.13. Comparación de distribuciones con LOD . . . . .	65
6.14. Comparación de distribuciones con LOD en Log10 . . . . .	66
6.15. Comparación de distribuciones sin LOD en Log10 . . . . .	67

# Introducción

---

Este trabajo surge debido a un problema de pérdida de información en el laboratorio de vinculación científica entre la Facultad de Medicina de la UNAM y el INMEGEN. A lo largo de más de 4 años se midió la concentración de aprox. 20 citocinas (proteínas involucradas en la señalización inmune) en las paredes del útero de aprox. 1000 mujeres embarazadas con y sin parto pretérmino, para intentar asociar esas concentraciones con la condición mencionada. Debido a la alta variación en los niveles de estas citocinas y a que se medían todas en un mismo ensayo, hubo muchas citocinas para las cuales entre 5 y el 30% de las mujeres tenían concentraciones mas bajas de lo que la sensibilidad del instrumento de medición permitía analizar, es decir su límite de detección. Tomar estos datos como perdidos sería un error, pues nos están desprovistos de información.

Es así como se arriba a un problema que abunda en el bioanálisis, donde se analizan sustancias de tejidos orgánicos a los que en general se les llama analitos. Más generalmente, se puede decir que los límites de detección conllevan la existencia de mediciones que no dan una información cuantitativa, sino cualitativa, de la concentración de un analito. A este tipo de datos se les llama datos censurados o datos debajo del límite de detección (**LOD** *limit of detection*). En muchos estudios, gran parte de los datos obtenidos son censurados debido a las limitaciones de la instrumentación y la variabilidad de las moléculas analizadas, lo que hace que se ignore información que está contenida en los datos y que podría influir sobre los resultados o diagnósticos que se hacen a partir de ellos.

En esta tesis se ha buscado explicar los factores que dan origen al límite de detección, con ayuda de herramientas que el autor aprendió a usar en su carrera de físico, como el modelaje matemático. Esto para poder informar a una estrategia de rescate de los datos censurados. Por otro lado, se hace una búsqueda y valoración de herramientas estadísticas para imputar los datos censurados o considerarlos. A pesar de que la estrategia se enfoca en métodos estadísticos para su cometido, el entendimiento de las causas de estos límites, con ayuda de modelos matemáticos, ha guiado la selección de las herramientas estadísticas.

Acorde a esta idea, en los **Antecedentes** se hace un recuento de las tecnologías actuales para cuantificar proteínas y su historia. Esto con el objetivo de colocar el tipo



## 1. INTRODUCCIÓN

---

de ensayo que se estudia en contexto y recalcar sus limitaciones y ventajas. Luego, se explica el funcionamiento del inmunoensayo *multiplex*, que se usó en el caso de estudio y, una vez que el lector esté familiarizado con los procesos que suceden en estos ensayos, se ofrece una explicación teórica de dichos procesos, orientada a explicar y estimar la sensibilidad del inmunoensayo. Finalmente, se habla un poco del tipo de proteína presente en el caso de estudio y su influencia en la naturaleza de los datos que arroja.

Con el contexto ya clarificado, se procede a escribir el capítulo de **Objetivos e Hipótesis**. Luego, en el capítulo **Caso de estudio**, se exponen las condiciones bajo las cuales se hizo el experimento, los métodos y los instrumentos usados. A continuación, en el capítulo **Herramientas estadísticas**, se presenta un panorama de las herramientas que se consideraron y se explica la teoría detrás de las que se implementaron. En **Resultados y Discusión** se presentan los resultados de aplicar dichas herramientas y se discute su efectividad. Finalmente, en **Conclusiones** se resume lo que se hizo y valora la estrategia de recuperación de los datos censurados.

# Antecedentes

---

## Breve historia de los inmunoensayos *multiplex*

La capacidad de analizar ha acompañado al ser humano desde que tiene uso de la razón. Se podría decir que los primeros actos de representación, como las pinturas rupestres, son, en parte, un análisis de la experiencia humana. En el análisis se separa una entidad en sus componentes, como lo indica su raíz griega. Esta está compuesta por *ανα* (ana) que significa *sobre algo* o *enteramente*, y *λυσις* (lysis) que es disolución o separación. Es decir, lo que trata sobre (engloba a) la separación de algo.

A lo largo de la historia se han tratado de delimitar, cada vez con mayor precisión, los componentes o las partes que informan sobre el estado o funcionamiento de un organismo. Información que concierne tanto a la biología como a la medicina. En la primera disciplina se trata de entender el funcionamiento de todo organismo vivo, tanto en su operación normal como patológica, que es lo que estudia el subcampo de la medicina. En la era moderna, el análisis ha llegado a tal punto que ya se comprende cuales son los componentes fundamentales que forman a todo ser vivo. El estudio de estos bloques fundamentales y sus leyes interseca con dominios de la ciencia más fundamentales, como la química y la física.

De esta manera, la búsqueda por entender las causas de una patología o la fisiología de un organismo, ha llevado a que se desarrollen técnicas analíticas cada vez más exactas. El tipo de técnica que concierne a esta tesis cae dentro del campo del **bioanálisis**, una disciplina en la que se miden componentes químicos y macromoléculas en un organismo vivo.

El crecimiento en esta rama de la ciencia ha sido impulsado por la necesidad de cuantificar los efectos de drogas en el cuerpo humano. Se podría decir que una de las primeras técnicas usadas por el bioanálisis fue la del ensayo **colorimétrico**. Técnica donde se usa un compuesto químico, que al reaccionar con el analito crea un derivativo con color y, según la intensidad de ese color, se puede estimar la concentración del analito.

Las características más importantes para medir la eficacia de estos métodos son la

**especificidad**, la **precisión**, la **sensibilidad** y el **rango dinámico**. La **especificidad** se refiere a la seguridad que se tiene de que el analito que se está detectando es efectivamente el que se quiere detectar. La **colorimetría** tenía muy baja especificidad, pues el reactivo podía hacer reacción con muchas sustancias además del analito. En el caso particular de las drogas, no se podía diferenciar entre la droga misma y sus metabolitos (Hill, 2009).

La **precisión** se refiere a la fineza de la escala con la que se mide la concentración, es decir el orden de magnitud que hay entre una medición y la siguiente más grande que sea significativamente diferente. Tal fineza es dictada por la variación o porcentaje de error de los replicados. Esta propiedad está ligada a la **sensibilidad** que se define como la mínima cantidad detectable que se diferencia de cero. Por ejemplo, la medición con intensidad de color pudo perder algo de subjetividad con un fotómetro, pero variaba mucho según las condiciones del experimento, como la difusión, y por eso no es muy sensible.

La primera técnica que logró una mejora de estos factores fue la **cromatografía**, desarrollada alrededor de 1900 por Mikhail Tsvet. La cromatografía se aprovecha de que cada componente de una mezcla se mueve a diferentes velocidades en un medio absorbente o adsorbente, para poder separarlos y así analizarlos según su color (lo que le da el nombre) u otras propiedades. No fue hasta el advenimiento de la cromatografía de capa delgada que se pudo usar esta técnica para todo tipo de fluidos biológicos. Gran parte del avance de esta técnica se debe al trabajo conceptual del químico Archer John Proter Martin y el bioquímico Richard Laurence Millington Synge, quienes ganaron el premio nobel de química en 1952, por la invención de la cromatografía de partición. Ellos pusieron las bases que propiciaron un gran avance en las técnicas cromatográficas.

En los primeros ensayos de **cromatografía**, la característica que más se optimizó fue la especificidad, pero es importante recalcar que la cromatografía por sí misma solo es una técnica de separación de componentes, para la cuantización de las cantidades de estos elementos, hizo falta integrar técnicas analíticas, como la espectrometría de masas, para verdaderamente tener un análisis cuantitativo. La cromatografía de gas permitió que se separaran moléculas de menor tamaño y luego se inventó la cromatografía líquida de alto rendimiento o *high performance liquid chromatography* (HPLC), que, junto con un espectrómetro de masas como detector (HPLC-MS), pudo clasificar y cuantizar con mayor precisión y rango. La espectrometría fue inventada por Eugen Goldstein en 1886 y perfeccionada por Arthur Jeffrey Dempster y F.W. Aston en 1918 y 1919 respectivamente. Un impedimento de esta técnica es que requiere que se ionicen las moléculas a analizar, lo que dificulta la detección de macromoléculas.

Otra gama de métodos bioanalíticos de gran precisión, son los basados en **electroforesis**, pero estos solo se pueden usar para sustratos específicos que sean susceptibles a campos eléctricos. Además de que solo se pueden aplicar a muestras muy pequeñas.

Las limitaciones de los métodos arriba descritos imposibilitaban la detección de macromoléculas y estas ya estaban siendo consideradas como agentes terapéuticos alrededor de 1950. Para ampliar el rango de moléculas detectables, hacía falta que se desarrollara la técnica de los **inmunoensayos**. Los **inmunoensayos** son una prueba

química en la que se mide la concentración de una molécula o macromolécula (analito) con ayuda de anticuerpos o antígenos, los cuales se conectan con una parte específica del analito, llamada epítipo y que están marcados para emitir una señal. En general, siempre hay uno o más anticuerpos (o antígenos) detectores y una molécula marcadora, con la cual se cuantifica al analito. Como primera descubridora de este tipo de técnica, la física Rosalinda Yalow ganó el premio nobel de medicina en 1977. Ella describió, junto con el médico Solomon Berson en los 1950's, a los **radioinmunoensayos** (RIA) (Yalow, 1960). Los usaron para detectar insulina endógena en plasma, pero el premio se les dio por el desarrollo de RIA para hormonas de péptidos.

Los radioinmunoensayos permitieron la detección de un nuevo campo de analitos, pero su ejecución era costosa y complicada, pues requería de anticuerpos con radioisótopos como marcadores. Un gran prosélito del entendimiento de las bases de los inmunoensayos fue el biofísico Roger Ekins, quien también desarrolló un RIA al tiempo de Yalow, pero para la detección de tiroxina en plasma humano (Ekins, 1960). Además, fue gracias a el uso de enzimas como marcadores que los inmunoensayos ganaron gran popularidad. A diferencia del RIA, en este ensayo se mide la intensidad del color que surge de la reacción entre las enzimas y su sustrato, el cual es añadido después de hacer un lavado para remover cualquier enzima que no esté atada a un anticuerpo ligado al analito. El **ELISA** (*Enzyme Linked Immunosorbent Assay*) (Lequin, 2005) y el EIA (*Enzyme Immunoassay*) fueron inventados independientemente, alrededor de 1960. No conllevaban ningún riesgo, costaban menos y eran más fáciles de ejecutar que un RIA. Este hito marcó el paso a la eficientización de los inmunoensayos. Su ejecución dejó de ser un arte que podía ser ejecutado por pocos y pasó a estandarizarse, trayendo consigo una industria de productos que ofrecen soluciones prefabricadas.

Otro peldaño en el avance hacia la automatización de estas pruebas, fue el de la gestación de los inmunoensayos no competitivos. Anteriormente, en los competitivos, se usaba una baja cantidad fija de anticuerpo mezclada con un volumen conocido de analito etiquetado y con la muestra que contiene el analito a medir. La concentración del analito etiquetado y ligado a los anticuerpos es, debido a la ley acción de masas, inversamente proporcional a la concentración de analito desconocida, por lo que se podía medir la concentración desconocida a partir la señal que emiten los analitos de volumen conocido. Fue gracias al avance en el diseño de anticuerpos monoclonales que empezaron a desarrollarse los ensayos no competitivos. A diferencia de los ensayos competitivos, donde no se requiere un alta cantidad de anticuerpos, los no competitivos utilizan un exceso de anticuerpos para marcar directamente al analito. De esta manera se llega más rápido a la saturación y se puede hacer una medida en menos tiempo. La manera más común de lograr esto, es teniendo un anticuerpo de captura que reconoce al analito y otro de reporte que tiene acoplado un elemento que emite una señal. Estos inmunoensayos son de tipo *sandwich* y además, permiten tener más especificidad, pues se requiere el reconocimiento del analito en dos partes diferentes para que se emita una señal.

Ekins hizo grandes avances en el modelaje matemático y la estadística de los inmunoensayos, con contribuciones como el ensayo de ligado de analito ambiente (Ekins,

1989). Demostró que la disminución de la concentración de anticuerpo de captura llevaba a una mayor sensibilidad y postuló un límite teórico para el límite de detección (LOD) de los inmunoensayos, en el orden de los fg/ml. Basado en eso, desarrolló los principios para y propuso el primer inmunoensayo *multiplex*, es decir un ensayo en el que se pueden medir muchos analitos en una misma muestra. Lo que caracterizaba a este ensayo, llamado de microarreglo, es que tiene varios elementos de captura (como anticuerpos) localizados en pequeñas regiones organizadas en un arreglo sobre una superficie sólida, llamados *microspots*.

Otra técnica coetánea y muy popular, fue ideada por los expertos en citometría de flujo, Horan y Wheeless, cuando hicieron un recuento de las posibles aplicaciones de la citometría de flujo (Horan and Wheeless, 1977). Concibieron una manera de usar las microperlas que ya se usaban en citometría de flujo, como la superficie de captura de la proteína. Al principio, se usaron microesferas de distintos tamaños y luego de diferentes colores para poder detectar múltiples analitos. Para emitir una señal, se añaden anticuerpos receptores, acoplados a un compuesto que emita una señal. Seguido a esto, las perlas pueden ser interrogadas, una a una, en tubos muy delgados, usando procesos de citometría de flujo.

Estos avances establecieron una base sólida para los inmunoensayos *multiplex*, desde la cual se han desarrollado muchas mejoras y adiciones que optimizan diferentes aspectos del ensayo.

## Tecnologías actuales para detección de proteínas

Con ayuda de la nanotecnología se pueden imprimir microarreglos que tienen una densidad de 106 puntos por  $mm^2$  (Wild, 2013). Algunas plataformas que implementan esta tecnología hoy en día, son la *Meso Scale Discovery Platform* y el *Quansys system*. Los equipos de *Meso Scale* no presumen analizar la mayor cantidad de analitos por placa (hasta 4), pero son muy rápidos en la lectura (hasta 70 seg por placa) y tienen mayor rango dinámico (6 logs) y sensibilidad (hasta 0.05 pg/ml) que los inmunoensayos de microperlas, gracias a el uso de electroquimioluminiscencia para detectar las señales.

Los equipos de *Luminex*<sup>®</sup>, son inmunoensayos de perlas magnéticas, que hacen uso de citometría de flujo o fotografía CCD, para detectar las señales. El equipo más eficiente en esta gamma es el *Flexmap 3D*<sup>®</sup>, el cual tiene la capacidad de diferenciar hasta 500 tipos diferentes de señal en una sola muestra. Si se usa a su máxima capacidad, el *Flexmap 3D*<sup>®</sup> puede hacer hasta 14400 mediciones por hora, comparada con las 96 que se hacen en una placa de ELISA convencional. Es gracias a esta alta eficiencia que se disminuye el costo de hacer un análisis de muchas proteínas y esto, a su vez, es la causa de que las plataformas de *Luminex*<sup>®</sup> sean populares. En el caso de estudio de esta tesis se uso la plataforma *MAGPIX*<sup>®</sup>, la cual cuesta aproximadamente un cuarto del precio del *Flexmap 3D*<sup>®</sup>, tiene un rango promedio de 3.5 logs y productividad suficiente como para ser más barato que ELISA, si se quieren medir más de 5 analitos a la vez. Además, el uso de microesferas magnéticas permite que se automaticen los pasos de

lavado y que aumente la reproducibilidad.

Una desventaja del *multiplex*, es que puede haber reactividad cruzada o no específica de los anticuerpos, lo que reduce la especificidad y aumenta el ruido de fondo, disminuyendo así también la sensibilidad. Teóricamente, los inmunoensayos de microarreglo presentan estos fenómenos en menor medida, ya que los anticuerpos están inmovilizados en zonas localizadas.

Los equipos que hacen uso de la microfluídica para automatizar los inmunoensayos, como el *Gyrolab* o *Ella*, requieren menos preparación, tienen disminuidas distancias de difusión y menores volúmenes de reacción. En el caso de *Ella* se compran placas en las que viene el reactivo integrado, mientras que en *Gyrolab*, se colocan los reactivos aparte y el equipo los mezcla. Para cada ensayo se compra un cartucho en forma de CD para *Gyrolab* o de placa para *Ella*, que cuenta con microcanales. La muestra se reparte en esos canales, se captura una cantidad exacta y se transfiere a una cámara donde hay perlas que capturan los analitos. Finalmente, se excitan las perlas y el fluoróforo de reporte con LED's para cuantificar la concentración del analito. Como cada canal es específico para un analito, no hay reactividad cruzada y la estructura microfluídica hace que disminuya el ruido de fondo lo que aumenta la replicabilidad y la sensibilidad. Desafortunadamente, debido a que solo se usa un canal por analito, la capacidad de medir muchos analitos en paralelo es baja. Otra ventaja de estos ensayos es que usan muy poco volumen de muestra, de 20 a 1000 nL para *Gyrolab* y de 2.5 a 25  $\mu$ L, para *Ella*.

A pesar de que se puede eliminar la reactividad cruzada en inmunoensayos *multiplex* con microfluídica, se hace más difícil aumentar el número de analitos simultáneos que se pueden detectar. Sin embargo, hay métodos que tienen la capacidad de mezclar todos los analitos que puedan detectar por separado en una sola corrida, sin que haya reactividad cruzada. Estos son los ensayos basados en espectrometría de masas y no hacen uso de etiquetado, como lo mayoría de los inmunoensayos. Para poder cuantificar proteínas, se les parte con enzimas en lugares específicos para dejar péptidos proteoespecíficos que identifican únicamente a la proteína. Como cada péptido tiene diferente masa y se ioniza diferente, se pueden detectar cuantos péptidos se quieran mientras se tenga un estándar de referencia de ellos. Una desventaja es que los equipos más sensibles hacen uso de cromatografía líquida en cuyas columnas solo se pueden depositar volúmenes de orden de nL. Por lo tanto, se tiene menos masa del analito y sube el límite de detección. Se han hecho esfuerzos para bajarlo, enriqueciendo la concentración de los péptidos a analizar con ayuda de microperlas magnéticas cubiertas de anticuerpos que se enlazan a los péptidos de interés, pero no se logran alcanzar niveles de los inmunoensayos de perlas convencionales (Anderson et al., 2009).

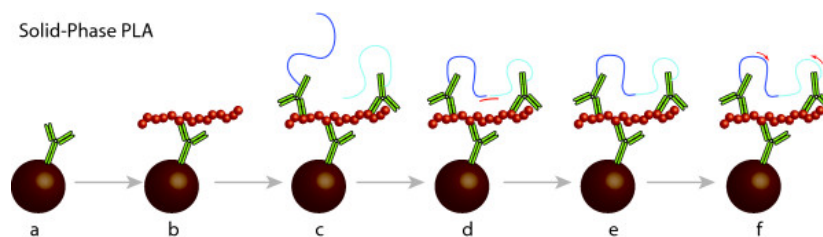
Otras técnicas libres de etiqueta, son las basadas en biosensores. Las más conocidas son las de interferometría de biocapa, o *bio-layer interferometry* (BLI) y resonancia de plasmón de superficie, o *surface plasmon resonance* (SPR). En BLI se emite luz blanca a la superficie del biosensor, que tiene una capa de anticuerpos acoplados con la que reaccionan las molécula a analizar. El reflejo de la luz en la capa de anticuerpos y la de analitos crea un patrón de interferencia que se desfasa, a medida que cambia

## 2. ANTECEDENTES

---

la densidad de la capa de analitos. Estos cambios se pueden medir en tiempo real y la medición en equilibrio informa sobre la concentración del analito. Los biosensores se exponen a la muestra, siendo sumergidos en los 96 o 384 pozos de una placa. El *Octet RED96e*, es uno de los instrumentos que implementa este proceso. En un sistema de SPR, se usa microfluídica para hacer pasar la muestra por canales, donde se encuentran los biosensores. Estos biosensores consisten de una capa de metal, con los anticuerpos de captura inmovilizados sobre ella y, del otro lado, un prisma. A través del prisma se dispara un láser y, a cierto ángulo, hay resonancia con los plasmones de la superficie metálica. Esta resonancia, hace que disminuya la intensidad del rayo reflejado y, como el ángulo de resonancia cambia con la densidad de moléculas acopladas a los anticuerpos, se puede obtener una medida de la cantidad de analito acoplado. Existe una gran variedad de marcas que venden equipos de SPR, una de las más conocidas es *Biacore*. Gracias a que las mediciones de estos biosensores se hacen en tiempo real, se pueden obtener parámetros cinéticos de la afinidad y disociación de los anticuerpos, lo cual es importante para la búsqueda de anticuerpos con alta especificidad. Sin embargo, la sensibilidad de estos ensayos no es muy alta, debido a el ligado no específico en la superficie del biosensor. Los sistemas con SPR son los que han alcanzado menores LOD, con ayuda de nanopartículas para bloquear la interacción no específica, han llegado a detectar concentraciones en el orden de 10 pg/ml (Martinez-Perdiguero et al., 2014) y, en otro estudio donde se amplificó la señal con nanopartículas de oro con ayuda de detección tipo *sandwich*, llegaron a medir concentraciones de hasta 10 fg/mL (Hendriks et al., 2018). Sin embargo, esto requiere mucho tiempo de preparación y la capacidad de *multiplex* es baja. Además, el costo del equipo y los biosensores es alto.

Una manera de aumentar la sensibilidad de los inmunoensayos, es amplificando la señal emitida, como el ELISA que fue el primero en hacerlo. Uno de tales ensayos, cuyo nombre en inglés es *Solid Phase Proximity Ligand Assay*, hace uso de la técnica de **PCR** (*Polymerase Chain Reaction*) en tiempo real, para amplificar la señal. Con ayuda de una fase sólida de anticuerpos que inmovilizan al analito, se pueden hacer lavados para eliminar contaminantes. Una vez lavada la muestra, se añade un par de anticuerpos que también están diseñados para ligarse con el analito, además de que están conectados a unas cadenas complementarias de oligonucleótidos (Véase Fig. 2.1), que al estar cerca, dan la base para empezar una reacción en cadena de polimerasa. La detección por medio de 3 anticuerpos y los lavados, ayudan a que el ensayo tenga mayor especificidad. La amplificación también le brinda una gran sensibilidad a este método, además de que lo hace resistente a efectos de matriz (Darmanis et al., 2011). Se puede aumentar la velocidad de amplificación y la especificidad del ensayo, haciendo uso de procesos como el **RCA** (*Rolling Circle Amplification*) o el **PEA** (*Proximity Extension Assays*) en lugar de la PCR. Estos ensayos de proximidad, tienen la ventaja de que se pueden usar con cualquier equipo de PCR de tiempo real y han alcanzado muy buenos niveles de sensibilidad y especificidad. Por ejemplo, los ensayos *ProQuantum* de *ThermoFisher* tienen sensibilidad de hasta 0.001pg/ml, un rango dinámico de 5 logs y solo necesitan de 2-5  $\mu$ L por ensayo. Sin embargo, no han sido adaptados para *multiplex*.

**Figura 2.1:**

En el SP-PLA, se hace uso de una fase sólida de anticuerpos que inmovilizan al analito, lo que permite que se pueden hacer lavados para eliminar contaminantes. Una vez lavada la muestra, se añade un par de anticuerpos que también están diseñados para ligarse con el analito, además de que están conectados a unas cadenas complementarias de oligonucleótidos, que al estar cerca dan la base para empezar reacción en cadena de polimerasa. Fuente:

(Kamali-Moghaddam et al., 2010)

Como se puede notar, cada estrategia para mejorar los inmunoensayos trae consigo algunas ventajas y otras desventajas. Aun así, en casi todas las ramas se han hecho avances y correcciones que emparejan el grado de sensibilidad. El equipo automatizado y con *multiplex* más sensible que se ha encontrado en esta investigación es el *Simoa* (*Single molecule array*), un equipo para ELISA's tipo sandwich que aprovecha la ya existente tecnología de microesferas magnéticas. Para cada ensayo se usan discos prefabricados con nanotecnología, que tienen 24 microarreglos de aprox. 216,000 pozos con volumen de  $40\text{ fL}$ , en donde cabe exactamente una microesfera magnética. El proceso de catalización de enzima y sustrato sucede en los pozos, lo que permite que se digitalice la medición de la señal por medio de un conteo de los pozos con enzimas activas. Solo los pozos que presentan un crecimiento de intensidad de fluorescencia se cuentan como activos. Según los diseñadores del método (Rissin et al., 2011), la estadística de Poisson estima que si la fracción de pozos activos es menor a 0.7, el número de enzimas conectadas por microesfera es aproximadamente 1. Por lo que se puede calcular la concentración a partir de un conteo de molécula única, cuando hay bajas concentraciones y, si la concentración es alta, se usa la intensidad de fluorescencia. Esta combinación permite una sensibilidad subfemtomolar y un rango dinámico de 4 logs. El sistema permite *multiplex* de hasta 10 analitos, y mide 66 muestras por hora. Además, el costo por ensayo es mucho más bajo que el de otras tecnologías de ultrasensibilidad, pues hace uso de los reagentes convencionales de un ELISA.

A continuación (Tabla 2.1) se presenta una tabla donde se pueden comparar más rápidamente las tecnologías actuales. En algunos campos hace falta información, sobre todo en los precios, pues es necesario hacer un cotización oficial para obtener los datos.

En perspectiva, el equipo usado en este estudio (*MAGPIX*<sup>®</sup>), tiene un rango dinámico promedio de 3.5 logs, una sensibilidad promedio de 0.1 pg/ml y un bajo costo por resultado cuando se usa el *multiplex*. El consumo de muestra esta en el rango medio, con un mínimo de  $20\ \mu\text{L}$ . Es moderadamente rápido, pues analiza 96 muestras en aprox 1h y el costo del equipo esta en el rango medio bajo, siendo que vale aprox \$35,000 USD.



## 2. ANTECEDENTES

	Magpix	Luminex 200	Flexmap 3D	Gyrolab xPlore
Sensibilidad (LOD) en pg/mL	0.1	0.1	0.1	0.1
Rango dinámico en logs	3.5	3.5	4.5	4
Tests/hora	96*50	128*100	288*500	112*3
Tiempo de preparación	2h	2h	2h	
Tamaño de muestra mínimo	20-200 µL	20-200 µL	10-200 µL	20-1000 nL
Analitos por pozo/canal	50	100	500	1
Equipo nuevo en USD	\$30,000	\$60,000	\$140,000	
	ProQuantum	ELISA	Octet Red96e	MESO Quick Ple
Sensibilidad (LOD) en pg/mL	0.05	10	10	0.05
Rango dinámico en logs	5	2	4.5	5
Tests/hora	1	96	192	96 en 70 seg
Tiempo de preparación	1h	4h	20 min	2h
Tamaño de muestra mínimo	2-5 µL	50 -100 µL	180-220 µL	150 uL
Analitos por pozo/canal	1	1	1	10
Equipo nuevo en USD	\$12,000	\$8,000	\$200,000	
	Singulex Erenna	Simoa-HD1	Siscapa	
Sensibilidad (LOD) en pg/mL	0.001	0.0001	5	
Rango dinámico en logs	4	4	4	
Tests/hora	384	96*4	24	
Tiempo de preparación	1h	20 min		
Tamaño de muestra mínimo	5-100 µL	1-100 µL	10-100 µL	
Analitos por pozo/canal	1	4	150 sin cruce	
Equipo nuevo en USD				

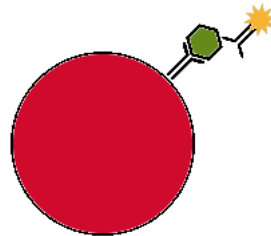
**Tabla 2.1:** Comparación de diferentes características de diferentes equipos en el mercado que usan las tecnologías expuestas en este capítulo. En algunos campos hace falta información, sobre todo en los precios, pues es necesario hacer un cotización oficial para obtener los datos.

## Inmunoensayo luminescente con perlas magnéticas en MAGPIX®

A continuación, se describe más a fondo el funcionamiento del equipo que se usó para el caso de estudio.

Como ya se mencionó en el capítulo anterior, el *MAGPIX*® es un equipo para correr inmunoensayos *multiplex* heterogéneos y hace uso de microesferas. Para poder diferenciar una amplia gama de analitos (en este caso hasta 50), se tiñen interiormente las microesferas magnéticas con dos tintes en diferentes concentraciones (en este caso se usaron *MAGPLEX*<sup>TM</sup>-C). El proceso de tinte se hace con fluoróforos y usa tecnología patentada *xMAP*®. Si se cubre cada perla con un reactivo ( anticuerpos o antígenos), se pueden detectar tantos reactivos en cada muestra, como diferentes microesferas se tengan. La superficie de las esferas tiene un diámetro de 6.5 µm y esta cubierta de aproximadamente 100 millones de grupos de carboxilo, a los cuales se conectan los anticuerpos detectores, en un proceso de activación y mezcla. Hay tres métodos comunes para hacer la reacción con el analito; el competitivo, el indirecto y el tipo *sandwich*. El caso que se estudiará es el de *sandwich*, en el que hay un anticuerpo de captura que reacciona con un epítipo del analito y otro detector que reacciona con otro epítipo y que da la intensidad de la señal (Véase Fig. 2.2 ). La necesidad de una doble detección para poder medir la señal, hace que suba la especificidad.

Para analizar una muestra, se le mezcla con una solución preparada de microesferas.



Capture Sandwich

**Figura 2.2:** En un inmunoensayo tipo *sandwich*, el analito (en verde) se liga a uno de los anticuerpos fijos. Luego, se añaden anticuerpos biotinilados que se conectan con el analito y, finalmente, se añade un fluoróforo receptor como ficoeritrina, para poder registrar la cantidad de analito ligado.

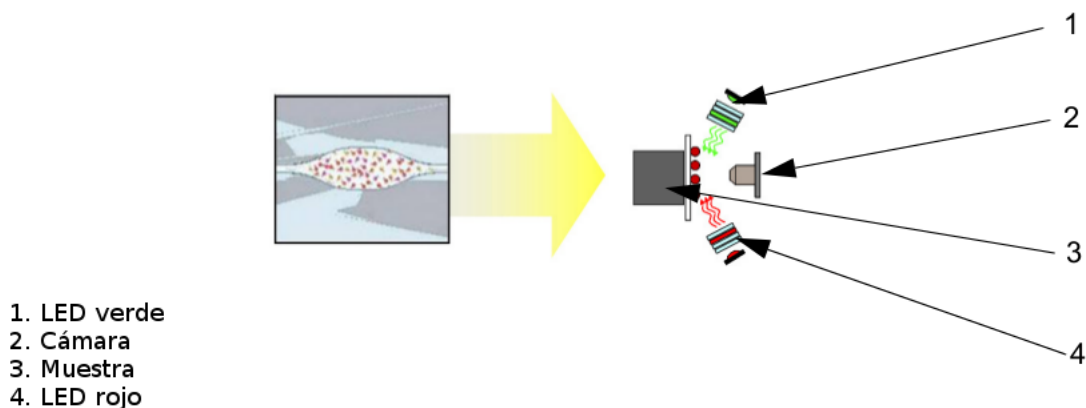
Fuente: Xmap Cookbook

Después de que las esferas hayan tenido tiempo para conectarse con la muestra, se añade una solución con anticuerpos detectores y biotinilados. La biotina anexada al anticuerpo facilita el acoplamiento de un fluoróforo de reporte, en un periodo posterior de incubación. En la mayoría de los ensayos *Luminex*<sup>®</sup> se usa streptavidina acoplada al fluorocromo ficoeritrina, más conocido como Streptavidin PE.

### Captura de las señales

A continuación sigue el proceso fluídico, en el que se succionan las microesferas pozo por pozo al compartimiento de la cámara (Veáse Fig. 2.3). Ahí, las microesferas magnéticas son atraídas a una monocapa, y son irradiadas por un LED verde y uno rojo. Mientras están siendo irradiadas, se les toma una foto con CCD. El LED rojo excita las microesferas y en el análisis de imagen se usan filtros CL1 Y CL2 para caracterizar cada esfera y descartar los aglomerados. El verde excita la molécula receptora y mide la intensidad fluorescente, la cual está relacionada con cuantas proteínas se ligaron. Finalmente, las perlas son expulsadas y se procede a succionar el siguiente volumen.

A comparación del método de ELISA, donde los anticuerpos principales están pegados a una superficie, el pequeño tamaño de las microesferas y el hecho de que estén flotando en una solución aumenta la tasa de reacción. La tasa de reacción alta propicia a que la dinámica sea parecida a la de una solución (Matson, 2009), lo que implica que la distribución estadística de las reacciones tiende a ser gaussiana o uniforme en menos tiempo que en un ELISA. Es por esto que se suele usar la media de las mediciones de cada analito en cada pozo como valor representante para la intensidad fluorescente (**MFI** *Mean fluorescent Index*). Adicionalmente, se usa un pozo en donde no hay ninguna muestra como señal de fondo, la cual se le resta al MFI y así se obtiene el **Net MFI**.



**Figura 2.3:** Compartimiento de la cámara digital en el *MAGPIX*®. Las esferas magnéticas atraídas a una monocapa, son estimuladas por un LED rojo que identifica el analito. El LED verde excita las moléculas receptoras y cuantifica la concentración de analito. Las señales emitidas son capturadas por una cámara para hacer análisis de imagen.

Fuente: HCYTOMAG-60K Info Sheet

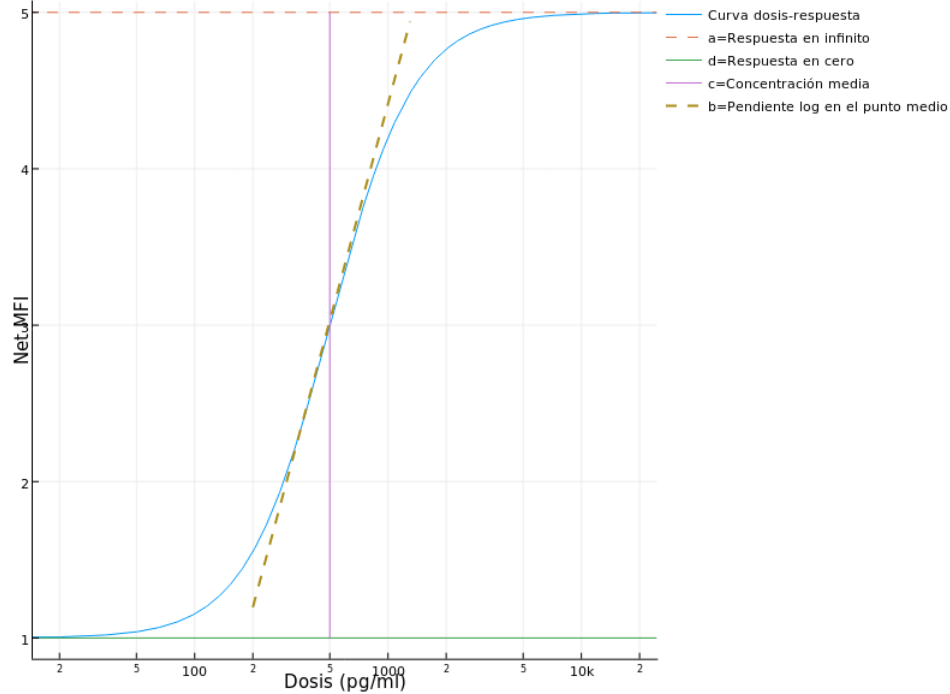
### Curva de dosis-respuesta

Para inferir las concentraciones desde el Net MFI se han utilizado diferentes modelos a lo largo de la historia de los inmunoensayos. El objetivo es obtener una curva de dosis-respuesta que de una relación funcional entre la respuesta, la cual en este caso es el Net MFI, y la dosis, que en este caso es la concentración. Para esto, los factores principales a considerar en este tipo de modelos son:

- Que siempre habrá un comportamiento asintótico hacia los límites de detección.
- Que las concentraciones tiene una tasa de crecimiento logarítmico.
- Que en el rango medio, en escala logarítmica, la relación es aproximadamente lineal.
- La función es monótona, es decir, no hay inversión de la pendiente.

Un modelo que no considera la saturación en las concentraciones límite, es una regresión lineal en escala semilogarítmica. Este modelo tendría altos niveles de error en los extremos, pero puede ser usado en los casos en los que el sustrato no vaya a saturar nunca a los reactivos, ni que haya muy pocas reacciones (Finney and Phillips, 1977). Si la variabilidad en los polos es muy alta, también se puede recurrir a este modelo económico e ignorar los datos en los polos.

El espacio **log-logit** es un espacio con escala logarítmica en los dos ejes y normalizado para que su rango vaya de 0 a 1 en la variable dependiente. Su nombre viene de la distribución logística, la cual tiende asintóticamente a un valor superior y uno



**Figura 2.4:** Visualización de parámetros en la curva log4p. Se muestran rectas que se desplazan o, en el caso de la pendiente, rotan a medida que varían los parámetros.

inferior en su variable dependiente. Trabajar con rectas en el espacio log-logit, resolvió el problema de saturación en los polos, pero estas rectas exigen que haya simetría en las concentraciones altas y bajas (Rodbard et al., 1968). La función logística reúne todas estas características (Rodbard et al., 1978). En su forma general, la curva puede ser caracterizada por 4 parámetros con la siguiente ecuación:

$$\text{Net MFI}(D) = d + \frac{a - d}{1 + \left(\frac{D}{c}\right)^b} \quad (2.1)$$

Donde  $D$  es las dosis y la contribución de los parámetros se menciona en la gráfica 2.4.

Estas curvas de calibración se deben calcular usando al menos cuatro estándares de concentración. Para esto se reservan pozos en la placa para diferentes concentraciones de todos los sustratos. Las concentraciones deben de estar uniformemente esparcidas en todo el rango de detección y deben de ser conocidas de antemano, para poder hacer el ajuste. Se puede disminuir el número de parámetros si se indican las respuestas para la concentración mínima y la máxima de antemano, lo cual reduce el tiempo del ajuste, pero con las computadoras de hoy en día el cambio no es muy notorio.

En muchos inmunoensayos la tasa de crecimiento en rangos altos de concentración es diferente a la de valores bajos. Debido a esto, se ha añadido un parámetro más que

rompe la simetría alrededor del punto de concentración media (Rodbard et al., 1978). La fórmula es, entonces, esta (D es dosis):

$$\text{Net MFI}(D) = d + \frac{a - d}{[1 + (\frac{D}{c})^b]^g} \quad (2.2)$$

En la figura 2.5 se puede ver un ejemplo de como este nuevo parámetro hace que la curva se más asimétrica.

Añadir ese parámetro permite que se puedan reportar valores mucho más cercanos a los umbrales de detección, mientras el ajuste de la curva sea bueno. Además de que aumenta mucho la recuperación del valor de los estándares con los que se arma la curva, sobre todo para los estándares menores cuando  $g < 1$ , pues se refleja. Otro aspecto importante es que si las concentraciones estándar está sobre todo en una zona cerca de las asíntotas de la curva, el modelo de 4 parámetros es mejor se promedia mas ruido estadístico por tener mas grados de libertad al eliminar parámetros (Gottschalk and Dunn, 2005c).

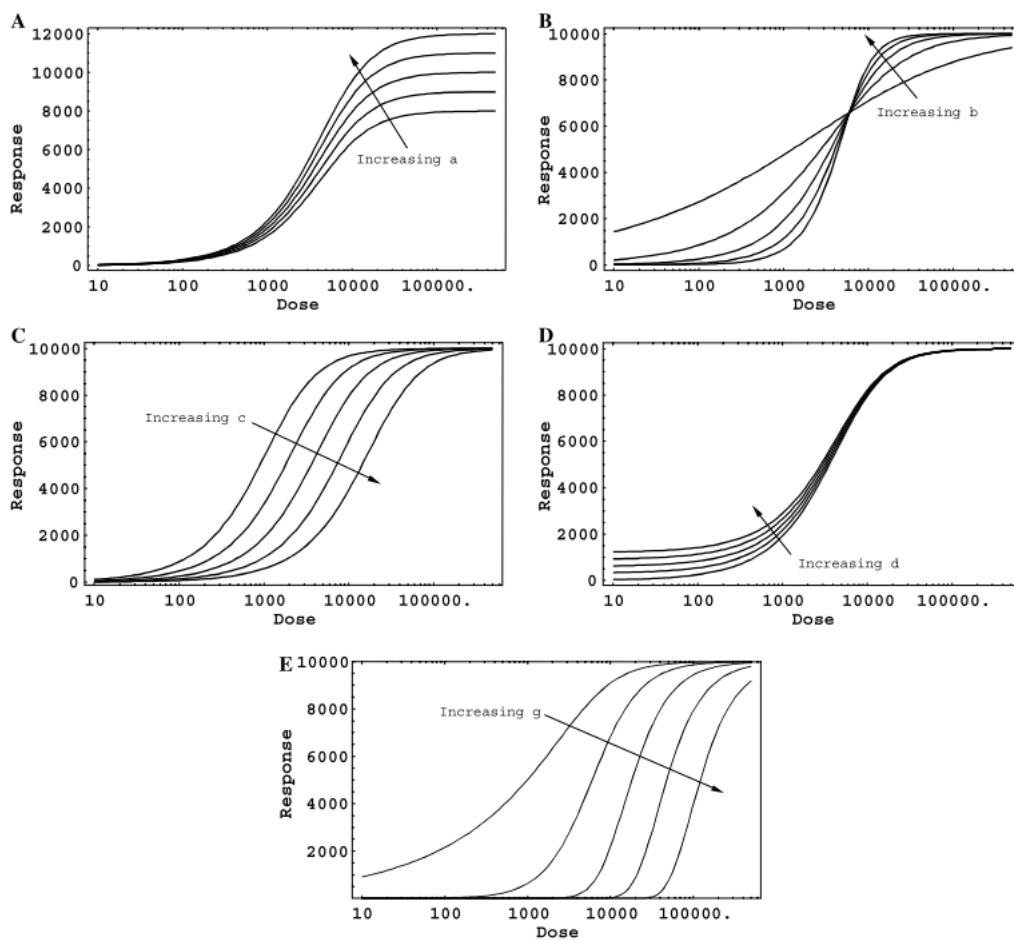
El ajuste de la curva de calibración a los estándares es muy delicado, hay muchas situaciones en las que, si se varían ciertos parámetros, el error de ajuste se mantiene constante, como es el caso con  $b$  y  $g$ , los cuales determinan conjuntamente la pendiente de la curva (Gottschalk and Dunn, 2005c). Dicho contratiempo, junto con la variabilidad de la respuestas en el pasado habían impedido que se usara la curva de 5 parámetros.

En general, se usa un método de mínimos cuadrados para el ajuste, pero es de mucha importancia añadir pesos a la suma de cuadrados para tomar en cuenta que la variabilidad en los valores menores contribuye mucho más que en los mayores. El software *Xponent*<sup>®</sup> ofrece la capacidad de usar pesos del tipo  $1/Respuesta^2$ . Sin embargo, hay estudios (Davidian et al., 1988; Finney and Phillips, 1977; Raab, 1981) que han demostrado que usar pesos basados en la varianza de la respuesta (obtenida de una acumulación histórica de cálculos de curvas o de replicados) disminuye la suma de errores cuadrados.

### Ejecución del ensayo por software

El software *Xponent*<sup>®</sup> se puede configurar para reflejar el diseño de la placa y las condiciones de control de calidad. Existe la posibilidad de guardar la configuración en **protocolos** que pueden ser reutilizados en ocasiones posteriores. Cada corrida de ensayo es guardada como un *'batch'* y se puede observar en tiempo real el resultado del análisis de cada pozo. Si no se cumplen ciertas condiciones, como que el número de perlas detectadas sea mayor a un número asignado, el pozo se muestra como amarillo en un diagrama de la placa. A partir de esa información se puede comprobar que la muestra haya sido preparada correctamente.

Para cada placa se reservan 6 pozos (o 12 cuando se usaron replicados) a los que se añaden cantidades conocidas de las citocinas a analizar, para poder calcular la curva de calibración. En general se usan las siguientes concentraciones (3.2, 16, 80, 400, 2000,



**Figura 2.5:** (A) - (E) Efectos de cambiar los parámetros a,b,c,d,g respectivamente, manteniendo todos los otros fijos. El punto de intersección de las curvas en (B) es cuando la respuesta es igual a c

Fuente: Gottschalk, P. G., & Dunn, J. R. (2005). The five-parameter logistic: A characterization and comparison with the four-parameter logistic. *Analytical Biochemistry*, 343(1), 54–65

10000  $pg/ml$ ). Además de reservarse un pozo para medir la señal de fondo y poder restarla a los valores de medición.

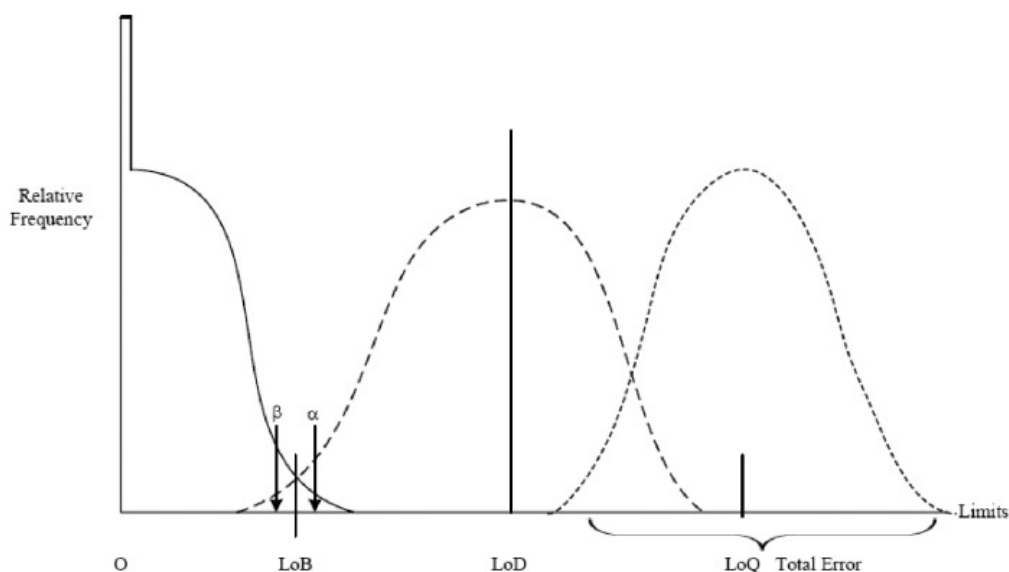
Terminada la lectura, se pueden exportar reportes que contienen los resultados del ajuste de las curvas (junto con gráficas) y las concentraciones detectadas. También es posible exportar un archivo Excel que contiene los parámetros de las curvas de ajuste, las Net MFI detectadas, las concentraciones inferidas, entre otros. Las respuestas que se salen de las asíntotas de la curva de calibración, se reportan como fuera del umbral de detección. Esto se muestra reportando un valor menor al de la concentración más baja, o más alta, de los estándares (por ej.  $< 3.2 pg/mL$  o  $> 10,000 pg/mL$ ).

### Modelo teórico de la sensibilidad en inmunoensayos

Como se ha visto en la sección anterior, la concentración de proteína analizada se infiere a partir de la intensidad de fluorescencia que detecta una CCD en el *MAGPIX*<sup>®</sup>. De manera general, se habla de una señal y se define al límite de detección (**LOD**) como la concentración correspondiente a la mínima señal que se puede diferenciar del ruido de fondo. Para determinarlo se hace uso del límite de señal de fondo (*Limit of Blank* **LOB**), el cual se puede definir como el promedio de la señal de fondo más 1.65 por su desviación estándar (Armbruster and Pry, 2008). Partiendo de esto se puede definir al LOD como el LOB más 1.65 por de la desviación estándar de la señal de fondo. No hay que confundir esta propiedad con el límite de cuantificación **LOQ**, que se define como el valor a partir del cual el porcentaje de error es menor a una cantidad deseada, normalmente de 10 a 20 %. A diferencia del LOD, que delimita cuando se puede afirmar que hay presencia de analito, el LOQ define un umbral de confianza cuantitativa. En la Fig. 2.6 está representada una distribución común de estos límites. La determinación del LOQ es más difícil, ya que se necesita saber la variación de las concentraciones cerca del límite de detección. Debido a la complejidad en el modelaje de estas variaciones desde principios físicos, se usan normalmente métodos estadísticos que dependen de replicas de las mediciones y que ignoran estos principios. Al final de esta sección se hablará un poco de una posible estrategia estadística.

Los factores fisicoquímicos que contribuyen principalmente a la sensibilidad de un inmunoensayo a temperatura, volumen, pH y presión constante, son la afinidad y disociación de los anticuerpos, el ligado no específico, la reactividad cruzada y la interferencia de anticuerpos.

El ligado entre anticuerpo y antígeno sucede debido a la acumulación de muchos enlaces no covalentes del tipo Van der Waal, iónico, hidrofóbico o de hidrógeno. La energía necesaria para liberar estos enlaces es del orden de  $10 \frac{kJ}{mol}$ . Por lo que el agua, cuya energía cinética a  $37^{\circ}C$  es en del mismo orden de magnitud en promedio, hace que estos enlaces sean de muy corta duración (Reverberi and Reverberi, 2007). Es por eso que la reacción antígeno-anticuerpo es reversible y se le puede asociar un constante tanto a la tasa de asociación como de disociación. La tasa de asociación está limitada por la difusión de los reagentes en la solución a ser menor que  $10^{-9} 1/(Ms)$ . Naturalmente,



**Figura 2.6:** Esquema de la distribución de los límites en las detección de señales.  $\alpha$  y  $\beta$  representan significancia estadísticas de decisión sobre si hay una señal o no

Fuente: Armbruster, D. A., & Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist. Reviews*, 29 Suppl 1(Suppl 1), S49-52

para un inmunoensayo lo mejor es tener una baja tasa de disociación y una alta de asociación. Esto se logra con el diseño de anticuerpos y la mayoría de los anticuerpos usados en ensayos *multiplex* son monoclonales ( que son producidos por una misma célula, por lo que su epítipo no varía), pero no es tan importante que el anticuerpo de captura sea monoclonal porque, aunque se capturen otras proteínas, si el anticuerpo de reporte es monoclonal, no se enlazarán al antígeno y no será reportado. Un buen diseño de anticuerpos es lo único que puede ayudar a aminorar el problema de reactividad cruzada, que es cuando los anticuerpos reaccionan con otra molécula que no es la buscada.

Existen algunos modelos para las interacciones entre antígeno y anticuerpo en la literatura. En estos modelos se llega a considerar la cinética de la reacción (Chang et al., 2012), se puede usar geometría fractal (Sadana and Ram, 1994) o integrales de convolución (Klenin et al., 2005) para considerar la geometría de los antígenos ligados a la fase sólida. El que consideramos puede considerarse el más simple posible, pues asume que las reacciones han alcanzado el equilibrio y que la distribución de los anticuerpos es homogénea, hecho que puede asumirse gracias a el uso de las perlas suspendidas.

El ligado no específico se refiere a cuando el anticuerpo de reporte se pega a la superficie sólida del inmunoensayo heterogéneo. En el caso específico de las microesferas, estas tienen poros del orden 100 a 200 angstroms (Angeloni et al., 2013), lo que les permite capturar a los anticuerpos o a los analitos. Este fenómeno contribuye al ruido de fondo de la señal, pero puede ser reducido con agentes de bloqueo que se adhieren a estos poros para impedir que otras moléculas lo hagan.



Finalmente, la interferencia de anticuerpos sucede en muestras de suero o con muchos anticuerpos endógenos. Sucede cuando estos anticuerpos se unen al analito o a los anticuerpos del ensayo. Se usan diluyentes diseñados para bloquear los anticuerpos endógenos, del tipo de muestra en la que se trabaja. Por ejemplo, en el caso de estudio se uso un diluyente que venía incluido en un *kit* para medir citocinas humanas.

En aras de elucidar los fenómenos físicos que influyen la sensibilidad de un inmunoensayo, se presenta a continuación un modelo teórico para la estimación del LOD. Afortunadamente, esta medida se acerca mucho al LOQ en los inmunoensayos de perlas magnéticas, a pesar de no considerar los errores en la optimización de la curva de dosis-respuesta. En este tipo de ensayos no se presenta la aumentada incongruencia entre el LOD y el LOQ que tienen los inmunoensayos los cuales pueden detectar señales de una sola molécula. Incongruencia que tiene su raíz en el ruido de disparo inherente a todo medidor de señales eléctrico cuando mide pocos fotones (Rissin et al., 2010).

El modelo que usaremos viene de (Wild, 2013, p. 52) y está inspirado en los estudios de Jackson y Ekins (Jackson and Ekins, 1986). Es un modelo simple que separa las fuentes de error debido a la ejecución y definición de los instrumentos, de la relación que hay entre la concentración del analito y la señal. Primero, calcula la señal a partir de la cual se sabe que hay presencia de analito y luego, hace uso de la ley de acción de masas para calcular la concentración de analito para la cual se mide esa señal.

A continuación se presenta una breve demostración de la ley de acción de masas, para elucidar como actúan las suposiciones que se hacen sobre la fórmula y para familiarizarse con ella.

### Ley de acción de masas

La ley de acción de masas fue desarrollada a través de principios fenomenológicos, entre 1864 y 1879 por Waage y Guldberg. Se basa en dos premisas. La primera es que la razón de ocurrencia de las reacciones es proporcional a la actividad o concentración de los reagentes. Este principio aplica sobre todo, para reacciones de primer orden (e.d. reacciones con átomos elementales). Mientras que, en el caso de acoplamiento de antígeno-anticuerpo, la constante de proporcionalidad varía y se mantiene solo bajo ciertas condiciones (Rich, 2008). La segunda premisa, que es más una condición que un principio, es que la reacción sea reversible y haya alcanzado un equilibrio químico. Haciendo uso de estas dos premisas se puede demostrar, con ayuda de la termodinámica, que la tasa de asociación es igual a la de disociación. Expresado matemáticamente esto es

$$K_{eq} = \frac{k_a}{k_d} = \frac{[S]^\sigma [T]^\tau \dots}{[A]^\alpha [B]^\beta \dots} \quad (2.3)$$

Donde  $K_{eq}$  es la constante de equilibrio,  $k_a$  es la constante de asociación,  $k_d$  la de disociación, las letras mayúsculas en el numerador son las concentraciones molares de los reagentes antes y las de abajo después de la reacción, y las letras griegas son coeficientes

estequiométricos. En la literatura no hay una convención de si los productos iniciales deben de estar en el denominador o el numerador, por lo que a veces la constante de equilibrio es el inverso de la que aparece en unos artículos.

Para demostrar esta relación se parte la energía libre de Gibbs en reacciones químicas, que se usa para reacciones químicas porque en ellas es más fácil mantener la temperatura y la presión constantes:

$$dG = Vdp - SdT + \sum_{i=1}^k \mu_i dN_i \quad (2.4)$$

Donde  $dN_i$  es el cambio en el número de reagentes y  $\mu_i$  es el potencial químico del reagente  $i$ . Sobre el cual ya se afirmó que es proporcional a la concentración del reagente:

$$\mu_i = \mu_i^0 + RT \ln[A_i] \quad (2.5)$$

Con  $R = 8.314 \frac{J}{Kmol}$  la constante de los gases y  $[A_i]$  la concentración molar del sustrato.

Un principio fundamental de la termodinámica es que los mínimos de potenciales termodinámicos implican un equilibrio. Como queremos encontrar un equilibrio en la tasa de ocurrencia de las reacciones podemos expresar a  $dN_i = \mp \nu_i d\xi$ , siendo que  $d\xi$  representa la ocurrencia de una reacción. El caso más representativo de reacción y el que normalmente se considera para la ley de masas es el siguiente  $\nu_1 A_1 + \nu_2 A_2 \rightleftharpoons \nu_3 A_3 + \nu_4 A_4$ . Si hacemos el cambio de variable a  $d\xi$ , cuando ocurre una reacción, el número de moléculas disminuye para las iniciales y aumenta para el producto final. Por eso se debe de añadir un menos a el diferencial de las moléculas iniciales.

Haciendo uso de la nueva variable  $\xi$  se puede definir al equilibrio térmico como cuando el cambio en la energía libre de Gibbs (a presión y temperatura constante) debido a las reacciones es 0:

$$\left(\frac{dG}{d\xi}\right)_{p,T} = 0 \quad (2.6)$$

Si se deriva el potencial de Gibbs, se sustituye a  $dN_i$  por  $d\xi$  y se pide que la solución haya alcanzado el equilibrio térmico, se puede hacer el desarrollo algebraico siguiente:

$$\begin{aligned} \left(\frac{dG}{d\xi}\right)_{p,T} &= \nu_3 \mu_3 + \nu_4 \mu_4 - \nu_1 \mu_1 - \nu_2 \mu_2 \\ \text{Por (2.5)} &= \nu_3 \mu_3^0 + \nu_4 \mu_4^0 - \nu_1 \mu_1^0 - \nu_2 \mu_2^0 \\ &\quad + RT(\ln[A_3]^{\nu_3} + \ln[A_4]^{\nu_4} - \ln[A_2]^{\nu_2} - \ln[A_1]^{\nu_1}) \\ 0 &= \Delta G^0 + RT \ln\left(\frac{[A_3]^{\nu_3} [A_4]^{\nu_4}}{[A_1]^{\nu_1} [A_2]^{\nu_2}}\right) \\ \Delta G^0 &= RT \ln K_{eq} \end{aligned} \quad (2.7)$$

Donde  $\Delta G^0$  es el cambio de energía libre de Gibbs estándar y es intrínseco a los reagentes, es decir es constante. De arriba se puede ver que el cociente en el logaritmo

debe ser constante y es precisamente la constante de equilibrio. De tal manera queda demostrada la ley de masas para reacciones reversibles, en equilibrio y cuyo potencial químico es proporcional a la concentración molar.

### Error del instrumento óptico

El otro elemento que hace falta para proponer una estimación del LOD, es el error intrínseco al instrumento de medición. Este depende de la sensibilidad del CCD y el brillo de la molécula de etiquetado (en este caso R-PE). El manual del instrumento *MAGPIX*<sup>®</sup> indica que puede detectar a partir de 700 moléculas de R-PE por microesfera. En los siguientes párrafos se explican brevemente los factores físicos que influyen en este umbral. Lamentablemente, no se cuenta con los suficientes datos técnicos del instrumento como para hacer un estimado razonable desde un modelo, pero se puede usar el reportado por el instrumento.

Si se exponen las influencias al error de medición del instrumento en el orden en que suceden, la primera sería el brillo de la molécula de reporte. Este es el producto del rendimiento cuántico  $\epsilon$ , que mide el porcentaje de fotones irradiados con respecto a los fotones absorbidos y el coeficiente de atenuación  $QE$ , que mide cuanta luz se absorbe en cierta área, con cierta concentración y tiene unidades  $\frac{1}{M \text{ cm}}$ . Una vez emitida la luz, la cantidad de fotones que llegan al CCD depende de la longitud del camino recorrido, el índice de refracción de la muestra y el tiempo de integración de la imagen. Además, en el caso de uso de perlas, su orientación con respecto a la cámara puede influir sobre el brillo, pues la moléculas sobre la perla que estén en el lado opuesto al CCD casi no contribuirán.

Como el CCD convierte fotones en señal eléctrica, está sujeto al ruido de disparo que es un fenómeno cuántico causado por la naturaleza estocástica de la radiación atómica y que causa que el tiempo entre fotones emitidos sea inherentemente aleatorio. Mientras más fotones se capturen, menor es el error debido a este efecto, pues el número de fotones emitidos sigue una distribución de Poisson con desviación estándar igual a la raíz del número de fotones capturados. Por otro lado, la eficiencia cuántica es intrínseca al material fotoeléctrico que se usa y es la probabilidad de que un fotón con cierta longitud de onda emita un fotoelectrón. Además, puede contribuir la corriente oscura que es la señal espontánea que se genera en cualquier diodo en su región de agotamiento, por fluctuaciones térmicas, entre otras cosas. Se genera aún cuando no esté siendo irradiada ninguna luz al detector. Dicho efecto disminuye con la temperatura de operación y también se puede calibrar la cámara con una imagen oscura para sustraer la señal. También, debido a que se detectan voltajes muy bajos y se usan potenciales mas grandes para la operación de la cámara, puede haber mucha interferencia. Es por eso que mientras mejor sea el diseño de los circuitos integrados, menor ruido habrá. Finalmente, una contribución muy pequeña al error proviene del transporte de la señal analógica a una señal digital. En este proceso surgen errores como el de reseteo, el de Johnson o el de parpadeo. Existen técnicas como el muestreo doble correlacionado para eliminar estas fuentes de ruido (Donati, 2000).

Como se puede notar, existen tantos factores que añaden ruido al sistema, que una estimación correcta de la sensibilidad del CCD, sólo se puede hacer con métodos experimentales (Ortyn et al., 2006).

### Estimación del LOQ y LOD y la influencia de parámetros en el LOD

El límite de detección (**LOD**), se define como la mínima concentración a partir de la cual se puede diferenciar de la señal de fondo. La mínima concentración de ficoeritrina (R-PE) que puede detectar el CCD (700 moléculas de R-PE por perla según el folleto técnico del *MAGPIX*<sup>®</sup>), vendría dando la concentración de anticuerpos de detección ligados  $[Ab2]_b$ . Partiendo de ese dato se quiere calcular la mínima concentración detectable  $[Ag1]_t$  de analito para ciertas constantes de equilibrio y concentraciones de anticuerpo, basándose en la ley de acción de masas.

Para obtener la concentración total de un anticuerpo antes de que se establezca su reacción reversible, se puede partir de la Ec. 2.3 en la siguiente forma:

$$K_{eq} = \frac{[Ag - Ab]_b}{[Ag]_f [Ab]_f} \quad (2.8)$$

Además, se debe usar la conservación de masa  $[Ab]_t = [Ab]_f + [Ab]_b$  (igual para  $Ag$ ) y el hecho de que  $[Ag - Ab]_b = [Ab]_b = [Ag]_b$  para llegar a la siguiente expresión. Donde el subíndice 't' significa total, 'b' es *bound* e.d. ligado y 'f' es *free* e.d. libre;  $[Ag2]_t$  es la concentración de proteína capturada en los anticuerpos,  $[Ab2]_t$  es la concentración de anticuerpo de detección y  $[Ab2]_b$  es la concentración de proteína ligada al anticuerpo de detección.

$$[Ag2]_t = \frac{[Ab2]_b}{([Ab2]_t - [Ab2]_b)K_{2eq}} + [Ab2]_b \quad (2.9)$$

Usando esta ecuación, se puede obtener la concentración total de antígeno capturado por las perlas ( $[Ag2]_t$ ) necesario, para que haya cierta cantidad de antígeno ligado a un anticuerpo de reporte ( $[Ab2]_b$ ). Es curioso notar que la dinámica que se asume con estas fórmulas exige que haya una concentración total de antígeno infinita para que todos los anticuerpos se hayan ligado a un antígeno. Además, mientras mayor sea la constante de equilibrio, menos antígeno se necesitará para que la cantidad de anticuerpo ligado se acerque a la total. Naturalmente, en la vida real esto no es así, pero como estamos asumiendo equilibrio esto no debería de pasar. Este comportamiento va a dar lugar a muchas singularidades que no reflejan el caso real, pero en las que se habría roto el equilibrio así que no habría porque considerarlas. Más adelante se explica su comportamiento geométrico en las gráficas.

Luego, se puede proceder a preguntar por la concentración inicial de antígeno que tuvo que haber si se ligó cierta cantidad de anticuerpo de detección (e.d.  $[Ag2]_t = [Ab1]_b$ ). Esto vendría siendo la sensibilidad o  $[Ag1]_t$ .

$$[Ag1]_t = \frac{[Ag2]_t}{([Ab1]_t - [Ag2]_t)K_{1eq}} + [Ag2]_t \quad (2.10)$$

## 2. ANTECEDENTES

---

Sustituyendo la Ec. 2.9 en la Ec. 2.10 se obtiene una fórmula para la sensibilidad en la que se pueden variar algunos parámetros.

$$[Ag1]_t = [Ab2]_b \left( \frac{1}{[Ab2]_t - [Ab2]_b} + \frac{1 + K2_{eq}([Ab2]_t - [Ab2]_b)}{K1_{eq}(K2_{eq}([Ab2]_t - [Ab2]_b)([Ab2]_b - [Ab1]_t) - [Ab2]_b)} \right) \quad (2.11)$$

Para poder visualizar comportamientos cercanos a los del laboratorio, se hizo un estimado de los valores de ciertos parámetros, basándose los ensayos ejecutados en el caso de estudio. La concentración de anticuerpos de captura  $[Ab1]_t$ , se puede obtener partiendo de que debe haber un agotamiento de anticuerpos al acoplarse con las esferas y de que los anticuerpos miden de 6 a 15 nm de ancho y 10-20 nm de alto. Los autores de (Chang et al., 2012) observaron experimentalmente que, en promedio, se acopla un anticuerpo por cuadrado de lado de 15 nm. Como en el *xMap Cookbook* se informa que las perlas magnéticas tienen  $6.5\mu m$  de diámetro, su área de superficie es  $A = 4\pi r^2 = 133\mu m^2$ . Lo cual quiere decir que hay aprox. 589,921 anticuerpos por perla y como cada anticuerpo tiene dos epítomos, se tiene el doble de sitios de ligado. De los reportes del software *XPonent* en el laboratorio, se dedujo, usando el máximo número de perlas contadas por analito, que debe de haber aprox. 350 perlas por  $50\mu L$ . Debido al tamaño del fluoróforo R-PE, su reacción con el anticuerpo es 1 a 1, así que el numero de anticuerpos ligados debe ser igual al numero de moléculas de R-PE. Usando la concentración de perlas calculada anteriormente (350 perlas por  $50\mu L$ ), podemos afirmar que la concentración de sitios de captura  $[Ab1]_t$  es aprox.  $1.5 ng/mL$  o  $2 \times 10^{-12} M$ .

Otro factor importante que se puede estimar con valores reales es la cantidad de anticuerpo de detección ligado  $[Ab2]_b$ . Primero se asume que las constantes de equilibrio tienen un valor de  $10^{12} \frac{1}{M}$  y que hay altas concentraciones de anticuerpo de captura  $[Ab1]_t$  en el orden de  $10^{-8} M$ . Luego, según el manual de uso del *MAGPIX*<sup>®</sup> se deben de contar al menos 50 perlas para que el resultado sea estadísticamente significativo. Ahora, asúmase que el antígeno se liga a 50 perlas en un volumen de  $50\mu L$ , que es la menor cantidad de volumen que se puede añadir a cada pozo. Esto equivale a  $10^6$  perlas por litro. Por otro lado, cuando se analizó la variación de la FI en las muestras de fondo, se observó que el coeficiente de variación (CV) está alrededor de 27%. Excepto por algunas perlas fuera de rango, la FI sigue una distribución normal. Si se asume un relación lineal entre el FI y el número de R-PE's y se quiere que el 95% de las perlas tenga menos de 700 moléculas, el promedio de moléculas por perla debería ser tal, que si se le suma dos veces la desviación estándar sea igual a 700. Tomando en cuenta el CV de arriba, esto nos da un promedio de 467 moléculas por perla. Ahora, si se multiplica la concentración de perlas por 467 anticuerpos de reporte por perla, esto nos da  $467 \times 10^6$  anticuerpos por litro, o (si se divide por el numero de Avogadro)  $0.77 fM$ . Esto equivale a que  $[Ab2]_b = 0.12 \frac{pg}{mL}$  aprox., si se toma en cuenta que los anticuerpos pesan aprox.  $150 KDa$ , peso que no varía mucho pues su variación sucede en los epítomos.

Con ayuda de estos valores se puede hacer una gráfica en la que se varía la concentración de anticuerpo de reporte y su constante de equilibrio (Véase Fig. 2.7). Se

puede notar que hay un punto donde se alcanza un llano y a pesar de que se aumente la constante de equilibrio o el anticuerpo de reporte, no hay mejoras considerables en la sensibilidad. Lo mismo pasa si se varía la constante de equilibrio del anticuerpo de captura, pero, a niveles bajos, el límite de detección aumenta mucho más precipitadamente que cuando se varía la otra constante de equilibrio. Según Landry et al. (2012), las constantes de equilibrio de los anticuerpos usados en inmunoensayos van de  $10^{-9}$  a  $10^{-12}$ . Al variar ambas también se llega a un llano.

Un factor siempre presente es el del ligado no específico. Este fenómeno ocurre sobre todo en los anticuerpos detectores, pues la variedad de las fuerzas en los epítomos de los anticuerpos permite que se conecten a diferentes arreglos moleculares, cuya cantidad aumenta mientras mas rica sea la matriz (o medio) de la muestra (Véase (Mendel and Mendel, 1985) para una descripción mas detallada de las fuerzas). Se puede asumir que, mientras no se esté cerca de la saturación, el ligado no específico depende linealmente de la concentración de anticuerpo de detección total. Por lo que se podría aproximar la cantidad de anticuerpo de detección ligado como  $[Ab2]_b + nsb \times [Ab2]_t$  en la Ec. 2.9. Sin embargo, sustituir esto implicaría que los anticuerpos ligados al antígeno se disocian a la misma tasa que los ligados a cualquier otro lugar. Esto no es cierto, pues los anticuerpos son mucho más específicos a su antígeno, pero se puede disminuir un poco la constante de equilibrio y que eso compense la menor especificidad de las otras reacciones.

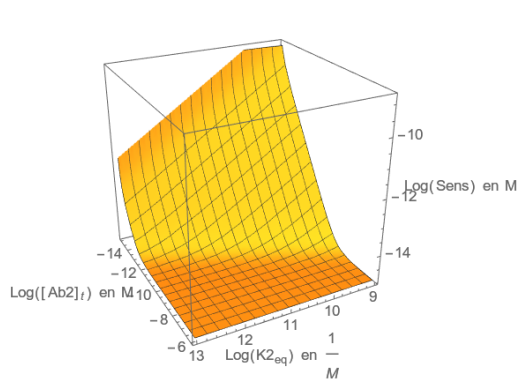
Una vez añadido este término, la sensibilidad ya no tiende asintóticamente a un valor mínimo si no que tiene un mínimo, pues si se añade demasiado anticuerpo de captura, el ligado no específico aumenta. En las gráficas de la Fig. 2.8 se puede ver como la concentración de anticuerpo de detección es la mas importante para optimizar la sensibilidad, la concentración del anticuerpo de captura y la constante  $K1_{eq}$  no importan pero, a medida que se disminuye  $K2_{eq}$ , el mínimo de sensibilidad se encuentra en concentraciones mayores de anticuerpo de detección.

Las singularidades que pueden aparecer en este tipo de expresión son varias. La mas obvia es la que ya se comentó anteriormente, que sucede cuando la concentración total de anticuerpo es igual a la del ligado. Otra es por ej. si se varían las concentraciones totales de anticuerpo, el divisor del segundo sumando de la Ec. 2.11 se hace 0 para una ecuación irracional de segundo grado, pero que para los valores del dominio que usamos se comporta linealmente. Esto se puede ver en la gráfica 2.8a, donde hay una línea recta vacía con artefactos rodeándola y las singularidades se hacen más presentes pues el factor de ligado no específico “aumenta” la concentración de anticuerpo ligado. También hay singularidades en regiones hiperbólicas cuando se varían las constantes de equilibrio contra alguna concentración total y cuando se varían las dos variables de equilibrio, pues por ej. el dividendo del segundo sumando de la Ec. 2.11 se hace 0 cuando  $K1_{eq}$  y  $K2_{eq}$  cumplen una ecuación hiperbólica.

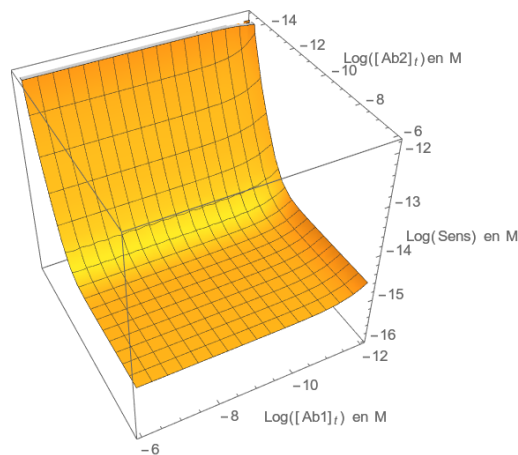
Las Fig. 2.7 y 2.8 ilustran la influencia de los diferentes parámetros en el LOQ. Un estimado para la mejor sensibilidad que se podría obtener con este modelo se puede calcular considerando los siguientes parámetros.

Estas gráficas sirve sobre todo para visualizar como actúan los parámetros sobre

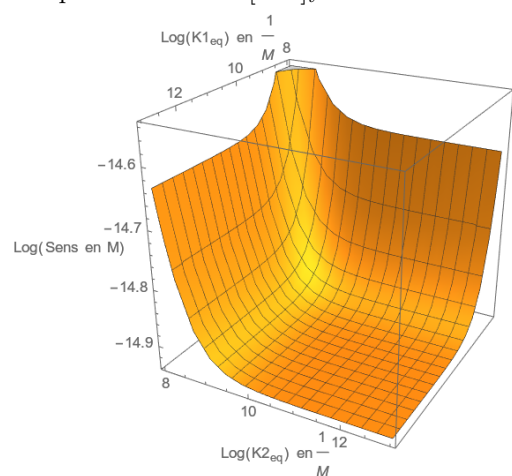
## 2. ANTECEDENTES



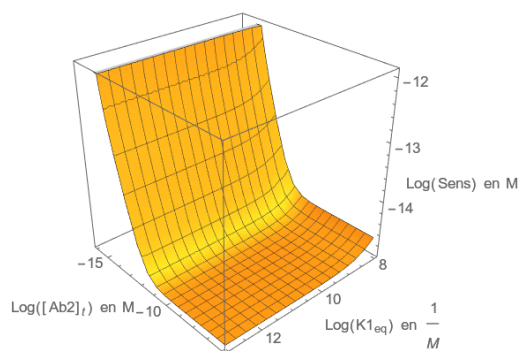
(a) Se varían la constante de equilibrio de la reacción entre anticuerpo de detección y el antígeno capturado  $K2_{eq}$  y la concentración total de anticuerpo de detección  $[Ab2]_t$ .



(b) Se varían la la concentración total de anticuerpo de captura  $[Ab1]_t$  y la concentración total de anticuerpo de detección  $[Ab2]_t$ .

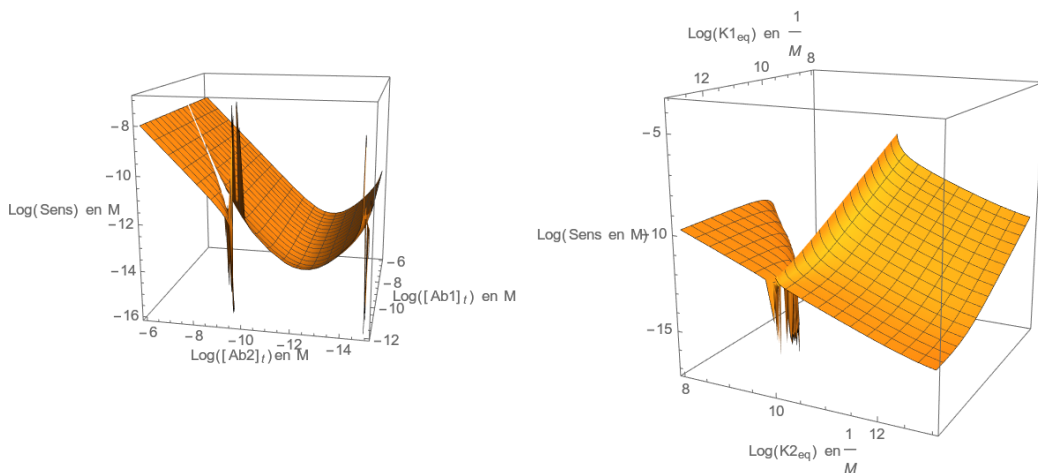


(c) Se varían la constante de equilibrio de la reacción entre anticuerpo de detección y el antígeno capturado  $K2_{eq}$  y la de reacción entre anticuerpo de captura y antígeno total  $K1_{eq}$ .



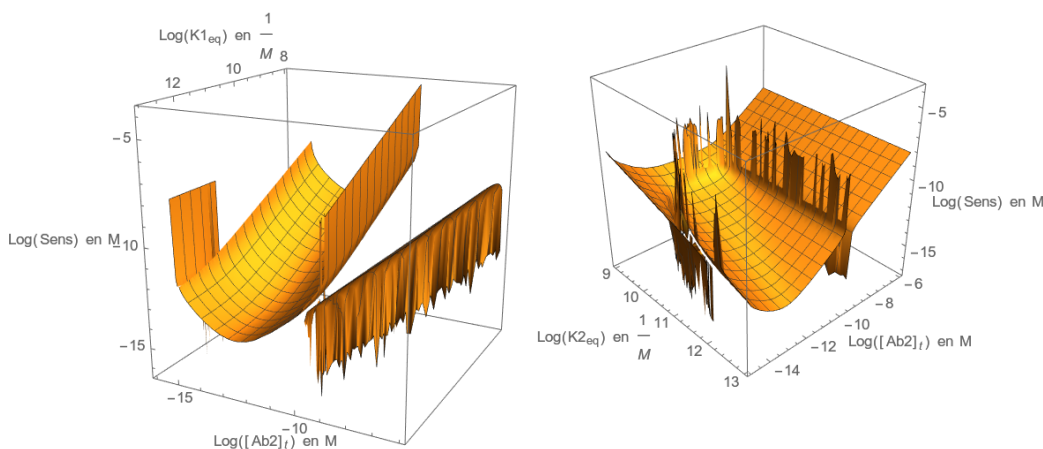
(d) Se varían la la concentración total de anticuerpo de detección  $[Ab2]_t$  y la constante de equilibrio del anticuerpo de captura  $K1_{eq}$ .

**Figura 2.7:** Límite de detección en molares, según un modelo en el que solo se toma en cuenta la ley de acción de masas, con la Ec.2.11. Cuando no se varían los valores para las parámetros, estos valen  $[Ab1]_t = 210^{-12}M$ ,  $[Ab2]_t = 10^{-8}M$ ,  $K1_{eq} = K2_{eq} = 10^{12} \frac{1}{M}$  y  $[Ab2]_b = 810^{-16}M$



(a) Se varían al concentración total de anticuerpo de captura  $[Ab1]_t$  y la concentración total de anticuerpo de detección  $[Ab2]_t$ .

(b) Se varían la constante de equilibrio del anticuerpo de captura  $K1_{eq}$  y la del anticuerpo de detección  $K2_{eq}$ .



(c) Se varían la la concentración total de anticuerpo de detección  $[Ab2]_t$  y la constante de equilibrio del anticuerpo de captura  $K1_{eq}$ .

(d) Se varían la constante de equilibrio de la reacción entre anticuerpo de detección y el antígeno capturado  $K2_{eq}$  y la concentración total de anticuerpo de detección  $[Ab2]_t$ .

**Figura 2.8:** Límite de detección en molares, según un modelo en el que se toma en cuenta la ley de acción de masas y el ligado no específico NSB con un valor de 0.1%. Los cortes abruptos en las gráficas *a*, *c*, *d* se deben a la singularidad que surge al dividir entre 0 y que se explican arriba. Cuando no se varían los valores para las parámetros, estos valen  $[Ab1]_t = 210^{-12}M$ ,  $[Ab2]_t = 10^{-13}M$ ,  $K1_{eq} = K2_{eq} = 10^{12} \frac{1}{M}$  y  $[Ab2]_b = 810^{-16}M$



## 2. ANTECEDENTES

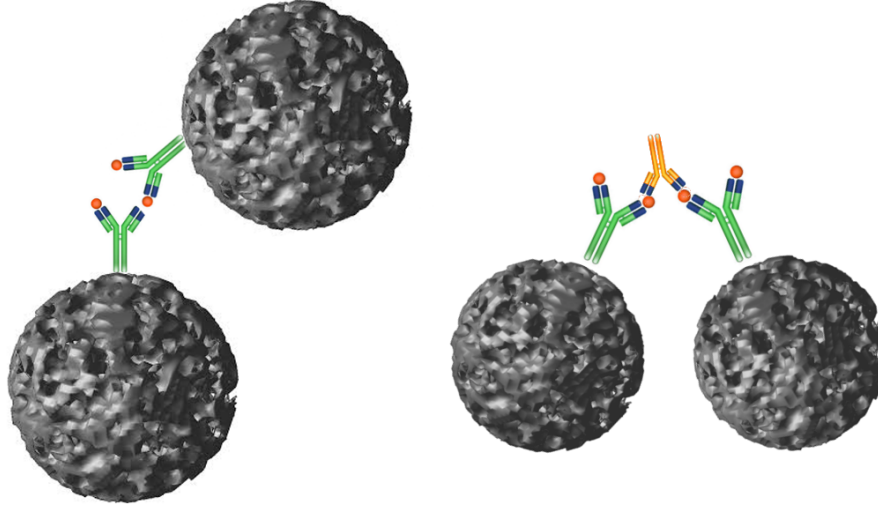
---

la sensibilidad, pero no tanto para ver el mínimo mas precisamente. Si no se toma en cuenta el ligado no específico (**nsb**, *non specific binding*) y se usan los valores determinados con datos del caso de estudio el  $LOD = 0.18 \frac{pg}{mL}$  = el cual no cambia casi nada comparado con la concentración de anticuerpos de captura ligados  $[Ab2]_b$ , pues se asume que se tiene muchísimo anticuerpo de captura ( $[Ab1]_t = 10^{-8}$ ) y que las constantes de equilibrio son las mejores que se conocen para anticuerpos monoclonales ( $10^{12} \frac{1}{M}$ ). Cuando se toma en cuenta un **nsb** de 1 %, el LOD es  $4.6 \frac{pg}{mL}$  y con 0.1 % es  $0.6 \frac{pg}{mL}$ .

En la ficha técnica del inmunoensayo preparado que se usó en el caso de estudio (Cytokine et al., 2013), se reportan límites de detección que llegan a ser tan bajos como  $0.4 \frac{pg}{mL}$  (para IL-8 ) y tan altos como  $26.3 \frac{pg}{mL}$  (para VEGF). Como se puede ver, esta variabilidad puede depender sobre todo de las constantes de equilibrio y la propensión al ligado no específico. Las concentraciones de los anticuerpos se pueden optimizar fácilmente. Además, otro factor muy influyente es el de el surgimiento de aglomerados de perlas. Este fenómeno no tiene que suceder por un ligado no específico, mas bien sucede por el hecho de que los anticuerpos tienen dos epítomos. Si pasan casos como los ilustrados en la Fig. 2.9, la fuerza del separador magnético no es suficiente como para que no se separen las perlas en los pasos de lavado o, las mediciones son descartadas en el paso de análisis de la imagen del CCD. Esto hace que se disminuya la sensibilidad del ensayo, pues se necesita más analito para generar una señal porque se pierde gran parte, pero también afecta la precisión.

Una observación interesante que surgió al calcular este límite teórico de detección fue que el límite de detección es muy sensible a la concentración de perlas que se añaden al sustrato. Además, a partir de cierto umbral, no importa que se aumente la concentración de anticuerpo de captura ( el cual está ligado a las perlas), por lo que en concentraciones bajas, es de mucho provecho disminuir el número de perlas para que haya más moléculas de R-PE por anticuerpo. En muchos laboratorios se usan paquetes prefabricados de perlas para los inmunoensayos en los que se optimizan los parámetros para obtener la mejor sensibilidad y rango dinámico. Es por esto que se escoge una concentración de perlas fija, la cual no proveyó la empresa de *MAGPIX*<sup>®</sup>, *Luminex*<sup>®</sup>, pues declamaron que es información patentada. Sin embargo, para concentraciones de analitos muy bajas, sería posible reducir la concentración de perlas para aumentar la fluorescencia de cada una. Por otro lado, se debe de considerar que la pérdida por aglomeración de perlas puede hacer que no se tenga suficientes perlas como para tener 50 por medición. Además, a pesar de que aumente la fluorescencia, se tendrían que hacer estudios para determinar el límite de detección, como por ej. por medio de perfiles de precisión, sobre los cuales se explica mas adelante.

Hasta ahora solo se ha tratado el límite de detección (LOD), que exclusivamente se refiere a cuando se puede diferenciar una señal de la señal de fondo, sin determinar un nivel de confianza. Como se ha explicado en el capítulo de funcionamiento de *MAGPIX*<sup>®</sup>, para poder determinar una concentración partiendo de la intensidad fluorescente se ajusta una curva logística de 5 parámetros a unos estándares. El ajuste de esta curva también aporta un grado de error. Es por esto que para poder determinar el límite de cuantificación (LOQ), en el que se determina un nivel de confianza para el



**Figura 2.9:** Posibles arreglos de ligados en los que se pueden armar aglomerados de perlas magnéticas, los cuales son descartados por el instrumento o son removidos en los procesos de lavado.

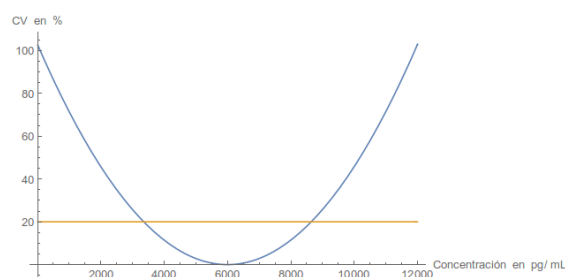
mínimo valor detectable, es necesario hacer uso de estadística.

En general, en el desarrollo de bioensayos no se generan intervalos de confianza para los LOQ. Sin embargo, hay algunos métodos desarrollados para poder dar esta información como los de (Gottschalk and Dunn, 2005a) y (Holstein et al., 2015). Además, los fabricantes de inmunoensayos de Milliplex<sup>®</sup>, hacen uso de el programa *StatLIA* de *Brendan Technologies*, el cual ejecuta un proceso parecido al de (Gottschalk and Dunn, 2005a), para obtener un intervalo de confianza del LOQ, mismo que enseñan en el folleto técnico de sus inmunoensayos preparados. Existe además un paquete de R, llamado *nCal* (Fong et al., 2013), que también usa técnicas parecidas para generar intervalos de confianza.

En este caso de estudio se hizo un programa en el lenguaje *Julia* (Bezanson et al., 2017), que se basaba en el artículo (Gottschalk and Dunn, 2005a), para poder obtener un perfil de precisión de las concentraciones medidas. Un perfil de precisión es una función la cual regresa un coeficiente de variación (e.d. desviación estándar entre promedio) para cada concentración y normalmente se ve como en la Fig. 2.10. En este programa se usaron simulaciones Monte Carlo de los valores de la concentraciones estándar para poder propagar las incertidumbres en la función logística de 5 parámetros, pues un propagación analítica requeriría de demasiados datos o poder de computo. Sin embargo, no se contaba con los suficientes datos para generar una distribución de las FI de los estándares. Para esto sería necesario usar réplicas de los estándares en el mismo ensayo o hacer una acumulación de datos históricos de ensayos en condiciones parecidas, pero ninguno de los dos casos fue factible.

## 2. ANTECEDENTES

---



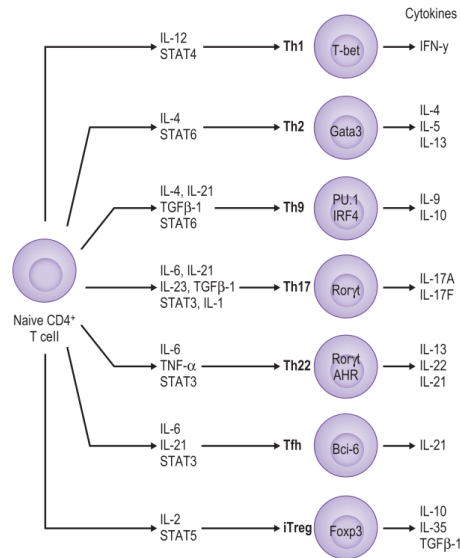
**Figura 2.10:** Posible gráfica de la función del perfil de precisión. Se dibuja una línea en la precisión de 20 %, que es un valor común para escoger el límite de cuantificación

## Citocinas

El tipo de proteínas cuantificadas en el caso de estudio de esta tesis, es el de las citocinas. Estas son proteínas producidas por células, que son de acción extracelular y sirven para la señalización celular, principalmente del sistema inmunológico, pero también de las células derivadas de la hematopoyesis (proceso de formación, desarrollo y maduración de las células que circulan en la sangre) y otras con menor frecuencia. Dicha señalización puede causar diferenciación (Véase por ej. Fig.2.11 ), mitosis, quimiotaxis (que la célula se transporte a un lugar) y otra gamma de procesos fisiológicos que se emplean al combatir virus, cerrar una herida y otros insultos al cuerpo. A diferencia de las hormonas que mantienen la homeostasis, son activadas mayormente en reacción a un ataque al cuerpo o un desorden, el cual ayudan a reparar, por lo que, en general, su presencia es transitoria o en bajas concentraciones. Además, raramente ejecutan su función lejos de la célula que las produce, es decir son autocrinas o paracrinas, a diferencia de las hormonas, que son producidas en lugares muy localizados y actúan en todo el cuerpo (endocrinas). Sin embargo, su diferenciación exacta con las hormonas sigue siendo un tema de debate (O'Shea et al., 2019). El tamaño de las citocinas es normalmente del orden de 5-30  $KDa$  (Thomson and Lotze, 2003), aunque hay algunas excepciones como la fractalquina que puede medir hasta 80 $KDa$ . Son producidas en su mayoría por macrófagos y linfocitos pero también por células endoteliales, fibroblastos, mastocitos y células estromales.

La necesidad de inmunoensayos con un amplio rango dinámico se acentúa cuando se quiere estudiar la señalización inmunológica, pues las concentraciones de estas proteínas en el cuerpo humano varían desde los  $fg/mL$  hasta los  $\mu g/mL$  o alrededor de 10 ordenes de magnitud. Las concentraciones bajas son debido a su uso para señalización local, pero también pueden causar su propia producción, lo que puede llevar a que su cantidad crezca de manera exponencial, incluso causando 'tormentas de citocinas' las cuales pueden ser muy dañinas para los tejidos.

Las citocinas transmiten señales a las células al acoplarse a complejos de proteínas llamados receptores de citocinas, los cuales están compuestos de varias proteínas que ocupan las regiones extracelular, citosólica (interior de la célula) e intermembranal. Una



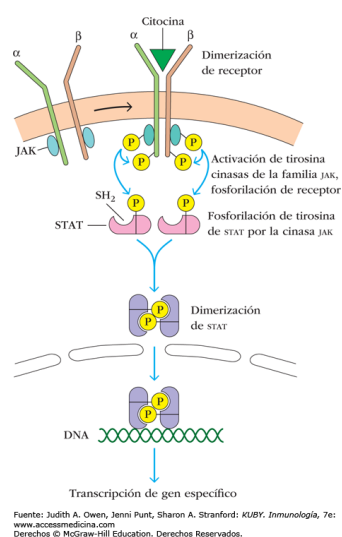
**Figura 2.11:** Citocinas y STAT (*Signal Transducer and Activator of Transcription*) involucrados en la diferenciación de la célula T.

Fuente: O'Shea, J. J., Gadina, M., & Siegel, R. M. (2019). Cytokines and Cytokine Receptors. In *Clinical Immunology* (p. 127–155.e1)

vez que una citocina se liga a un receptor se detona una cascada de señalización que llega hasta el núcleo. Los primeros actores son proteínas que reaccionan con los receptores, por ej. las janus quinasas **JAK**'s, las cuales, a su vez, son detectadas por complejos de proteínas que transducen al núcleo y funcionan como factores de transcripción, como por ej. los **STAT**'s (*Signal Transducer and Activator of Transcription*). En la Fig.2.12 se puede ver un esquema de la cascada de transducción.

En general, la actividad de las citocinas en las células se puede atribuir a alteraciones de la expresión genética. No obstante, la manera en que actúan es muy variada y el conjunto de formas en las que ellas que actúan ha sido llamada la 'filosofía molecular de la acción de citocinas' (Thomson and Lotze, 2003, p. 12). Esto es porque las células siempre están expuestas a una gran variedad de citocinas *in vivo*, las cuales arman un red de interacciones de diferente tipo. Además de la **pleiotropia** y la **redundancia** mencionadas arriba, las citocinas pueden tener un efecto **sinérgico** o **antagónico** sobre sus funciones, lo que quiere decir que una acción de cierta citocina puede amplificarse si otra citocina está presente o que cierta función solo se ejecute si dos citocinas trabajan al unísono, pero también que un acción que ejecuta una citocina se anule al estar presente otra. Si estas interacciones suceden al mismo tiempo, también puede haber interacciones con retraso en el tiempo que pueden ser denominadas como **cascadas de citocinas** donde una citocina aumenta o disminuye la producción de otra citocina. La intensidad de la acción de las citocinas también puede ser modulada por las mismas citocinas a través de la **transmodulación de los receptores**, que es que se ajuste la expresión de receptores para citocinas, pero también a través de la **transeñalización**

## 2. ANTECEDENTES



**Figura 2.12:** Esquema simplificado de la transducción de la señal de un receptor de citocina dimérico

Citocina	Función	Citocina	Función
IL-2	Proliferación, citotoxicidad, producción de IFN- $\gamma$	MCP-1	Quimioatáctica, atrae monocitos, T-memoria, dendritos
sIL-2Ra	Parte extracelular del receptor de IL2	IP-10	Quimioatáctica, atrae monocitos, macrófagos, T, dendritos, NK
IL-4	Proliferación, diferenciación Th2	IFN- $\gamma$	Activación, sube presentación de antígenos
IL-12p40	Segmento de IL12 que bloquea receptores	MIP-1 $\alpha$	Activan granulocitos, síntesis de citocinas proinflamatorias
IL-10	Inhibidora de síntesis de citocinas	MIP-1 $\beta$	
IL-1Ra	Antagonista de IL-1	Eotaxin	Quimioattractoras de eosinófilos
IL-8	Reacción inmunitaria, quimiotáctica para neutrófilos	VEGF	Facto de crecimiento vascular endotelial, angiogenico
IL-12p70	Diferenciación, proliferación, quimiotaxis	IL-17	Inflamación en el endotelio

**Tabla 2.2:** Funciones más comunes de las citocinas usadas en el caso de estudio

**de receptores**, donde se ajusta la intensidad con la que un receptor transmite la señal al interior de la célula.

Debido a esta acción en conjunto, los ensayos multiplex son muy útiles para investigar estados fenotípicos, pues se pueden deducir perfiles funcionales, por ej. un perfil de acción proinflamatorio cuando hay presencia de varias citocinas asociadas a esa función y pocas citocinas antiinflamatorias.

Las citocinas que se escogieron para este análisis cubren un amplio rango de funciones, facilitando así la interpretación de varios posibles perfiles. En la siguiente tabla se puede ver un resumen de sus funciones mas comunes.

# Objetivo e Hipótesis

---

## Planteamiento del problema

En las ciencias biomédicas se descubren cada vez más biomarcadores y técnicas para medirlos que aumentan en eficacia. Una de éstas es la técnica de ensayos *multiplex* que permite medir varios analitos en un mismo sustrato, además de hacer muchas mediciones a la vez. Con la cantidad de mediciones aumenta también el número de datos censurados y estos llegan a constituir un porcentaje considerable de las observaciones. Los datos censurados ofrecen información cualitativa de la medición a pesar de no dar un valor concreto. Por eso, considerarlos en un análisis puede hacer la diferencia entre encontrar una diferencia en poblaciones o una correlación y no hacerlo. En este estudio, la censura viene de un límite de detección. Con ayuda de la ciencias físicas se puede entender porque surgen estos límites de detección y usar ese conocimiento para apoyar el análisis de los datos. Además, con ayuda de métodos estadísticos, se puede recuperar la información cualitativa que brindan los datos censurados.

## Objetivo

El objetivo de esta tesis es explicar la base teórica de los valores debajo del límite de detección en inmunoensayos luminiscentes del tipo *multiplex* y evaluar técnicas estadísticas que permitan rescatar esos valores perdidos para su análisis.

## Objetivos particulares

1. Definir las bases físicas del límite de detección de ensayos *multiplex* luminiscentes considerando la base del principio de funcionamiento interno del instrumento y el manejo estadístico de los resultados.

### 3. OBJETIVO E HIPÓTESIS

---

2. Investigar maneras de evitar que los datos con valores muy pequeños o muy altos sean censurados o de reducir la incertidumbre de estos datos.
3. Buscar formas de integrar la información cuantitativa de los valores censurados en diferentes modelos de análisis estadístico.
4. Probar la eficiencia de los métodos investigados con datos de una población de mujeres a las que se le midió la concentración de 20 citocinas.

### Hipótesis

Con ayuda de un conocimiento teórico que explique los límites de detección en los ensayos *multiplex* luminiscentes y con técnicas estadísticas que integren los datos censurados, se puede obtener nueva información sobre los datos adquiridos en estos ensayos.

## Caso de estudio

---

El estudio se hizo en una cohorte de 946 mujeres, 496 de las cuales permanecieron después de aplicar los criterios de eliminación. 9.6% de ellas tuvo parto pretérmino (Nacimiento con menos de 37 semanas o 259 días cumplidos). Se les extrajo exudado cervicovaginal, en máximo 8 visitas que fueron de las 3 a las 42 semanas después de la última menstruación. Por supuesto, el proyecto fue aprobado por el Comité de Ética e Investigación del INMEGEN y los respectivos hospitales.

### **Criterios de inclusión**

Las muestras de exudado cervicovaginal se obtuvieron de pacientes embarazadas, a las cuales se les consultó su última fecha de menstruación. Se incluyeron todas las pacientes que al momento de la toma de la muestra tuvieran edad gestacional entre 3 y 42 semanas. Además, se solicitó por escrito su aceptación a participar en el protocolo de investigación.

### **Criterios de exclusión y eliminación**

Se excluyeron las pacientes con demostración clínica de infección cervicovaginal y se eliminaron las que resultaron positivas a infección luego de un cultivo microbiológico del exudado cervicovaginal. Las pacientes que refirieron haber tenido relaciones sexuales en las 24 h previas también se excluyeron. Además, se excluyeron las que tenían complicaciones médicas u obstétricas mayores, como diabetes mellitus, diabetes gestacional o infección genitourinaria. Finalmente, se eliminaron pacientes con menos de 20 semanas de gestación.



### Toma de las muestras

Las muestras se tomaron con un hisopo de dacrón estéril y un tubo de transporte con 1 mL de amortiguador (*Tris-base 0.05M, NaCl 0.15M, BSA 1%, Tween-20 al 0.1%*). El hisopo se frotó en el fondo del saco de Douglas, auxiliándose con un espejo vaginal lubricado con solución salina, e inmediatamente después se sumergió el hisopo en la solución amortiguadora de extracción, durante el transporte se almacenó a 4°C. El extracto del tubo se centrifugó a 2000 *rpm \* g*, a 4°C durante 15 minutos, para precipitar el material particulado y se colectó el sobrenadante; la muestra se conservó a -20°C, hasta su análisis.

### Medición y captura de datos

Se capturaron datos de las mujeres, como la fecha de la última menstruación (la cual se tomó como fecha de inicio de embarazo), su elegibilidad por edad, la fecha de nacimiento del bebé, el peso del bebé al nacer y su género en un tabla en Excel. A cada una de las pacientes se les asignó un número de identificación. Además se estuvo comprobando si se cumplía alguno de los siguientes 14 diagnósticos:

1. Término normal. Nacimiento con 37 semanas o 259 días cumplidos, sin complicaciones médicas.
2. Pretérmino. Nacimiento con menos de 37 semanas o 259 días cumplidos.
3. Óbito. Feto de más de 12 semanas de gestación muerto antes de su expulsión o extracción.
4. Aborto. Expulsión del producto de la gestación durante las primeras 20 semanas de gestación.
5. Huevo muerto retenido. Producto de la gestación muerto de menos de 12 semanas.
6. Muerte fetal temprana. Producto de la gestación muerto entre 12 y 19 semanas de gestación.
7. Baja DG. Pacientes en las que se establece el diagnóstico de Diabetes gestacional. Se utilizaron los criterios sugeridos por la American Diabetes Association ([American Diabetes Association, 2003](#)).
8. RCIU. Peso fetal por debajo del límite inferior de la curva de peso (percentila 10) en función de la edad gestacional. ([Pittaluga P. et al., 2002](#))
9. Miomas. Tumores uterinos benignos de origen muscular.

10. RPM. Ruptura prematura de membranas. Salida espontánea de líquido amniótico por una solución de continuidad en ausencia de contracciones uterinas y cambios cervicales. Puede ser de término o pretérmino.
11. DPP. Desprendimiento prematuro de placenta. Separación parcial o total de la placenta normalmente insertada en la cavidad uterina después de la semana 20 de gestación y antes del tercer período del trabajo de parto.
12. Defecto congénito. Grupo heterogéneo de enfermedades con cromosomopatía o malformaciones congénitas.
13. Falla de trabajo de parto. Trabajo de parto que no progresa al tercer período.
14. Oligohidramnios. Disminución del volumen de líquido amniótico corregido por edad gestacional.

A cada mujer se le trataron de medir las siguientes citocinas por medio de inmunoensayos *multiplex*: IL-1 $\alpha$ , IL-1 $\beta$ , IL-6, TNF- $\alpha$ , IFN- $\gamma$ , IL-17, IL-12p40, IL-12p70, IL-2, IL-8, IL-4, IL-10, IL-1Ra, sIL-2Ra, IP-10, MCP-1, MIP-1 $\alpha$ , MIP-1 $\beta$ , Eotaxin, VEGF.

Para medir estas concentraciones, se usó la plataforma *MAGPIX*<sup>®</sup> con software *xPONENT*<sup>®</sup> y diferentes placas con citocinas preparadas, como la *HYCTOMAG-60K* de *MILLIPLEX*<sup>®</sup> con 21 citocinas.



**Figura 4.1:** Imágen del equipo para inmunoensayos de citometría de flujo y perlas magnéticas *MAGPIX*<sup>®</sup> de *Luminex*<sup>®</sup>.

La plataforma *MAGPIX*<sup>®</sup> ejecuta inmunoensayos luminiscentes en modalidad *multiplex* con perlas magnéticas.

El manual de las perlas magnéticas (*HCYTOMAG-60K*) contiene las concentraciones mínimas que se pueden detectar para cada analito y la concentración más dos desviaciones estándar. La estimación de estos valores se hizo por medio del software de análisis de inmunoensayos *MILLIPLEX*<sup>®</sup> *Analyst 5.1*. Un programa que usa la variación en los estándares de varios ensayos con las mismas condiciones, para simular

#### 4. CASO DE ESTUDIO

---

muchas curvas estándar y así obtener un perfil de precisión para las concentraciones. De este perfil se infiere la mínima concentración detectable.

Usando estos datos, se puede ajustar la tabla que regresa el software Luminex para sólo reportar valores que no estén fuera del umbral de detección.

Además, se fueron recolectando los resultados de los inmunoensayos en una base de datos en Excel en la cual se ingresó el número de la paciente, el número de visita de la paciente, la fecha en la que se tomó la muestra y si fue término o pretérmino. Usando la fecha de último período de menstruación, se infirieron los días de gestación (DG) y las semanas de gestación (SDG).

La considerable cantidad de datos procesados invita a que se cometan muchos errores. Es por eso que se diseñaron programas en R para confirmar la limpieza de los datos. Se programó una rutina para la base de datos de los exudados y una para la de las pacientes. Ambos aseguraban que el formato de la fecha fuera el correcto y generaban un reporte escrito de los errores encontrados. Además, la rutina para los exudados validó que el formato de la entrada que indicaba el número de paciente y la visita fueran correctos, que los DG y los SDG no superaran los 420 días y que tuvieran un mismo formato, que el número de visita fuera congruente con la fecha de medición y que los niveles de concentración de citocina tuvieran el formato correcto.

Una vez importadas y limpiadas estas bases de datos, se pudo aplicar un algoritmo que calculara los DG (días de gestación) usando el último período menstrual y que confirmara que todas las pacientes en la tabla de exudados se encontraran en la de información del paciente.

# Herramientas estadísticas

---

La práctica más común es no reportar los valores censurados. Una decisión que implica perder la información cuantitativa que ofrecen los datos, pues se sabe que están debajo, o por encima, de cierto límite. Es por esto que se hizo una búsqueda de métodos estadísticos que permitieran integrar la información cuantitativa que tienen los datos censurados. En particular, se escogieron métodos que se adaptaran a análisis estadísticos que predigan una variable categórica binaria o que comparen poblaciones.

Además, la mayoría de los análisis multivariados requieren de observaciones con todas las variables para considerarlas en el análisis. Por lo que también se evaluaron técnicas para insertar los valores perdidos por otras causas.

## Panorama

A continuación se presenta un resumen de todos los métodos que se consideraron, seguido por un ahondamiento de los que se implementaron.

La manera más simple y directa de considerar los valores bajo el umbral, es imputando el límite de detección ( $LOD$ ) entre 2 o entre  $\sqrt{2}$ . La suposiciones en el caso de  $LOD/2$  son que la distribución de los  $LOD$  es cuadrada con promedio y media  $LOD/2$  (una distribución altamente oblicua con desviación geométrica estándar mayor a 3). Cuando se sustituye por  $LOD/\sqrt{2}$ , la distribución es triangular y tiene media  $LOD/\sqrt{2}$ . La inconveniencia de esta técnica es que muchas veces afecta la varianza de la distribución de los datos a una no realista.

El primer modelo considerado fue el Tobit. Desarrollado por James Tobin ([Tobin, 1958](#)), como una modificación de la distribución logit, hace uso de regresión logística para probar una hipótesis y trunca la variable dependiente para los datos sin medición. Sin embargo, solo sirve para cuando hay una acumulación de valores que tiende al límite de detección. No es el caso de los datos obtenidos, pues estos pueden estar muy fuera del rango y lejos del límite de detección. Además, la variable censurada en nuestro caso es la independiente.

Otra técnica considerada fue la de Kaplan-Meier ([Kaplan and Meier, 1958](#)), la

cual permite construir una función de distribución acumulada (CDF) usando datos censurados. El problema de este método es que asume que los datos censurados están distribuidos aleatoriamente a través de todo el rango. Los datos censurados en los extremos contribuyen lo mismo que en una CDF sin ajustar.

Si la distribución empírica se parece a una distribución conocida, se pueden simular los datos faltantes. Sin embargo, ajustar los datos a una distribución sin truncar haría que las colas sean más cortas de lo que en realidad son. Tal dificultad se puede sobrellevar ajustando los datos a una distribución truncada, lo cual se hace con ligeros cambios a la fórmula matemática de la distribución y utilizando algoritmos de optimización para ajustarse a la nueva fórmula. Este método se implementó en R y en el capítulo de resultados se puede ver su eficacia.

Debido a la naturaleza longitudinal de los datos y a las acentuadas variaciones en las concentraciones de citocinas, es difícil ajustar las poblaciones a una distribución conocida. En estos casos se pueden implementar métodos no paramétricos, en particular los derivados de los rangos logarítmicos. El uso de rangos permite que los valores censurados se asignen a un rango empatado, manteniendo así la información cualitativa que brindan. De esta manera, se pueden obtener diferencias entre poblaciones que toman en cuenta los valores censurados. En este estudio se usó el estadístico Mann-Whitney para encontrar los periodos en los que las citocinas tuvieran mayores diferencias entre las distribuciones de poblaciones de término y pretérmino.

Por otra parte, en los datos con los que se trabajó existen muchos datos perdidos. La alta cantidad de datos permite que se pueda usar uno de los algoritmos de *machine learning* con mayor poder de predicción, pero poca información sobre la estructura de los datos, el algoritmo de bosques aleatorios (*Random Forest*). Con ayuda de este algoritmo se imputaron los datos faltantes y no se perdieron los casos en los que algunas de las variables tenían valores perdidos.

Una posible alternativa a este método que es muy comúnmente usada en la bioinformática (por ej. en (Shah et al., 2017)) es la imputación por vecinos más cercanos (*k-Nearest Neighbours*).

## Comparación de distribuciones con la prueba de Mann-Whitney

Una manera muy rudimentaria de considerar los valores bajo el umbral de detección en un análisis estadístico es con el uso de comparación de distribuciones por métodos de rangos no paramétricos. Este acercamiento tiene la ventaja de poder usarse en una gran variedad de datos, pues no se restringe a una familia de funciones parametrizadas. Sin embargo, se limita a solo comparar distribuciones, lo que excluye hacer predicciones o buscar correlaciones.

La naturaleza misma de los métodos de rango integra los valores debajo de un umbral. Pues estos se acumulan en un solo rango, eliminando así las diferencias en el

tipo de variable con el que se trabaja. Es importante notar, sin embargo, que cuando hay muchos valores censurados se tendrán muchos empates por lo que cambian las premisas del método que se usa para comparar.

Para detallar el tratamiento de estas modificaciones, se explica la teoría detrás de la prueba de Mann-Whitney (MW).

### El estadístico de Mann-Whitney

Sean  $x_i \in X$  con  $i \in 1, \dots, n$  y  $y_j \in Y$  con  $j \in 1, \dots, m$  las dos poblaciones que se quieren comparar. Asígnese primero que ninguno de los valores se repite. En este caso, se le puede asignar a cada observación un rango, dándole el valor de 1 a la menor de todas, 2 a la siguiente menor y así sucesivamente. El estadístico de Wilcoxon de rangos sumados,  $W$ , es la suma de los rangos que están en el grupo con menos elementos. Sin embargo, se cumple que si se define a  $W_X$  o  $W_Y$  como la suma de los rangos en el grupo  $X$  o  $Y$  entonces

$$W_X + W_Y = \frac{(n+m)(n+m+1)}{2}$$

Por lo que cualquier resultado para alguno de ellos, se puede calcular del otro.

El acercamiento que se toma para ver si las poblaciones difieren, es el de rechazar la hipótesis nula, que dice que la distribución de las poblaciones es la misma. Si las distribuciones fueran iguales, la distribución de los valores que puede tomar  $W$  por pura aleatoriedad, bajo la condición de que los grupos tengan los tamaños  $n, m$ , se puede construir desde las posibilidades que hay de tomar  $n$  valores de  $n+m = N$  posibles valores o  $\binom{n+m}{m}$ . Esta es la distribución nula  $F_W(x|n, m)$ .

El estadístico de Mann-Whitney  $U$  simplifica la obtención con permutaciones exactas de la distribución nula. En lugar de considerar la suma de los rangos, se define a

$U_{XY}$  = El número de parejas tales que  $x_i < y_j$

$U_{YX}$  = El número de parejas tales que  $y_i < x_j$

Sin presencia de empates, se cumple que  $U_{XY} + U_{YX} = n \times m$ , por lo que se puede escoger cualquiera de los dos como estadístico  $U$ . La distribución  $F_U(x|n, m)$  es igual a la de  $W$  gracias a que

$$U_{XY} = W_Y - \frac{n(n+1)}{2}$$

$$U_{YX} = W_X - \frac{m(m+1)}{2}$$

El cálculo de  $F_U(x|n, m)$  se facilita considerando relaciones de recurrencia. Sea  $Fr_U(u|n, m)$  la frecuencia con la que una  $X$  precede a un  $Y$  un número  $u$  de veces, entonces se cumple

$$Fr_U(u|n, m) = Fr_U(u-m|n-1, m) + Fr_U(u|n, m-1)$$

## 5. HERRAMIENTAS ESTADÍSTICAS

---

Considerando que, bajo la hipótesis nula, hay  $\binom{N}{n}$  posibles maneras de asignar los rangos a las poblaciones, la distribución de  $U$  es la siguiente

$$F_U(x|n, m) = \binom{N}{n} Fr_U(x|n, m)$$

Usando esta fórmula y la Ec. 5.2.1, se puede ver que el número de maneras en las que se puede obtener la suma de rangos  $w$  es:

$$Fr_W(w|n, m) = Fr_W(w - m - n|n - 1, m) + Fr_W(w|n, m - 1)$$

Así se puede obtener  $F_W$  por recurrencia, que es más eficiente que usar tablas.

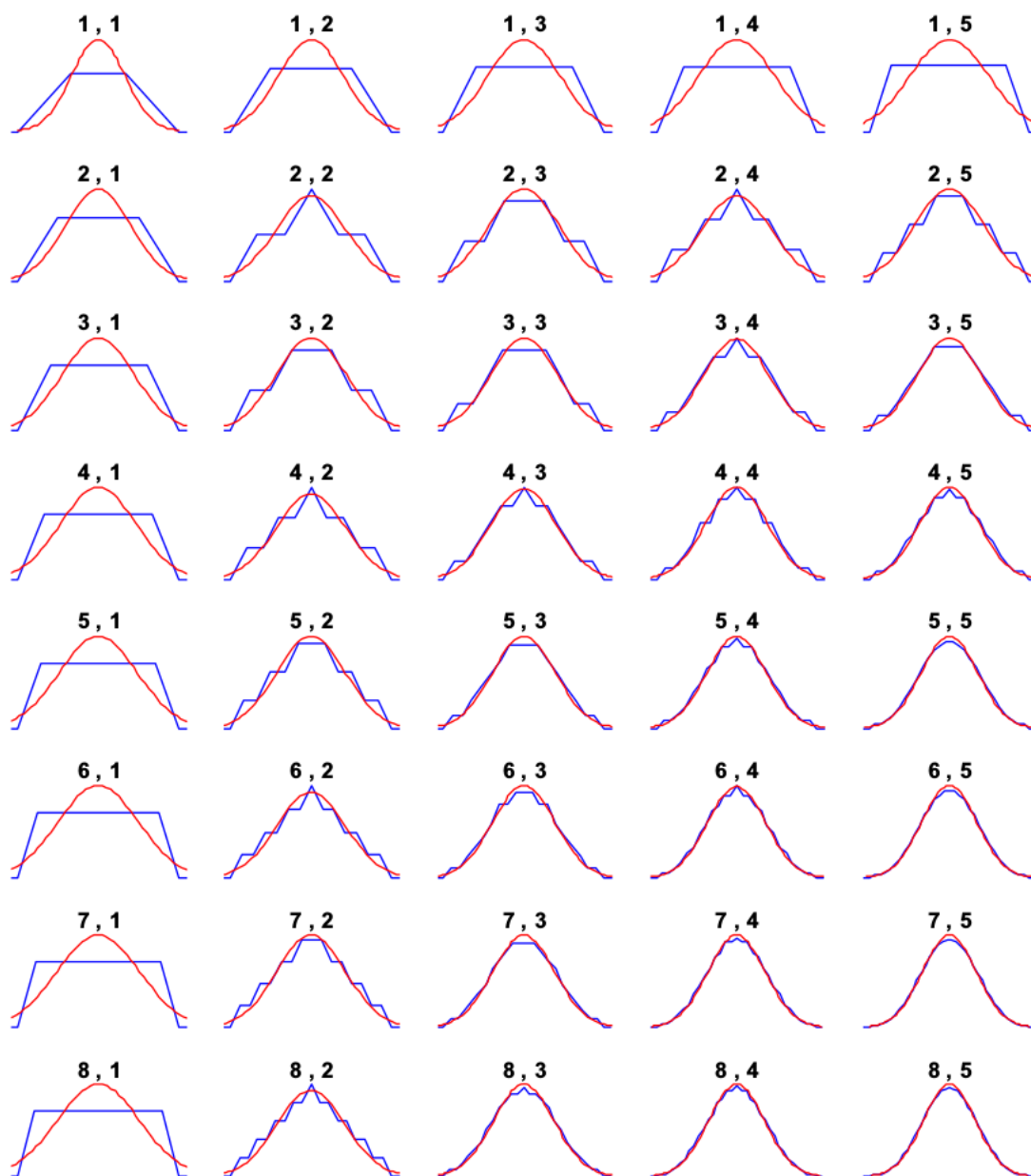
Cuando las poblaciones son muy grandes la recurrencia es ineficiente. Por eso se aproxima con una distribución a la que se tiende asintóticamente. Para poder asegurar que la aproximación escogida es la adecuada, se deben cumplir dos condiciones. Primero, se comprueba, calculando los valores exactos, que se tiende a la distribución límite. Segundo, se prueba un teorema de convergencia para cuando el tamaño de la muestra tiende a infinito. Esto solo sugiere que la distribución del estadístico se aproxima a la teórica, para grandes  $n$ , pero no lo prueba, pues el teorema de límite afirma que sucede en algún momento, no dice cuando.

Para poder hacer la aproximación a la normal se debe de calcular el valor esperado y la varianza que adquiere la suma de rangos por su naturaleza. Se puede ver que son las siguientes:

$$E(W) = N(N + 1)/2$$
$$\text{Var}(W) = \frac{n(N - n)(N + 1)}{12}$$

Con ayuda de las relaciones de recurrencia se puede calcular el valor exacto de la distribución y se le puede comparar con la aproximación. Se puede ver en la Figura 5.1 que la distribución exacta tiende rápidamente a la normal. Hay algunos detalles como que, si el tamaño de uno de los grupos es muy pequeño, la distribución solo puede tomar pocos valores. Otro aspecto a considerar es que en las colas la aproximación tiende a ser más imprecisa, es por eso que se hace una corrección de continuidad integrando hasta uno de los posibles valores de  $W$ , para aproximarse más al área del histograma discreto con el que se calcula la probabilidad exacta.

Una de las versiones más usadas del teorema de límite central dice que la distribución del promedio de  $n$  variables independientes y de idéntica distribución tiende a una normal cuando  $n \rightarrow \infty$ . Debido a que el estadístico es una suma de rangos, tiene sentido aplicar el teorema a este estadístico, pero se debe de usar una versión más general. En este caso, la distribución de las variables no tiene que ser idéntica. Además, se suman los rangos de sólo una de las dos poblaciones. Lo cual es como escoger aleatoriamente con reemplazo, un procedimiento que hace que las variables sean dependientes. A continuación, se presenta una traza del teorema necesario como es presentado en (Lehmann, 1975, p.64).



**Figura 5.1:** Aproximaciones a la normal para diferentes tamaños de población  $N$  y de grupos  $n$  y  $m$

Fuente: [Bellera et al. \(2010\)](#)



**Teorema 5.1** Sean  $\Pi_1, \dots, \Pi_N$  una serie de poblaciones tal que  $\Pi_i$  tiene  $i$  elementos. Llámese  $v_{Ni}$  a estos elementos y sea  $v_n$  el promedio de los valores de la población. Tómese una muestra de cada una de estas poblaciones, tal que la que proviene de la población  $\Pi_N$  tenga tamaño  $n$  y sus elementos sean  $A_{N1}, \dots, A_{Nn}$ . Además, sea

$$S_N = \sum_{i=1}^n A_{Ni}$$

Entonces  $S_N^2 = \frac{(S_N - E(S_N))}{\sqrt{\text{Var}(S_N)}}$  es asintóticamente normal a  $\mathcal{N}(0, 1)$  si se cumple que

$$\lim_{N \rightarrow \infty} \frac{\max(v_{Ni} - v_N)^2}{\sum_{i=1}^N (v_{Ni} - v_N)^2} = 0$$

En el caso de MW sin empates la población son los rangos, es decir los valores de 1 hasta  $N$ , y es fácil ver que se cumple la desigualdad. Nótese que el teorema considera que la suma esta compuesta de una muestra de la población, lo que da la dependencia a esos datos. Esta dependencia se puede superar en una prueba, si se define una población hecha de variables aleatorias que toman un valor de la población  $\Pi_N$  con cierta probabilidad y 0 en otro caso. De tal manera que, al hacer la suma, se obtiene lo mismo que cuando se escogió la muestra con reemplazo. Usando esa población se puede usar el teorema para variables independientes, para demostrar que el estadístico  $W$  es asintóticamente normal.

### Ajustes para empates

El método más común para romper los empates es el del rango medio. Consiste en asignar a todos los rangos empatados el promedio de los rangos que hubieran tenido si no estuvieran empatados. Esto hace que cambie la distribución de  $W$  a una a la que se le asignará el nombre  $W^*$ . La mejor manera para compensar por este cambio es rehacer el cálculo de la distribución de manera exacta, pero hay demasiadas combinaciones para muestras grandes, por eso también se hicieron ajustes a la aproximación normal. En este caso, el valor esperado no cambia y la varianza se ajusta un poco dejando los siguientes parámetros para la normal:

$$E(W) = N(N + 1)/2$$

$$\text{Var}(W) = \frac{mn(N + 1)}{12} - \frac{mn \sum_{i=1}^e (d_i^3 - d_i)}{12N(N - 1)}$$

Donde  $d_i$  es el número de veces que se repite el rango  $i$ , siendo  $e$  la cantidad de rangos únicos. Usando el Teorema 5.1, se puede llegar a un criterio para cuando la distribución converge a una normal que considera empates:

$$\exists a < 1 \text{ tal que } \max\left(\frac{d_i}{N}\right) < a \text{ cuando } N \rightarrow \infty$$

Esto quiere decir que mientras no haya demasiados valores iguales, la suma convergerá.

### Comparación con otros métodos

Las razones por las que se escoge este método son:

1. Las poblaciones son de diferentes tamaños así que se tienen que usar pruebas de dos muestras, no pareadas. Lo cual excluye todos los métodos de una muestra como el Wilcoxon de rangos con signo, el Tarone-Ware, Peto y Flemington-Harrington (Kleinbaum and Klein, 2005).
2. El estadístico de Baumgartner-Weiss-Schindler (BWS) da mejores resultados, pues ubica diferencias tanto en varianza, como promedios y kurtosis (Baumgartner et al., 1998). Sin embargo, el error tipo I aumenta con esta prueba si se tienen muchos empates (Neuhauser, 2003). Se ha propuesto una alternativa usando permutación exacta, pero el número de muestras es muy grande para calcular todas las posibilidades. También existen métodos para no tener que considerar todas las posibilidades pero su implementación no es fácil y, en muchos casos, no hay pruebas formales de su poder estadístico.
3. Los datos que se analizaron no cumplen las condiciones para que la aproximación normal pueda ser usada.

### Simulaciones

Para justipreciar la efectividad de este tipo de pruebas con datos de la naturaleza de los que serán analizados aquí, se corrieron algunas simulaciones con la paquetería “coin” de R, la cual usa el ajuste mencionado en la sección anterior para la prueba de Mann-Whitney. Primero, se buscó la tasa de error tipo I. La cual, en este caso, se refiere a cuando no se debería rechazar la hipótesis nula (**H0**) (e.d. que las distribuciones son la misma) pero si se rechazó. Se hicieron N simulaciones y se calculó la tasa de error entre todas ellas. Esto se logró generando datos de una misma distribución gamma (con parámetros: forma = 2, escala = 2) o normal (promedio = 0, desv = 1), pero una de ellas con cierto porcentaje de los datos mínimos igualados al máximo de los mínimos. Segundo, la tasa de error tipo II, que es cuando se debería de rechazar H0 (e.d. que las distribuciones son diferentes) pero no se hace. Esto, simulando dos muestras de datos de distribución gamma o normal, pero con un promedio una unidad menor, la varianza igual y con cierto porcentaje de valores censurado (Gamma: Dist 1: forma = 2, escala = 2; Dist 2: forma = 9/8, escala = 8/3)(Normal: Dist 1: promedio = 0, desv = 1; Dist

## 5. HERRAMIENTAS ESTADÍSTICAS

---

	N	TypeI_norm	TypeII_norm	TypeI_gamma	TypeII_gamma
1	10	0.06	0.44	0.05	0.83
2	50	0.04	0	0.07	0.41
3	100	0.07	0	0.05	0.09
4	300	0.07	0	0.05	0
5	500	0.06	0	0.07	0
6	1000	0.04	0	0.04	0
7	1500	0.08	0	0.06	0

(a) 0% de datos censurados

	N	TypeI_norm	TypeII_norm	TypeI_gamma	TypeII_gamma
1	10	0.06	0.44	0.06	0.77
2	50	0.05	0	0.04	0.35
3	100	0.06	0	0.07	0.13
4	300	0.14	0	0.15	0
5	500	0.23	0	0.27	0
6	1000	0.47	0	0.43	0
7	1500	0.65	0	0.65	0

(b) 5% de datos censurados

	N	TypeI_norm	TypeII_norm	TypeI_gamma	TypeII_gamma
1	10	0.06	0.44	0.05	0.83
2	50	0.08	0.01	0.1	0.41
3	100	0.13	0	0.17	0.1
4	300	0.52	0	0.5	0
5	500	0.77	0	0.79	0
6	1000	0.97	0	0.98	0
7	1500	0.99	0	1	0

(c) 10% de datos censurados

	N	TypeI_norm	TypeII_norm	TypeI_gamma	TypeII_gamma
1	10	0.06	0.44	0.05	0.83
2	50	0.15	0.01	0.15	0.43
3	100	0.28	0	0.36	0.11
4	300	0.89	0	0.9	0
5	500	1	0	0.98	0
6	1000	1	0	1	0
7	1500	1	0	1	0

(d) 15% de datos censurados

**Tabla 5.1:** tasa de errores tipo I para simulaciones de la misma distribución, una sin datos censurados y otra con diferente porcentaje de datos censurados. Además, tasa de errores tipo II, en simulaciones de dos muestras con porcentaje de datos variados, trasladando el promedio de una de las dos por una unidad y manteniendo la varianza igual.

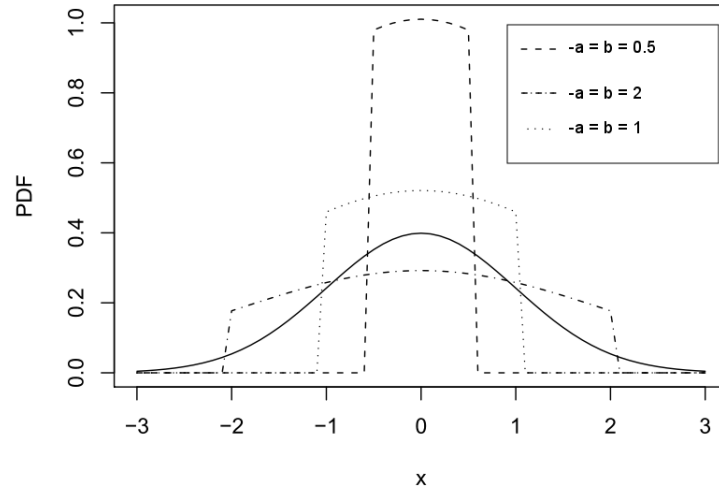
2: promedio = 1, desv = 1). Los resultados se pueden ver en las Tablas 5.1.

Como se puede notar, la tasa de error tipo I es consistente con la el criterio de  $p < 0.05$  cuando no se censuraron los datos.

Sin embargo, cuando el número de muestras supera los 300 y el 15% de los datos está censurado, la tasa de error tipo I se hace 1. Es decir, en todos los casos se estima que las distribuciones son diferentes a pesar de haber sido generados de la misma distribución. Sin embargo, los datos debajo del límite podrían ser otros y no se debería de esperar que esta prueba adivine la tendencia de los datos no censurados. En consecuencia, lo único que se puede afirmar en el caso de que esta prueba no estime una diferencia significativa cuando no se consideran datos bajo el umbral y luego, al considerarlos si la estime, es que hay una cantidad considerable de datos bajo el límite de detección. Por eso es importante complementar esta prueba con una imputación, si es posible.

Por otro lado, los errores del tipo II disminuyen con el crecimiento del número de muestras. Esto es congruente con lo anterior y se puede concluir que si había un diferencia sin empates y, al considerar empates, dejó de haberla, en verdad hay una semejanza. Aunque es muy poco probable que esto pase.

Es por esto que, aunque se logre recuperar un poco de información con este método, se eligió complementarlo con la imputación con distribuciones pues mientras m[as datos se tenga y mas datos censurados haya, mayor será la probabilidad de que esta prueba falle, a pesar de considerar los valores bajo el umbral de detección.



**Figura 5.2:** Distribuciones normales truncadas en  $a, b$  con media 0 y desviación estándar 2.

Fuente: Nadarajah, S., Kotz, S. (2006). R Programs for Computing Truncated Distributions. Journal of Statistical Software

## Imputación con distribuciones truncadas

Cuando la distribución de los datos a analizar tiene un límite de detección, se presenta una interrupción abrupta en esta. Si se quiere ajustar una curva a estos datos, se deben de hacer ajustes a las fórmulas de densidad de probabilidad (*probability density function pdf*) y densidad acumulada (*cummulative density function cdf*), para después ajustarse a estas curvas usando (por ej.) el método de estimación por máxima verosimilitud (**MLE**).

Al truncar la distribución, esta se debe normalizar para recuperar el rango truncado. Esto hará que la curva se traslade en el eje de la probabilidad como en la Fig. 5.2.

Sean  $g(x)$  la pdf original y  $G(x)$  la cdf original, entonces se les puede truncar de la siguiente manera:

$$f(x) = \begin{cases} \frac{g(x)}{G(b)-G(a)}, & \text{si } a \leq x \leq b \\ 0, & \text{otro caso} \end{cases} \quad (5.1)$$

$$F(x) = \frac{G(\max(\min(x, b), a)) - G(a)}{G(b) - G(a)}$$

La distribución de probabilidad gamma se puede ver como una generalización de la distribución exponencial, la cual representa la distribución de los intervalos entre dos eventos en un proceso de Poisson homogéneo sobre la línea real (es decir estocástico). Un proceso de Poisson se puede ver como una distribución de puntos sobre la semirecta real positiva, donde cada punto representa un evento y en un intervalo  $[x, x + dx]$  hay en promedio  $\lambda \times dx$  puntos, magnitud a la que se llama intensidad. Además, la ocurrencia

de un evento es independiente de los otros. En un proceso gamma se mantiene la independencia pero no la homogeneidad, es decir la intensidad es ahora  $\frac{k}{x} \exp(-\theta x) dx$ . En esta caso la distribución marginal para un tiempo  $t$ , es decir la distribución de los intervalos entre eventos, es  $p_\gamma(t)$  y la acumulada es  $P_\gamma(t)$  con:

$$\begin{aligned} p_\gamma(t) &= \frac{1}{\Gamma(k)\theta^k} t^{k-1} e^{-t/\theta} \\ P_\gamma(x) &= \frac{1}{\Gamma(k)} \int_0^{t/\theta} t^{k-1} e^{-t} dt \end{aligned} \quad (5.2)$$

Donde  $k$  es el parámetro de forma y  $\theta$  es el de escala inversa o razón. Estos caracterizan la curva por medio del valor esperado  $E[X] = k\theta$  y la varianza  $Var(X) = k\sigma^2$ .

Se escogió esta distribución pues su cola derecha es más larga, en el sentido de que decrece aritméticamente en lugar de exponencialmente como la gaussiana.

Para hacer un ajuste por MLE, necesitamos la verosimilitud  $\mathcal{L}$  en logaritmo. Asumiendo que las muestras que tenemos cumplen que  $a \leq x_i \leq b$ :

$$\begin{aligned} \ln \mathcal{L} &= \sum_{i=1}^N \ln f(x_i | k, \sigma) \\ &= \sum_{i=1}^N \left[ \ln \left( \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}} \right) - \ln(P_\gamma(b) - P_\gamma(a)) \right] \\ &= \sum_{i=1}^N \left[ -\ln(\Gamma(k)) - k \ln(\sigma) - \ln(P_\gamma(b) - P_\gamma(a)) + (k-1) \ln(x_i) - x_i/\sigma \right] \\ &= -N(\ln(\Gamma(k)) + k \ln(\sigma) + \ln(P_\gamma(b) - P_\gamma(a))) + (k-1) \sum_{i=1}^N \ln(x_i) - x_i/\sigma \end{aligned}$$

Los parámetros que maximizan esta función son los mismos que maximizan la probabilidad de obtener los valores de la muestra. Usando la información sobre el promedio y la varianza se pueden proponer valores iniciales para una optimización. En particular, si  $\sigma = k/4$ , la densidad mantiene una forma de campana y usando el promedio de la muestra  $\bar{X}$ , se puede proponer que  $k = 2\sqrt{\bar{X}}$ .

Una vez obtenidos los parámetros se pueden producir valores en el intervalo  $(a, b)$ , primero generando valores aleatorios de una distribución uniforme en el intervalo  $(P_\gamma(a), P_\gamma(b))$  y luego evaluándolos en la función de cuantiles (o inversa de la función acumulada) sin truncar  $P_\gamma^{-1}(x)$ . Estos valores tienen la ventaja de mantener la varianza de la distribución a la que se ajustaron, a diferencia de cuando solo se imputa un mismo valor siempre.

## Imputación con random forest

Además de valores debajo del umbral, también hay una alta cantidad de valores perdidos por otras razones que no tienen la información cualitativa que tienen los LODs. Simplemente descartar observaciones con valores perdidos no suele ser una práctica razonable, ya que se puede perder información valiosa y comprometer el poder inferencial (Enders, 2010). Incluso se puede causar sesgo de selección en algunos casos. Además, la eliminación de observaciones con valores perdidos puede dar lugar a que queden pocas observaciones, cuando hay un gran número de variables predictivas que contienen valores perdidos, especialmente para datos de alta dimensión. Para ello, se han desarrollado muchos métodos de imputación estadística para los datos faltantes. Sin embargo, en configuraciones de datos de gran dimensión y gran escala, como problemas genómicos, de neuroimagen, o en nuestro caso de ensayos *multiplex*, muchos de estos métodos no funcionan.

Los algoritmos de bosques aleatorios (*random forests*) (RF) de Breiman (Breiman, 2001), son un enfoque atractivo para la inserción de datos faltantes para configuraciones de *big data*. Tienen las siguientes propiedades deseables: manejan tipos mixtos de datos faltantes; abordan las interacciones y la no-linealidad; se puede escalar a dimensiones muy altas; y se pueden configurar direcciones cuando hay más variables que observaciones. Fueron originalmente propuestos por Breiman y se implementaron por primera vez en el paquete de R de Random Forests (Liaw and Wiener, 2002).

Los árboles de decisión son la base de los Random Forests. Dependiendo de las variables que se quieren predecir, se pueden usar dos tipos de árboles. Los árboles de *clasificación* son para cuando la variable a predecir es nominal y los de *regresión* son para cuando el predictor es cuantitativo. A continuación se explicará el algoritmo usado en los árboles de regresión, pues en nuestro caso de estudio las variables perdidas son cuantitativas.

El algoritmo que se utilizó y que se puede encontrar en Breiman (2001) se puede resumir como sigue:

- 1 Dividimos el espacio predictor, esto es, el conjunto de posibles valores para  $X_1, \dots, X_p$  en  $J$  disjuntas regiones  $R_1, \dots, R_J$ .
- 2 Para cada observación que cae en la región  $R_j$ , hacemos la misma predicción, que es el promedio de las respuestas para las observaciones en la región  $R_j$ .

Por ejemplo, supóngase que en el Paso 1 se obtienen dos regiones:  $R_1, R_2$ , y el promedio de las respuestas sobre las observaciones en la primera región es 10, mientras que el promedio de respuesta en las observaciones que se encuentran en la segunda región es 20. Entonces, para una observación dada  $X = x$ , si  $x \in R_1$  vamos a predecir el valor de 10, y si  $x \in R_2$  vamos a predecir el valor de 20.

Para construir las regiones  $R_1, \dots, R_J$ , se construyen rectángulos de  $n$ -dimensiones o cajas (aunque la geometría de las regiones podría tener cualquier forma). El objetivo es encontrar los rectángulos  $R_1, \dots, R_J$  que minimizan la suma del error de predicción

(RSS):

$$\sum_{j=1}^J \sum_{i \in R_j} (y_i - \hat{y}_{R_j})^2,$$

donde  $\hat{y}_{R_j}$  es el promedio de las observaciones en la  $j$ -ésima caja.

No es computacionalmente factible el considerar cada posible partición de la región en  $J$  cajas. Por ello, se toma un enfoque *top-down*, que se llama *binary splitting recursivo*. Se inicia en la parte superior del árbol y sucesivamente se parte el espacio predictor; cada partición es indicada vía dos nuevas ramas del árbol. Supóngase que se quiere predecir con la variable  $X_j$  y que el punto de corte es  $s$ , entonces se parte la región  $R_1(j, s) := \{X|X_j < s\}$  y  $R_2(j, s) := \{X|X_j \geq s\}$ .

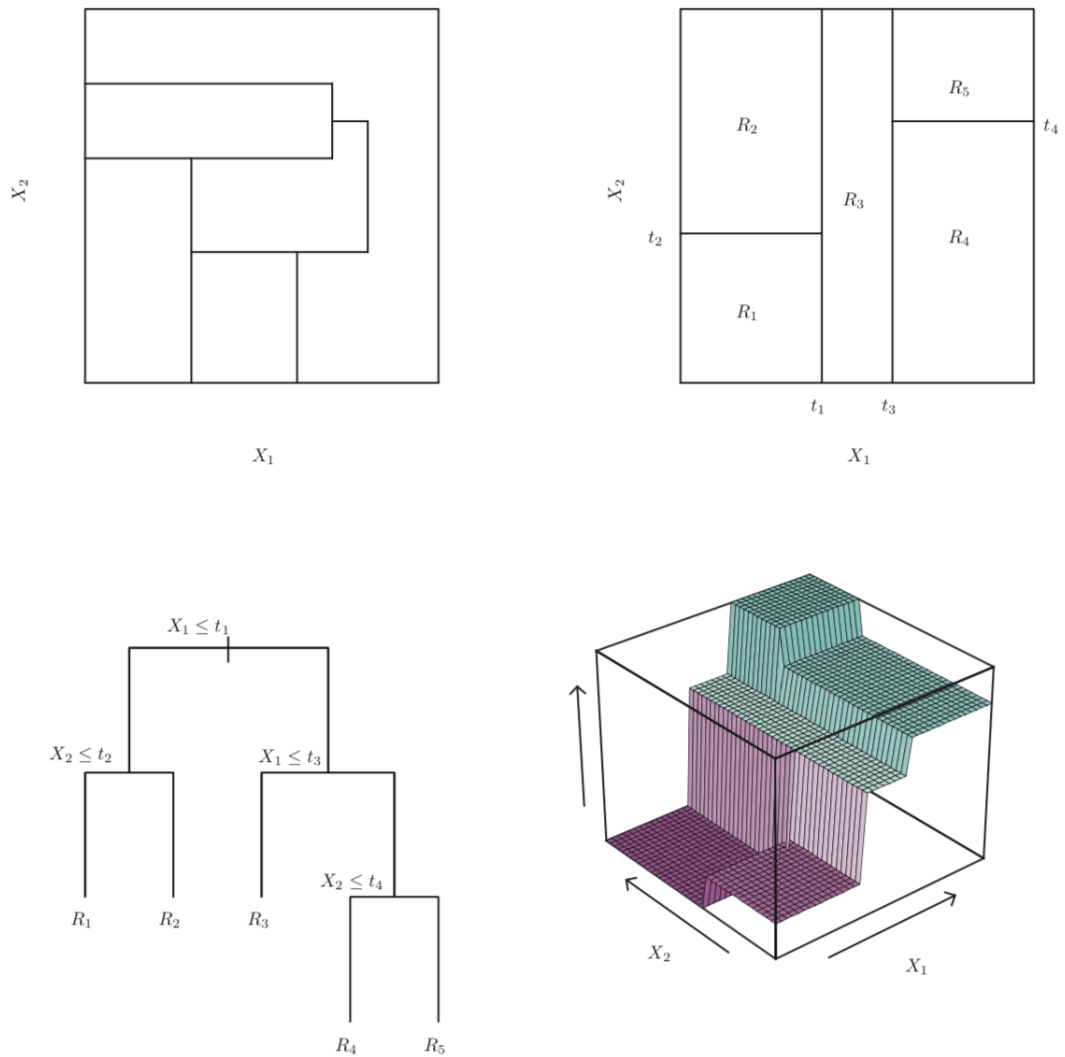
Normalmente, se consideran todos los predictores  $X_1, \dots, X_p$ , y todos los posibles valores de corte  $s$  para cada una de las  $X_j$ . Pero se elige el predictor y el punto de corte tal que el árbol resultante tiene el menor RSS. Esto es, se busca el  $j$  y la  $s$  que minimizan :

$$\sum_{i: x_i \in R_1(j, s)} (y_i - \hat{y}_{R_1})^2 + \sum_{i: x_i \in R_2(j, s)} (y_i - \hat{y}_{R_2})^2,$$

donde  $\hat{y}_{R_a}$  es el promedio de respuesta en  $R_a(j, s)$  para  $a = 1, 2$ .

La innovación de RF consiste en tomar muchos de éstos árboles, una técnica llamada *bagging*, y, en el caso de regresión, regresar el promedio de todas las predicciones. No obstante, si se tomara la decisión sobre todos los predictores en cada nodo y se permitiera que los árboles crecieran mucho, el algoritmo tomaría demasiado tiempo. Es por eso que se toma solo cierto número de nodos en cada decisión (normalmente esto se especifica como parámetro *mtry*), los cuales se escogen aleatoriamente y de manera independiente, pero con la misma distribución, en todos los árboles (*bootstrapping*). Breiman demostró que esto hace que el error de generalización, el cual depende del poder de cada árbol, converga a cierto valor. Gracias a esto se evita que la predicción se ajuste demasiado a los datos con los que se entrena el árbol, un fenómeno llamado *overfitting*.

Existen otros algoritmos que también usan árboles de poca profundidad y acumulan sus predicciones para hacer una mejor predicción. Uno de ellos se llama *Adaboost* y usa pesos (armados de RSS obtenidos anteriormente) para elegir cuales predictores usar en cada decisión. Sin embargo, Breiman demostró que su algoritmo se acerca al error de generalización de *Adaboost*, pero es más robusto en presencia de ruido.



**Figura 5.3:** En la parte superior izquierda vemos una partición que no podrá resultar de un *binary splitting* recursivo. A su derecha vemos el output de un ejemplo bidimensional para un *binary splitting* recursivo. En la parte inferior izquierda observamos el árbol que corresponde a la partición del *binary splitting*. A su derecha observamos una gráfica de la superficie de predicción correspondiente al árbol.

Fuente: James et al. (2013)





# Resultados y Discusión

## Análisis exploratorio

Una vez acatados los criterios de inclusión, exclusión, eliminación y después de haber hecho una limpieza con ayuda de los algoritmos, quedaron 946 pacientes de las cuales 497 tenían al menos una entrada en la tabla de niveles de concentración de citocinas.

	Citocina	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD
1	IL.1a	9.4	44.0	698.2	2248.6	2275.2	10000.0	3373.5
2	IL.1b	0.8	5.0	67.8	1635.4	1162.5	10000.0	3219.5
3	IL.6	0.9	3.6	12.4	457.9	43.5	10000.0	1910.0
4	TNFa	0.7	1.2	4.7	96.5	12.8	10000.0	443.0
5	INFg	0.8	1.4	4.0	78.7	17.3	4947.9	257.5
6	IL.17	0.7	0.7	2.3	268.3	8.4	10000.0	1496.2
7	IL.12.p40.	7.4	7.4	10.3	535.3	28.7	10000.0	2078.3
8	IL.12.p70.	0.6	0.9	2.9	69.9	11.9	2041.3	202.7
9	IL.2	1.0	1.4	4.5	56.2	17.3	3077.6	208.5
10	IL.8	0.6	366.3	1635.0	4006.8	10000.0	10000.0	4195.5
11	IL.4	4.5	4.5	8.1	302.7	63.0	10000.0	1164.0
12	IL.10	1.1	1.4	5.0	248.3	19.6	10000.0	1374.3
13	IL.1Ra	8.3	61.5	4701.2	4859.5	10000.0	10000.0	4153.0
14	sIL.2Ra	11.2	11.2	20.0	410.4	55.2	10000.0	1571.3
15	IP.10	8.6	88.1	276.2	1556.5	774.0	10000.0	3136.7
16	MCP.1	1.9	30.5	102.3	1064.1	360.4	10000.0	2596.4
17	MIP.1a	2.9	5.4	16.1	544.8	68.1	10000.0	1965.5
18	MIP.1b	3.0	12.0	27.4	440.2	73.2	10000.0	1640.9
19	Eotaxin	4.0	10.4	20.6	457.1	38.5	10000.0	1649.7
20	VEGF	26.3	45.9	156.7	1815.7	473.1	10000.0	3475.7

	Citocina	Debajo	Encima	NAs
1	IL.1a	13.3	0.0	25.9
2	IL.1b	7.5	0.0	21.6
3	IL.6	5.2	0.0	15.1
4	TNFa	10.5	0.1	19.3
5	INFg	0.6	43.2	10.7
6	IL.17	14.5	0.9	30.3
7	IL.12.p40.	3.3	19.1	35.9
8	IL.12.p70.	7.2	62.2	7.2
9	IL.2	2.9	0.0	20.3
10	IL.8	0.0	43.3	16.3
11	IL.4	0.0	43.3	22.4
12	IL.10	1.9	0.1	21.8
13	IL.1Ra	28.2	0.0	43.9
14	sIL.2Ra	1.3	19.1	29.1
15	IP.10	9.2	19.1	11.7
16	MCP.1	5.6	0.0	7.3
17	MIP.1a	0.1	0.1	15.2
18	MIP.1b	0.1	0.1	7.6
19	Eotaxin	0.5	19.1	10.9
20	VEGF	0.1	0.0	33.3

(a) Distribución de las concentraciones de cada citocina para toda la cohorte en todos los tiempos

(b) Porcentaje de valores por debajo y encima del límite de detección y de los cuales no se tienen datos.

**Tabla 6.1:** Resumen de las concentraciones

En la tabla 6.1a se puede ver un resumen de la distribución de las concentraciones a lo largo de todo los tiempos y con toda la población de cada citocina. Además, en

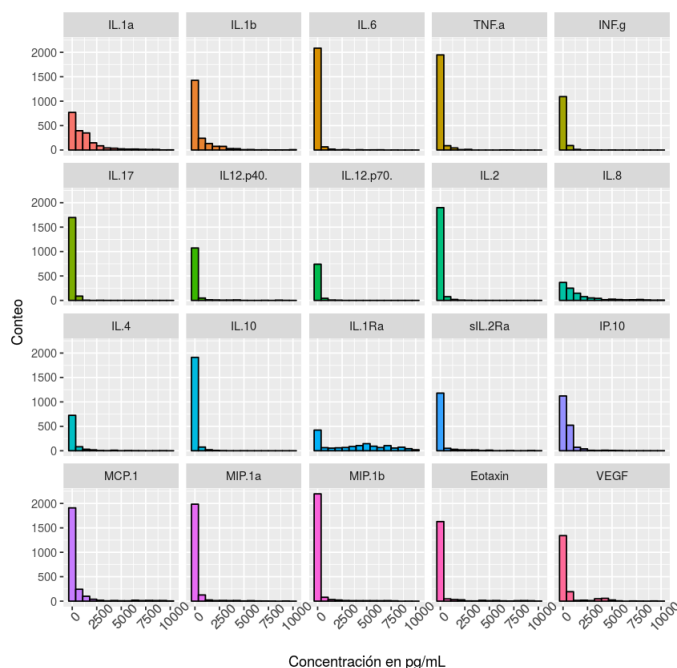
## 6. RESULTADOS Y DISCUSIÓN

la tabla 6.1b, se puede apreciar el porcentaje de datos debajo y por encima del límite de detección, como también el porcentaje de valores no obtenidos. Es importante notar que el valor mínimo en la mayoría de los casos, es el límite de detección indicado en el manual de las perlas magnéticas. El valor máximo se escogió como 10,000 pg/ml.

Como se puede ver, hay un gran cantidad de valores debajo del límite de detección. Los valores por encima pueden ser recuperados usando las muestras que se tienen en congelación y diluyéndolas. También se puede apreciar que hay una cantidad considerable de datos perdidos o no documentados.

Continuando con las diferencias entre las distribuciones de los dos casos, unos descriptivos superficiales para cada citocina entre término y pretérmino se pueden observar en el conjunto de tablas 6.2.

Se puede ver en los histogramas de la Fig. 6.1 que, a pesar de que el rango de valores que toman las citocinas es muy grande, la mayoría de los datos es menor a 1000pg/mL. En cambio, en los histogramas para cada citocina en escala logarítmica y sin datos censurados, se puede ver una distribución aproximadamente unimodal en la mayoría de los casos, pero con colas largas.



**Figura 6.1:** Histogramas en escala normal y sin LODs, para todas las citocinas analizadas.

Si se añaden los datos censurados, usando la convención de que se les impute el valor  $\text{LOD}/\sqrt{2}$  y 10,000 en los superiores, se puede ver en cuales citocinas es mayor el porcentaje de casos.

	Citocina	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD
1	IL.1a	9.4	49.7	692.3	2225.8	2252.9	10000.0	3353.0
2	IL.1b	0.8	4.6	62.7	1496.8	1043.3	10000.0	3058.6
3	IL.6	0.9	3.6	12.1	388.9	40.3	10000.0	1732.2
4	TNFa	0.7	1.2	4.7	85.6	12.5	6833.2	385.3
5	INFG	0.8	1.5	4.0	72.7	17.0	4947.9	252.6
6	IL.17	0.7	0.7	2.3	278.2	8.4	10000.0	1536.3
7	IL12.p40.	7.4	7.4	10.3	462.4	28.3	10000.0	1919.6
8	IL.12.p70.	0.6	0.9	2.8	66.0	11.1	2041.3	198.3
9	IL.2	1.0	1.4	4.5	53.0	17.0	3077.6	203.5
10	IL.8	0.6	315.6	1440.9	3842.0	10000.0	10000.0	4172.2
11	IL.4	4.5	4.5	8.2	270.2	60.3	10000.0	1083.0
12	IL.10	1.1	1.4	5.0	205.9	18.9	10000.0	1224.9
13	IL.1Ra	8.3	53.0	4478.4	4624.9	10000.0	10000.0	4103.6
14	sIL.2Ra	11.2	11.2	20.2	362.6	55.0	10000.0	1460.1
15	IP.10	8.6	87.5	272.4	1350.2	722.6	10000.0	2874.8
16	MCP.1	1.9	32.0	102.3	936.7	340.7	10000.0	2377.6
17	MIP.1a	2.9	5.2	15.9	498.2	66.1	10000.0	1862.6
18	MIP.1b	3.0	12.0	27.9	390.6	72.6	10000.0	1515.8
19	Eotaxin	4.0	10.7	20.9	407.7	38.5	10000.0	1536.4
20	VEGF	26.3	44.5	151.5	1640.6	410.7	10000.0	3296.9

(a) Distribución de las concentraciones de cada citocina para casos término

	Citocina	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD
1	IL.1a	9.4	9.4	827.4	2557.9	2832.6	10000.0	3637.2
2	IL.1b	0.8	23.2	430.0	3515.4	10000.0	10000.0	4514.7
3	IL.6	0.9	4.3	21.3	1393.9	87.2	10000.0	3399.3
4	TNFa	0.7	1.1	3.7	245.1	14.5	10000.0	908.4
5	INFG	0.8	1.4	2.9	154.7	217.0	1416.3	303.7
6	IL.17	0.7	0.7	2.9	135.5	8.6	10000.0	768.3
7	IL12.p40.	7.4	7.4	11.0	1377.4	44.5	10000.0	3306.9
8	IL.12.p70.	0.6	0.7	4.0	105.8	126.3	1484.1	237.4
9	IL.2	1.0	2.5	5.4	99.4	31.0	1685.3	263.8
10	IL.8	2.2	1980.9	7122.9	6118.7	10000.0	10000.0	3927.9
11	IL.4	4.5	4.5	7.2	718.7	361.2	10000.0	1879.4
12	IL.10	1.1	1.1	3.9	823.9	28.7	10000.0	2614.2
13	IL.1Ra	8.3	7224.9	10000.0	8041.9	10000.0	10000.0	3459.4
14	sIL.2Ra	11.2	11.2	16.5	958.6	74.4	10000.0	2458.8
15	IP.10	8.6	99.0	364.9	3926.2	10000.0	10000.0	4684.3
16	MCP.1	1.9	20.0	97.1	2791.5	7847.8	10000.0	4288.1
17	MIP.1a	2.9	5.8	19.9	1177.8	113.0	10000.0	2970.2
18	MIP.1b	3.0	10.9	21.9	1114.1	89.7	10000.0	2757.3
19	Eotaxin	4.0	6.0	16.3	1024.2	38.7	10000.0	2560.7
20	VEGF	26.3	79.1	377.5	4203.7	10000.0	10000.0	4753.4

(b) Distribución de las concentraciones de cada citocina para casos pretérmino.

	Citocina	Debajo	Encima	NAs
1	IL.1a	11.1	13.1	0.0
2	IL.1b	10.4	7.1	0.0
3	IL.6	11.7	2.9	0.0
4	TNFa	19.4	0.0	0.0
5	INFG	0.0	0.0	43.5
6	IL.17	0.0	0.0	1.0
7	IL12.p40.	0.0	0.0	20.1
8	IL.12.p70.	0.0	0.0	63.5
9	IL.2	20.6	0.0	0.0
10	IL.8	0.0	0.0	43.5
11	IL.4	0.0	0.0	43.5
12	IL.10	19.3	1.5	0.0
13	IL.1Ra	16.4	25.1	0.0
14	sIL.2Ra	0.0	0.0	20.1
15	IP.10	0.0	0.0	20.1
16	MCP.1	0.0	0.0	0.0
17	MIP.1a	12.2	3.1	0.0
18	MIP.1b	5.6	1.6	0.0
19	Eotaxin	0.0	0.0	20.1
20	VEGF	20.2	12.1	0.0

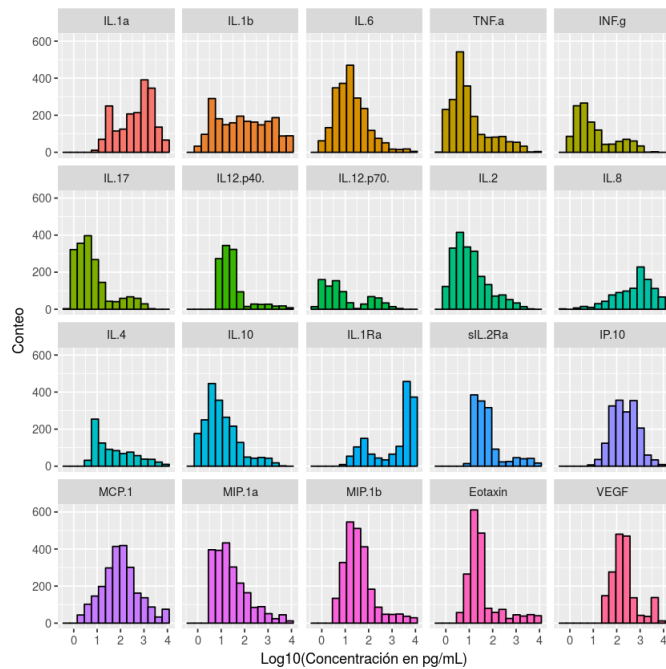
(c) Porcentaje de valores por debajo y encima del límite de detección y de los cuales no se tienen datos. Casos término.

	Citocina	Debajo	Encima	NAs
1	IL.1a	33.5	16.2	0.0
2	IL.1b	6.1	31.8	0.0
3	IL.6	8.9	13.4	0.0
4	TNFa	17.9	0.6	0.0
5	INFG	0.0	0.0	39.7
6	IL.17	26.3	0.6	0.0
7	IL12.p40.	0.0	0.0	6.1
8	IL.12.p70.	0.0	0.0	45.3
9	IL.2	16.2	0.6	0.0
10	IL.8	0.0	0.0	40.2
11	IL.4	0.0	0.0	40.2
12	IL.10	28.5	7.3	0.0
13	IL.1Ra	6.7	69.3	0.0
14	sIL.2Ra	0.0	0.0	5.6
15	IP.10	0.0	0.0	5.6
16	MCP.1	3.4	24.0	0.0
17	MIP.1a	5.0	8.9	0.0
18	MIP.1b	6.7	6.1	0.0
19	Eotaxin	0.0	0.0	5.6
20	VEGF	0.0	0.0	0.6

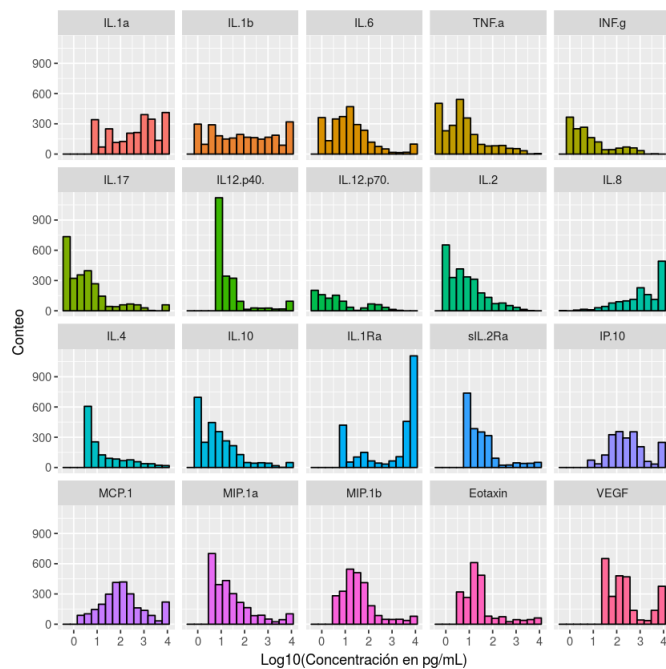
(d) Porcentaje de valores por debajo y encima del límite de detección y de los cuales no se tienen datos. Casos pretérmino

Tabla 6.2: Resumen de las concentraciones, término vs pretérmino

## 6. RESULTADOS Y DISCUSIÓN



**Figura 6.2:** Histogramas en escala log10 y sin LODs, para todas las citocinas analizadas.



**Figura 6.3:** Histogramas en escala log10 con LODs, para todas las citocinas analizadas.

## Imputación con Random Forest

Se usó el paquete “caret” para el lenguaje de programación R, el cual ofrece una interfaz muy completa para entrenamiento de clasificación y regresión. Se puede elegir que se corra el entrenamiento muchas veces, variando el número de variables aleatorias que se consideran para hacer el *splitting* en cada decisión. Para decidir con que número de parámetros se obtuvo la mejor, existen muchos métodos pero los dos más conocidos para bosques aleatorios son *k-fold cross validation* (**CV**) y *out of bag* (**oob**) *estimate*. Los dos métodos hacen uso de datos con los que no se tomó la decisión del *splitting*, para hacer una recursión y estimar que tan buen ajuste tiene. En el **oob**, se usan los datos que quedaron fuera debido a la elección aleatoria de las variables y en **CV** se corre el entrenamiento varias veces, variando los datos que se usan para entrenar y para probar. Se puede escoger entre dos métricas para decidir el número de variables óptimo; el error promedio de los cuadrados (*root mean square error* o *RMSE*) o el coeficiente  $R^2$ . Si el número de árboles aleatorios es muy grande, **oob** tiende a **CV** (Janitza and Hornung, 2018).

En este caso se usó CV y como métrica RMSE, para no tener que usar tantos árboles y porque con pocos árboles se obtuvieron buenos resultados. En la mayoría de los casos se usaron 5 CV's y 600 árboles, pues se observó que al aumentar la cantidad de árboles no hubo mucha mejora en la capacidad predictiva. Sin embargo, para la citocina IL-8, con esos valores, el coeficiente de determinación  $R^2$  era del orden 0.5, por eso se usaron 3 repeticiones de CV's con 10 iteraciones cada uno y 400 árboles. Esto subió el  $R^2$  a 0.94. Además se hizo uso del paquete 'doMC' para correr el entrenamiento en paralelo, haciendo uso de 7 núcleos del procesador. En promedio cada entrenamiento duró 50 segundos, cuando se exploraba un rango de 6 números de predictores.

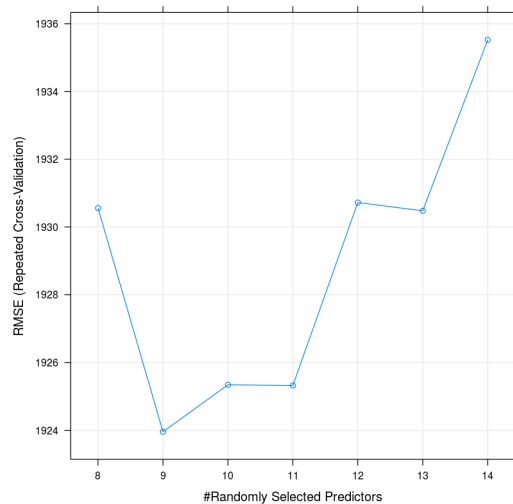
Una vez que se corre la función 'train', el modelo es guardado en una variable, desde la cual se puede extraer toda la información. En particular se puede generar un gráfico del **RMSE** asociado a cada número de variables. Se buscó que el punto escogido se comporte como un mínimo, es decir que se encuentre en un pozo como el de la Fig. 6.4.

Debido a que IL-12p70 tenía aprox. 60 % de datos perdidos, se consideró que eran demasiados y se descartó a esta citocina para el entrenamiento. Además, como la función para entrenar a los bosques aleatorios solo trabaja con las filas completas se imputó el promedio en cualquiera de los valores perdidos. Sin embargo, cualquier citocina con mas de 20 % de valores perdidos no se usó para la predicción. Adicionalmente, los LOD's se mantuvieron en su valor límite y se creó una variable que dividió los días de gestación en cuartiles, la cual también se uso para entrenar los árboles. Sin embargo, está variable no representó ninguna mejora considerable en el poder de predicción de los bosques.

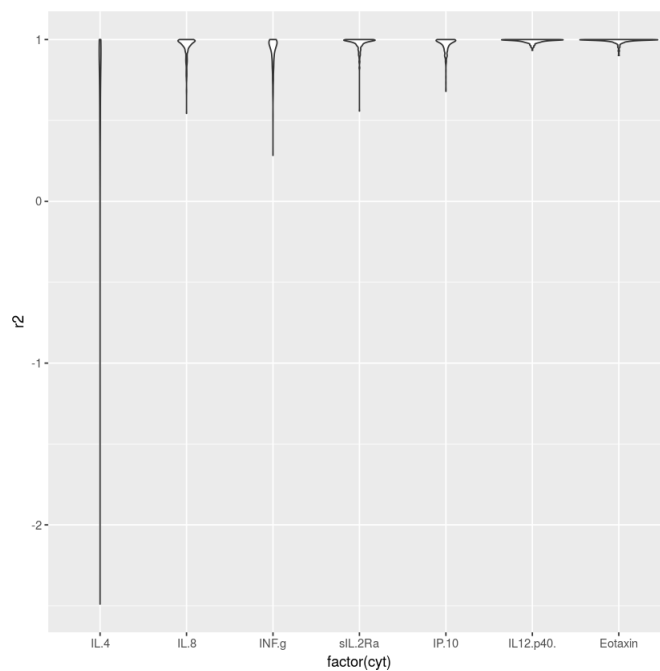
En cada una de las predicciones, se separó la muestra en 80 % de los valores para entrenamiento y 20 % para prueba. Después de entrenar los bosques se hicieron 300 pruebas con diferentes particiones de los datos y se evaluó la predictibilidad del modelo con el coeficiente de determinación. Las gráficas de violines en la Fig. 6.5 muestran la distribución de las  $R^2$ . Cuando la  $R^2$  es menor a 0, quiere decir que los datos predichos son peores que si se predijera el promedio para todos los datos.

## 6. RESULTADOS Y DISCUSIÓN

---



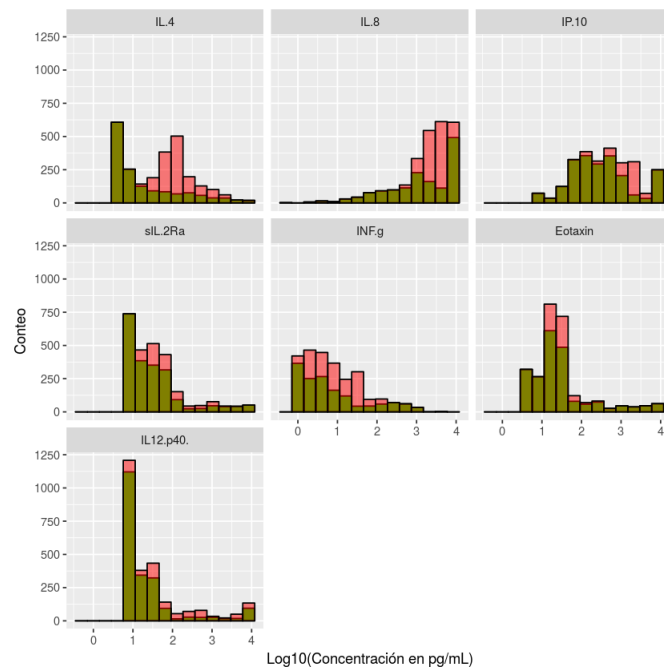
**Figura 6.4:** Número de variables para el binary splitting aleatorio contra RMSE (raíz del error promedio de los cuadrados) en el entrenamiento para predecir la citocina IL-8 por regresión con bosques aleatorios. Se eligieron 9 variables automáticamente.



**Figura 6.5:** Distribución de coeficientes de determinación  $R^2$  de la regresión con las predicciones de RF y para cada una de las citocinas.

Como se puede ver, de las citocinas que se eligieron para imputar, todas excepto IL-4 predicen decentemente.

Otra manera de asegurarse de que las predicciones tienen sentido es ver como se transforma el histograma de los valores de la citocina cuando se imputan los valores. En la Fig. 6.6 se pueden ver los valores sin imputar en verde y los imputados en rojo. Es posible observar que en general se sigue la tendencia excepto para IL-4, lo cual no sorprende tanto, pues el poder de predicción fue muy bajo. A pesar de que se usaran mucho más árboles, o más CV's con repeticiones, la tendencia no mejoraba.



**Figura 6.6:** Comparación de histogramas con y sin RF imputado.

## Imputación con distribuciones truncadas

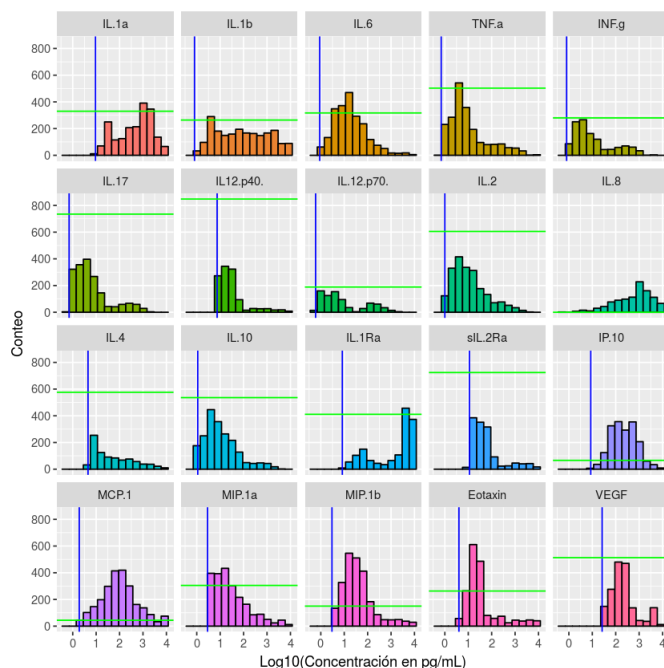
Para ejecutar este ajuste se usó el lenguaje de programación R, junto con el paquete “ggplot2” para graficar y “reshape” para ajustar el formato de las bases de datos tal que fueran más fáciles de visualizar. El truncado de la distribución se programó en el mismo *script* partiendo de los principios matemáticos expuestos en la sección 5.3.

Como se vio en el análisis exploratorio ( y se puede ver en la Fig. 6.7), los datos se asemejan más a una distribución unimodal cuando se usa escala logarítmica. Existen casos, como el de IL12, donde aparece una figura bimodal, la cual puede haber surgido porque los datos se tomaron a lo largo del embarazo y pueden representar dos etapas con funcionamiento distinto. Por eso, para ajustarse a esos datos habría que usar



## 6. RESULTADOS Y DISCUSIÓN

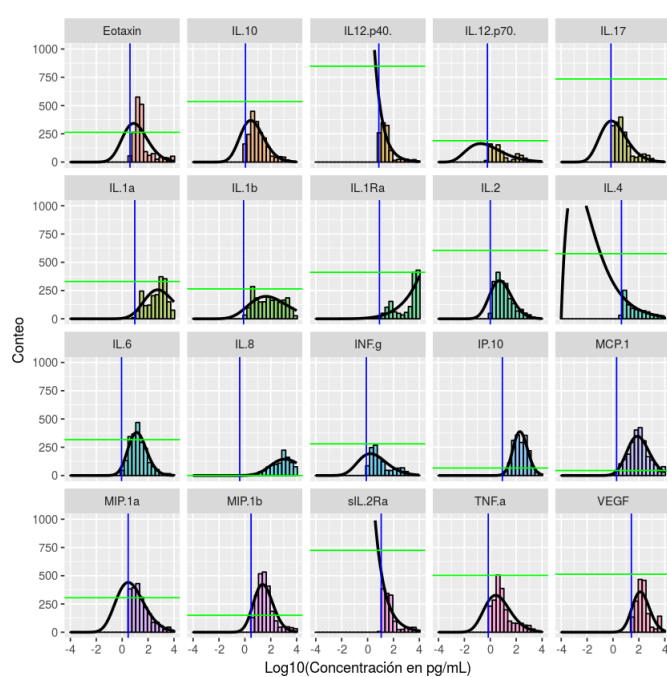
distribuciones bimodales y el número de parámetros aumentaría, lo que aumenta los máximos locales y hace menos óptimo el ajuste. En la Fig. 6.7 se puede observar que el número de datos censurados es congruente con la tendencia de la distribución. Si el LOD está cerca del pico de la distribución, la prevalencia es muy alta. Si está en los extremos de la cola, la prevalencia es baja.



**Figura 6.7:** Histogramas en escala log10, el LOD en línea azul y el número de datos en el LOD en línea verde. Para todas las citocinas analizadas.

Una complicación que surge al transformar los datos a escala logarítmica es que ya no están acotados por el 0, pueden ir de  $-\infty$  a  $\infty$ . Esto se resolvió trasladando la función gamma tal que la concentración mínima que pueda haber de citocinas esté donde estaba el 0. Tomando en cuenta que el peso de una citocina es  $150 \text{ kDa}$  y considerando 1000 citocinas en  $50 \mu\text{L}$ , se tendría una concentración de  $3 \times 10^{-4} \frac{\text{pg}}{\text{mL}}$ . Por lo que se trasladó a la función gamma 4 unidades. Los resultados de la optimización bajo estas condiciones se pueden ver en la Fig. 6.8.

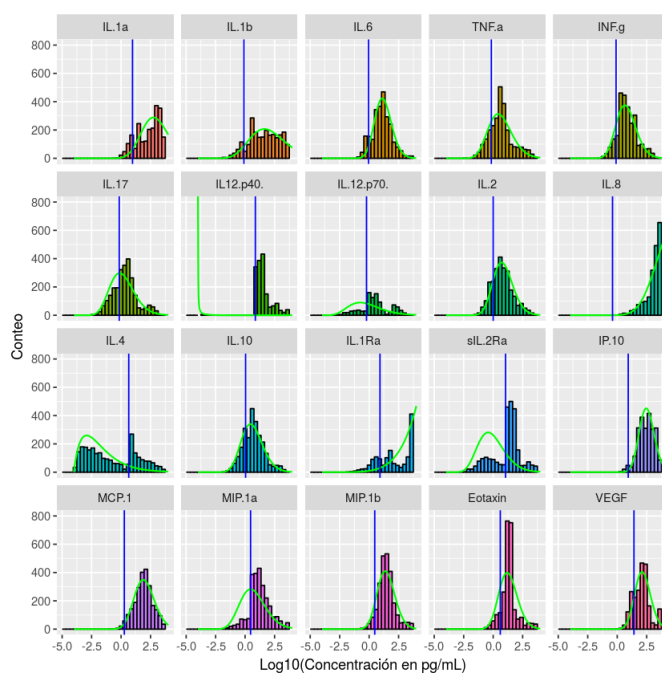
Como se puede notar, en la mayoría de los casos, la curva de la distribución ajustada se alinea muy bien con el histograma. También se puede apreciar que en muchos de estos casos casi no hay asimetría, cuyo valor estadístico solo depende del parámetro de forma y está dado por  $\frac{2}{\sqrt{k}}$ , lo que también se puede apreciar en la propiedad que dice que cuando  $k$  es grande la distribución tiende a la normal. Las excepciones tienen dos razones para su desajuste. En la citocina **IL.1Ra**, se ve claramente una bimodalidad y la distribución termina ajustándose a la cola de una gamma con su promedio en valores muy altos. En cambio, con **sIL-2Ra** y **IL-12p40**, el crecimiento es tan pronunciado



**Figura 6.8:** Histogramas de las concentraciones, el LOD en línea azul y el el número de datos en el LOD en línea verde. Se graficó además la distribución gamma truncada con parámetros obtenidos de optimizar con MLE

## 6. RESULTADOS Y DISCUSIÓN

que la gamma tiende a ser exponencial, lo que causaría que solo se imputen valores en el pico de la distribución. En estos casos se renunció a imputar datos, pues un ajuste a una bimodal conlleva mas variabilidad por el número de parámetros y en los otros dos casos hay pocos valores como para vislumbrar la tendencia de los datos. En efecto, se puede notar que en los casos de **sIL2Ra** y **IL12p40**, la mayoría de los datos está debajo del límite de detección.

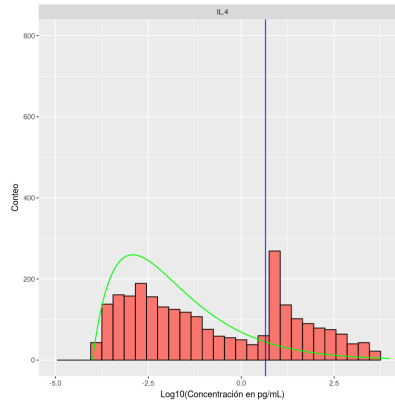


**Figura 6.9:** Histogramas de las concentraciones reales e imputadas, el LOD en línea azul.

Se graficó además la distribución gamma con parámetros obtenidos de optimizar con MLE

Una vez imputados los datos, se puede apreciar en la Fig. 6.9 que en la mayoría de los casos no descartados, la cantidad de datos faltantes concuerda con la densidad estimada por la optimización. Sin embargo, a pesar de que las concentraciones de **IL-4** se ajustaban bien a la cola de una gamma muy asimétrica, al imputar una baja cantidad de datos en las regiones de mayor densidad de la distribución, se desequilibran las proporciones. Esto no necesariamente significa que la imputación sea incorrecta, también puede querer decir que hicieron falta mas mediciones. Dicha especulación es apoyada por el hecho de que la citocina IL-4 tiene un 43.3 % de datos perdidos. Debido a esto se decidió imputar los valores perdidos partiendo de la distribución aproximada (Fig. 6.10 ). Sin embargo, a pesar de que el pico de la distribución sube un poco, la cola sigue quedándose muy por debajo de la densidad de los datos reales. Esto levanta la sospecha de que esta distribución no refleja una situación real, pues debería de haber demasiados valores bajo el límite de detección como para que el ajuste sea bueno. Por eso tampoco se consideró como candidato a imputación. Es probable que

una distribución con mayor curtosis se aproxime mejor a estos datos.



**Figura 6.10:** Histograma de las concentraciones reales e imputadas para LODs inferiores y datos perdidos, el LOD en línea azul en el caso de IL4. Se graficó además la distribución gamma con parámetros obtenidos de optimizar con MLE

Otro posible rescate se logra usando los datos imputados con bosques aleatorios. Al imputar estos datos antes de la optimización y luego imputar los censurados se obtuvo lo que se ve en la Fig. 6.11 .

Se puede observar que el comportamiento de sIL-2Ra, tiene más sentido que cuando no se imputaron los valores perdidos. A pesar de que no hay suficientes valores debajo del límite como para que se siga la tendencia. En el caso de IL-8, no hay mucho que considerar pues no hay valores debajo del umbral. En IL-4, el comportamiento cambia pero sigue siendo igual de inverosímil.

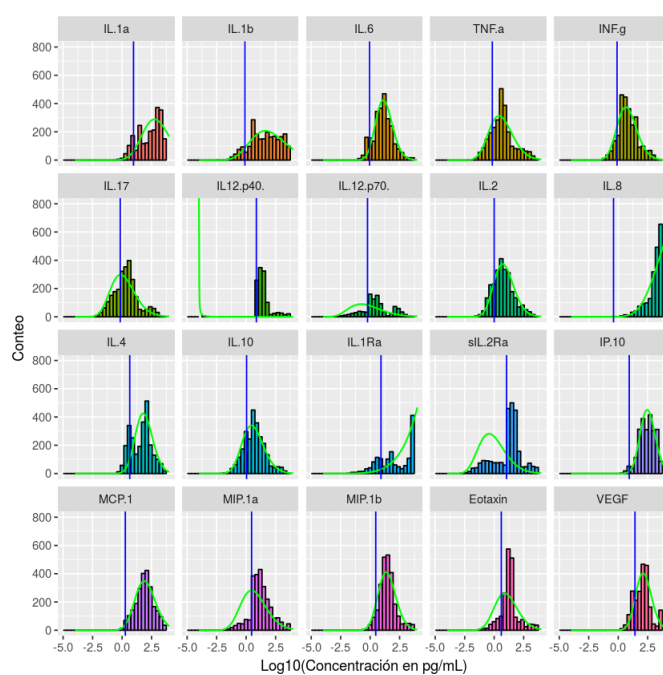
## Comparación de distribuciones con Mann-Whitney

Con ayuda del paquete “coin” de R, que ofrece pruebas de permutación para el problema general de independencia de muestras, se aplicó la prueba de Mann-Whitney con corrección para empates en su aproximación normal.

Desde que Kolmogorov propuso su prueba no paramétrica en 1933 (Kolmogorov, 1933), se abrió un paradigma en el que se tradujo a los estadísticos de las pruebas de rangos a una medida entre las dos poblaciones en la distribución acumulativa empírica. La prueba de Mann-Whitney se traduce a el promedio de las distancias entre las dos distribuciones. En la Fig. 6.12 se pueden ver graficadas las distribuciones empíricas acumulativas de las mediciones provenientes de mujeres que sufrieron parto pretérmino en rojo y las normales en negro sin tomar en cuenta los LODs. Se eligió mostrar esto y no una aproximación de la densidad de probabilidad pues en estas se debe de asumir un valor específico para los LODs, lo que hace que aparezca un pico no representativo

## 6. RESULTADOS Y DISCUSIÓN

---



**Figura 6.11:** Histogramas de las concentraciones con los valores perdidos imputados por RF y los censurados, por gamma después de haber imputado con RF. El LOD en línea azul. Se graficó además, la distribución gamma con parámetros obtenidos de optimizar con MLE.

en las gráficas. En el eje horizontal se presenta el valor  $p$ , cuya hipótesis nula es que las distribuciones son iguales.

Luego se hizo la misma comparación, pero considerando los límites de detección (Fig. 6.13).

Se puede notar que hay un gran cambio en las citocinas que presentan una diferencia estadística. En particular, las citocinas **IL-1a**, **IL-4**, **IL-12p70**, **sIL2Ra** dejaron de presentarla, mientras que **Eotaxin**, **IL-2**, **IL-6**, **MIP-1a**, **IL-1b** la ganaron. Intuitivamente, mientras más separadas estén las curvas de la distribución empírica, más diferencia significativa existe. Sin embargo, en el caso de **MCP-1** con LODs, la distancia entre las distribuciones es muy grande, pero el estadístico aumenta su valor. Esta presunta contradicción se aclara al visualizar la distribución empírica en escala logarítmica (Fig. 6.14 y 6.15). Estos hechos corroboran la intuición que se tenía.

## Comparación

En esta sección se busca comparar si los dos métodos propuestos para recuperar los datos censurados dan las mismas predicciones y si no, en que difieren. Además de evaluar el efecto de la imputación por RF con estos métodos. Para esto se corrieron las pruebas de Mann-Whitney con y sin imputación y también pruebas Baumgartner-Weiss-Schindler (BWS) con los datos imputados por la distribución gamma. La prueba BWS considera la varianza en los datos para calcular su estadístico y se ha mostrado que es más sensible a oblicuidad y curtosis (Veáse (Baumgartner et al., 1998)), pero pierde mucho poder estadístico con los empates si se usan los métodos convencionales para resolverlos, como el de rangos medios (Neuhauser, 2003).

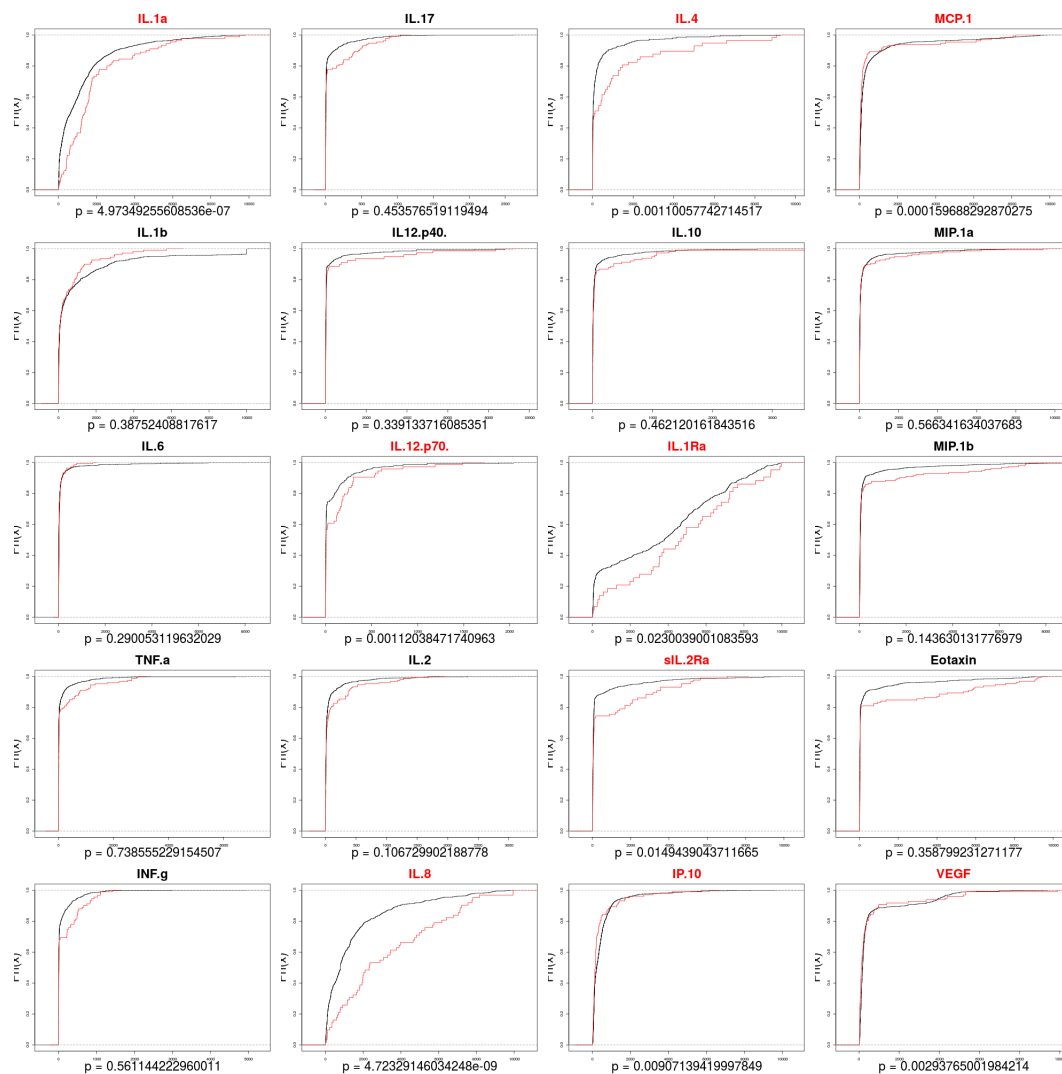
En la Tabla 6.3 se insertaron los estadísticos para los casos mencionados. Es importante notar que se quitaron los valores del límite de detección superior, pues estos no fueron imputados con la distribución gamma, para evaluar la efectividad de los métodos con menos ruido.

Se puede notar que todas las citocinas que exhibieron una diferencia estadística con el método de Mann-Whitney dejando los datos censurados iguales, también la presentaron cuando se imputaron con la distribución gamma. No se nota una tendencia de si para alguno de los dos casos es más grande o más pequeña la diferencia.

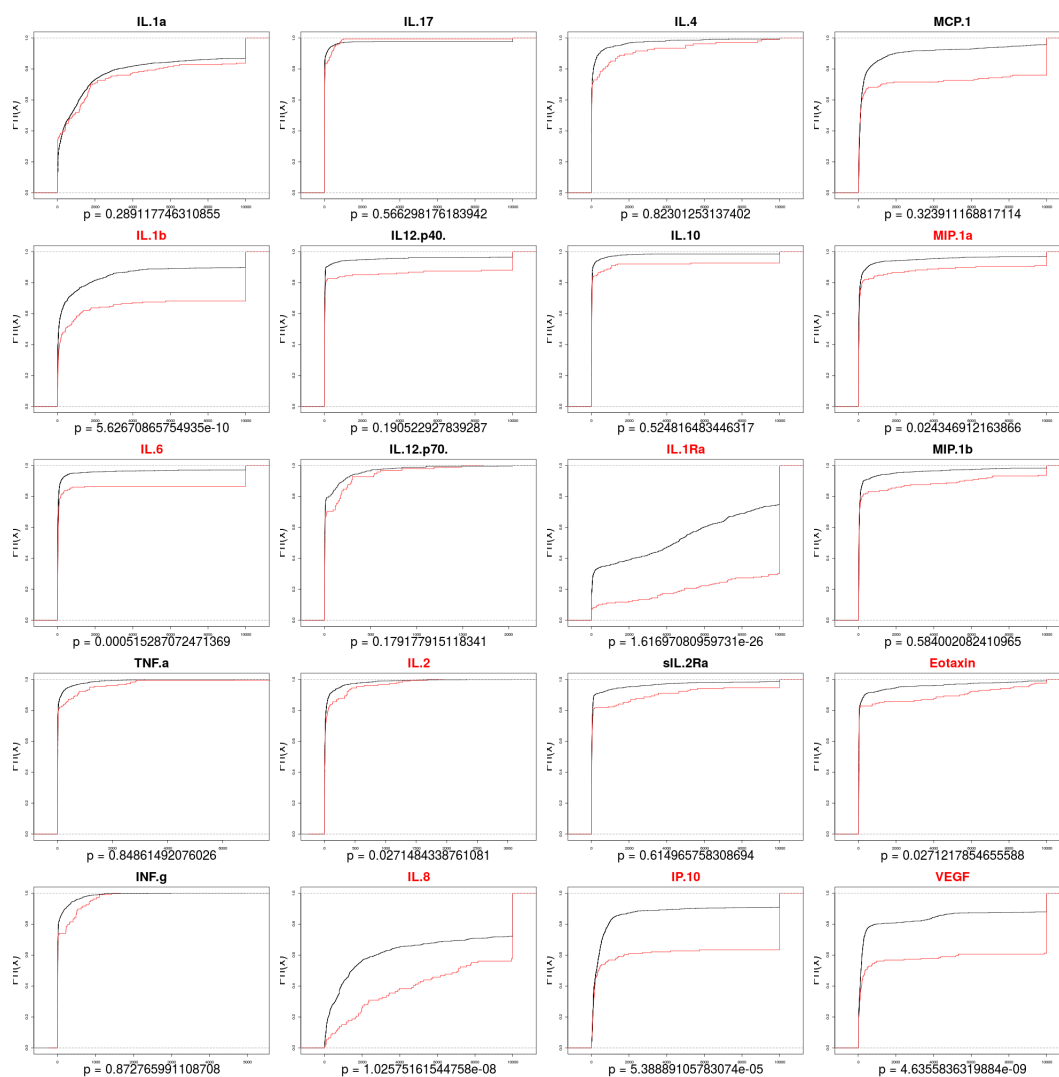
El estadístico BWS no se utilizó en los casos en los que se usó el mismo valor para los LOD's, pues no existe un buen ajuste para rangos empatados (Neuhauser, 2003). Por otro lado, se aprecia que este estadístico es mucho más sensible a los cambios en las distribuciones. Cuando no se consideraron los LOD's, todas las citocinas excepto MIP-1 $\alpha$  e IL-12p40 tenían  $p < 0.05$ . Al imputarse los valores debajo del umbral con la distribución gamma, TNF- $\alpha$  y VEGF, dejaron de ser estadísticamente diferentes.

Al imputar valores perdidos con RF, solo en Eotaxin dejaron de diferir las distribuciones, en el caso en que no se consideraron los LOD's. Pero con la prueba BWS, se mantuvieron las diferencias. Esto seguramente se debe a que la prueba de BWS es demasiado sensible y aunque las distribuciones cambien un poco, todavía contarán

## 6. RESULTADOS Y DISCUSIÓN



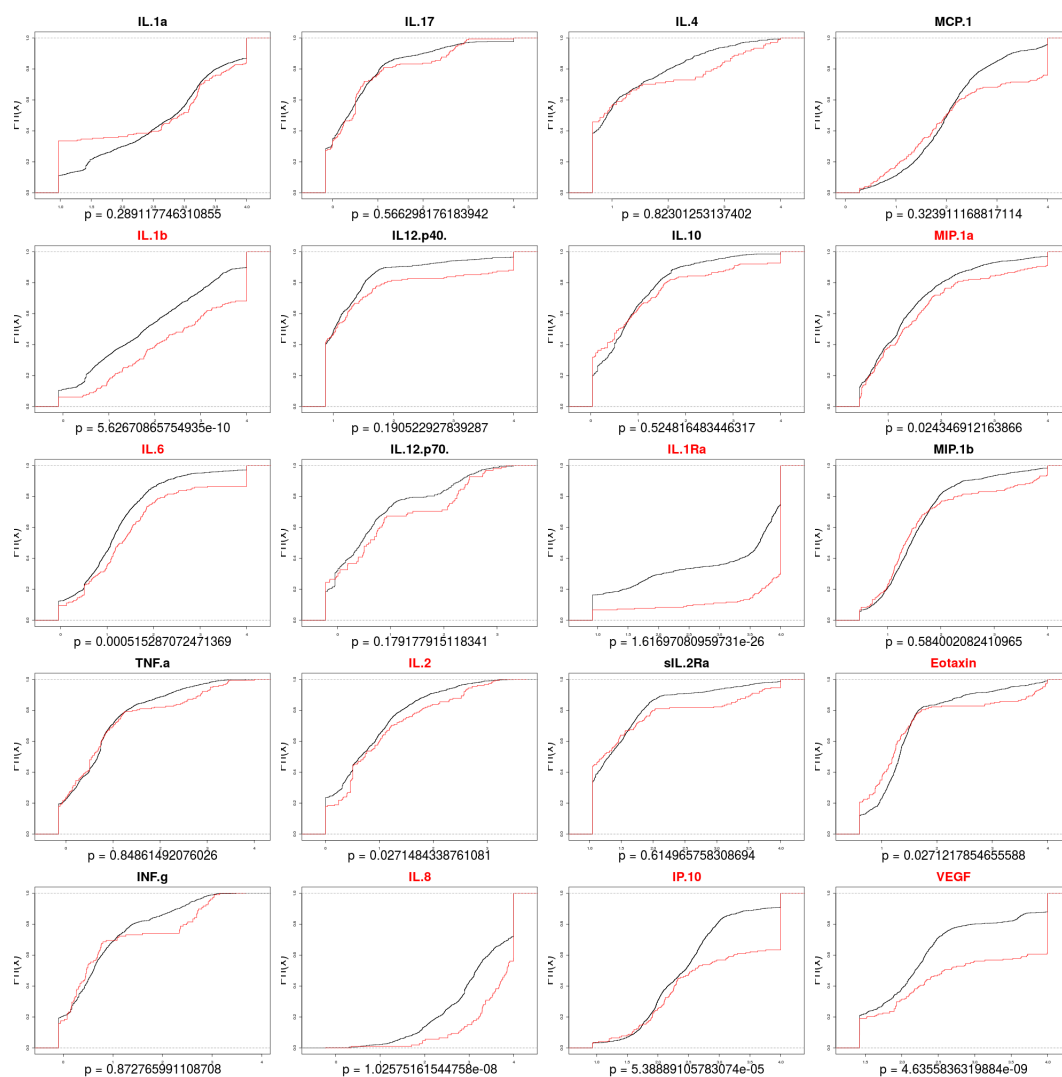
**Figura 6.12:** Distribuciones empíricas acumulativas de las mediciones que fueron término en negro y pretérmino en rojo, sin considerar LODs. En el eje horizontal, la significancia estadística del test de Mann-Whitney. Las citocinas con nombre en rojo cumplen que  $p < 0.05$ .



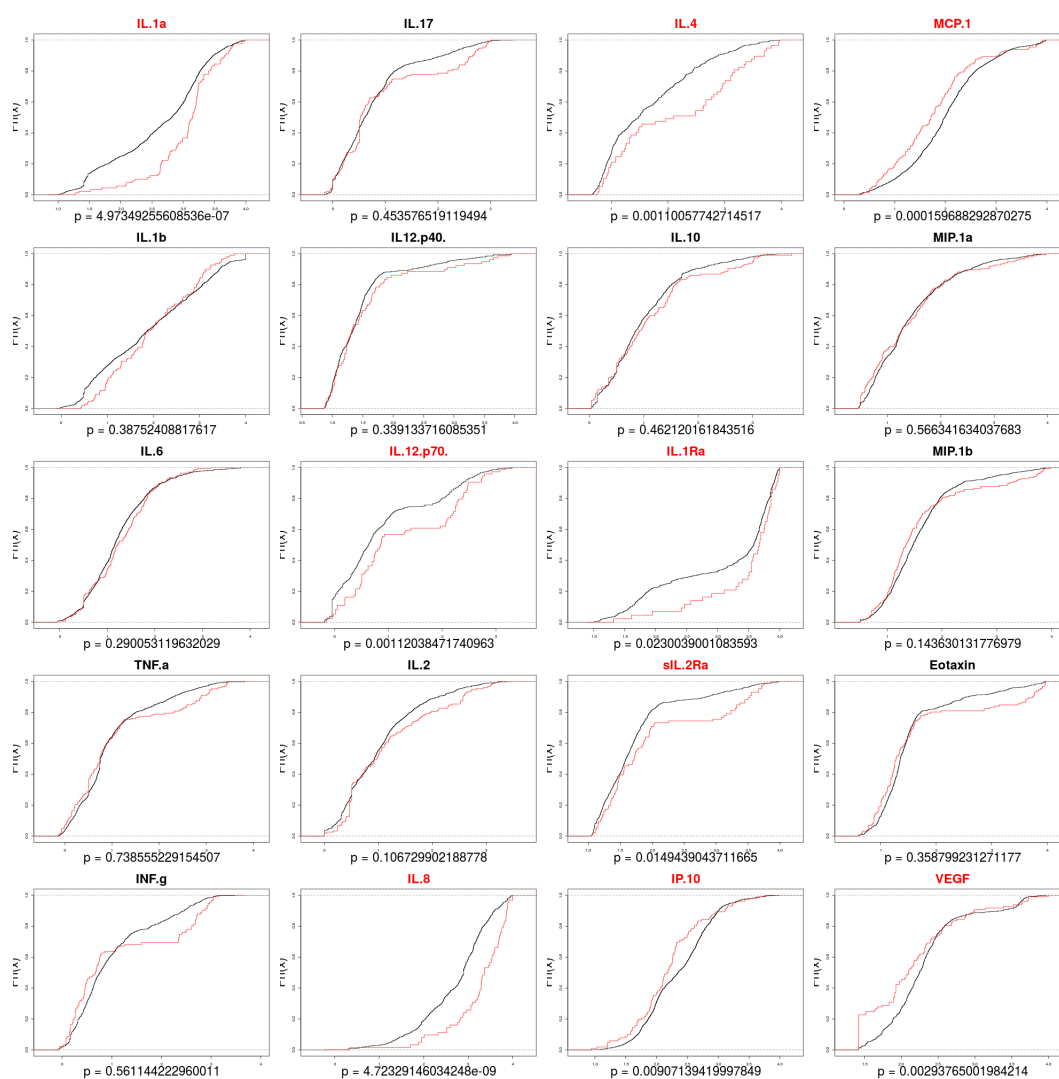
**Figura 6.13:** Distribuciones empíricas acumulativas de las mediciones que fueron término, en negro y pretérmino en rojo, considerando LODs. En el eje horizontal, la significancia estadística del test de Mann-Whitney asintótico. Las citocinas con nombre en rojo cumplen que  $p < 0.05$ .



## 6. RESULTADOS Y DISCUSIÓN



**Figura 6.14:** Distribuciones empíricas de las mediciones que fueron término en negro y pretérmino en rojo, considerando LODs. En el eje horizontal, la significancia estadística del test de Mann-Whitney asintótico. Las citocinas con nombre en rojo cumplen que  $p < 0.05$ . Las concentraciones en escala  $\log_{10}$ .



**Figura 6.15:** Distribuciones empíricas de las mediciones que fueron término en negro y pretérmino en rojo, sin considerar LODs. En el eje horizontal, la significancia estadística del test de Mann-Whitney asintótico. Las citocinas con nombre en rojo cumplen que  $p < 0.05$ . Las concentraciones en escala  $\log_{10}$ .

## 6. RESULTADOS Y DISCUSIÓN

	MW_NOL	BW_NOL	MW_L	MW_L_G	BW_L_G	MW_NOL_RF	BW_NOL_RF	MW_L_G_RF	BW_L_G_RF
IL.1a	4.97E-07	6.08E-08	2.65E-02	2.14E-02	8.45E-07	4.97E-07	6.08E-08	3.44E-02	6.31E-06
IL.1b	1.27E-01	4.04E-03	7.73E-02	7.98E-02	2.17E-02	1.27E-01	4.04E-03	7.82E-02	2.17E-02
IL.6	2.89E-01	5.25E-02	2.15E-01	2.20E-01	1.48E-01	2.89E-01	5.25E-02	2.16E-01	1.33E-01
TNFa	7.31E-01	9.69E-03	9.48E-01	8.35E-01	5.81E-02	7.31E-01	9.69E-03	8.37E-01	5.50E-02
INFg	5.61E-01	8.64E-04	8.73E-01	7.81E-01	1.57E-02	8.35E-02	1.54E-02	7.96E-02	2.33E-02
IL.17	2.99E-01	8.73E-09	3.08E-01	4.79E-01	1.48E-02	2.99E-01	8.73E-09	2.96E-01	1.69E-02
IL.12.p40.	3.39E-01	2.55E-01	5.24E-01	3.68E-01	1.00E+00	3.25E-01	1.70E-01	1.61E-02	1.00E+00
IL.12.p70.	1.12E-03	0.00E+00	1.79E-01	2.24E-01	6.95E-04	1.12E-03	0.00E+00	1.69E-01	1.02E-03
IL.2	1.87E-01	5.21E-03	2.71E-02	1.61E-02	4.14E-03	1.87E-01	5.21E-03	2.85E-02	9.93E-03
IL.8	4.72E-09	2.24E-09	4.72E-09	4.72E-09	2.24E-09	8.01E-12	1.12E-09	8.01E-12	1.12E-09
IL.4	1.15E-03	3.84E-04	8.70E-01	1.88E-01	1.42E-02	1.15E-03	3.84E-04	3.79E-01	7.89E-02
IL.10	1.02E-01	1.47E-05	4.10E-02	4.35E-02	2.50E-04	1.02E-01	1.47E-05	5.73E-02	4.24E-04
IL.1Ra	2.30E-02	3.40E-02	1.11E-01	1.15E-01	2.04E-01	2.30E-02	3.40E-02	1.06E-01	1.86E-01
sIL.2Ra	1.49E-02	8.77E-03	1.46E-01	1.55E-01	0.00E+00	1.59E-01	2.39E-02	4.15E-03	1.72E-04
IP.10	9.07E-03	3.36E-03	7.59E-03	7.59E-03	3.34E-03	6.15E-04	1.93E-04	3.95E-04	1.45E-04
MCP.1	9.55E-05	3.85E-05	3.25E-05	3.26E-05	1.42E-05	9.55E-05	3.85E-05	3.17E-05	1.34E-05
MIP.1a	5.66E-01	7.65E-02	3.07E-01	3.30E-01	3.72E-01	5.66E-01	7.65E-02	3.05E-01	2.61E-01
MIP.1b	1.38E-01	1.06E-02	1.01E-01	9.41E-02	1.21E-02	1.38E-01	1.06E-02	1.01E-01	1.52E-02
Eotaxin	3.59E-01	6.72E-03	9.98E-03	1.15E-02	2.78E-04	9.77E-02	6.01E-04	4.92E-04	2.98E-06
VEGF	9.73E-03	6.48E-06	8.76E-01	9.90E-01	6.21E-02	9.73E-03	6.48E-06	7.83E-01	2.56E-01

**Tabla 6.3:** Valores p de la comparación de las distribuciones entre término y pretérmino, para diferentes combinaciones de todos los métodos usados. MW = Mann-Whitney, BW = Baumgartner-Weiss-Shindler, RF = Imputados con bosques aleatorios, G = Imputado con gamma, L = Con LOD's, NOL = Sin LODs. Las celdas verdes claras cumplen que  $p < 0.05$  y las rojas no se consideran por redundancia o porque el ajuste o imputación no tenía sentido.

como diferentes.

Combinar todos los métodos considerados en esta tesis mantuvo los resultados iguales en 3 de los 4 casos, pero también brindó nueva información en el caso sIL-2Ra. Citocina para la cual no había diferencia estadística con el test de Mann-Whitney pero al imputar los datos perdidos con RF, se pudo hacer un mejor ajuste de distribución truncada que llevo a que ya no difirieran tanto las distribuciones entre término y pretérmino.

Además, se pudo apreciar que cuando hubo demasiados datos perdidos y además censurados, el rescate de los datos no fue efectivo, como fue el caso de IL-4.

## Conclusiones

---

Gran parte de la investigación médica y biológica trata con las concentraciones de ciertos agentes en el cuerpo humano. Una de las tecnologías más utilizadas para cuantificar concentraciones de proteínas y otras moléculas de tamaños parecidos, es la del inmunoensayo. A pesar de que el rango de concentraciones que pueden medir los equipos para ejecutar inmunoensayos ha aumentado últimamente, en casos como el de las citocinas, la variación sigue siendo tan grande que inevitablemente habrá datos fuera del umbral de detección. La ejecución de un inmunoensayo es algo costoso y requiere de mucho tiempo, es por esto que es importante saber cuales son los parámetros que influyen en éste límite de detección y como se puede recuperar la información cualitativa que ofrecen los datos fuera del umbral.

La explicación del porque de los límites de detección puede hacerse con ayuda de modelos físicos de la dinámica de los anticuerpos, además de otros factores como los aparatos ópticos. Por otro lado, con el uso de herramientas estadísticas se puede recuperar cierta información sin tener que hacer el gasto correr otro inmunoensayo.

Tomando en cuenta esta problemática, se uso esta tesis para explicar el funcionamiento de los inmunoensayos *multiplex* de perlas magnéticas. Se presentó un poco de su historia y se pusieron en contexto con las tecnologías actuales para medición de concentraciones de proteínas.

Para explicar el origen de los límites de detección se presentó un modelo básico desde principios físicos, e.d. con la ley de acción de masas, para modelar el límite de detección de este tipo de inmunoensayos, en específico para el instrumento *MAGPIX*<sup>®</sup>. Con ayuda de este modelo se concluyó que un límite teórico esta alrededor de  $0.1 \frac{pg}{mL}$ . Además, al variar los parámetros en los cálculos se notó que el límite de detección para las condiciones del *MAGPIX*<sup>®</sup> es muy sensible a la concentración de perlas magnéticas. Por eso se propone que, con tal de no bajar de cierta concentración de anticuerpo de captura, se puede disminuir la concentración de perlas magnéticas para poder detectar menores concentraciones.

Luego, se presentó el caso de estudio con el cual se probarían las herramientas estadísticas para recuperar la información de los datos censurados. El cual se trataba de las concentraciones de citocinas en el exudado cervicovaginal de mujeres embarazadas

## 7. CONCLUSIONES

---

con parto normal y pretérmino.

Con el fin de recuperar información a partir de herramientas estadísticas se decidió usar imputación con distribuciones truncadas y también el ajuste de empates de la prueba de Mann-Whitney (aunque este no dio buenos resultados). A parte, se usaron bosques aleatorios para imputar los valores perdidos pero no censurados y para ver si esta imputación ayudaba a la imputación de valores por distribución truncada.

Se pudo observar que, efectivamente, la consideración de los valores imputados en las pruebas de Mann-Whitney usando un solo valor para todos los datos censurados, produjo diferencias estadísticas en citocinas diferentes a las que la presentaron sin considerar los valores censurados. Sin embargo, como se comprobó en la sección de simulaciones (5.2.4), la mayoría de estas diferencias se deben a la sola presencia de empates. La distribución de los datos podría haber sido la misma.

Al correr la imputación por distribución gamma truncada, hubo un buen ajuste a la mayoría de los datos de las citocinas. Sin embargo, las citocinas **sIL2Ra**, **IL-1Ra**, **IL-4** no presentaron un buen ajuste, algunas porque eran muy pocos los datos y otras porque la distribución era bimodal. Una vez imputados los valores perdidos con bosques aleatorios, se observó una mejora en el ajuste a la citocina sIL-2Ra. A pesar de que no había suficientes datos debajo del límite como para que se siguiera la tendencia de la curva.

Finalmente, se hizo una comparación de diferentes combinaciones de estos métodos por medio de sus valores estadísticos  $p$ . Las mismas citocinas que presentaron un diferencia estadística usando un mismo valor para los LOD's, presentaron una diferencia cuando se imputaron estos valores usando el ajuste a una distribución.

Es por esto que se concluye que la consideración de los datos censurados por medio de los métodos de pruebas de Mann-Whitney e imputación con distribuciones truncadas puede proveer mas información sobre el comportamiento de un fenómeno. Sin embargo, el que se puedan usar estas herramientas depende mucho del tipo de distribuciones en los datos que se tengan. En particular es importante que las distribuciones no sean bimodales y que no haya más de 50 % de datos perdidos.

---

Apéndice A

# Código/Manuales/Publicaciones

---

## **Apéndice**

Apéndice



# Bibliografía

---

- American Diabetes Association (2003). Gestational diabetes mellitus. *Diabetes care*, 26 Suppl 1(suppl 1):S103–5. [34](#)
- Anderson, N. L., Jackson, A., Smith, D., Hardie, D., Borchers, C., and Pearson, T. W. (2009). SISCAPA peptide enrichment on magnetic beads using an in-line bead trap device. *Molecular & cellular proteomics : MCP*, 8(5):995–1005. [7](#)
- Angeloni, S., Cordes, R., Dunbar, S., Garcia, C., Gibson, G., Martin, C., and Stone, V. (2013). *xMAP® Cookbook*. 1 edition. [17](#)
- Armbruster, D. A. and Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *The Clinical biochemist. Reviews*, 29 Suppl 1(Suppl 1):S49–52. [16](#)
- Azadeh, M., Gorovits, B., Kamerud, J., MacMannis, S., Safavi, A., Sailstad, J., and Sondag, P. (2018). Calibration Curves in Quantitative Ligand Binding Assays: Recommendations and Best Practices for Preparation, Design, and Editing of Calibration Curves. *The AAPS Journal*, 20(1):22.
- Baumgartner, W., Weiß, P., and Schindler, H. (1998). A nonparametric test for the general two-sample problem. *Biometrics*, 54(3):1129–1135. [43](#), [63](#)
- Bellera, C. A., Julien, M., and Hanley, J. A. (2010). Normal Approximations to the Distributions of the Wilcoxon Statistics: Accurate to What N? Graphical Insights. *Journal of Statistics Education*, 18(2). [41](#)
- Berg, A. and Politis, D. N. (2007). Density Estimation of Censored Data with Infinite-Order Kernels. *arXiv*, pages 1–18.
- Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, 59(1):65–98. [27](#)
- Blanche, P., Latouche, A., and Viallon, V. (2015). Time-dependent AUC with right-censored data: A survey. *Lecture Notes in Statistics*, 215:239–251.



## BIBLIOGRAFÍA

---

- Breidinger, S. and Woolf, E. (2017). Regulated Bioanalysis—Historical Aspects and Key Concepts Related to Its Use. In *Regulated Bioanalysis: Fundamentals and Practice*, pages 87–102. Springer, Cham.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32. [47](#)
- Brendan Technologies, I. (2012). Weighting : The Good , The Bad and The Ugly.
- Bůžková, P., Brown, E. R., and John-Stewart, G. C. (2010). Longitudinal data analysis for generalized linear models under participant-driven informative follow-up: An application in maternal health epidemiology. *American Journal of Epidemiology*, 171(2):189–197.
- Chang, L., Rissin, D. M., Fournier, D. R., Piech, T., Patel, P. P., Wilson, D. H., and Duffy, D. C. (2012). Single molecule enzyme-linked immunosorbent assays: theoretical considerations. *Journal of immunological methods*, 378(1-2):102–15. [17](#), [22](#)
- Commo, F. and Bot, B. M. (2015). N-Parameter Logistic Regression [R package nplr version 0.1-4]. pages 1–11.
- Conover, J. W. (1973). On Methods of Handling Ties in the Wilcoxon Test. *Journal of the American Statistical Association*, 68(344):985–988.
- Cytokine, H., Magnetic, C., and Panel, B. (2013). Human Cytokine / Chemokine Magnetic Bead Panel 96 Well Plate Assay. [26](#)
- Darmanis, S. (2011). *Solid-phase Proximity Ligation Assays - High-performance and multiplex protein analyses*. PhD thesis, Uppsala University.
- Darmanis, S., Nong, R. Y., Väneliid, J., Siegbahn, A., Ericsson, O., Fredriksson, S., Bäcklin, C., Gut, M., Heath, S., Gut, I. G., Wallentin, L., Gustafsson, M. G., Kamali-Moghaddam, M., and Landegren, U. (2011). ProteinSeq: High-Performance Proteomic Analyses by Proximity Ligation and Next Generation Sequencing. *PLoS ONE*, 6(9):e25583. [8](#)
- Davidian, M., Carrol, R. J., and Smith, W. (1988). Variance functions and the minimum detectable concentration in assays. *Biometrika*, 75(3):549–556. [14](#)
- Davis, D., Zhang, A., Etienne, C., and Huang, I. (2002). Principles of Curve Fitting for Multiplex Sandwich Immunoassays. *BioRad Bulletin*, Technical:1–4.
- de Jager, W., Bourcier, K., Rijkers, G. T., Prakken, B. J., and Seyfert-Margolis, V. (2009). Prerequisites for cytokine measurements in clinical trials with multiplex immunoassays. *BMC immunology*, 10:52.
- Deng, H. and Wickham, H. (2011). Density Estimation In R. *Electronic publication*, (September):17.

- 
- Donati, S. (2000). *Photodetectors : devices, circuits, and applications*. Prentice Hall PTR. 20
- Dudley, R. A., Edwards, P., Ekins, R. P., Finney, D. J., McKenzie, I. G., Raab, G. M., Rodbard, D., and Rodgers, R. P. (1985). Guidelines for immunoassay data processing. *Clinical Chemistry*, 31(8):1264–1271.
- Dupage, M. and Bluestone, J. A. (2016). Harnessing the plasticity of CD4+ T cells to treat immune-mediated disease. *Nature Reviews Immunology*, 16(3):149–163.
- Efron, B. (1967). The two sample problem with censored data. In *Proceedings of the fifth Berkeley symposium on*, pages 831–853.
- Ekins, R. (1991). Immunoassay design and optimization. In *Principles and Practice of Immunoassay*, pages 96–153. Palgrave Macmillan UK, London.
- Ekins, R. P. (1960). The estimation of thyroxine in human plasma by an electrophoretic technique. *Clinica Chimica Acta*, 5. 5
- Ekins, R. P. (1989). Multi-analyte immunoassay. *Journal of pharmaceutical and biomedical analysis*, 7(2):155–68. 5
- Ellington, A. A., Kullo, I. J., Bailey, K. R., and Klee, G. G. (2010). Antibody-based protein multiplex platforms: technical and operational challenges. *Clinical chemistry*, 56(2):186–93.
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press. 47
- Finney, D. J. and Phillips, P. (1977). The Form and Estimation of a Variance Function, with Particular Reference to Radioimmunoassay. *Applied Statistics*, 26(3):312. 12, 14
- Fong, Y., Sebestyen, K., Yu, X., Gilbert, P., and Self, S. (2013). nCal: An R package for non-linear calibration. *Bioinformatics*, 29(20):2653–2654. 27
- Fong, Y., Wakefield, J., De Rosa, S., and Frahm, N. (2012). A Robust Bayesian Random Effects Model for Nonlinear Calibration Problems. *Biometrics*, 68(4):1103–1112.
- Gillespie, B. W., Chen, Q., Reichert, H., Franzblau, A., Hedgeman, E., Lepkowski, J., Adriaens, P., Demond, A., Luksemburg, W., and Garabrant, D. H. (2010). Estimating Population Distributions When Some Data Are Below a Limit of Detection by Using a Reverse Kaplan-Meier Estimator. *Epidemiology*, 21(Supplement):S64–S70.
- Gottschalk, P. G. and Dunn, J. R. (1994). Fitting Brendan’s Five-Parameter Logistic Curve. *Bio-Rad Bulletin*, (Tech Note 3022).
- Gottschalk, P. G. and Dunn, J. R. (2005a). Determining the error of dose estimates and minimum and maximum acceptable concentrations from assays with nonlinear dose - Response curves. *Computer Methods and Programs in Biomedicine*, 80(3):204–215. 27
-

## BIBLIOGRAFÍA

---

- Gottschalk, P. G. and Dunn, J. R. (2005b). Measuring Parallelism, Linearity, and Relative Potency in Bioassay and Immunoassay Data. *Journal of Biopharmaceutical Statistics*, 15(3):437–463.
- Gottschalk, P. G. and Dunn, J. R. (2005c). The five-parameter logistic: A characterization and comparison with the four-parameter logistic. *Analytical Biochemistry*, 343(1):54–65. [14](#)
- Gran, J., Wasmuth, L., Amundsen, E., Lindqvist, B., and Aalen, O. (2009). Growth rates in epidemic models: application to a model for HIV/AIDS progression. *Statistics in medicine*, 28(July 2006):221–239.
- Hand, D. J. (2009). Measuring classifier performance: a coherent alternative to the area under the ROC curve. *Machine Learning*, 77(1):103–123.
- Harel, O., Perkins, N., and Schisterman, E. F. (2014). The Use of Multiple Imputation for Data Subject to Limits of Detection. *Sri Lankan Journal of Applied Statistics*, 5(4):227–246.
- Hendriks, J., Stojanovic, I., Schasfoort, R. B. M., Saris, D. B. F., and Karperien, M. (2018). Nanoparticle Enhancement Cascade for Sensitive Multiplex Measurements of Biomarkers in Complex Fluids with Surface Plasmon Resonance Imaging. *Analytical Chemistry*, 90(11):6563–6571. [8](#)
- Hill, H. (2009). Development of bioanalysis: a short history. *Bioanalysis*, 1(1):3–7. [4](#)
- Hodges, J. and Lehmann, E. (2005). *Basic Concepts of Probability and Statistics*. Society for Industrial and Applied Mathematics, Philadelphia.
- Holland, P. W. (2002). Two Measures of Change in the Gaps Between the CDFs of Test-Score Distributions. *Journal of Educational and Behavioral Statistics*, 27(1):3–17.
- Holstein, C. A., Griffin, M., Hong, J., and Sampson, P. D. (2015). Statistical Method for Determining and Comparing Limits of Detection of Bioassays. *Analytical Chemistry*, 87(19):9795–9801. [27](#)
- Horan, P. K. and Wheelless, L. L. (1977). Quantitative single cell analysis and sorting. *Science (New York, N.Y.)*, 198(4313):149–57. [6](#)
- Hornung, R. W. and Reed, L. D. (1990). Luminex xMAP Technology Bring your biomarkers to life with the established leaders in instrumentation and software. *Applied Occupational and Environmental Hygiene*, 5(1):46–51.
- Howson, C. P., Kinney, M. V., and Lawn, J. E. (2012). Born too soon: the global action report on preterm birth. Technical report, New York, NY.
- Irizarry, R. A. and Love, M. I. (2015). *Data Analysis for the Life Sciences*. Leanpub.

- 
- Jackson, T. M. and Ekins, R. P. (1986). Theoretical limitations on immunoassay sensitivity. Current practice and potential advantages of fluorescent Eu<sup>3+</sup> chelates as non-radioisotopic tracers. *Journal of immunological methods*, 87(1):13–20. 18
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer Texts in Statistics. Springer New York, New York, NY, 7 edition. 49
- Janitza, S. and Hornung, R. (2018). On the overestimation of random forest’s out-of-bag error. *PLOS ONE*, 13(8):e0201904. 55
- Jin, Y., Hein, M. J., Deddens, J. A., and Hines, C. J. (2011). Analysis of lognormally distributed exposure data with repeated measures and values below the limit of detection using SAS. *Annals of Occupational Hygiene*, 55(1):97–112.
- John W. Pratt (2014). Remarks on Zeros and Ties in the Wilcoxon Signed Rank Procedures. *Journal of the American Statistical Association*, 54(287):655–667.
- John Wiley & Sons, I. (2014). Two-Sample Problem, Baumgartner-Weiss-Schindler Test Tests. (1):1–3.
- Kaeding, M. (2015). *Bayesian Analysis of Failure Time Data Using P-Splines*. Springer Fachmedien Wiesbaden, Wiesbaden.
- Kamali-Moghaddam, M., Pettersson, F. E., Wu, D., Englund, H., Darmanis, S., Lord, A., Tavoosidana, G., Sehlin, D., Gustafsdottir, S., Nilsson, L. N. G., Lannfelt, L., and Landegren, U. (2010). Sensitive detection of A $\beta$  protofibrils by proximity ligation - relevance for Alzheimer’s disease. *BMC Neuroscience*, 11(1). 9
- Kanzow, C., Yamashita, N., and Fukushima, M. (2005). Levenberg-Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *Journal of Computational and Applied Mathematics*, 173(2):321–343.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Source Journal of the American Statistical Association*, 53(282):457–481. 37
- Kay, C., Nix, A. B. J., Kemp, K. W., Rowlands, R. J., Richards, G., Groom, G. V., Griffiths, K., and Wilson, D. W. (1986). Dose-interpolation of immunoassay data: Uncertainties associated with curve-fitting. *Statistics in Medicine*, 5(2):183–192.
- Khardani, S. and Semmar, S. (2014). Nonparametric conditional density estimation for censored data based on a recursive kernel. *Electronic Journal of Statistics*, 8(2):2541–2556.
- Kleinbaum, D. G. and Klein, M. (2005). *Survival analysis : a self-learning text*. Springer. 43
-

## BIBLIOGRAFÍA

---

- Klenin, K. V., Kusnezow, W., and Langowski, J. (2005). Kinetics of protein binding in solid-phase immunoassays: Theory. *Journal of Chemical Physics*, 122(21). 17
- Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuar.* 61
- Kupcova Skalnikova, H., Cizkova, J., Cervenka, J., and Vodicka, P. (2017). Advances in Proteomic Techniques for Cytokine Analysis: Focus on Melanoma Research. *International journal of molecular sciences*, 18(12).
- Landry, J. P., Fei, Y., and Zhu, X. (2012). Simultaneous Measurement of 10,000 Protein-Ligand Affinity Constants Using Microarray-Based Kinetic Constant Assays. *ASSAY and Drug Development Technologies*, 10(3):250–259. 23
- Lee, M., Kong, L., and Weissfeld, L. (2012). Multiple imputation for left-censored biomarker data based on Gibbs sampling method. *Statistics in Medicine*, 31(17):1838–1848.
- Lehman, S. Y. (1961). Exact and Approximate Distributions for the Wilcoxon Statistic with Ties. *Journal of the American Statistical Association*, 56(294):293.
- Lehmann, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. HOLDEN-DAY, INC., San Francisco. 40
- Lequin, R. M. (2005). Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). *Clinical chemistry*, 51(12):2415–2418. 5
- Liaw, A. and Wiener, M. (2002). Classification and Regression by randomForest. *R news*, 2(December):18–22. 47
- Life Technologies (2014). Moving from ELISA technology to the MAGPIX ® system for multiplex sample analysis.
- Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*.
- Luminex (2010). Software User Manual.
- Luminex (2011). MAGPIX Hardware Installation and User Manual.
- Martinez-Cambolor, P., Carleos, C., and Corral, N. (2012). Sobre el estadístico de Cramér-von Mises. *Rev. Mat*, 19(1):89–101.
- Martinez-Perdiguero, J., Retolaza, A., Bujanda, L., and Merino, S. (2014). Surface plasmon resonance immunoassay for the detection of the TNF $\alpha$  biomarker in human serum. *Talanta*, 119:492–497. 8
- Matson, R. S. (2009). *Microarray Methods and Protocols*. CRC Press. 11
- Mendel, C. M. and Mendel, D. B. (1985). 'Non-specific' binding. The problem, and a solution. *The Biochemical journal*, 228(1):269–72. 23

- 
- Meraz Cruz, N., Beltrán, J., Hernández, M. A., Bermejo Martínez, L., and Vadillo Ortega, F. (1994). Interleucina 1-alfa y prostaglandina E2 como mediadores moleculares en ruptura prematura de membranas / Interlukin-1-alfa and prostaglandin E2 underlying mechanism of premature rupture of membranes. *Perinatol. reprod. hum*, 8(4):222–6.
- Milliplex (2013). Improved analysis of multiplexed biomarker quantitation data with.
- Nadarajah, S. and Kotz, S. (2006). R Programs for Computing Truncated Distributions. *Journal of Statistical Software*, 16(August):1–8.
- Nadeau-Vallée, M., Obari, D., Quiniou, C., Lubell, W. D., Olson, D. M., Girard, S., and Chemtob, S. (2016). A critical role of interleukin-1 in preterm labor. *Cytokine & Growth Factor Reviews*, 28:37–51.
- Neuhauser, M. (2003). A note on the exact test based on the Baumgartner – Weiss – Schindler statistic in the presence of ties. *Computational Statistics & Data Analysis*, 42:561–568. [43](#), [63](#)
- Neuhäuser, M. (2011). *Nonparametric Statistical Tests: A Computational Approach*. CRC Press.
- Noether, G. E. (1987). Sample Size Determination for Some Common Nonparametric Tests. *J American Statistical Association*, 82(April 2013):645–647.
- Ortyn, W. E., Hall, B. E., George, T. C., Frost, K., Basiji, D. A., Perry, D. J., Zimmerman, C. A., Coder, D., and Morrissey, P. J. (2006). Sensitivity measurement and compensation in spectral imaging. *Cytometry Part A*, 69A(8):852–862. [21](#)
- O’Shea, J. J., Gadina, M., and Siegel, R. M. (2019). Cytokines and Cytokine Receptors. In *Clinical Immunology*, chapter Cytokines, pages 127–155.e1. Content Repository Only! [28](#)
- Pittaluga P., E., Díaz A., V., Mena N., P., and Corvalán V., S. (2002). Curva de crecimiento intrauterino para prematuros entre 23 a 36 semanas de edad gestacional. *Revista chilena de pediatría*, 73(2):135–141. [34](#)
- Pratt, J. W. and Gibbons, J. D. (1981). *Concepts of Nonparametric Theory*. Springer Series in Statistics. Springer New York, New York, NY.
- Raab, G. M. (1981). Estimation of a Variance Function, with Application to Immunoassay. *Applied Statistics*, 30(1):32. [14](#)
- Raab, G. M. (1983). Comparison of a logistic and a mass-action curve for radioimmunoassay data. *Clinical Chemistry*, 29(10):1757–1761.
- Rathore, D. P. S. (2007). Trends in the methods of measurement in analytical chemistry. *Exploration and Research for Atomic Minerals*, 17(February 2007):145–149.
-

## BIBLIOGRAFÍA

---

- Reverberi, R. and Reverberi, L. (2007). Factors affecting the antigen-antibody reaction. *Blood transfusion*, 5(4):227–40. [16](#)
- Rice A, J. (2007). *Mathematical Statistics and Data Analysis*. Duxbury, 3rd edition.
- Rich, R. R. (2008). *Clinical immunology : principles and practice*. Mosby/Elsevier. [18](#)
- Rigobon, R. and Stoker, T. M. (2007). Estimation with censored regressors: Basic issues. *International Economic Review*, 48(4):1441–1467.
- Rissin, D. M., Fournier, D. R., Piech, T., Kan, C. W., Campbell, T. G., Song, L., Chang, L., Rivnak, A. J., Patel, P. P., Provuncher, G. K., Ferrell, E. P., Howes, S. C., Pink, B. A., Minnehan, K. A., Wilson, D. H., and Duffy, D. C. (2011). Simultaneous Detection of Single Molecules and Singulated Ensembles of Molecules Enables Immunoassays with Broad Dynamic Range. *Analytical Chemistry*, 83(6):2279–2285. [9](#)
- Rissin, D. M., Kan, C. W., Campbell, T. G., Howes, S. C., Fournier, D. R., Song, L., Piech, T., Patel, P. P., Chang, L., Rivnak, A. J., Ferrell, E. P., Randall, J. D., Provuncher, G. K., Walt, D. R., and Duffy, D. C. (2010). Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nature Biotechnology*, 28(6):595–599. [18](#)
- Ritz, C., Baty, F., Streibig, J. C., and Gerhard, D. (2015). Dose-Response Analysis Using R. *PLOS ONE*, 10(12):e0146021.
- Rodbard, D., Munson, P. J., and De Lean, A. (1978). Improved curve-fitting, parallelism testing, characterization of sensitivity and specificity, validation, and optimization for radioligand assays. In *Radioimmunoassay and related procedures in medicine 1977*. [13](#), [14](#)
- Rodbard, D., Rayford, P. L., Cooper, J. A., and Ross, G. T. (1968). Statistical quality control of radioimmunoassays. *The Journal of clinical endocrinology and metabolism*, 28(10):1412–1418. [13](#)
- Sadana, A. and Ram, A. B. (1994). Fractal Analysis of Antigen-Antibody Binding Kinetics: Biosensor Applications. *Biotechnology Progress*, 10(3):291–298. [17](#)
- Sadler, W. A. and Smith, M. H. (1985). Estimation of the response-error relationship in immunoassay. *Clinical Chemistry*, 31(11):1802–1805.
- Schwenk, J. M., Lindberg, J., Sundberg, M., Uhlén, M., and Nilsson, P. (2007). Determination of Binding Specificities in Highly Multiplexed Bead-based Assays for Antibody Proteomics. *Molecular & Cellular Proteomics*, 6(1):125–132.
- Secretaría de Salud (2013). Programa de Acción Específico Salud Materna y Perinatal 2013-2018. page 100.

- 
- Shah, J. S., Rai, S. N., DeFilippis, A. P., Hill, B. G., Bhatnagar, A., and Brock, G. N. (2017). Distribution based nearest neighbor imputation for truncated high dimensional data with applications to pre-clinical and clinical metabolomics studies. *BMC Bioinformatics*, 18(1):1–13. [38](#)
- Sotnikov, D. V., Zherdev, A. V., and Dzantiev, B. B. (2017). Mathematical modeling of bioassays. *Biochemistry (Moscow)*, 82(13):1744–1766.
- Thomson, A. W. and Lotze, M. T. (2003). *The Cytokine Handbook*. Number v. 1-2 in The Cytokine Handbook. Academic Press, 4 edition. [28](#), [29](#)
- Tighe, P. J., Ryder, R. R., Todd, I., and Fairclough, L. C. (2015). ELISA in the multiplex era: Potentials and pitfalls. *PROTEOMICS - Clinical Applications*, 9(3-4):406–422.
- Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26(1):24. [37](#)
- Wild, D. (2013). *The Immunoassay Handbook: Theory and applications of ligand binding, ELISA and related techniques*. [6](#), [18](#)
- Wingren, C. and Borrebaeck, C. A. (2007). Progress in miniaturization of protein arrays—a step closer to high-density nanoarrays. *Drug Discovery Today*, 12(19-20):813–819.
- Won, J.-H., Goldberger, O., Shen-Orr, S. S., Davis, M. M., and Olshen, R. A. (2012a). Significance analysis of xMap cytokine bead arrays. *Proceedings of the National Academy of Sciences of the United States of America*, 109(8):2848–53.
- Won, J.-H., Goldberger, O., Shen-Orr, S. S., Davis, M. M., and Olshen, R. A. (2012b). Significance analysis of xMap cytokine bead arrays (Annex). *Proceedings of the National Academy of Sciences*, 109(8):2848–2853.
- Woolley, C. F., Hayes, M. A., Mahanti, P., Douglass Gilman, S., and Taylor, T. (2015). Theoretical limitations of quantification for noncompetitive sandwich immunoassays. *Analytical and Bioanalytical Chemistry*, 407(28):8605–8615.
- Yalow, R. (1960). Immuno-assay of endogenous plasma insulin in man. *J. clin. Invest*, 39. [5](#)
- Yamanishi, C. D., Chiu, J. H. C., and Takayama, S. (2015). Systems for multiplexing homogeneous immunoassays. *Bioanalysis*, 7(12):1545–1556.
- Zhao, Y. D., Rahardja, D., and Qu, Y. (2008). Sample size calculation for the Wilcoxon – Mann – Whitney test adjusting for ties. *STATISTICS IN MEDICINE*, 27(September 2006):462–468.