

**UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE SALUD PÚBLICA**



COMPARACION DE DISTINTAS MEDICIONES DEL
CAMBIO EN EL INDICE DE MASA CORPORAL EN
LA INFANCIA EN SU PREDICCION DE LA ADIPOSIDAD
A LOS CINCO AÑOS

RODRIGO RETAMAL YERMANI

**TESIS PARA OPTAR AL GRADO DE
MAGISTER EN BIOESTADISTICA**

**Director de Tesis: Prof. SERGIO ALVARADO O.
Cotutor: Prof. CAMILA CORVALAN A.**

2011

Santiago, Septiembre de 2011

Agradecimientos

Quiero agradecer al profesor Sergio Alvarado, bajo cuya guía se realizó esta Tesis, por su inestimable apoyo y comprensión. Especialmente quiero agradecer a la profesora Camila Corvalán quien fue la impulsora de este trabajo y quien me prestó apoyo y comprensión en todo momento durante su realización.

A la comisión revisora, profesores Rosa Montaña y Gabriel Cavada, quienes apoyaron con su crítica la elaboración del presente trabajo.

Al Departamento de Antropología de la Universidad de Chile, por medio de su director, profesor Mauricio Uribe, por prestarme todo el apoyo y tiempo necesarios para la consecución de la presente Tesis.

Índice

Agradecimientos	2
RESUMEN	5
ABSTRACT	6
1. INTRODUCCIÓN	7
1.1. El Problema de la Obesidad infantil	9
1.2. La obesidad como fenómeno mórbido	10
1.3. Consecuencias de la obesidad infantil	11
1.4. Prevención.....	14
2. PREGUNTA DE INVESTIGACIÓN.....	22
3. OBJETIVO GENERAL.....	22
4. OBJETIVOS ESPECÍFICOS.....	23
5. HIPÓTESIS DE TRABAJO	23
6. MATERIAL	24
7. MÉTODOS	25
7.1. Análisis de datos.....	26
7.2. Variables de estudio	26
7.3. Variable respuesta.....	27
7.4. Variables predictivas	28
7.5. Análisis exploratorio de datos.....	29
7.6. Modelación estadística	29
7.7. Validación de los resultados obtenidos	34
7.8. Consentimiento ético	36
8. RESULTADOS	37
8.1. Exploración de datos	37
8.3. Construcción de los modelos predictivos.....	42
8.4. Validación de los modelos predictivos	59
9. DISCUSION.....	63
9.1 Modelación por GLM.....	64

9.2 Modelación por GAM	65
9.3 Validación de los modelos.....	67
10. CONCLUSIONES	72
11. BIBLIOGRAFÍA	73
ANEXO 1: CASOS FICTICIOS PARA EJEMPLIFICAR REGRESION GAM UTILIZANDO DBMI	82
ANEXO 2: GRAFICOS DE CONCORDANCIA ENTRE FMI OBSERVADO Y PREDICHO	83
ANEXO 3: GRAFICOS DE DISPERSION CON RECTA DE AJUSTE ENTRE FMI OBSERVADO Y PREDICHO	86

RESUMEN

El monitoreo temprano del crecimiento infantil puede servir para la detección temprana de niños en riesgo de presentar alteraciones de salud futuras como obesidad y sus complicaciones. Existe evidencia de que existirían periodos “críticos” en los que desviaciones de lo que se considera normal serían relevante, sin embargo en estos resultados puede influir la forma en la que se expresan los cambios de antropometría y los modelos estadísticos que se utilizan para evaluar las relaciones.

Por ende, el propósito del presente estudio fue comparar tres tipos de mediciones de cambio del Índice de Masa Corporal (IMC sin transformaciones, IMC estandarizado y percentil del IMC) y dos estrategias estadísticas diferentes (Modelos Lineales Generalizados y Modelos Aditivos Generalizados) utilizados para predecir el Índice de Masa Grasa ($\text{grasa}/\text{talla}^2 = \text{IMG}$). Para esto se contó con las mediciones del IMC a los seis y 24 meses de edad y el IMG a los 60 meses en una muestra de 379 niños chilenos (201 femeninos y 178 masculinos) del Estudio de Crecimiento y Obesidad (ECO), 2009.

Los resultados muestran que las mediciones transformadas del IMC (puntaje z y percentil) obtuvieron mejores estimaciones del IMG en comparación con el IMC sin transformar. Por otra parte, ningún modelo estadístico presentó mejores resultados que otro, sino que ambos pueden ser utilizados complementariamente, ya que mientras que los modelos lineales aportan con ecuaciones predictivas del IMG, los modelos aditivos informan sobre valores específicos de ganancia de IMC que contribuyen notoriamente al aumento del IMG.

En conclusión, el presente estudio muestra que la mejor forma de expresar los cambios del IMC para detectar oportunamente niños en riesgo de presentar obesidad temprana es estandarizándolos (ya sea como percentil o puntaje z) y que para evaluar estas asociaciones se pueden utilizar modelos lineales y/o aditivos dependiendo de cual sea el objetivo del análisis.

ABSTRACT

The early monitoring of child growth might be useful for early detection of children in risk to present future health alterations like obesity and their complications. There is evidence that would exist “critical” periods in which the deviations of what is considered normal would be relevant, however in their results might affect how the changes in anthropometry are expressed and the statistical models used to evaluate the relationships.

Therefore, the aim of the present study was to compare three types of measurements of change of Body Mass Index (BMI without transformations, standardized BMI and centile BMI) and two different statistical models (Generalized Linear Models and Generalized Additive Models) using to predict Fat Mass Index ($\text{fat}/\text{height}^2=\text{FMI}$). Measurements of BMI at six and 24 month-old and FMI at 60 month were used in a 379 Chilean children sample (201 female and 178 male) obtained from the Growth and Obesity Study (ECO).

The results show that transformed measurements of BMI (z-score and centile) had better estimations of FMI compared to the BMI without transformations. Besides, no statistical model gave better results than the other, but both could be used complementarily, since while the linear models give predictive equations, the additive models inform about specific values of gain of BMI that contribute notoriously to increasing of FMI.

In conclusion, the present study shows that the best manner to express the changes of BMI to detect opportunely children in risk of have early obesity is standarizing it (centile or z-score) and for evaluating these associations would be useful apply linear models or additives, depending the objective of the analysis.

1. INTRODUCCIÓN

El objetivo principal del presente estudio es contribuir a clarificar las metodologías estadísticas necesarias para identificar patrones de crecimiento infantil que se asocien a mayor adiposidad y riesgos de salud futuros. Hasta el momento, uno de los hallazgos más relevantes al respecto muestra que la ganancia excesiva de adiposidad, medida indirectamente como Índice de Masa Corporal (peso en kg/ talla en m^2 ; en inglés *Body Mass Index*; en adelante BMI) en los primeros dos años de vida, sería un predictor importante de la composición corporal y de los niveles metabólicos en edad preescolar y en etapas futuras (Baird *et al.*, 2005; Leunissen *et al.*, 2009). Sin embargo, los cambios del BMI pueden expresarse de diferentes formas: como cambios del BMI crudo, como cambios de percentiles del BMI (en adelante PBMI), como cambios porcentuales del BMI (en adelante %BMI) o de los puntajes estandarizados del BMI (en inglés *BMI-for-age z-score*; en adelante BAZ), los cuales se realizan de acuerdo a un grupo de referencia. La relevancia en la selección correcta del indicador de BMI estriba en que el grado de dispersión de la variable varía según el indicador. Como resultado de estas diferencias en dispersión, observamos que la magnitud de las asociaciones entre los diferentes indicadores de BMI y una determinada variable respuesta (como por ejemplo la grasa corporal) también son variables. Por otra parte, las ganancias de peso y talla (cuya relación es la que expresa el BMI) no son siempre proporcionales en este periodo de la vida. En un primer momento, la ganancia de peso supera a la ganancia de talla, lo que se traduce en un mayor BMI, mientras que en una segunda etapa la ganancia de peso cae manteniéndose estable la ganancia de talla, lo que se traduce en una disminución del BMI. En una tercera instancia, la ganancia de peso nuevamente excede la velocidad de ganancia de talla con el consiguiente aumento del BMI en lo que se ha denominado “rebote adiposo”. Dado este patrón de variabilidad del BMI en este periodo, se plantea que el uso de modelos que permitan mayor “flexibilidad” al modelamiento de los datos podría mejorar la predictibilidad por sobre los modelos lineales.

Es por esto que la presente Tesis de Magíster en Bioestadística tiene dos propósitos. El primero es comparar tres formas de medición de cambio del BMI entre los seis y 24 meses de edad en la predictibilidad de la adiposidad corporal, medida a través del FMI a los cinco años, utilizando la información disponible del Estudio de Crecimiento y Obesidad (Proyecto ECO: Corvalán *et al.* 2009, Kain *et al.* 2009). El segundo objetivo es comparar el rendimiento de dos procedimientos analíticos distintos sobre estos mismos datos: Modelos Lineales Generalizados (en inglés *Generalized Linear Models*; en adelante GLM) y Modelos Aditivos Generalizados (en inglés *Generalized Additive Models*; en adelante GAM). El seguimiento de la cohorte ECO sigue actualmente en curso (edad actual de los niños= 10 años), por lo que los resultados de esta Tesis podrían ser utilizados para evaluar de mejor forma las relaciones entre crecimiento infantil y adiposidad en estos niños y de esta forma, contribuir a identificar potenciales periodos de la vida temprana susceptibles de intervención para prevenir obesidad infantil en nuestro país. En este sentido, la presente Tesis pretende ser un aporte a nivel epidemiológico en el área de estudio de crecimiento y salud.

En este primer capítulo se describen los Antecedentes del estudio. El segundo capítulo plantea el Problema de Investigación de la presente Tesis, enmarcado desde un enfoque biomédico y estadístico. El tercer y cuarto capítulos exponen el Objetivos General y los Objetivos Específicos respectivamente. El quinto capítulo plantea las Hipótesis de trabajo que guiaron el presente trabajo. El sexto capítulo describe el Material utilizado, en tanto que el séptimo capítulo expone los Métodos usados en el presente estudio. El octavo capítulo presenta los resultados obtenidos en el presente trabajo, en tanto que en el noveno capítulo se realiza la Discusión de los resultados obtenidos. Finalmente, el décimo capítulo presenta las Conclusiones de la presente Tesis de Magíster.

1.1. El Problema de la Obesidad infantil

La obesidad infantil ha llegado a ser la enfermedad nutricional más prevalente en los países desarrollados, incrementándose notoriamente durante las últimas tres décadas (Dietz, 2002). La rapidez de estos cambios en un tiempo tan corto ha llevado a tratarlo como un fenómeno epidémico (Cameron *et al.*, 2003). Hasta hace poco tiempo, la obesidad infantil era vista como un problema estético, en donde los mayores riesgos asociados eran las eventuales consecuencias de su persistencia en la adultez (Dietz, 2002). Sin embargo, actualmente se han observado riesgos significativos en la salud asociados a la obesidad infantil, tales como la aparición temprana de complicaciones metabólicas y cardiovasculares, maduración acelerada y problemas psicológicos importantes, entre otros (Dietz, 2002).

Una definición simple de obesidad es el exceso de grasa corporal (Cole y Rolland-Cachera, 2002). El criterio para estimar sobrepeso y obesidad se asocia al exceso de adiposidad sobre algún punto de corte especificado previamente y basado idealmente en asociación a factores de riesgo, tales como la morbilidad actual y la mortalidad actual y tardía. Debido a la naturaleza altamente dinámica de la infancia y la adolescencia, la definición y punto de corte del sobrepeso y la obesidad deben ser específicos a la edad, a diferencia de los individuos adultos (Cole y Rolland-Cachera, *op. cit.*). A esa edad, la gran mayoría de las veces se utilizan criterios estadísticos, basado en medidas porcentuales o desviaciones estándar para definir el exceso de grasa corporal, dada la dificultad de realizar estudios para definir puntos de corte.

En Chile, la obesidad constituye el principal problema nutricional de la población infantil (Burrows *et al.*, 2001). Kain *et al.* (2007), analizando las bases de datos antropométricas de los asistentes a jardines infantiles de la Junta Nacional de Jardines Infantiles (JUNJI) durante la última década, estimaron que la prevalencia de la obesidad en preescolares ($PBMI \geq 95$; NCHS/CDC 2000 www.cdc.org), alcanza el 18%, especialmente en el segmento de tres a cinco años de edad. Datos

del año 2008 de la Junta Nacional de Auxilio Escolar y Becas (JUNAEB), muestran una prevalencia de obesidad (peso para la talla ≥ 2 desviaciones estándar; OMS 2007) por sobre el 20% en niños y niñas al momento de ingresar al colegio (http://www.junaeb.cl/prontus_junaeb/site/artic). Estas prevalencias superan casi dos veces lo reportado en otros países latinoamericanos (de Onis, 2000) y son comparables a las reportadas en países desarrollados (Wang y Lobstein, 2006). Además, el sostenido incremento de la obesidad y el sobrepeso, además del impacto en morbilidad en nuestra población y sus efectos económicos en nuestro país, motivan un estudio exhaustivo sobre este fenómeno.

1.2. La obesidad como fenómeno mórbido

La obesidad infantil está determinada por factores genéticos, conductuales, ambientales-sociales y su interacción entre ellos (Martínez, 2000). Si bien diversos estudios han mostrado que la obesidad no puede explicarse únicamente por factores ambientales (Schousboe *et al.*, 2003; Santos *et al.*, 2005; Bell *et al.*, 2005), es evidente que el notorio aumento de la prevalencia de obesidad en los últimos años no es atribuible únicamente a factores genéticos, sino que se debe más bien a cambios socio-culturales que han inducido una mayor ingesta de alimentos altamente energéticos y a un progresivo descenso de la actividad física (Hill, *et al.*, 2003; Stein y Colditz, 2004). Por ello, son precisamente los factores conductuales y socio-culturales aquellos en los que se ha centrado mayormente las acciones en prevención y control de la obesidad, además de la posibilidad de ser modificados, a diferencia de los factores genéticos que no pueden ser modificados. Dentro de los factores de riesgo modificables de obesidad infantil se encuentran un peso al nacimiento bajo los 2500g o sobre 4000g, obesidad materna pre-gestacional y ganancia de peso excesiva durante el embarazo, diabetes gestacional, tabaquismo durante el embarazo, menor duración de la lactancia materna e introducción temprana de sólidos, edad más temprana del rebote de adiposidad, menor nivel socioeconómico, etnicidad, consumo elevado de alimentos de alto contenido

energético y horas de sueño (siesta y sueño nocturno), sólo por mencionar los más nombrados en la literatura (Agras *et al.*, 2004; Reilly *et al.*, 2005).

En relación a la dieta, los factores que han generado mayor impacto en el aumento de la prevalencia de la obesidad infantil son la enorme disponibilidad de alimentos altamente energéticos y la reducción del tiempo de los padres en la preparación de alimentos saludables (Lobstein *et al.* 2004). Es posible que también contribuya la aplicación de estándares internacionales de recomendación energética que no estén basados en estudios de gasto energético, provocando una sobreestimación de alrededor de un 20% del aporte de energía requerido por los niños en diferentes situaciones (ej. FAO/OMS/UNU 1985, FAO 2004).

En relación a la actividad física, se ha observado en países desarrollados que en la actualidad los niños caminan menos que los niños de décadas anteriores (Gavarry *et al.*, 2003), en tanto que se ha mostrado en diversos trabajos el rol de la televisión como un factor de riesgo de prevalencia de obesidad (Crespo *et al.*, 2001; Saelens *et al.*, 2002).

1.3. Consecuencias de la obesidad infantil

Si bien los problemas cosméticos o de estética física son de poca importancia clínica, pueden tener un impacto negativo en la imagen corporal de los niños e interfieren en su diario vivir. En ambos sexos, la obesidad es muy desfigurativa (Zwiauwer *et al.*, 2002). En relación al exceso de tejido adiposo, los problemas más complejos son la pseudo-ginecomastia y el pseudo-hipogenitalismo. Las mamas parecen ser prominentes en las niñas, en tanto en los niños se observa una aparente hipoplasia de los genitales y ginecomastia, indicando problemas hormonales. La aparente hipoplasia se debe a que los genitales son “enterrados” por la grasa abdominal inferior, pareciendo desproporcionadamente más pequeños en consideración con el tamaño total del niño. En la obesidad abdominal, el

abdomen se comporta como un péndulo, aparecen estrías blancas o púrpuras en el abdomen y alrededor de las mamas. En las niñas obesas, el hirsutismo a menudo está presente. Por otra parte, se ha observado una alta proporción de niños obesos con *Acantosis nigricans*, una rara condición cutánea que ocurre en asociación con una amplia variedad de patologías (Zwiauer *et al.*, 2002).

Entre los desórdenes hormonales, se ha documentado una alta actividad de andrógenos en niñas con sobrepeso y se observa una correlación positiva entre el índice de andrógeno libre y la razón peso/talla (Wabitsch *et al.*, 1995) y entre los niveles de testosterona y el BMI (Malecka-Tendera *et al.*, 1998). Respecto a los problemas ortopédicos, los cartílagos inmaduros, las placas de crecimiento sin fusionar y los huesos cartilaginosos sin desarrollo completo, contribuyen a una variedad de desórdenes ortopédicos, siendo particularmente las más reconocidas, la enfermedad de Blount y la epifisiolisis (*Split femoral epiphysis*, Zwiauer *et al.*, 2002).

En cuanto a las enfermedades gastrointestinales, las altas concentraciones de enzimas hepáticas son comunes en niños y jóvenes obesos (Kinugasa *et al.*, 1984; Noguchi *et al.*, 1995), así como el hígado graso (Moran *et al.*, 1983; Frelut *et al.*, 1995,1996) y la elevada proporción de colesterol excretado por la bilis (Shaffer y Small, 1977).

Respecto a las enfermedades respiratorias y del sueño, la obesidad infantil contribuye a su aparición de muchas maneras. El sobrepeso se asocia significativamente a individuos asmáticos (Luder *et al.*, 1998), mientras que el asma es un factor de riesgo en niños obesos con bajo nivel de actividad física (Gennuso *et al.*, 1998). Pruebas de función pulmonar muestran una alta incidencia de obstrucción pulmonar benigna (*mild obstructive lung*) en niños obesos, lo que puede predisponer al síndrome de hipoventilación adulta (Zwiauer *et al.*, 2002). Por otra parte los desórdenes del sueño parecen tener relación con el grado de obesidad. En la mayoría de los niños afectados, los problemas se manifiestan básicamente

por períodos de ronquidos y episodios del síndrome de apnea del sueño, encontrándose un tercio de esta patología en niños con obesidad (Mallory *et al.*, 1989).

Respecto a los problemas neurológicos, se ha observado que el *Pseudotumor Cerebri* (hipertensión intracraneal benigna) se asocia a obesidad, hallándose este síndrome sobre un 50% de niños obesos (Zwiauwer *et al.*, 2002).

En cuanto a los desórdenes inmunológicos, algunos estudios muestran que los infantes presentan infecciones respiratorias más frecuentes que lo esperado (Tracy *et al.*, 1971) o bien son más severas y duraderas (Caroli y Chandra, en prensa). La alta mortalidad de niños obesos se ha asociado a la hiperinsulinemia, la cual reduce la citotoxicidad mediada por linfocitos (Koffler *et al.*, 1991) y la leptina, la cual modula la función inmune de las células T (Lord *et al.*, 1998).

Además, la obesidad infantil está asociada a una variedad de desórdenes metabólicos y endocrinos, los cuales son precursores de condiciones asociadas a un incremento en la morbilidad y mortalidad en el adulto, tales como el riesgo cardiovascular, aterosclerosis coronaria, hiperinsulinemia, hiperandrogenia, resistencia a la insulina, diabetes *mellitus* y dislipidemia, entre otras. Se ha observado que la edad de la pubertad entre niños europeos y estadounidenses ha ido disminuyendo sostenidamente durante el último siglo, en paralelo con el incremento del promedio de grasa corporal, posiblemente asociado al incremento en las concentraciones de leptina. Por lo tanto, la maduración temprana parece ser un determinante de la obesidad tardía (Zwiauwer *et al.*, 2002).

Por otra parte, los problemas sicosociales son comunes en los niños obesos y varían ampliamente en severidad. Niños miembros de familias obesas, presentan pocos problemas psicológicos (Doherty y Harkaway 1990). Sin embargo, al empezar la escuela y especialmente durante la adolescencia, los niños obesos deben enfrentar una sociedad que los percibe como estúpidos (Zwiauwer *et al.*, 2002),

siendo en algunos casos más rechazados que otros niños con discapacidades severas (Richardson *et al.*, 1961). Los niños obesos en esta época presentan a menudo una baja autoestima y dificultades en las relaciones con niños no obesos (Isnard-Mugnier *et al.*, 1993; Wardle *et al.*, 1995). Consecuentemente, los niños y adolescentes obesos muestran altos niveles de ansiedad y una imagen corporal disturbada (Kimm *et al.*, 1991).

Respecto a las enfermedades cardíacas, en un estudio basado en la cohorte de Boyd Orr, los niños que presentaron un percentil sobre 75 en BMI presentaron mayor riesgo en mortalidad por enfermedades isquémicas cardíacas que aquellos ubicados entre los percentiles 25 y 49 (Franks *et al.*, 2010; Gunnell *et al.*, 1998). Recientemente, un estudio con cerca de 5000 niños americanos-indios demostró que la obesidad infantil se asocia con mayor muerte antes de los 55 años de vida (Zheng *et al.*, 2000).

Muchos de estos problemas se mantienen en la adultez, presentando consecuencias más complejas en los sujetos afectados, debido a que la obesidad es un factor de riesgo de muchas de las patologías comentadas anteriormente (Zwiauer *et al.*, 2002).

1.4. Prevención

Debido a lo anterior, hay muchas razones para prevenir la obesidad en niños. Básicamente, la prevalencia de la obesidad es alta y sigue aumentando, la obesidad infantil conduce a la obesidad adulta, limita la actividad física, está asociada con desventajas sicosociales, es factor de riesgo de muchas enfermedades y es difícil de tratar exitosamente (Lissau *et al.*, 2002).

El tratamiento efectivo en los niños afectados y la prevención de la obesidad en niños susceptibles debe convertirse en una política prioritaria. Como ha sucedido

con otras enfermedades, las modificaciones ambientales pueden representar las intervenciones más efectivas, económicas y duraderas. Por otra parte, debe desarrollarse y evaluarse la infraestructura necesaria para promover la actividad física (Dietz, 2002).

Sin embargo, la prevención también presenta sus desafíos. La epidemia de la obesidad aún no es vista con la urgencia que demanda. Los pediatras se encuentran pobremente equipados para tratar la obesidad y los métodos de atención primaria aún se encuentran en fase experimental. Deben establecerse métodos efectivos para mantener el peso en los niños que ganan peso muy rápido y reducir el peso en aquellos con sobrepeso (Dietz, 2002).

La primera fase de prevención de la obesidad es su detección. Con este fin, se han generado distintos indicadores que detectan en mayor o menor grado, la obesidad temprana.

1.4.1. BMI y grasa corporal

Como se dijo anteriormente, la obesidad es un exceso de grasa corporal, por lo que el BMI intenta estimar la grasa corporal presente en el cuerpo. El punto crucial, entonces es la estimación de la grasa corporal. Se asume que la composición corporal está dividida en grasa y libre de grasa. La medición ideal de obesidad debiera considerar evaluar directamente la presencia, cantidad y distribución del tejido graso. Para su medición se ha recurrido a diferentes técnicas, tales como el uso de medidas antropométricas, hidrodensitometría, dilución isotópica y absorciometría de energía dual por rayos X. La adiposidad regional y visceral se ha medido por tomografía computada, resonancia magnética o DXA (Must *et al.* 1992; Must y Anderson, 2005). Si bien estas técnicas de laboratorio ofrecen mediciones más precisas que la antropometría, la cual se verá a continuación, ninguna de ellas es ampliamente utilizada, debido a términos de costo, requerimientos técnicos,

invasividad, necesidad de cooperación del sujeto y en algunos casos (como la tomografía computada), a riesgos en la salud. Otros métodos, como la impedancia bioeléctrica, han ganado un amplio uso, debido a que no son invasivos y son rápidos. Por esta razón, tradicionalmente se han utilizado indicadores antropométricos, tales como el peso, la talla, los pliegues cutáneos y la circunferencia de la cintura, entre otros, que son mínimamente invasivos y de bajo costo, con el fin de evaluar indirectamente la presencia de obesidad. En el caso particular del peso y la talla, estas variables se utilizan vinculadas. Algunos índices que las relacionan son la relación peso/talla, el BMI o índice de Quetelet (Garrow y Webster, 1985), el cual se utiliza también en adultos y el índice de Rohrer (peso/talla³; Cole, 2002).

La validez estadística del BMI ha sido estudiada numerosamente. Se ha mostrado que este índice ajusta satisfactoriamente el peso para la estatura en la mayoría de los grupos etarios, sobre todo en niños (Poskitt, 2000; OMS 1995). Sin embargo, algunos autores afirman que este indicador presenta debilidades predictivas, ya que no es un indicador directo de la adiposidad, por lo que no refleja su variabilidad total, además de no distinguir claramente entre masa grasa y masa magra, lo que disminuye su valor epidemiológico (Cole, 2006; Wells, 2002). Debido a lo anterior, en diferentes estudios se ha preferido estimar el FMI en vez del BMI, debido a que los factores de riesgo asociados a la obesidad se derivan del exceso de grasa más que del peso corporal, particularmente de la grasa abdominal y visceral (Gasteyger y Temblay, 2002; Lamarche, 1998).

Además, para proveer una definición aceptable de sobrepeso y obesidad, deben tenerse dos cosas presentes: la población de referencia y el punto de corte del BMI. En relación a la población de referencia, existen varias referencias nacionales e internacionales. Dentro de las referencias internacionales que presentan información sobre el BMI en la niñez, se encuentra la *International Obesity Taskforce* (IOTF), *Centers for Disease Control and Prevention* (CDC) y la Organización Mundial de la Salud (OMS 2006 y 2007). La IOTF en 1997 organizó

un encuentro para definir sobrepeso y obesidad infantil. Utilizando datos longitudinales, se establecieron puntos de corte para cada edad (PBMI para la edad) que corresponden al valor BMI de 25 kg/m² para sobrepeso y 30 kg/m² para obesidad en la adultez (Cole *et al.*, 2000; Bellizi y Dietz, 1997). Otras de la referencias utilizadas corresponden a la CDC, que considera como obesidad el valor BAZ (PBMI= 95) y la nueva referencia de la OMS que se denomina “estándar” de crecimiento, ya que se basa en la distribución de crecimiento de una muestra de niños sanos representantes de los cinco continentes, alimentados de pecho y luego de acuerdo a normas de alimentación complementaria (Kuczmarski *et al.* 2000; WHO, 2006).

En general, el sobrepeso y la obesidad se definen sobre la base del peso relativo a la estatura, usando puntos de corte a partir del BMI (Cole *et al.*, 2000). En adultos, los valores de BMI 25 y 30 kg/m² han sido utilizados para categorizar el sobrepeso y la obesidad respectivamente (Cole, *op. cit.*). Sin embargo, en niños, debido a la falta de un punto de corte objetivo y científico, se ha debido recurrir al uso del PBMI (Cole *et al.*, 2000), estableciéndose puntos de cortes arbitrarios y disímiles entre investigadores, lo que impide el análisis comparativo (Cole, 2006). La IOTF identifica un PBMI de 91 como signo de sobrepeso y un PBMI de 98 como signo de obesidad (Cameron *et al.*, 2006). En Estados Unidos, en cambio, utilizan datos basados en tablas nacionales, en donde un PBMI de 85 es criterio de “peligro de obesidad” y un PBMI de 95 como criterio de obesidad (Cameron *et al.*, 2006). La OMS recomienda el uso de más dos desviaciones estándar del BMI para la edad, para definir sobrepeso y más de tres desviaciones estándares para definir obesidad. Sin embargo en la literatura científica habitualmente se utiliza el criterio de mayor de una desviación estándar para sobrepeso y mayor a dos desviaciones estándar para obesidad, lo que actualmente se utiliza en población chilena (MINSAL, 2007).

Otra alternativa intermedia entre utilizar mediciones directas de tejido graso y usar antropometría es desarrollar ecuaciones antropométricas (que consideran varios

indicadores antropométricos a la vez), a partir de mediciones directas de composición corporal, como el uso de con agua marcada, impedanciometría, etc. Este caso, sin embargo, también presenta dificultades de definir el mejor indicador, la población de referencia y sus puntos de corte. Tradicionalmente, se ha utilizado como indicador el porcentaje de masa grasa (masa grasa medida en kilogramos/peso corporal total medido en kilogramos). Sin embargo, recientemente se ha mostrado que éste es un indicador sesgado de la grasa corporal en niños, aún en presencia de ajustes, que arrojan estimaciones poco precisas, debido posiblemente a que no es un indicador independiente de la grasa corporal (Wells, 2002). Debido a lo anterior, muchas investigaciones han optado por el uso del FMI (Van Itallie *et al.*, 1990), ya que presenta una serie de beneficios predictivos y de ajustes, estimándose como el mejor indicador de la grasa corporal en niños de uno a once años de edad (Wells *et al.*, 2002).

1.4.2. Cambios de BMI en la niñez

En los niños, el exceso de grasa corporal o adiposidad se asocia a los cambios de adiposidad normal durante el crecimiento, lo que agrega aún mayor complejidad a su estudio. Para definir adiposidad, primero se necesita establecer qué se entiende por adiposidad normal para cada edad y luego establecer el punto de corte respecto a la identificación de obesidad. Idealmente, un *Gold Standard* se realizaría a partir de una cohorte, permitiendo su identificación temprana. Este punto de corte depende principalmente de la edad del niño, por lo que es necesario el conocimiento de la relación de la adiposidad entre los niños, además del patrón individual en cada uno de ellos.

Debido al carácter dinámico del crecimiento y desarrollo que afecta a niños y jóvenes, es necesario un seguimiento durante todas sus etapas, lo que ha motivado el desarrollo de estudios de trayectorias del peso o del BMI durante la infancia y adolescencia, con el fin de lograr predecir la obesidad tardía. El seguimiento del

BMI desde el nacimiento a la adultez es crucial para entender y detectar oportunamente la obesidad.

También muchos estudios se han enfocado en la prevención de la ganancia de peso, permitiendo al niño llegar a ser más delgado a medida que aumenta de estatura. Con esta finalidad, es necesario estimar el cambio de peso a corto plazo (Barlow *et al.*, 2002). Como el BMI ha sido uno de los indicadores más utilizados para medir la adiposidad (Zemel *et al.*, 1997; Ellis, 2000), este podría ser utilizado como indicador del cambio de adiposidad. Los cambios del BMI pueden ser expresados como cambios del BMI crudo, como cambio porcentual (%BMI), como cambio del BMI estandarizado por edad (BAZ) o por cambio de los percentiles (PBMI) (Law *et al.*, 1992; Zannolli and Morgese, 1996; Pietrobelli *et al.*, 1998; Schroeder *et al.*, 1999; Ong *et al.*, 2000; Maynard *et al.*, 2001; Cameron *et al.*, 2003).

Sin embargo, no existen muchos estudios que comparen la variabilidad de estas diferentes mediciones en los niños dependiendo del grado de adiposidad. Es posible que la dispersión de los datos varíe, dependiendo del indicador de cambio del BMI que se utilice y por ende, también varíe el grado de asociación que tenga este predictor sobre variables respuestas futuras. Por ejemplo, el cambio del %BMI es más grande que el BMI en niños obesos en comparación con niños delgados. Por otra parte, las curvas de percentiles de las tablas de BMI (Kuczmarsky *et al.*, 2000), están notoriamente más apartadas en los percentiles más altos, debido a la asimetría de la distribución de BMI. Por ejemplo, en niños de siete años de edad, el rango entre el percentil 90 a 97 es de 2 kg/m^2 , mientras que el percentil tercero a décimo es 0.5 kg/m^2 , es decir, aproximadamente cuatro veces mayor. Además, los cambios en la escala de percentil son mucho mayores en la cola de distribución del BMI que en el centro. Para niños de siete años de edad, siete puntos percentiles cerca de la mediana (percentiles 46.5 a 53.5) cubren un rango del BMI de 0.25 kg/m^2 , la mitad en comparación con el percentil tercero a décimo y un octavo en comparación con el rango percentil 90-97. Por lo tanto, la selección del indicador de

BMI que utilizamos para evaluar crecimiento tiene especial relevancia en poblaciones en que tengamos altas prevalencias de desnutrición u obesidad.

Con el propósito de hallar cuál es la medición más apropiada para medir el cambio en la adiposidad, es necesario conocer cómo el BMI varía en el tiempo en niños normales (Cole *et al.*, 2005). Un importante principio es que la medida a corto plazo dentro de cada niño debería ser la misma, sin importar la adiposidad de cada niño, por lo cual este principio podría ser empleado como medida de ganancia. Sin embargo, este principio no se aplica para las distintas mediciones del BMI, ya que %BMI y BAZ atenúan el cambio del BMI en los niños más obesos y la ganancia del PBMI atenúa los cambios del BMI tanto en niños obesos como no obesos, reduciendo la variabilidad de estas mediciones y generando posibles problemas de sesgo, que podrían sobrestimar o subestimar la adiposidad de la población de estudio.

Al respecto, Cole y colaboradores (2005) realizaron un estudio en una muestra de 135 niños italianos de kindergarten, a quienes midieron el BMI en tres oportunidades, entre los 29 y 68 meses de edad en intervalos regulares de nueve meses. Su propósito era comparar cuál es el mejor indicador del cambio del BMI (BMI, %BMI, BAZ y PBMI) a corto plazo, respecto a su consistencia a lo largo del espectro de adiposidad. Comparando y correlacionando las desviaciones estándares calculadas para cada medición de cambio del BMI dentro y entre niños obesos y no obesos, estos autores observaron que las desviaciones estándares intra-niño de BAZ y PBMI eran significativamente menores en niños obesos que en no obesos, mientras que las desviaciones estándares de BMI y %BMI eran similares en ambos grupos. Además, la desviación estándar intra-niño de BAZ, PBMI y en menor medida del %BMI estaba significativa e inversamente correlacionadas con BAZ, a diferencia del BMI, el cual no presentaba significación estadística.

Los autores concluyeron que aunque BAZ es óptimo para estimar la adiposidad, no es necesariamente el mejor indicador del cambio en la adiposidad a corto plazo; mejores alternativas serían BMI o %BMI. En este sentido, el principal hallazgo de estos autores fue el percatarse que ciertas mediciones del BMI (como BAZ) dependen del estado de adiposidad inicial, en tanto que BMI y %BMI no dependen de ella. Por otra parte, PBMI es al parecer menos sensible para medir el cambio en adiposidad, particularmente en niños obesos.

Las implicancias prácticas de este estudio radican en que el cambio en la adiposidad a corto plazo en un contexto clínico podría ser medido tanto por el BMI como por el %BMI, en tanto que PBMI es al parecer la medida menos útil. Por otra parte, BAZ es menos útil que lo ideal, ya que su variabilidad depende sobre el estatus de adiposidad inicial, aunque en la práctica está altamente correlacionado con las otras mediciones.

Según los autores, estos resultados destacan la importancia de utilizar métodos relativamente estables para estimar el cambio en la adiposidad en el seguimiento a corto plazo de los niños en riesgo de obesidad. Sin embargo, enfatizan la necesidad de realizar otros estudios con una muestra más amplia, sobre diferentes rangos etarios y períodos de seguimiento, con el fin de desarrollar recomendaciones prácticas sobre el mejor modo de estimar el cambio en la adiposidad en programas de tratamiento de obesidad infantil.

No obstante estos estudios, no se han realizado mayores investigaciones concernientes a las ventajas y desventajas que posee cada una de estas mediciones sobre el cambio en la adiposidad y en la predicción de la obesidad infantil. Por lo tanto, las preguntas que persisten son ¿Qué implicancias tiene la selección del indicador de cambio de BMI en la predicción de grasa corporal, particularmente, en países con alta desnutrición u obesidad?

Por otra parte, es posible que otras metodologías estadísticas puedan presentar mejores ajustes para variables en las que no se conoce su naturaleza paramétrica. En este sentido, los Modelos Aditivos Generalizados (*Generalized Additive Models*) es actualmente una novedosa estrategia de modelación que presenta mayor flexibilidad en comparación con los modelos lineales generalizados. En este sentido ¿La modelación mediante GAM podría entregar mejores predicciones del FMI en comparación con GLM?

2. PREGUNTA DE INVESTIGACIÓN

Como fue comentado recientemente, es importante conocer qué mediciones del cambio de BMI pueden ser las más útiles para estimar el FMI e identificar tempranamente aquellos niños con mayor riesgo de presentar obesidad infantil en el futuro. En este caso en particular, ¿Las diferentes mediciones de cambio del BMI entre los seis y 24 meses pueden predecir la obesidad a los 5 años? En caso de que fuera posible lo anterior, ¿Cuál de las diferentes mediciones del cambio de BMI (BMI, PBMI y BAZ), es el mejor estimador del FMI?

Por otra parte, si se aplican diferentes estrategias estadísticas, ¿Podrá alguna de ellas presentar mayor ventaja en predecir el FMI por sobre otra?

Por lo tanto, el presente estudio pretende ser un aporte a nivel epidemiológico en el área de estudio de crecimiento y salud.

3. OBJETIVO GENERAL

Comparar tres formas de medición del cambio del BMI (BMI, PBMI y BAZ), entre los seis y 24 meses de vida, en la estimación del FMI de los mismos sujetos a los cinco años (60 meses) de edad mediante dos estrategias diferentes de modelación.

4. OBJETIVOS ESPECÍFICOS

1. Comparar las estimaciones del FMI, realizadas a partir de las mediciones de ganancia del BMI (BMI, PBMI y BAZ), mediante los enfoques GLM y GAM, con el fin de sugerir qué medición es la que mejor predice el FMI, calculado a los cinco años de edad.
2. Comparar las estimaciones del FMI realizadas a partir de las mediciones de ganancia del BMI (BMI, PBMI y BAZ), mediante los enfoques GLM y GAM, con el fin de sugerir qué enfoque (GLM o GAM) es el que mejor predice el FMI calculado a los cinco años de edad.
3. Evaluar si estas respuestas son diferentes por sexo

En este sentido, el presente estudio pretende ser un aporte a nivel epidemiológico clínico en términos de clarificar cuál es la mejor forma de expresar los cambios de BMI en etapas tempranas de la vida y qué modelos deben usarse para explorar las relaciones entre los cambios de BMI tempranos y adiposidad futura.

5. HIPÓTESIS DE TRABAJO

El análisis se realizará en consideración de dos hipótesis de investigación:

- Las distintas formas de medir el incremento del BMI entre los seis a 24 meses se diferencian en sus ajustes y predictibilidad sobre el FMI medido en los mismos niños a los cinco años de edad (60 meses).

Con esto se espera que algunas mediciones del incremento del BMI estimen mejor el FMI respecto a otras mediciones.

- Las distintas estrategias estadísticas utilizadas presentan diferencias notorias en sus resultados sobre la predictibilidad del FMI.

Con esto se espera que cierta estrategia estadística sea más útil que otras en la predictibilidad del FMI a partir de la ganancia de las distintas mediciones del BMI.

6. MATERIAL

Se cuentan con información de los registros de una cohorte retrospectiva-prospectiva de niños nacidos el año 2002 que participaron en el estudio desarrollado en el proyecto FONDECYT 1060785 (investigadora responsable Juliana Kain). En este proyecto, durante el año 2006 se seleccionaron niños pertenecientes a la Junta Nacional de Jardines Infantiles (JUNJI) en seis comunas del área Sur de Santiago. El programa está caracterizado por proveer educación y alimentación a niños ubicados en niveles socioeconómicos medio-bajos y bajos de Chile. Del total de niños asistentes a los jardines en esas zonas durante al menos dos años, se seleccionaron los niños que poseyeran un peso al nacer entre los 2500g y 4500g de acuerdo a los registros médicos (n=1954 niños). Todos estos niños fueron invitados a participar, sin embargo 303 decidieron no participar y 455 fueron excluidos por poseer un peso de nacimiento inadecuado o por falta de información de crecimiento en los primeros años de vida (n= 421) y por ser gemelos (n=34) quedando la muestra final conformada por 1196 sujetos. No se hallaron diferencias en la edad, sexo, antropometría al momento del nacimiento, peso y estatura entre los niños que fueron enrolados dentro de la investigación, respecto de otros no enrolados (Kain *et al.* 2009). Las pérdidas de casos durante el período de seguimiento 2006 a 2009 se describen en la Tabla 1:

	2006	2007	2008	2009
Muestra Total	1195	1173	1150	1151
Fuera de Santiago	32	46	25	41
Ausencias	10	29	14	61
Exclusiones	8	5	7	1

Tabla 1. Descripción de pérdidas de casos durante el periodo de seguimiento (2006 a 2009).

De la muestra original (n=1196), se seleccionaron solamente aquellos casos que tuvieran mediciones del BMI entre cinco a siete meses y de 23 a 25 meses, así como el FMI entre 55 a 65 meses. Por lo tanto, la muestra de este estudio finalmente quedó constituida por 379 individuos (201 mujeres y 178 hombres). A su vez, esta muestra fue dividida en dos submuestras: una constituida por el 70% de los casos, la cual fue denominada entrenamiento y con la cual se realizaron las modelaciones estadísticas. La otra submuestra se constituyó con el 30% restante y se utilizó para la validación de los modelos generados a partir de la muestra de entrenamiento (Tabla 2).

Muestra	Mujeres	Hombres	Total
Entrenamiento	155	114	269
Validación	46	64	110
Total	201	178	379

Tabla 2. Composición de la muestra de entrenamiento y de validación desagregada por sexo.

7. MÉTODOS

Sobre esta muestra de niños se consideraron las medidas de peso y talla registradas entre los cinco a siete meses y entre los 23 a 25 meses. El peso de los niños fue obtenido utilizando una pesa pediátrica con una precisión de 10g, en tanto que la talla fue obtenida con una precisión de 0.1cm. La calidad de estas mediciones ha sido verificada en comparación con datos obtenidos de otras fuentes secundarias. Menos del 1% de las mediciones fueron excluidas de los análisis por ser consideradas *outliers* (Puntaje z del peso para la edad <-6 o >5 ; Puntaje z de la

talla para la edad <-6 o >6; BAZ <-5 o >5); las desviaciones estándar de los puntajes z se encontraron dentro del rango válido aceptado por la OMS.

Entre los 55 a 65 meses de edad se les registró el peso, talla, circunferencia de cintura y de cadera y cuatro pliegues corporales en cada uno de los niños a partir de métodos estandarizados (Gutiérrez-Gómez *et al.*, 2009), labor realizada por dos nutricionistas entrenados pertenecientes al proyecto FONDECYT 1060785 (La confianza intra-interobservador fue confirmada para el peso y la talla con una correlación intraclase ≥ 0.75). Con estos datos se calculó el FMI, estimado a partir de la cantidad de tejido graso obtenido de la fórmula antropométrica (grasa en kg/estatura²).

7.1. Análisis de datos.

El diseño de análisis de la presente Tesis contempla estudiar el efecto que tienen las diferentes maneras de medir el cambio de BMI (BMI, PBMI y BAZ) sobre el FMI, calculado en los mismos niños a los cinco años. Para ello se utilizaron dos estrategias analíticas distintas, descritas más abajo.

7.2. Variables de estudio

La Tabla 3 detalla las variables utilizadas en el presente estudio.

Variable	Descripción	Escala	Niveles
SEX	Sexo, expresión gonadal.	Binaria	0=Femenino; 1=Masculino
FMI	Índice de masa grasa.	Continua	Calculado entre los 55 a 65 meses de edad.
BMI6	Índice de Masa Corporal alcanzado a los seis meses de edad.	Continua	Calculado entre los cinco a siete meses de edad.

BMI24	Índice de Masa Corporal alcanzado a los 24 meses de edad.	Continua	Calculado entre los 23 a 25 meses de edad.
BAZ6	Puntaje z del Índice de Masa Corporal alcanzado a los seis meses de edad.	Continua	Calculado entre los cinco a siete meses de edad.
BAZ24	Puntaje z del Índice de Masa Corporal alcanzado a los 24 meses de edad.	Continua	Calculado entre los 23 a 25 meses de edad.
PBMI6	Percentil del Índice de Masa Corporal alcanzado a los seis meses	Continua	Calculado entre los cinco a siete meses de edad.
PBMI24	Percentil del Índice de Masa Corporal alcanzado a los 24 meses de edad.	Continua	Calculado entre los 23 a 25 meses de edad.
DBMI	Ganancia del Índice de masa corporal (BMI) entre seis y 24 meses de edad.	Continua	Diferencia entre BMI6 y BMI24.
DBAZ	Ganancia del puntaje z del Índice de masa corporal (BAZ) entre los seis y 24 meses de edad.	Continua	Diferencia entre BAZ6 y BAZ24.
DPBMI	Ganancia del percentil del Índice de masa corporal (PBMI) entre los 6 y 24 meses de edad.	Continua	Diferencia entre PBMI6 y PBMI24.

Tabla 3. Variables utilizadas en el presente estudio.

7.3. Variable respuesta

El Índice de Masa Grasa (FMI) de cada niño entre los 55 a 65 meses de edad fue utilizado en el presente estudio como variable respuesta. Éste fue calculado a partir de una ecuación predictiva de la grasa corporal previamente desarrollada en una muestra de niños pertenecientes a la JUNJI (3 a 4 años) y validada utilizando una dilución de deuterio (Kain *et al.*, 2007; Velásquez *et al.*, 2008). Esta ecuación utiliza la edad, el sexo, la circunferencia abdominal y los pliegues tricipital y escapular. Luego, la grasa corporal estimada se dividió por la estatura al cuadrado de cada

uno de los niños, con el fin de calcular el FMI. Cabe mencionar que este valor será considerado como el FMI observado para cada sujeto y será utilizado en comparaciones posteriores.

7.4. Variables predictivas

Las variables predictivas de la presente investigación son la ganancia del BMI, la ganancia de BAZ y la ganancia de PBMI, obtenidas a partir de la diferencia entre los seis y los 24 meses para cada niño. Debido a que %BMI puede ser utilizado de forma intercambiable con BMI, no fue utilizado en el presente trabajo.

El cálculo de la ganancia de las diferentes mediciones del BMI se realizó mediante un procedimiento denominado valores residualizados (en inglés *Residualized scores*), utilizado comúnmente en diferentes disciplinas científicas para medir el cambio (o ganancia) de una variable medida en dos momentos diferentes (DuBois, 1957, 1970; Keijzer-Veen, *et al.* 2005). Los puntajes residualizados se obtienen a partir de los residuos obtenidos de un análisis de regresión lineal en donde se emplea como variable respuesta la medición del post-test (BMI, BAZ y PBMI a los 24 meses) y como predictor el respectivo pre-test (BMI, BAZ y PBMI a los seis meses), mediante el siguiente procedimiento:

$$y = \beta_0 + \beta_1 X_1 + e \quad (1)$$

En donde el interés es obtener los valores sobre los errores residuales e para cada sujeto. El uso de estos valores residualizados radica en las siguientes consideraciones: no se encuentran correlacionados con el estado inicial, a diferencia de los valores de ganancia crudos (posttest – pretest), los cuales muestran una correlación negativa con los estados iniciales. Debido a que los sujetos no fueron medidos todos exactamente a la misma edad, sino que presentan algunas diferencias en días, no presentan un inicio común para todos los casos. En consecuencia, no se pueden utilizar métodos de análisis estadísticos que tienen como supuesto un punto de partida común.

En este caso, el uso de los valores residualizados se convierte en la estrategia ideal para medir el cambio ya que corrige la correlación entre la variable temprana y tardía, presenta mayor certeza y precisión en comparación con otras mediciones de cambio, no requiere el uso de ajustes y se pueden calcular términos de orden mayor a partir de ellos.

7.5. Análisis exploratorio de datos

Se realizaron de estadísticas exploratorias sobre cada una de las variables, tales como medidas de tendencia central y dispersión, además de la elaboración de procedimientos visuales, como el gráfico de caja y bigotes (*boxplot*). También se realizaron test de normalidad de Shapiro-Wilk, con el fin descartar normalidad en cada una de las variables.

7.6. Modelación estadística

a) Modelos Lineales Generalizados (GLM):

Uno de los procedimientos estadísticos utilizado para modelar la pregunta de investigación fue de acuerdo a los modelos lineales generalizados (GLM). En general, los GLM constituyen una familia de modelos que presentan una identidad común a ciertas clases de modelos de regresión (Hardin y Hilbe, 2001), que poseen una sola variable respuesta que distribuye de acuerdo a la familia exponencial de distribuciones de probabilidades. Reestructurando la relación entre el predictor lineal y el ajuste, se pueden modelar relaciones lineales que previamente eran consideradas no lineales. Los GLM presentan las siguientes características (Hardin y Hilbe *op. cit.*, p.9):

1. Un componente aleatorio para la respuesta y , que presenta una variable aleatoria continua x , con una función de probabilidad de la familia exponencial, de la forma

$$f(x|\theta) = h(x)c(\theta) \exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right) \quad (2)$$

2. Un componente lineal sistemático relacionado al predictor lineal, $\eta=x\beta$, al producto de la matriz de diseño x y a los parámetros β .
3. Una función de enlace diferenciable, conocida, monótona y uno a uno $g(\cdot)$ que relaciona al predictor lineal a los valores ajustados. Dado que la función es uno a uno, existe una función inversa que relaciona la respuesta media esperada $E(y)=\mu$, al predictor lineal, tal que $\mu=g^{-1}(\eta)=E(y)$.
4. La varianza puede cambiar con los predictores sólo como una función de la media.
5. Existe un solo algoritmo de mínimos cuadrados re ponderados iterativamente (en inglés *iteratively reweighted least squares* IRLS) que es suficiente para ajustar a todos los miembros de la clase.

En GLM, las observaciones de la variable dependiente y son independientes y normales, con una varianza constante σ^2 . Se asume que el valor medio de y puede depender de otras cantidades o predictores x_1, x_2, \dots, x_{p-1} . En la situación más simple se asume que esta dependencia es lineal y se escribe

$$y = \beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1} + e \quad (3)$$

Los modelos lineales generalizados especifican una relación entre la media de la variable aleatoria y y una función de la combinación lineal de los predictores. Esta generalización admite una especificación del modelo que permite la agregación de variables respuestas continuas y discretas, permitiendo además una descripción de la varianza como una función de la media (Hardin y Hilbe *op. cit.*, p.9).

En este estudio, se utilizó el FMI como variable respuesta, en tanto que la ganancia de cada una de las mediciones del BMI fue usada como variable predictiva, además del sexo y de la interacción entre ambas. De este modo, se realizaron tres modelos, cada uno con una de las distintas mediciones del BMI (BMI, BAZ y PBMI). Cada uno de los modelos puede ser escrito como sigue:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e \quad (4)$$

En donde X_1 corresponde a la ganancia de cada una de las mediciones del BMI, X_2 es la variable correspondiente al sexo, denominada SEX (0: femenino; 1: masculino) y X_3 corresponde a la interacción entre las dos variables anteriores. Estos análisis se realizaron utilizando la estrategia *Backward*, descartando las variables que no fueron significativas dentro del modelo.

Posterior a los análisis de regresión utilizando GLM, se revisaron los valores residuales con el fin de constatar violaciones del supuesto de normalidad (*gross violations to normality*), mediante la prueba de Shapiro-Wilk y la realización de histogramas de los residuos ajustados a una curva normal. También se realizaron las pruebas de homocedasticidad de Breusch-Pagan y Cook-Weisberg, rechazándose con un valor de probabilidad inferior a 0.05.

Además, con el fin de observar el ajuste entre la variable respuesta FMI y los predictores de ganancia de BMI, BAZ y PBMI, se realizaron gráficos de dispersión entre estas variables, ajustados a una recta de cuadrados mínimos.

b) Enfoque GAM

Este enfoque fue desarrollado por Hastie y Tibshirani (1990), el cual especifica una distribución y una función de enlace g relacionando a los valores esperados de la distribución al predictor m de variables y ajusta las funciones $f_j(X_j)$ para satisfacer:

$$g(E(Y)) = \beta_0 + \sum_{j=1}^p f_j(X_j) \quad (5)$$

Las funciones $f_j(X_j)$ pueden ser ajustadas utilizando medias paramétricas o no paramétricas, proveyendo el potencial para mejores ajustes en comparación con otros modelos. Además, las variables predictivas se incorporan en su estructura como parte de f_j y no directamente como realiza GLM.

Existen varios suavizadores (*smoothers*) que han sido desarrollados para estimar f_j , tales como los suavizadores polinomiales (*polynomial smoothers*, Wand y Jones, 1995; Fan y Gijbels, 1996), *regression splines* (Eubank, 1988, 1999), *smoothing splines* (Wahba, 1990; Green y Silverman, 1994) y *P-splines* (Ruppert, Wand y Carrol, 2003) y *Multivariate Adaptive Regression Splines* (MARS). Esta última estrategia fue la utilizada en la presente Tesis para modelar la pregunta de investigación.

Desarrollado por Jerome Friedman (1991), MARS es un tipo de técnica de regresión no paramétrica que puede ser visto como una extensión de los modelos lineales que automáticamente modela no linealidades e interacciones. Construye un modelo de regresión no lineal a partir del producto de funciones “base de suavizamiento” (*spline*):

$$g(\hat{f}(x)) = \sum_{j=1}^k c_j B_j(x) \quad (6)$$

En donde $B_j(x)$ son las funciones base y c_j es un coeficiente constante. Por lo tanto, el modelo es una suma ponderada de funciones base, en donde cada función base $B_j(x)$ puede tomar una de las siguientes formas:

1. Una constante, de valor uno (1), asociada únicamente al intercepto.
2. Una función bisagra de la forma:

$\max(0, x-c)$ o $\max(0, c-x)$, en donde c corresponde a un valor constante, denominado nodo.

3. Un producto de dos o más funciones bisagra. Estas funciones base pueden modelar la interacción entre dos o más variables.

Las funciones bisagra son la parte clave de los modelos MARS. MARS selecciona automáticamente las variables y los valores para los nodos y funciones bisagra.

Estas funciones incorporan en su estructura las variables predictivas que entran en la modelación como parte de una función y no directamente como lo realiza una regresión lineal, produciendo un modelo para la variable respuesta en estudio, que automáticamente selecciona las variables predictoras que aparecen en la ecuación final, las que se incorporan en las funciones base (Lewis y Stevens 1991). En el presente trabajo los modelos de regresión se realizaron utilizando quince funciones-base como máximo.

MARS construye los modelos en dos fases: *forward* y *backward*.

En la fase *forward* MARS comienza sólo con el intercepto que corresponde a la media de la variable respuesta y luego va adicionando repetidamente funciones base de a pares.

Debido a que esta fase construye modelos sobreajustados, la fase *backward* poda los modelos, removiendo las funciones base uno a uno. Debido a esto, para cada relación entre las variables respuesta y predictora, MARS genera varias ecuaciones predictivas que modelan esta relación, desde modelos con pocas funciones base a modelos con varias funciones base.

La fase *backward* utiliza el *Generalized Criterion Value* (GCV) para comparar cada uno de los modelos, con el fin de escoger el mejor modelo: los valores más bajos de GCV son los mejores. El GCV es una forma de regularización, el cual enfrenta la bondad de ajuste contra la complejidad del modelo.

Los modelos MARS presentan una serie de beneficios:

- Son más flexibles que GLM
- Son simples de entender e interpretar que otros modelos
- Son apropiados para datos continuos y categóricos
- Genera una selección automática de variables, incluyendo las variables importantes y excluyendo las poco importantes.
- Presentan una buena compensación del sesgo y la varianza en situaciones de no linealidad.

En el presente trabajo, el ajuste de los modelos MARS se realizó para diez modelos diferentes, los que presentaron distinto número de funciones-base seleccionadas. Finalmente, el modelo final se seleccionó en función del menor valor del *Generalized Criterion Value* (GCV).

La ganancia de cada una de las mediciones del BMI fue analizada en un modelo separado, con el fin de mantener una estructura similar a los modelos GLM, con el fin de servir posteriormente para la comparación de los resultados. De este modo, también se espera obtener tres ecuaciones predictivas del FMI dentro del enfoque GAM, correspondientes cada una de ellas a cada una de las mediciones del BMI (BMI, BAZ, PBMI).

En total, se obtuvieron seis ecuaciones predictivas del FMI, a partir de las tres mediciones del BMI (BMI, BAZ y PBMI) y de los dos enfoques estadísticos, GLM y GAM.

7.7. Validación de los resultados obtenidos

El presente trabajo utilizó el 70% de la muestra total (muestra de entrenamiento) para la construcción de cada uno de los seis modelos mencionados anteriormente (tres ecuaciones predictivas del FMI para cada una de las dos estrategias

analíticas, GLM y GAM), en tanto que el 30% restante se utilizó para la validación y comparación de estos seis modelos propuestos.

Al respecto, el presente estudio intenta buscar:

1. La medida de ganancia de cada una de las mediciones del BMI (BMI, BAZ y PBMI), que estime con mayor certeza y precisión el FMI observado a los cinco años de edad. Se entendió por certeza, la ausencia de sesgo sistemático en la estimación. En tanto, precisión se entendió como la ausencia de errores aleatorios en la estimación.
2. El enfoque estadístico (GLM o GAM) que estime con mayor certeza y precisión el FMI observado en estos niños a los cinco años de edad.

Con el fin de validar los resultados, se estimó el FMI en base a cada una de las ecuaciones predictivas del FMI generadas a partir de los seis modelos estadísticos propuestos, las que se compararon con el FMI observado. Ambas mediciones fueron obtenidas en la muestra de validación. Para comparar los valores estimados y observados del FMI en la muestra de validación, se utilizaron las siguientes pruebas estadísticas:

- a. Coeficiente de correlación de Pearson (*r Pearson*).
- b. Análisis de regresión entre la variable respuesta observada y la estimada. Este procedimiento se realiza sin intercepto (regresión a través del origen) y tiene una serie de propiedades útiles para ser utilizado como método de comparación entre procedimientos. Por una parte, se obtiene el coeficiente de determinación, denominado en el presente trabajo R^2_{NOC} , el cual fue utilizado para medir la certeza entre el FMI observado y el estimado. Por otra parte, el estimador de la pendiente se denomina $COEF_{NOC}$ el cual puede ser utilizado para

conocer si el FMI estimado subestima (valores menores a 1) o sobreestima (valores mayores a 1) el FMI observado.

- c. Coeficiente de correlación de concordancia de Lin (1989, 2000), que se utiliza para conocer las coincidencias de medidas continuas obtenidas por dos métodos o dos personas. En el presente trabajo fue abreviado como **LIN**. Este coeficiente combina las medidas de precisión y certeza, con el objetivo de determinar cuán lejos el dato observado se desvía de la línea de concordancia perfecta. En el presente trabajo también se utilizó el factor de corrección de sesgo, que hace referencia a la medida de la certeza en el coeficiente de concordancia de Lin (en el presente trabajo fue denominado como **CERTEZA**).
- d. La proporción del error medio absoluto (en inglés *Mean proportion of absolute error*, en adelante **MPAB**), el cual se calcula mediante la siguiente fórmula:

$$MPAB = \frac{1}{N} \sum_i \frac{|y_i - \hat{f}_q(x_i)|}{y_i} \quad (7)$$

En este caso, los valores **MPAB** cercanos a cero se consideran con menor error.

Se espera que estos estadísticos entreguen la información suficiente para comparar todos los modelos y procedimientos propuestos, con el fin de seleccionar la mejor medida de ganancia del BMI y el mejor método analítico.

7.8. Consentimiento ético

El protocolo de estudio del proyecto FONDECYT 1060785, al cual pertenece la muestra del presente trabajo, fue aprobado por el Director Ejecutivo del Programa JUNJI y el Instituto de Nutrición y Tecnología de los Alimentos (INTA) de la

Universidad de Chile. El consentimiento informado fue obtenido de todos los padres o apoderados de los niños.

8. RESULTADOS

8.1. Exploración de datos

La Tabla 4 muestra las estadísticas descriptivas de las variables del presente estudio en la muestra total. Las variables FMI, BMI6, BMI24, BAZ6 y BAZ24 presentan medias y medianas bastante similares en cada una de ellas. Además, se aprecia una notoria asimetría en todas las variables, destacando FMI, BMI6, BMI24, PBMI24 y DBMI. La curtosis de estas variables muestra que FMI y DBMI son notoriamente leptocúrticas, en tanto que BMI6, BMI24, BAZ6, BAZ24 y DBAZ presentan leptocurtosis leve; PBMI6, PBMI24 y DPBMI muestra leve platicurtosis.

Variable	N	Media	Mediana	Desviación estándar	RIC	Asimetría	Curtosis	Min	Max
FMI	379	3.99	3.78	1.28	1.34	1.35	5.58	1.75	9.17
BMI6	379	17.97	17.82	1.70	2.28	0.61	3.77	13.64	24.28
BMI24	379	16.90	16.78	1.46	1.87	0.71	3.77	13.70	22.31
BAZ6	379	0.51	0.44	1.06	1.43	0.28	3.29	-2.41	4.00
BAZ24	379	0.75	0.68	0.99	1.28	0.26	3.22	-2.11	4.01
PBMI6	379	63.19	67.00	27.99	47.39	-0.45	2.04	0.80	99.99
PBMI24	379	70.08	75.18	24.96	37.88	-0.75	2.62	1.74	99.99
DBMI	379	0.00	-0.15	1.17	1.54	0.87	4.88	-2.70	6.11
DBAZ	379	0.00	-0.07	0.80	1.08	0.46	3.71	-2.14	3.57
DPBMI	379	0.00	3.46	20.37	27.43	-0.39	2.84	-62.80	53.22

Tabla 4. Estadísticas descriptivas de las variables del presente estudio en la muestra total.

Las Figura 1 presenta los gráficos de caja y bigotes de las variables que se utilizaron en el presente estudio. En la variable BMI, el eje *y* representa los valores de peso/estatura²; en BAZ, el eje *y* representa los valores del puntaje *z* obtenido del cálculo de peso/estatura²; para la variable PBMI, el eje *y* representa los valores en percentiles de peso/estatura² y en FMI, el eje *y* representa los valores de grasa corporal/ estatura². En estos gráficos es notoria la presencia de casos *outliers* en

todas las variables, sobre todo en FMI, BMI6, BMI24 y DBMI. A diferencia, PBMI presenta escasos valores *outliers*, los que se encuentran sólo en la variable DPBMI.

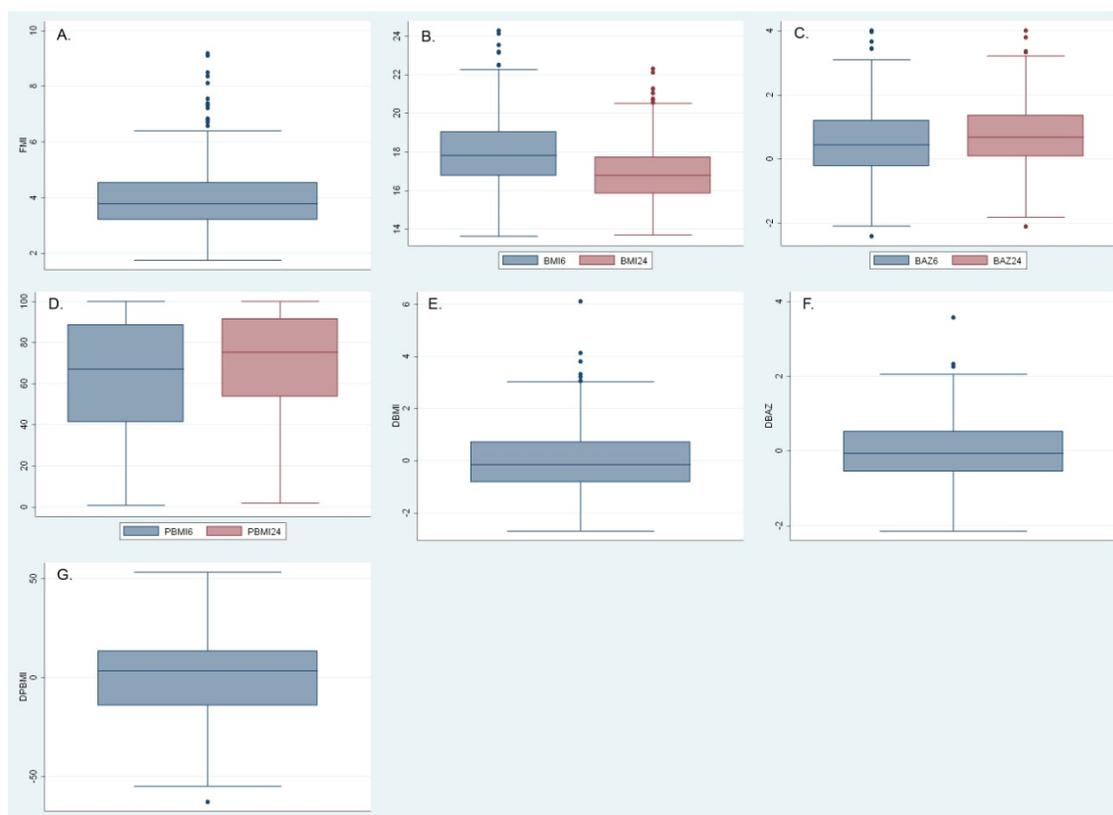


Figura 1. Gráficos de caja y bigotes de las variables del presente estudio en la muestra total. A. FMI; B. BMI6 y BMI24; C. BAZ6 y BAZ24; D. PBMI6 y PBMI24; E. DBMI; F. DBAZ; G. DPBMI.

La Tabla 5 muestra las estadísticas descriptivas de las variables en estudio para el sexo femenino. Se observa que la media de FMI es mayor a la muestra total, así como la media de BAZ24 y PBMI24. En tanto, la desviación estándar de FMI, BAZ6, BAZ24, PBMI6, PBMI24 y DBMI son menores a la muestra total. También se observan mayores diferencias entre medias y medianas de cada variable respecto a la muestra total.

Variable	N	Media	Mediana	Desviación		Asimetría	Curtosis	Min	Max
				estándar	RIC				
FMI	201	4.27	3.92	1.21	1.45	1.35	4.73	2.45	8.50
BMI6	201	17.80	17.55	1.72	2.34	0.69	3.92	13.64	24.28
BMI24	201	16.89	16.75	1.51	1.87	0.96	4.28	14.14	22.31
BAZ6	201	0.51	0.41	1.04	1.46	0.33	3.42	-2.41	4.00

BAZ24	201	0.83	0.77	0.97	1.29	0.55	3.43	-1.27	4.01
PBMI6	201	63.35	65.91	27.14	47.96	-0.40	2.07	0.80	99.99
PBMI24	201	71.89	77.94	23.31	37.54	-0.71	2.53	10.20	99.99
DBMI	201	0.00	-0.24	1.29	1.59	1.06	5.19	-2.70	6.11
DBAZ	201	0.00	-0.11	0.83	1.08	0.71	4.11	-2.00	3.57
DPBMI	201	0.00	4.02	20.43	28.98	-0.41	2.75	-62.80	44.42

Tabla 5. Estadísticas descriptivas del sexo femenino de la muestra total.

La Tabla 6 muestra las estadísticas descriptivas del sexo masculino. Hay mayor similitud con la muestra total respecto al sexo femenino, además de que las diferencias entre la media y la mediana no son tan notorias. También se observa mayor leptocurtosis en comparación con la muestra total.

Variable	n	Desviación							
		Media	Mediana	estándar	RIC	Asimetría	Curtosis	Min	Max
FMI	178	3.69	3.50	1.15	1.33	1.57	7.51	1.75	9.17
BMI6	178	18.16	18.11	1.66	2.12	0.55	3.72	14.65	24.11
BMI24	178	16.92	16.84	1.40	1.77	0.36	2.96	13.70	20.76
BAZ6	178	0.50	0.53	1.09	1.41	0.22	3.16	-2.09	3.97
BAZ24	178	0.67	0.65	1.02	1.30	0.00	2.85	-2.11	3.14
PBMI6	178	63.01	70.02	29.00	46.81	-0.49	1.99	1.83	99.99
PBMI24	178	68.04	74.05	26.62	38.68	-0.73	2.53	1.74	99.92
DBMI	178	0.00	-0.10	1.03	1.47	0.40	3.16	-2.33	3.22
DBAZ	178	0.00	-0.01	0.76	1.03	0.09	2.98	-2.14	2.05
DPBMI	178	0.00	3.17	20.37	25.96	-0.36	2.95	-54.89	53.22

Tabla 6. Estadísticas descriptivas de la muestra total segregada por sexo masculino.

La Tabla 7 y la Figura 2 describen las variables en submuestra de entrenamiento. En general se observa que las variables no presentan notorias diferencias en respecto a la muestra total.

Variable	N	Desviación							
		Media	Mediana	estándar	RIC	Asimetría	Curtosis	Min	Max
FMI	269	4.02	3.81	1.21	1.31	1.48	6.41	1.80	9.17
BMI6	269	17.96	17.82	1.70	2.23	0.66	4.12	13.64	24.28
BMI24	269	16.89	16.75	1.46	1.78	0.73	3.69	14.14	22.31
BAZ6	269	0.51	0.49	1.06	1.41	0.28	3.57	-2.41	4.00
BAZ24	269	0.75	0.70	0.98	1.26	0.31	3.09	-1.48	4.01
PBMI6	269	63.51	68.79	27.78	46.61	-0.49	2.10	0.80	99.99
PBMI24	269	70.00	75.80	24.91	37.09	-0.73	2.56	6.94	99.99
DBMI	269	0.00	-0.16	1.15	1.61	0.67	3.36	-2.74	4.09

DBAZ	269	0.00	-0.11	0.78	1.12	0.36	2.76	-2.02	2.30
DPBMI	269	0.00	3.92	20.20	30.69	-0.34	2.67	-62.93	55.95

Tabla 7. Estadísticas descriptivas de las variables del presente estudio en la submuestra de entrenamiento.

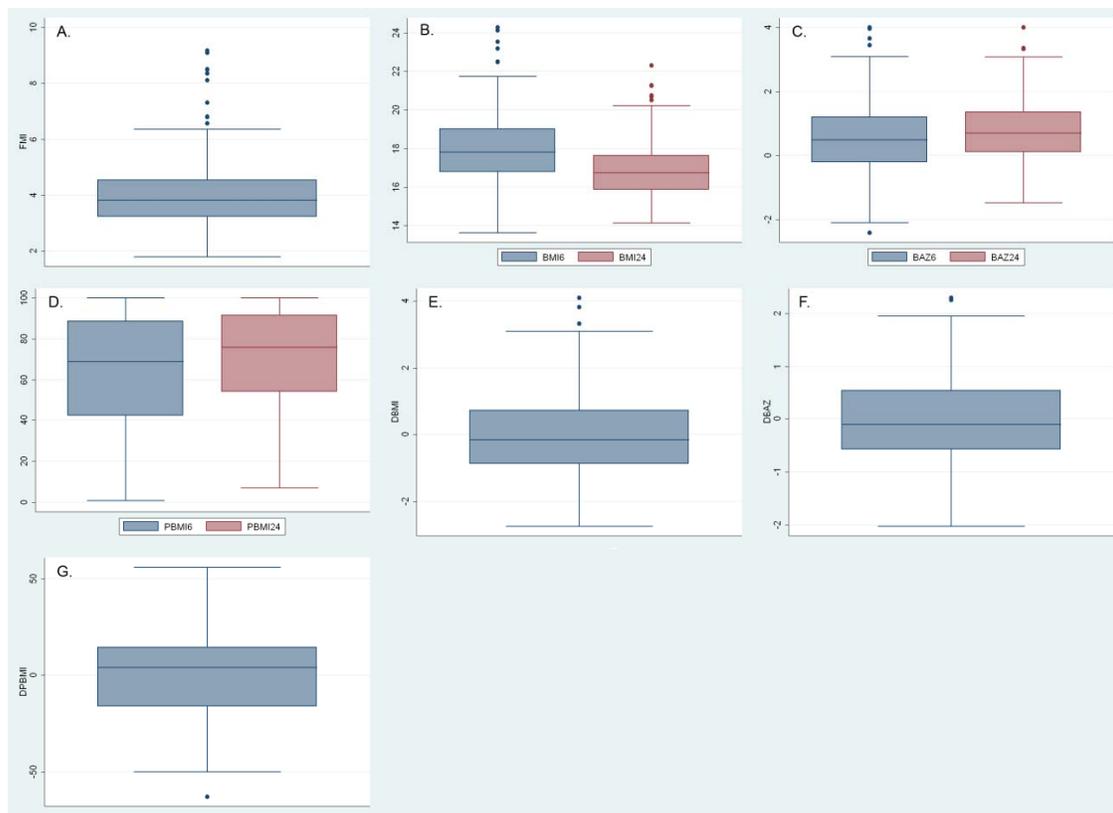


Figura 2. Gráficos *boxplot* de las variables en la submuestra entrenamiento. A. FMI; B. BMI6 y BMI24; C. BAZ6 y BAZ24; D. PBMI6 y PBMI24; E. DBMI; F. DBAZ; G. DPBMI.

La Tabla 8 y la Figura 3 describen las variables en estudios en la muestra de validación. También en esta submuestra no se observan marcadas diferencias respecto a la muestra total ni tampoco respecto a la submuestra de entrenamiento.

Variable	N	Media	Mediana	Desviación estándar	RIC	Asimetría	Curtosis	Min	Max
BMI	110	3.94	3.67	1.25	1.36	1.08	3.74	1.75	7.54
BMI6	110	17.99	17.74	1.70	2.42	0.48	2.94	14.75	23.14
BMI24	110	16.94	16.82	1.48	1.99	0.66	3.97	13.70	22.11
BAZ6	110	0.49	0.29	1.07	1.49	0.26	2.64	-1.53	3.44
BAZ24	110	0.76	0.66	1.02	1.37	0.15	3.47	-2.11	3.80
PBMI6	110	62.41	61.59	28.61	48.92	-0.34	1.91	6.30	99.97
PBMI24	110	70.27	74.54	25.20	39.72	-0.79	2.76	1.74	99.99
DBMI	110	0.00	-0.02	1.21	1.32	1.24	7.63	-2.47	6.02
DBAZ	110	0.00	-0.01	0.84	0.91	0.62	5.42	-2.23	3.56

DPBMI | 110 0.00 3.76 20.68 23.73 -0.51 3.36 -56.71 46.86
Tabla 8. Estadísticas descriptivas de las variables del presente estudio en la submuestra de validación.

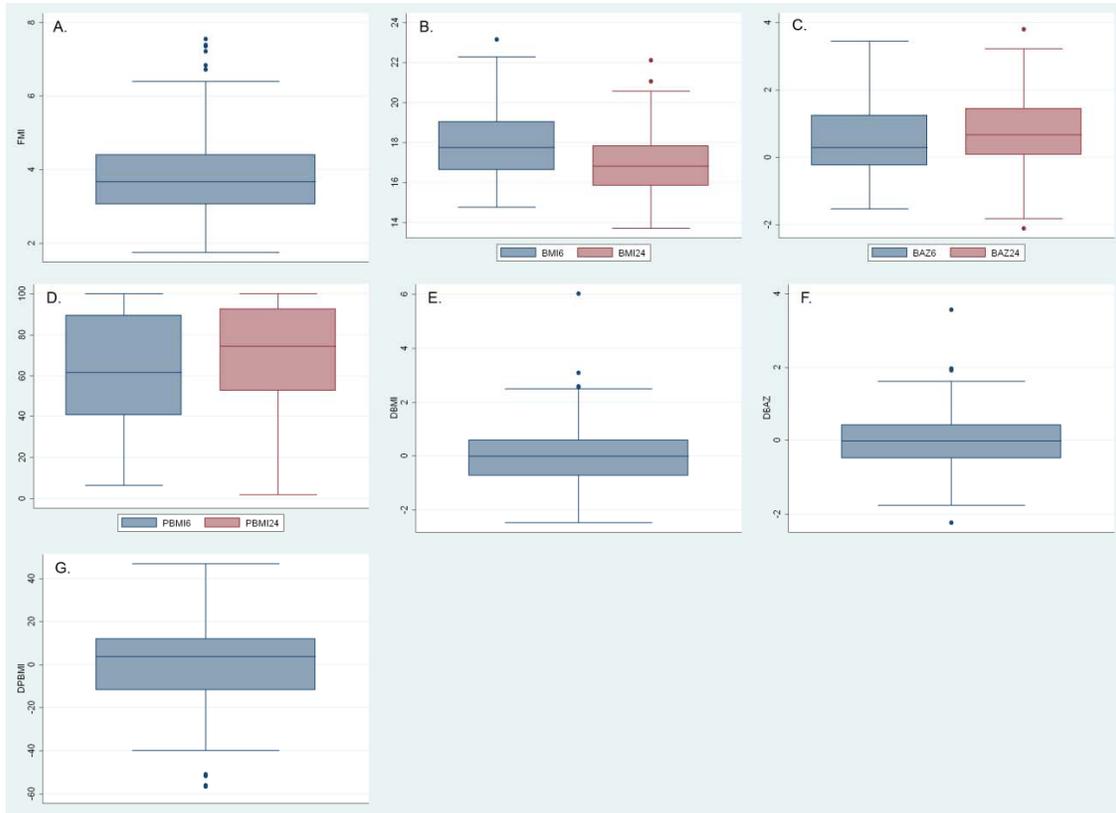


Figura 3. Gráficos *boxplot* de las variables en la submuestra validación. A. FMI; B. BMI6 y BMI24; C. BAZ6 y BAZ24; D. PBMi6 y PBMi24; E. DBMI; F. DBAZ; G. DPBMI.

Se observa que todas las variables, a excepción de BAZ6, BAZ24 y DBAZ no presentan una clara simetría. Con el fin de conocer si algunas de las variables presentaron distribución normal, se realizó la prueba de Shapiro-Wilk a cada una de ellas (Tabla 9). Como se observa, sólo las variables BAZ6 y BAZ24 tanto en la muestra total como en las submuestras no rechazan la hipótesis de normalidad.

MUESTRA TOTAL				ENTRENAMIENTO				VALIDACION			
Variable	Obs	W	Prob>z	Variable	Obs	W	Prob>z	Variable	Obs	W	Prob>z
BMI	379	0.91	0.00	BMI	269	0.90	0.00	BMI	269	0.90	0.00
BMI6	379	0.98	0.00	BMI6	269	0.97	0.00	BMI6	269	0.97	0.00
BMI24	379	0.97	0.00	BMI24	269	0.97	0.00	BMI24	269	0.97	0.00

BAZ6	379	0.99	0.10	BAZ6	269	0.99	0.08	BAZ6	269	0.99	0.08
BAZ24	379	0.99	0.17	BAZ24	269	0.99	0.09	BAZ24	269	0.99	0.09
DBMI	379	0.96	0.00	DBMI	269	0.97	0.00	DBMI	269	0.97	0.00
DBAZ	379	0.99	0.00	DBAZ	269	0.99	0.01	DBAZ	269	0.99	0.01
PBMI6	379	0.93	0.00	PBMI6	269	0.93	0.00	PBMI6	269	0.93	0.00
PBMI24	379	0.92	0.00	PBMI24	269	0.91	0.00	PBMI24	269	0.92	0.00
DPBMI	379	0.99	0.00	DPBMI	269	0.98	0.00	DPBMI	269	0.98	0.00

Tabla 9. Resultados de la prueba de Shapiro-Wilk para las variables en estudio.

8.3. Construcción de los modelos predictivos

8.3.1. Análisis de regresión lineal (GLM)

A continuación se presentan las modelaciones realizadas mediante los análisis de regresión lineal, entre la variable respuesta FMI y las variables predictivas ganancia de las distintas mediciones del BMI (DBMI, DBAZ y DPBMI), el sexo (SEX) y la interacción entre ambas variables (DBMI-SEX, DBAZ-SEX y DPBMI-SEX).

8.3.1.1. Regresión lineal utilizando DBMI, SEX e interacción DBMI-SEX

Este análisis de regresión arrojó los siguientes resultados (Tabla 10):

FMI	Coef.	Err. St.	P>t	[95% Conf. Interval]	
DBMI	0.46	0.07	0.00	0.32	0.59
SEX	-0.58	0.13	0.00	-0.83	-0.32
DBMI_SEX	0.03	0.12	0.81	-0.21	0.26
cons	4.26	0.08	0.00	4.09	4.43

Tabla 10. Resultados del análisis de regresión utilizando como predictores DBMI, SEX e interacción DBMI-SEX ($R^2=0.245$; $MSE=1.047$), sobre la variable respuesta FMI.

En esta tabla se observa que todas las variables predictivas son significativas, a excepción de la variable interacción DBMI-SEX. Debido a esto, se descartó esta última variable y se realizó un nuevo ajuste, quedando la ecuación predictiva como:

$$\text{FMI} = 4.26 + 0.46 \cdot \text{DBMI} - 0.58 \cdot \text{SEX}$$

Llama la atención que los coeficientes de la ecuación predictiva no variaron una vez removida la variable interacción DBMI-SEX y realizado nuevamente el análisis de regresión lineal. El coeficiente de determinación de este nuevo análisis de regresión fue de 0.247 y el error medio cuadrático de 1.04. La prueba de Shapiro-Wilk de los valores residuales de este modelo tuvo un valor W de 0.893 y un p-valor de 0.000, lo que rechaza el supuesto de normalidad de los residuos. El test de homocedasticidad de Breusch-Pagan y Cook-Weisberg de rechaza el supuesto de varianza constante (p-valor=0.0001). La Figura 4 muestra el histograma de los valores residuales de este análisis, ajustado a una curva normal. Este gráfico muestra que los residuos presentan una distribución bastante similar a la normal, aunque se encuentra truncada la cola de los valores negativos y más alargada la cola de los valores positivos.

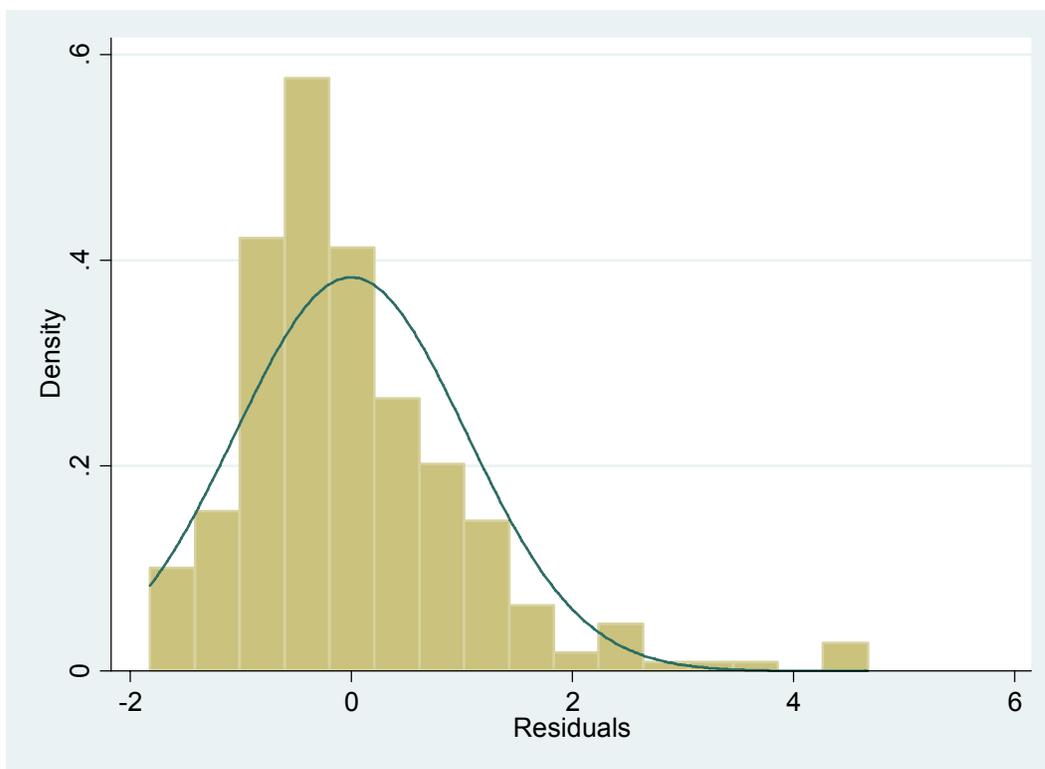


Figura 4. Histograma de los valores residuales del presente análisis.

El gráfico de dispersión entre FMI y BMI ajustado a una recta mínimo cuadrática se presenta en la Figura 5. Se aprecia en este gráfico que la relación entre FMI y DBMI presentan una relación lineal positiva y leve, con una amplia dispersión de los casos y la notoria presencia de casos *outliers* en los valores más altos de FMI (sobre ocho).

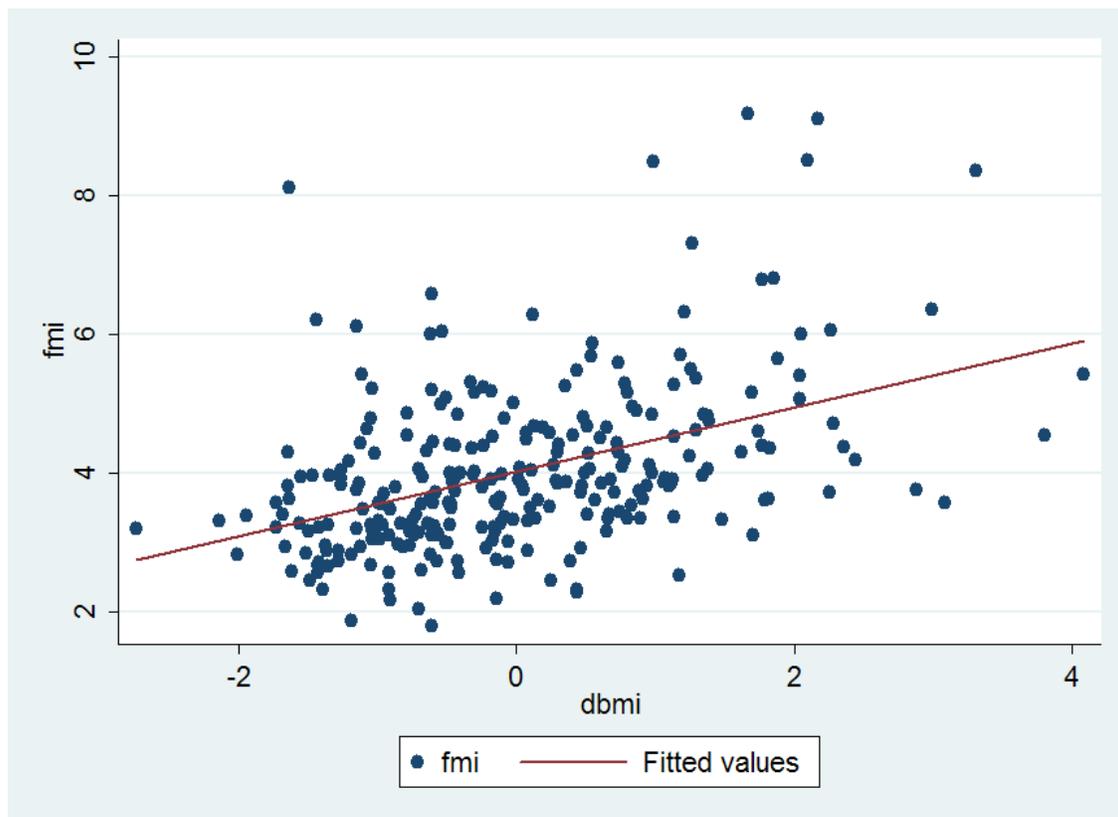


Figura 5. Dispersión y ajuste lineal entre FMI y DBMI.

8.3.1.2. Regresión FMI, DBAZ y sexo.

Los resultados del análisis de regresión utilizando como predictores DBAZ, SEX e interacción DBAZ-SEX sobre la variable FMI arrojaron los siguientes resultados (Tabla 11):

FMI	Coef.	Err. St.	P>t	[95% Conf. Interval]
DBAZ	0.69	0.11	0.00	0.49 0.90
SEX	-0.58	0.13	0.00	-0.83 -0.32
DBAZ_SEX	-0.05	0.17	0.76	-0.39 0.28
Cons	4.26	0.08	0.00	4.09 4.43

Tabla 11. Resultados del análisis de regresión utilizando como predictores DBMI, SEX e interacción DBMI-SEX ($R^2=0.238$; $MSE=1.052$) sobre la variable respuesta FMI.

Debido a que la variable interacción DBAZ-SEX no fue significativa, se decidió remover esta variable y se realizó nuevamente un análisis de regresión, quedando la ecuación predictiva como sigue:

$$\mathbf{FMI = 4.26 + 0.67*DBAZ - 0.57*SEX}$$

El coeficiente de determinación de este nuevo análisis fue de 0.24 y el error medio cuadrático de 1.05. En este caso, los coeficientes de la ecuación de regresión registraron leves cambios una vez removida la variable interacción DBAZ-SEX y realizado nuevamente un análisis de regresión. La prueba de Shapiro-Wilk realizada sobre los residuos de este análisis presentó valores W de 0.89 y un p-valor de 0.00, lo que rechaza el supuesto de normalidad de los residuos. El test de homocedasticidad de Breusch-Pagan y Cook-Weisberg rechaza el supuesto de varianza constante (p-valor=0.0002). La Figura 6 muestra el histograma de los valores residuales ajustados a una curva normal. Este gráfico muestra que los residuos presentan una distribución similar a la normal, aunque es notoria la asimetría hacia los valores más altos.

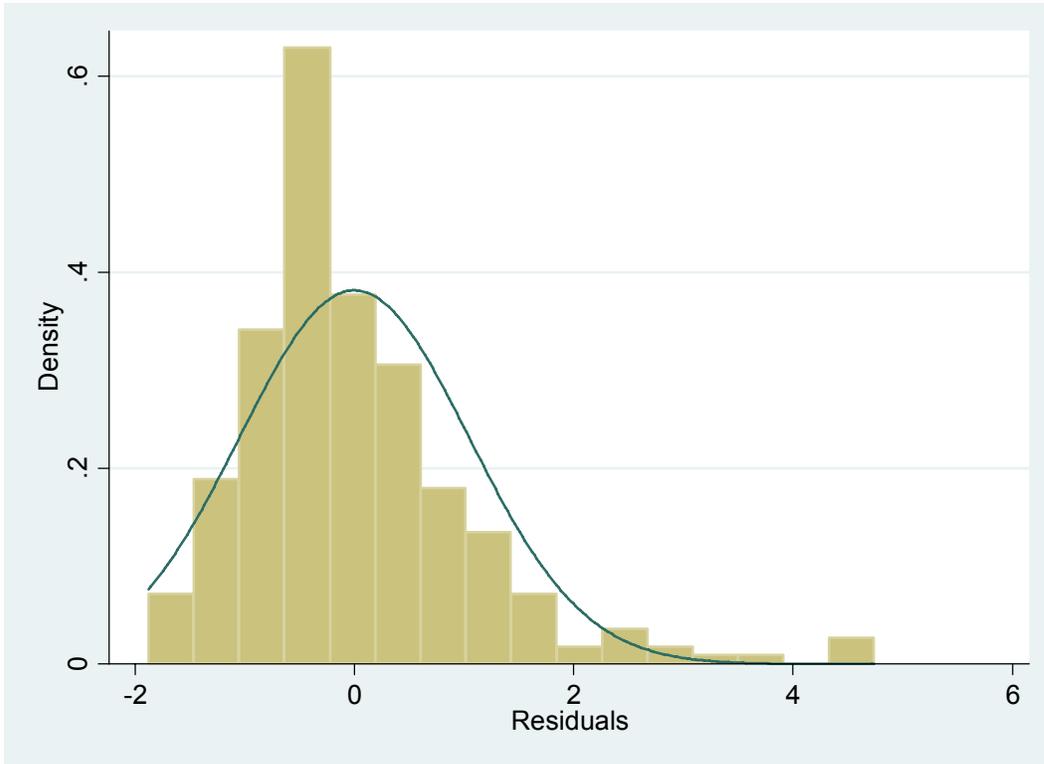


Figura 6. Histograma de los valores residuales.

El gráfico de dispersión entre FMI y DBAZ ajustado a una recta de regresión se presenta en la Figura 7. En este gráfico se observa una leve relación lineal positiva entre estas variables, con una marcada dispersión de los datos y la notoria presencia de casos *outliers* en valores más altos de FMI.

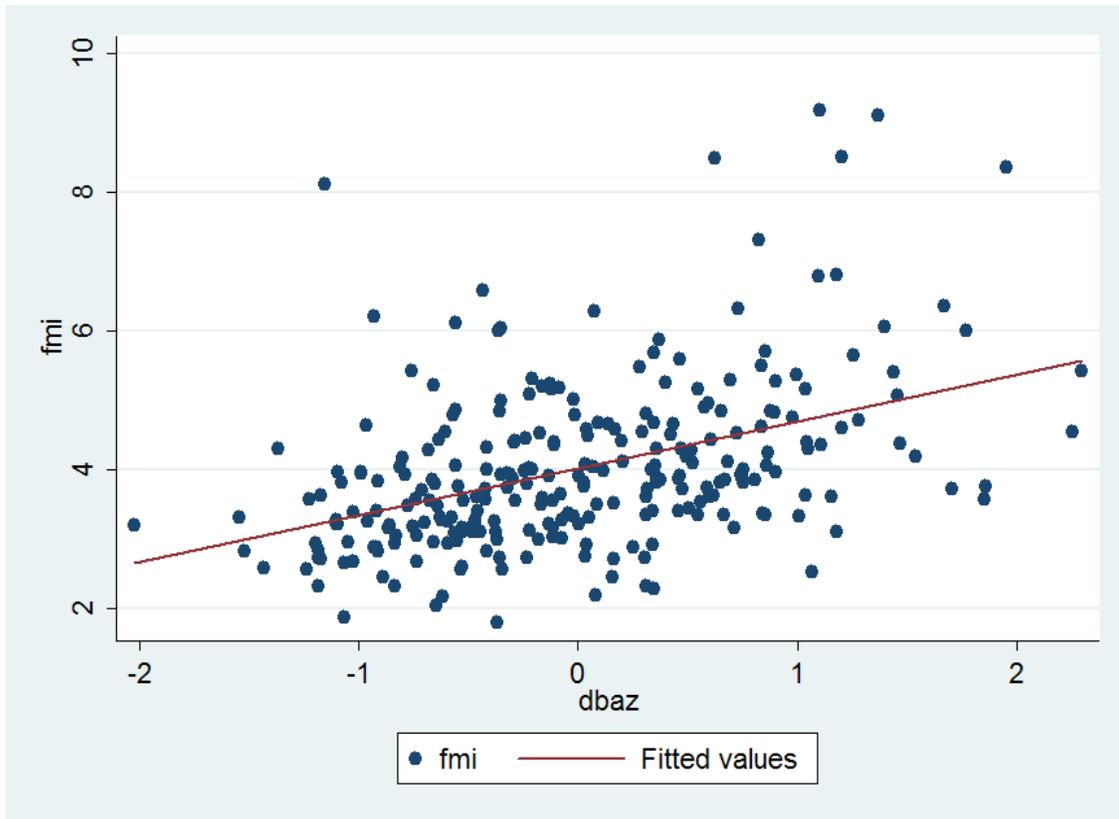


Figura 7. Dispersión y ajuste lineal entre FMI y DBAZ.

8.3.1.3. Regresión FMI, DPBMI y sexo.

La Tabla 12 muestra los resultados del análisis de regresión lineal entre FMI y DPBMI.

FMI	Coef.	Err. St.	P>t	[95% Conf. Interval]	
DPBMI	0.03	0.004	0.00	0.02	0.03
SEX	-0.58	0.13	0.00	-0.84	-0.32
DPBMI_SEX	-0.00	0.006	0.56	-0.02	0.01
Cons	4.26	0.09	0.00	4.09	4.43

Tabla 12. Resultados del análisis de regresión entre FMI y DBMI para el sexo masculino (R²=0.2104; MSE=1.07).

Debido a que la variable interacción DPBMI-SEX no fue significativa, se decidió retirar esta variable, quedando la ecuación predictiva como sigue:

$$\text{FMI} = 4.26 + 0.03 \cdot \text{DPBMI} - 0.58 \cdot \text{SEX}$$

El coeficiente de determinación de este nuevo análisis fue de 0.21 y el error medio cuadrático de 1.07. Similarmente al análisis de regresión lineal utilizando la variable DBMI, los coeficientes de la ecuación de regresión no variaron una vez removida la variable interacción DPBMI-SEX y realizado nuevamente el análisis de regresión. La prueba de Shapiro-Wilk de los residuos del anterior análisis presentó valores W de 0.97 y un p-valor de 0.00, lo que rechaza el supuesto de normalidad de los residuos. El test de homocedasticidad de Breusch-Pagan y Cook-Weisberg rechaza el supuesto de varianza constante (p-valor=0.0050). La Figura 8 muestra el histograma realizado sobre los valores residuales del presente análisis. Este gráfico muestra que los valores residuales presentan una distribución tendiente a la normalidad, aunque se observa asimetría hacia los valores más altos.

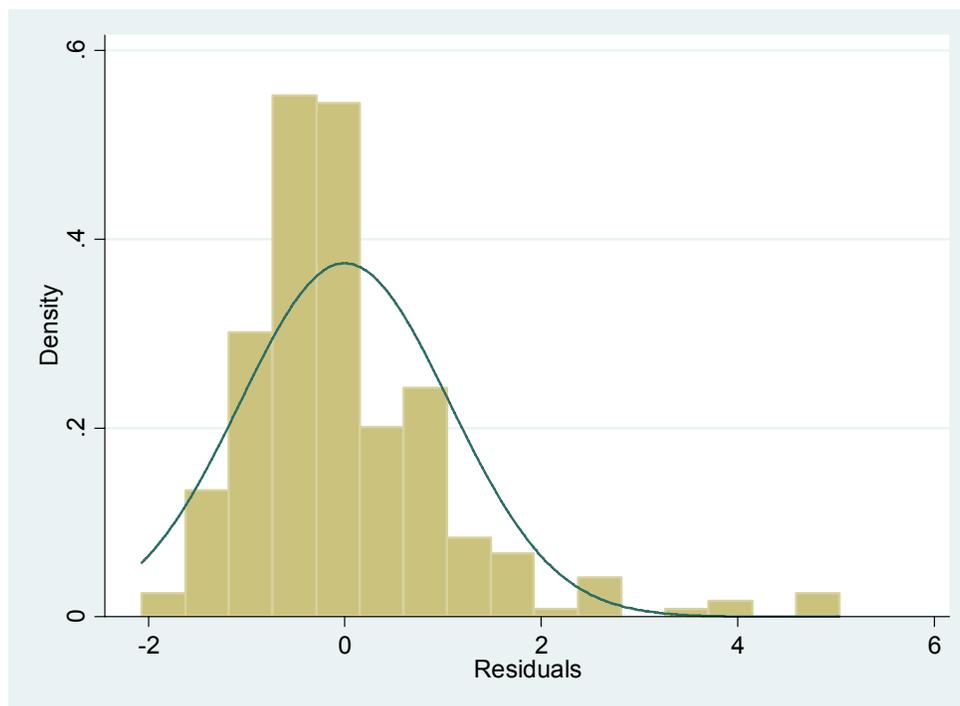


Figura 8. Histograma de los valores residuales.

El gráfico de dispersión entre FMI y DPBMI ajustado a una recta de regresión se presenta en la Figura 12. Se observa que existe una relación lineal positiva y leve entre estas variables, con una marcada dispersión de los casos y la notoria presencia de casos *outliers* en los valores FMI sobre ocho.

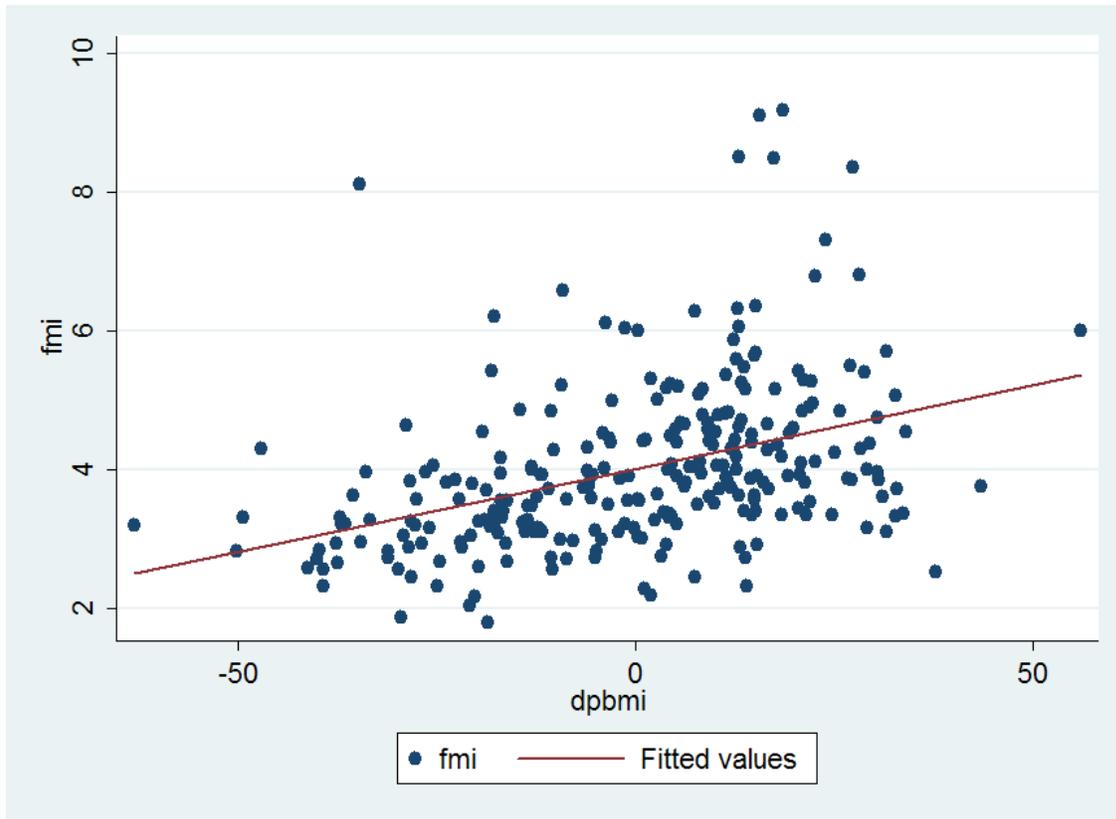


Figura 9. Dispersión y ajuste lineal entre FMI y DPBMI.

8.3.2. Análisis de regresión GAM

A continuación se presentan las modelaciones realizadas mediante la estrategia de Modelos Aditivos Generalizados (GAM), entre la variable respuesta FMI y las variables predictivas ganancia de las distintas mediciones del BMI (DBMI, DBAZ y DPBMI), el sexo (SEX) y la interacción entre ambas variables (DBMI-SEX, DBAZ-SEX y DPBMI-SEX).

8.3.2.1. Regresión FMI DBMI sexo e interacción.

La Tabla 13 muestra los resultados del modelo ajustado.

Modelo	1	2	3	4	5	6	7	8	9	10	11
Nº Funciones Base	1	2	3	4	5	6	7	8	9	10	11
Variables seleccionadas	DBMI	DBMI SEX									

Funciones base	BF3	BF3 BF15	BF3 BF11 BF15	BF3 BF10 BF11 BF15	BF3 BF10 BF11 BF12 BF15	BF3 BF10 BF11 BF12 BF13 BF15	BF3 BF8 BF10 BF11 BF12 BF13 BF15	BF3 BF4 BF8 BF10 BF11 BF12 BF13 BF15	BF3 BF4 BF8 BF10 BF11 BF12 BF13 BF15	BF1 BF3 BF4 BF6 BF8 BF10 BF11 BF12 BF13 BF15	BF1 BF3 BF4 BF5 BF6 BF8 BF10 BF11 BF12 BF13 BF15
GCV	1.192	1.182	1.197	1.173	1.174	1.203	1.220	1.244	1.274	1.307	1.343

Tabla 13. Tabla de resultados de funciones base y GCV del análisis MARS para la variables DBMI, SEX e interacción DBMI-SEX.

De ellos, el modelo con cuatro funciones base presentó el menor valor GCV, por lo que fue escogido. La ecuación predictiva resumida es la siguiente:

$$\mathbf{FMI = 3.58 + 0.58*BF3 - 6.09*BF10 + 2.05*BF11 - 0.44*BF15}$$

Hay que destacar que si bien son cuatro las funciones base que participan en esta ecuación, algunas de ellas presentan otras funciones base anidadas.

Las funciones base son las siguientes:

$$BF1 = \begin{cases} 0; & \text{si } DBMI = . \\ 1; & \text{si } DBMI \neq . \end{cases}$$

$$BF3 = \max(0, DBMI - (-0.92)) * BF1$$

$$BF5 = \max(0, SEX - (-0.0000000025)) * BF1$$

$$BF6 = \max(0, DBMI - (2.26)) * BF1$$

$$BF8 = \max(0, DBMI - (-1.05)) * BF1$$

$$BF10 = \max(0, SEX - (-0.0000000025)) * BF6$$

$$BF11 = \max(0, DBMI - (1.07)) * BF5$$

$$BF15 = \max(0, SEX - (-0.0000000025)) * BF8$$

Por lo tanto, la ecuación completa es:

$$FMI = 3.58 + 0.58 * [\max(0, DBMI + 0.92) * BF1] - 6.09 * [\max(0, SEX + 0.00000000025) * [\max(0, DBMI - 2.26) * BF1] + 2.05 * [\max(0, DBMI - 1.07) * \max(0, SEX + 0.00000000025) * BF1] - 0.44 * [\max(0, SEX + 0.0000000025) * \max(0, DBMI + 1.05) * BF1]$$

Los nodos de la presente ecuación son los valores DBMI -0.92, -1.05, 1.07 y 2.26.

Con el fin de explicar este modelo, se generaron once casos ficticios, los que se muestran en la Tabla 1 del Anexo 1. Estos contienen datos con valores negativos, cero, positivos y faltantes para la variable DBMI, en tanto que la variable SEX presenta valores cero para el sexo femenino, uno para el sexo masculino y datos faltantes. El resto de las variables muestran cada una de las funciones base mencionadas anteriormente (BF1, BF3, BF5, BF6, BF8, BF10, BF11 y BF15), además de la variable FMI estimada a partir de este modelo.

Supongamos que la variable DBMI no presenta un valor observable, *i.e.* es *missing*. Eso implica que la función base BF1 toma el valor cero (0). Luego, las funciones base BF3, BF5, BF6 y BF8 toman el valor cero. Ante esta situación, cualquier individuo, sin importar el sexo que tenga, presentará un FMI estimado de 3.58, correspondiente al valor del intercepto. Es decir, si DBMI es *missing*, el modelo MARS genera un valor a imputar, lo que muestra que la función BF1 es una función que permite imputar el valor en DBMI, ya que se encuentra anidada en cada función base.

Consideremos ahora que la variable DBMI presenta un dato observable, *i.e.* no es *missing*. En primer lugar, hay que hacer notar que esta función presenta cuatro nodos representado por las funciones base BF3, BF10, BF11 y BF15, los cuales serán denominados t1, t2, t3 y t4 con propósitos didácticos:

$$T1: 0.58 * [\max(0, DBMI + 0.92) * BF1]$$

$$T2: - 6.09 * [\max(0, SEXO + 0.0000000025) * [\max(0, DBMI - 2.26) * BF1]$$

$$T3: 2.05 * [\max(0, DBMI - 1.07) * \max(0, SEXO + 0.0000000025) * BF1]$$

$$T4: - 0.44 * [\max(0, SEXO + 0.0000000025) * \max(0, DBMI + 1.05) * BF1]$$

Recordemos el rango de DBMI para el sexo masculino y femenino:

$$DBMI = \begin{cases} \text{Masculino} & (-2.33; 3.22) \\ \text{Femenino} & (-2.7; 6.11) \end{cases}$$

Por lo tanto, dentro de ambos rangos de DBMI se hallan los cuatro nodos mencionados anteriormente, los cuales se multiplican por sus respectivos coeficientes, lo que se representa en la Figura 10:

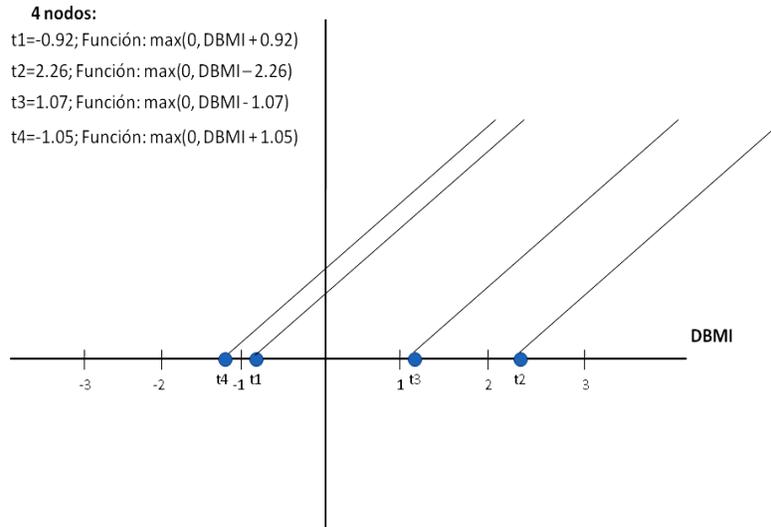


Figura 10. Representación de los cuatro nodos de la ecuación predictora del FMI a partir de DBMI.

Consideremos solamente $T1 = 0.58 * [\max(0, \text{DBMI} + 0.92) * BF1]$. T1 tiene anidada la función 1, dada por $[\max(0, \text{DBMI} + 0.92) * BF1]$, es decir, no considera el coeficiente. Entonces, Si $\text{DBMI} \leq -0.92$, entonces la función 1 toma valores cero ($[\max(0, -0.92+0.92)=0]$), sin importar el sexo que tenga el individuo. En cambio, si $\text{DBMI} > -0.92$, la función 1 aumenta (sin importar el sexo que tenga el individuo).

Por lo tanto, el FMI estimado mínimo y máximo para cada sexo serán los siguientes:

$$\text{FMI (estimado)}_{\text{mínimo|masculino}} = 3.58 + 0.58 * 0 = 3.58$$

$$\text{FMI (estimado)}_{\text{máximo|masculino}} = 3.58 + 0.58 * 3.22 + 0.92 = 5.98$$

$$\text{FMI (estimado)}_{\text{mínimo|femenino}} = 3.58 + 0.58 * 0 = 3.58$$

$$\text{FMI (estimado)}_{\text{máximo|femenino}} = 3.58 + 0.58 \cdot 6.11 + 0.92 = 7.66$$

Para el sexo masculino, los valores DBMI que se encuentran entre -2.7 y -0.92 van a entregar como resultados valores de FMI iguales a 3.58, en tanto que valores sobre 0.92 hasta 3.22 entregarán valores de FMI estimados que irán en aumento de acuerdo con el valor entregado por la función 1 y su correspondiente multiplicación por el coeficiente 0.58, entregando un valor FMI máximo de 5.98 (Figura 11).

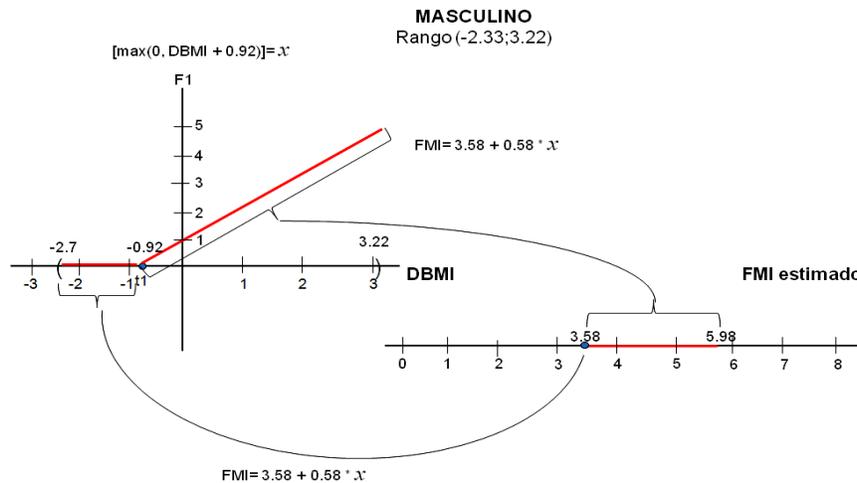


Figura 11. Esquema representativo de la función base BF3 y el intercepto para la estimación del FMI en el sexo masculino.

En tanto que para el sexo femenino, los valores DBMI que se encuentran entre -2.33 hasta -0.92 van a entregar como resultados valores de FMI iguales a 3.58, en tanto que valores sobre 0.92 hasta 6.11 entregarán valores de FMI estimados que irán en aumento de acuerdo con el valor entregado por la función 1 y su correspondiente multiplicación por el coeficiente 0.58, entregando un valor FMI máximo de 7.66 (Figura 12).

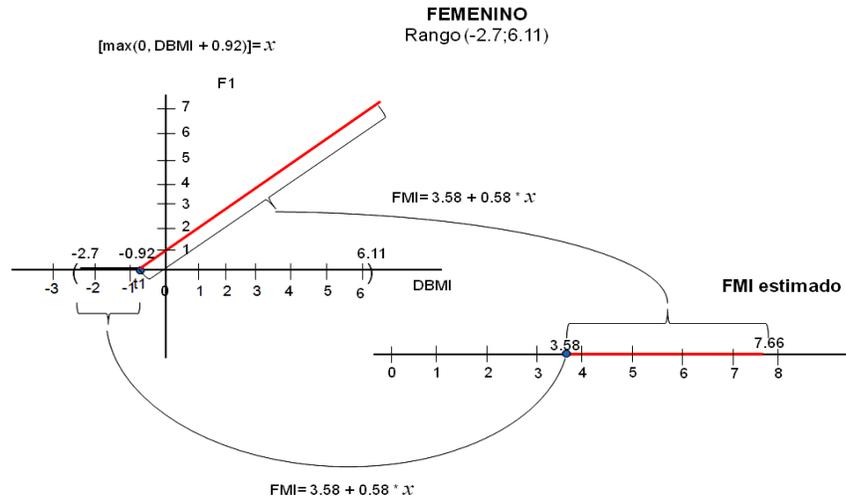


Figura 12. Esquema representativo de la función base BF3 y el intercepto para la estimación del FMI en el sexo femenino.

Consideremos ahora las funciones que presentan interacción con la variable SEX. La ecuación predictiva presenta tres funciones base en donde se encuentra asociada la variable SEX (BF5, BF10 y BF15). Es decir, si la variable SEX toma el valor cero (femenino), cada una de las funciones base se sumará a este valor, obteniendo como resultado el valor de las funciones base BF5, BF10 y BF15 igual a 0.0000000025, que es un valor cercano a cero, por lo que cualquier cálculo posterior, usando este valor como producto, presentará como resultados valores cercanos a cero. Respecto al sexo masculino (SEX=1), las funciones base BF5, BF10 y BF15 tendrán valores cercanos a uno. Lo anterior se puede resumir de la siguiente manera:

$$SEX = \begin{cases} 1 \Rightarrow \max(0; SEX + 0.0000000025) = 1.0000000025 \approx 1 \\ 0 \Rightarrow \max(0; SEX + 0.0000000025) = 0.0000000025 \approx 0 \end{cases}$$

Por ejemplo, la función base BF10 presenta anidada la función BF6. En caso de que la variable SEX sea cero (femenino), entonces la función base BF10 tendrá un valor cercano a cero, por lo que al multiplicarse por la variable BF6, también

presentará valores cercanos a cero. Por lo tanto, en el caso femenino, la función base BF10 no participa en la predicción del FMI, ni tampoco ninguna de las funciones base a excepción de BF3.

A diferencia, en el sexo masculino no sólo la función base BF3 participa en la estimación de FMI, sino que además lo hacen todas las funciones base restantes. Debido a que la variable SEX en el sexo masculino toma el valor uno (1), todas las funciones base que presenten esta variable como factor, sumarán el valor de 0.000000025, dando como resultado el valor 1.000000025, que es prácticamente uno.

Ahora, cada una de las funciones base que participan en el sexo masculino, lo harán de acuerdo al valor DBMI que adopte cada sujeto. Por ejemplo, si se cuenta con sujetos masculinos que presenten valores DBMI menores a -2.26 (Tabla 1, Anexo 1), sólo las funciones base que participan en la estimación del FMI son BF1 y BF5, la primera indicando que DBMI es un valor observado y la segunda indicando que se trata de un sujeto masculino. Por lo tanto, este sujeto sólo contará con el valor del intercepto para estimar el FMI.

Si se cuenta con sujetos con un valor DBMI igual a -1, las funciones base BF8 y BF15, participan en la estimación de FMI. Como BF8 se encuentra anidada dentro de BF15, ambas tendrán el mismo (0.05). Este valor se multiplica por la constante -0.44, obteniendo el valor de -0.022, el cual se sumará al intercepto:

$$3.58 + (-0.022) = 3.558 \approx 3.56$$

De este modo, el FMI estimado para este sujeto es cercano a 3.56. A medida que aumentan los valores de DBMI para los sujetos masculinos, las diferentes funciones base irán sumándose en su contribución en la estimación del FMI.

En el sexo masculino, los individuos son estimados sólo por el intercepto si hay casos con DBMI inferiores a -1.05. Es decir, si la ganancia de BMI es negativa en

aproximadamente una desviación estándar, FMI no se ve afectado. Si DBMI no excede el valor de 2.26, sólo participan las funciones base BF3, BF11 y BF15, en tanto que si superan este valor, se suma la función base BF10. En consecuencia, los casos masculinos con altos valores de ganancia de BMI entre los seis y 24 meses tendrán mayores valores de FMI estimado respecto al sexo femenino.

La Figura 13 muestra las relaciones de las interacciones de las funciones base asociadas a este análisis.

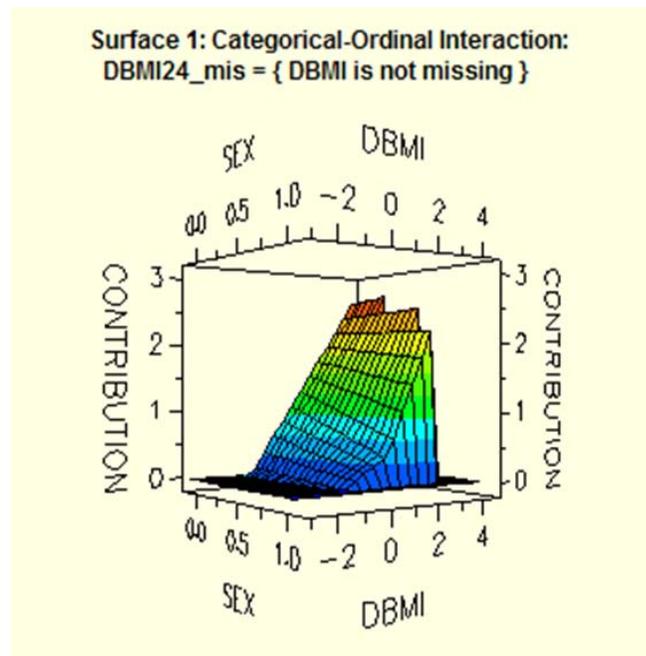


Figura 13. Superficie de interacción de funciones base de predictores seleccionados por MARS en el presente análisis.

8.3.2.2. Regresión FMI DBAZ sexo e interacción.

La Tabla 14 muestra los resultados del presente análisis.

Modelo	1	2	3	4	5	6	7	8	9	10	11
Nº Funciones Base	1	2	3	4	5	6	7	8	9	10	11
Variables seleccionadas	DBAZ	DBAZ SEX									

Funciones base	BF3	BF3 BF5	BF3 BF4 BF5	BF3 BF4 BF5 BF14	BF3 BF4 BF5 BF13 BF14	BF3 BF4 BF5 BF6 BF13 BF14	BF3 BF4 BF5 BF6 BF8 BF13 BF14	BF3 BF4 BF5 BF6 BF8 BF13 BF14 BF15	BF1 BF3 BF4 BF5 BF6 BF8 BF13 BF14 BF15	BF1 BF3 BF4 BF5 BF6 BF8 BF10 BF13 BF14 BF15	BF1 BF3 BF4 BF5 BF6 BF8 BF10 BF12 BF13 BF14 BF15
GCV	1.238	1.199	1.188	1.215	1.209	1.238	1.268	1.302	1.337	1.374	1.415

Tabla 14. Tabla de resultados de funciones base y GCV del análisis MARS para la variables DPBAZ, SEX e interacción DPBAZ-SEX.

El modelo con tres funciones base presentó el menor valor GCV, por lo que fue seleccionado.

La ecuación predictiva es la siguiente:

$$\mathbf{FMI = 4.24 + 0.91*BF3 - 0.48*BF4 - 0.55*BF5}$$

Las funciones base asociadas a esta ecuación son las siguientes:

$$BF1 = \begin{cases} 0; & \text{si } DBAZ = . \\ 1; & \text{si } DBAZ \neq . \end{cases}$$

$$BF3 = \max(0, DBAZ - (0.19)) * BF1$$

$$BF4 = \max(0, 0.19 - (DBAZ)) * BF1$$

$$BF5 = \max(0, SEX - (-0.0000000024)) * BF1$$

Por lo tanto, el modelo completo es:

$$FMI = 4.24 + 0.91*[\max(0, DBAZ - 0.19) * BF1] - 0.48* \max(0, 0.19 - DBAZ) * BF1] - 0.55*] * \max(0, SEX + 0.0000000024) * BF1]$$

Los nodos que participan en este modelo son los valores DBAZ 0.19 y -0.19.

Para ambos sexos se observa una relación positiva desde valores DBMI -2 hasta 0.19, pero cuando alcanza este valor, hay un aumento de la pendiente, debido a la interacción de la función base BF3.

La única diferencia que se observa entre ambos sexos es la función base BF5, la cual se asocia a la variable SEX, reduciendo en -0.55 el valor de FMI obtenido por las funciones base restantes en el sexo masculino respecto al femenino. Es decir, para valores idénticos de DBAZ, el valor estimado de FMI será menor en 0.55 en el sexo masculino respecto al femenino.

8.3.3.3. Regresión FMI DPBMI sexo e interacción.

La Tabla 15 muestra los resultados del presente análisis.

Modelo	1	2	3	4	5	6	7	8	9	10
Nº Funciones Base	1	2	3	4	5	6	7	8	9	10
Variables seleccionadas	DPBMI	DPBMI I SEX	DPBMI I SEX	DPBMI I SEX	DPBMI I SEX	DPBMI I SEX	DPBMI I SEX	DPBMI I SEX	DPBMI I SEX	DPBMI I SEX
Funciones base	BF3	BF3 BF5	BF3 BF5 BF12	BF3 BF5 BF12 BF14	BF3 BF5 BF8 BF12 BF14	BF3 BF5 BF8 BF10 BF12 BF14	BF1 BF3 BF5 BF8 BF10 BF12 BF14	BF1 BF3 BF5 BF6 BF8 BF10 BF12 BF14	BF1 BF3 BF5 BF6 BF8 BF10 BF11 BF12 BF14	BF1 BF3 BF4 BF5 BF6 BF8 BF10 BF11 BF12 BF14
GCV	1.254	1.196	1.214	1.202	1.223	1.248	1.275	1.305	1.337	1.372

Tabla 15. Tabla de resultados de funciones base y GCV del análisis MARS para la variables DPBMI, SEX e interacción DPBMI-SEX.

De estos modelos, el modelo con dos funciones base presentó el menor valor de GCV, por lo que fue seleccionado.

La ecuación predictiva es la siguiente:

$$\mathbf{FMI = 3.51 + 0.026*BF3 - 0.58*BF5}$$

En donde las funciones-base asociadas a esta ecuación son las siguientes:

$$BF1 = \begin{cases} 0; & \text{si } DPBMI = . \\ 1; & \text{si } DPBMI \neq . \end{cases}$$

$$BF3 = \max(0, DPBMI - (-28.49)) * BF1$$

$$BF5 = \max(0, SEX - (-0.00000000024)) * BF1$$

El nodo que participa en esta ecuación es el valor DPBMI de -28.49. En este caso, todos los valores de DPBMI menores a -28.49 presentan un FMI constante de 3.51, en tanto que valores de FMI mayores a -28.49 presentan un cambio del FMI debido a la interacción con la función base BF3, sin importar el sexo que presente el caso.

La función base BF5 se asocia a la variable SEX, por lo tanto, el hecho de ser masculino reduce el valor del FMI en -0.58 en comparación con el sexo femenino.

8.4. Validación de los modelos predictivos

La Tabla 16 presenta los resultados de las estadísticas utilizadas para comparar el FMI observado y el FMI estimado a partir de los modelos GLM y GAM realizado con las variables DBMI, DBAZ, DPBMI y sexo en la submuestra de validación.

MPAB es la proporción media del estimador de error absoluto (*Mean Proportion of Absolute Estimation Error*); R^2_{NOC} y $COEF_{NOC}$ son respectivamente el coeficiente de determinación y el coeficiente de la variable predictiva, calculados a partir de una regresión sin intercepto entre el FMI observado y el estimado por cada enfoque (GLM y GAM) y para cada medición del BMI (DBMI, DBAZ, DPBMI). $COEF_{NOC}$ no será utilizado como estimador de la precisión y certeza de los estimadores, sino que será utilizado para conocer si el estimador sobreestima o subestima la variable respuesta FMI observada.

LIN hace referencia al coeficiente de correlación de concordancia de Lin (1989, 2000), en tanto que **CERTEZA** es el factor de corrección de sesgo (*bias-correction factor*), que hace referencia a la medida de la certeza en el coeficiente de correlación de concordancia de Lin, en tanto que r **PEARSON** es el coeficiente de correlación de Pearson, el cual corresponde a la medida de la precisión en el coeficiente de correlación de concordancia de Lin (Tabla 16).

	DBMI;SEX		DBAZ;SEX		DPBMI;SEX	
	GLM	GAM	GLM	GAM	GLM	GAM
MPAB	0.228	0.241	0.233	0.224	0.231	0.228
R^2_{NOC}	0.920	0.916	0.920	0.920	0.922	0.923
$COEF_{NOC}$	0.997	0.996	0.996	0.999	0.998	1.001
LIN	0.301	0.245	0.312	0.295	0.308	0.290
CERTEZA	0.798	0.791	0.817	0.818	0.782	0.750
r PEARSON	0.376	0.309	0.382	0.361	0.394	0.387

Tabla 16. Resultados de los estimadores de comparación entre FMI observado y estimado.

Respecto a $COEF_{NOC}$, se observa que en general los modelos tienden a subestimar el FMI, a excepción del modelo en donde se utilizó la variable PBMI como predictor y GAM como estrategia analítica. De estos resultados se infiere que el FMI estimado se acerca notoriamente al FMI observado en la muestra de validación en todos los métodos, lo que contrasta con las Figuras 1 a 6 del Anexo 2, que muestran los gráficos de concordancia de los datos ajustados al eje mayor de los datos (línea verde) y al eje de la perfecta concordancia (línea roja). Como se observa en estas figuras, en ninguno de los modelos propuestos se logra una concordancia cerca de la línea de la perfecta concordancia, sino que en general están relativamente desviados. El modelo que más llama la atención es el modelo GAM, utilizando la variable DBMI, la cual presenta una gran acumulación de datos al inicio de la nube de puntos y que corresponde a valores que sólo se les pudo estimar el FMI a partir del intercepto.

Llama la atención que R^2_{NOC} presente valores tan altos en todas sus estimaciones a diferencia de los otros estadísticos de validación. Al respecto, se realizó una regresión entre los valores de FMI estimados y observados en la muestra de validación, agregando el intercepto, con el fin de observar qué valores de R^2 son observados, además de los valores de los coeficientes que arroja cada uno de estos análisis (Tabla 17).

GLM			GAM		
dBMI	dBAZ	dPBMI	dBMI	dBAZ	dPBMI

R^2	0.13	0.14	0.15	0.26	0.25	0.22
β_0	0.97	1.04	0.73	-0.17	1.03	1.02
β_1	0.76	0.74	0.82	1.04	-0.08	-0.04

Tabla 17. Resultados del análisis de regresión entre los valores estimados de FMI y los valores observados de FMI en la muestra de validación.

Los resultados muestran que los valores de R^2 son notoriamente más bajos en comparación con los análisis sin intercepto. Además, se advierte que dentro de los modelos GLM DBMI presenta menor valor de R^2 en comparación con DBAZ, en tanto que en las regresiones sin intercepto es DBAZ, el que presenta mayores valores de R^2 en comparación con DBMI. DPBMI en ambas regresiones es la medición que presenta mayores valores de R^2 . Con el fin de graficar la relación entre los valores observados y estimados del FMI utilizando las distintas mediciones del BMI se realizaron gráficos de dispersión ajustados a una recta de regresión (Anexo 3). Los valores R^2 obtenidos de una regresión lineal a partir de los valores de FMI estimados a partir de los modelos GAM contra el FMI observado en la muestra de validación, también presentan una importante disminución en comparación con los valores R^2 calculados mediante una regresión sin intercepto. En este caso, también ocurre un cambio respecto a cuál es la mejor medición del BMI, debido a que

Las Tablas 18 y 19 son una síntesis de la Tabla 16, con el fin de conocer qué medición del BMI es mejor para predecir el FMI de acuerdo a cada enfoque analítico, GLM o GAM. Esta tabla se realizó marcando cuál de las pruebas de validación entregó el mejor valor para cada uno de los enfoques. Mientras que el mejor valor de **MPAB** está asociado al puntaje más cercano a cero (ausencia de error), el resto de las pruebas considera el mejor valor a que se encuentra más cercano a la unidad. Por ejemplo, considerando únicamente los modelos GLM, la prueba **MPAB** favorece a DBMI, ya que presenta el puntaje más bajo, es decir, el menor error entre el valor observado y el estimado, en tanto que R^2_{NOC} favorece a

DPBMI, ya que presenta un valor de 0.922, valor más alto que 0.920, perteneciente tanto a DBMI como DBAZ.

Si se realizan predicciones de FMI con modelos GLM, las mediciones recomendadas serían DBAZ y DPBMI, ya que ambas presentan la misma cantidad de pruebas de validación a su favor, en tanto que DBMI presenta únicamente una medidas a su favor (**MPAB**).

GLM	DBMI	DBAZ	DPBMI
MPAB	X	.	.
R^2_{NOC}	.	.	X
LIN	.	X	.
CERTEZA	.	X	.
r PEARSON	.	.	X
TOTAL	1	2	2

Tabla 18. Resumen de los resultados de las validaciones organizados por medición del BMI sólo para los modelos GLM.

Por otra parte, los análisis con modelos GAM favorecen a DBAZ, debido a que presenta tres pruebas de validación a su favor (**MPAB**, **LIN** y **CERTEZA**), en tanto que DPBMI presenta dos (R^2_{NOC} y **r Pearson**) y DBMI ninguno (Tabla 18).

GAM	DBMI	DBAZ	DPBMI
MPAB	.	X	.
R^2_{NOC}	.	.	X
LIN	.	X	.
CERTEZA	.	X	.
r PEARSON	.	.	X
TOTAL	0	3	2

Tabla 19. Resumen de los resultados de las validaciones organizados por medición del BMI sólo para los modelos GAM.

Por otra parte, también es importante conocer qué estrategia estadística (GLM o GAM) es la recomendada para predecir mejor FMI. Esto puede evaluarse a partir de las mismas Tablas 18 y 19. Por ejemplo, si se cuenta con DBMI, es recomendable utilizar un modelo GLM, en tanto que si se cuenta con DBAZ es recomendable

utilizar un modelo GAM. Finalmente, para DPBMI ambos enfoques presentan la misma cantidad de estimadores a su favor, por lo que sería recomendable evaluar el propósito del uso de estos enfoques estadísticos para conocer cuál es el recomendable.

9. DISCUSION

A partir de las estadísticas descriptivas, la inspección de las variables no mostró diferencias notorias entre la muestra total y las submuestras de entrenamiento y validación. De esto se desprende que en la generación de las submuestras, los datos se segregaron homogéneamente entre ellas, evitado así posibles sesgos muestrales.

Por otra parte, se observó que las variables en estudio no presentan un comportamiento simétrico y rechazan la prueba de normalidad de Shapiro-Wilk, a excepción de las variables BAZ6 y BAZ24.

DBMI, DBAZ y DPBMI presentaron una media muy cercana a cero en la muestra y submuestras, debido a la forma en cómo se construyeron estas variables, que es mediante la obtención de los valores residualizados, descrito anteriormente, técnica utilizada comúnmente para medir el cambio de una medición tomada en dos tiempos diferentes (DuBois, 1957, 1970; Keijzer-Veen, *et al.* 2005).

La media de los percentiles del BMI y BAZ a los seis y 24 meses (BMI6, BMI24, BAZ6 y BAZ24) se encuentran sobre 50 y 0.5 respectivamente, indicando que la muestra presenta valores por sobre lo registrado internacionalmente (NCHS-WHO), sugiriendo la presencia de obesidad o al menos sobrepeso. Lo anterior ha sido descrito previamente para el universo de niños asistentes a JUNJI por Rojas y Corvalán (2010), por lo que en el presente estudio se considera como lo esperado para la población chilena de nivel medio-bajo socioeconómico.

9.1 Modelación por GLM

Cuando se construyeron los modelos predictivos mediante el enfoque GLM, se observó que las variables de interacción entre la ganancia de las diferentes mediciones del BMI y la variable SEX no resultaron ser significativas, por lo que se eliminaron de los análisis. Esto quiere decir que la relación entre la ganancia de las diferentes mediciones del BMI con el FMI a los cinco años no difiere por sexo.

En general, los coeficientes de determinación presentaron bajos valores, sin exceder en ninguno de ellos el valor 0.25. La variable SEX, en tanto, presentó valores negativos, lo que implica que el hecho de ser hombre (1) reduce en general la estimación del FMI respecto al sexo femenino (0). Lo anterior, debido a que en general el sexo femenino presenta mayor masa grasa y menor masa magra en comparación con el sexo masculino, lo que se ha observado ya a nivel fetal (Wells, 2007; Kirchengast, 2010) y que se refleja en las medidas de corpulencia.

Respecto a la revisión de los supuestos de los modelos lineales generalizados, se observó que los valores residuales de cada uno de estos modelos rechazaron la prueba de normalidad de Shapiro-Wilk. Sin embargo, en los respectivos histogramas de los valores residuales, no se observaron violaciones notorias de la normalidad. Debido a que en la presente Tesis se estaban buscando violaciones graves de normalidad (*gross violation of normality*), no se tomó finalmente en cuenta el rechazo de la prueba de normalidad de Shapiro-Wilk, debido a que es una prueba muy sensible de la normalidad. Por lo tanto, en el presente trabajo se asume que los modelos lineales generalizados obtuvieron buenos ajustes entre la variable respuesta (FMI) y sus variables predictivas (ganancia de las mediciones del BMI y SEX).

Además, se observó que los análisis de homocedasticidad realizados sobre cada uno de los modelos GLM rechazaron el supuesto de varianza constante, lo cual también es observable en los gráficos de dispersión entre la variable respuesta y las mediciones del BMI. Es posible que esto esté evidenciando que existen otras variables explicativas que no se incluyeron en el modelo, lo cual pueda estar influyendo en la varianza de los datos. Si bien esto es bastante plausible, en la presente investigación, la finalidad no fue hallar todas las variables que explican la variación del FMI, sino que únicamente fue conocer cuál de las distintas mediciones del BMI explica mejor la variación del FMI.

Por otra parte, la revisión de los gráficos de dispersión entre la variable respuesta y las mediciones del BMI dan cuenta de una relación lineal positiva, pero leve, una notoria dispersión de los datos en torno a las rectas de ajuste (Figuras 8, 10 y 12) y la presencia de cinco valores extremos para los modelos completos, tres correspondientes al sexo femenino y dos al sexo masculino, que se ubican en los valores más altos de FMI. Cuando se revisaron estos casos *outliers*, se constató que son casos obesos que se diferencian notoriamente del resto de la muestra, pero que no pueden ser descartados del análisis.

En conclusión, el análisis GLM muestra un comportamiento bastante similar en la predicción de la ganancia de las distintas mediciones del BMI, sobre todo entre las variables BAZ y PBMI.

9.2 Modelación por GAM

Los modelos GAM también presentaron características bastante similares entre ellos, las cuales se describen a continuación.

Los modelos GAM muestran que los valores estimados de FMI no se alteran respecto a ciertos valores de ganancia de las mediciones de BMI, tales como valores de ganancia negativa o cercana a cero (-0.92 para DBMI, 0.19 para DBAZ y

-28.59 para DPBMI). Es decir, existe un límite mínimo en donde los individuos no pierden FMI a pesar de tener una ganancia negativa o baja en las mediciones del BMI.

Por otra parte, si la ganancia de las distintas mediciones de BMI supera este valor “umbral”, la relación entre FMI y la ganancia de cada una de las mediciones del BMI aumenta linealmente y de forma sostenida, con algunas variaciones dependiendo del tipo de medición del BMI. Esto quiere decir que si la ganancia de BMI es suficiente, éste contribuye en el FMI estimado de modo constante.

Para ciertos casos, como fue observado en la construcción de los modelos DBMI y DBAZ, existe un segundo punto de relación entre FMI y las mediciones del BMI. En este caso, una ganancia más importante de las mediciones del BMI incrementa aún más la relación con FMI, es decir, hay una mayor contribución de las mediciones del BMI sobre el puntaje de FMI alcanzado.

De modo similar al observado en los modelos GLM, el sexo masculino presenta en general menor FMI estimado respecto al sexo femenino, dado el mismo valor de la medición del BMI.

También se observó en algunas de las mediciones del BMI, que existen otros nodos que generan cambios en la predicción del FMI a lo largo de la distribución de las variables predictivas, es decir, que a medida que se incrementa la ganancia de las diferentes mediciones del BMI, puede cambiar el FMI estimado, lo que se observa generalmente como un incremento del FMI a medida que aumentan los valores de la ganancia de las diferentes mediciones del BMI.

Al respecto, un caso notorio es el observado cuando se utilizó la variable DBMI en el sexo masculino. En este caso, se observaron cuatro nodos que interactuaron a medida que aumentan los valores de DBMI, generando cambios notorios en la estimación del FMI. Los nodos dan cuenta de un leve aumento del FMI estimado

asociado al DBMI, desde valores de ganancia negativa (-0.92) hasta niveles de ganancia positiva del BMI (1.07). Posteriormente, el FMI se incrementó aún más con valores DBMI por sobre 1.07. Finalmente, el coeficiente asociado a la función base BF15 da cuenta de una abrupta caída en la estimación del FMI en valores DBMI por sobre 2.26. Llama la atención este último cambio asociado a los más altos de DBMI, por lo que se tabularon los casos que presentaron este comportamiento (Tabla 20).

SEX	DBMI	FMI
1	2.53	5.38
1	3.04	3.38

Tabla 20. Casos masculinos con valores DBMI sobre dos.

Como puede observarse, sólo estos dos casos se utilizaron para construir la función base BF15 que es la que genera esta abrupta caída del FMI estimado. Es decir, la función base BF15 fue construida únicamente con estos dos valores, lo que constituye más bien un hecho casuístico más que una generalidad. Debido a lo anterior, no es recomendable considerar este fenómeno como biológicamente plausible, por lo que se descartó como un fenómeno que pueda generalizarse.

9.3 Validación de los modelos

9.3.1. Medición de la ganancia de BMI

Cuando se realizaron las pruebas de validación entre el FMI estimado a partir de cada uno de los seis modelos y el FMI observado en la muestra de validación (Tabla 16), se observó que los resultados de las validaciones presentaron coeficientes bastante similares entre ellos, hallándose sólo algunos casos con notorias diferencias, como DBAZ y DPBMI para los modelos GAM (*LIN* y *r Pearson*).

De todos modos, al revisar las mediciones del BMI dentro de cada modelo estadístico, se observó que tanto para GLM como para GAM, las mediciones que

obtuvieron la mayor cantidad de pruebas de validación fueron DBAZ y DPBMI. Por lo tanto, en el presente trabajo puede aceptarse que DBAZ y DPBMI serían las mediciones de ganancia del BMI más óptimas para predecir la adiposidad tardía.

Algunas de las características que hacen de DBAZ y DPBMI estimadores óptimos de ganancia de BMI para predecir la adiposidad tardía es que atenúan el cambio de BMI en los más obesos, eliminando la asimetría del BMI crudo. Posiblemente estas transformaciones no estén generando mayores sesgos, sino que al contrario, permiten un mejor análisis y presentan mejor asociación con FMI. Es posible que el BMI crudo presente la menor cantidad de estimaciones a su favor, porque las modelaciones estadísticas no consiguieron ajustar la asimetría que le es propia a su distribución. Es posible que la reducción de la variabilidad de las distribuciones de las mediciones transformadas de BMI consiga mejores ajustes que BMI crudo, tanto para los modelos GLM como GAM.

Estos resultados difieren respecto a los encontrados en Cole y cols. (2005). Hay que destacar que el trabajo de Cole y cols. (*op. cit.*) no puede compararse con el presente trabajo, debido a que los objetivos que persiguen ambos trabajos son disímiles: el objetivo de estos autores es medir a corto plazo los cambios en adiposidad en niños normales y obesos. Ellos concluyeron que la ganancia de BMI y de %BMI es mejor. Su motivación es saber qué indicador utilizar para saber si sujeto responde o no a un tratamiento. Lógicamente esto es particularmente útil en niños obesos.

A diferencia, en el presente trabajo, el objetivo es identificar cuál es la mejor forma de expresar los cambios del BMI para predecir adiposidad posterior dada nuestra prevalencia conocida de obesidad en la población. En este caso, es mejor utilizar mediciones que disminuyan la asimetría de las colas de la distribución. Con esto, se logra que los modelos predigan de mejor forma la realidad.

En el presente trabajo, la edad de los individuos de la muestra se ubica antes de la aparición del rebote adiposo (en inglés *adiposity rebound*), en un momento que contiene la aparición del primer *peak* de BMI. A diferencia, la muestra de Cole y cols. (*op. cit.*) se halla precisamente durante la aparición del rebote adiposo, es decir, en un momento más tardío que la del presente trabajo. Si bien ambos momentos son completamente diferentes, también comparten la cualidad de ser periodos de cambios dinámicos en la composición corporal de los niños. Son precisamente estas transiciones las que le confieren importancia en el desarrollo de la pregunta de investigación del presente trabajo. Cuál medición del BMI es la mejor predictora de la grasa corporal durante estos periodos transicionales es posiblemente más importante que conocer cuál es el mejor predictor durante periodos de crecimiento sostenido (Cole *op. cit.*, p.424). Posiblemente, las notorias fluctuaciones del BMI que suceden durante este periodo pueden afectar las estimaciones, como fue observado, por ejemplo, en los resultados del coeficiente de determinación en los modelos GLM, los cuales no superan un valor de 0.25. De todos modos, es importante notar que en el presente trabajo no se analizaron eventuales transformaciones cuadráticas o cúbicas de las variables predictivas, las cuales podrían aportar mayor flexibilidad a las curvas, dando quizás mejores ajustes respecto a FMI.

También, es posible que otras variables predictivas pueden explicar la variabilidad de la variable respuesta. En este sentido, hay que destacar que en el presente trabajo se intentó predecir el FMI a los cinco años a partir de la ganancia de las distintas mediciones del BMI entre los seis y 24 meses, edades bastante alejadas de la edad de la variable respuesta. Es posible que los cambios asociados a la adiposidad durante este periodo no logren una buena predicción.

Un punto que merece atención fueron los resultados del R^2_{NOC} en comparación con el R^2 considerando el intercepto. Mientras que en el primer caso se observan valores muy altos, en el segundo caso los valores disminuyen drásticamente,

además de observarse que el intercepto en muchos casos se distancia de cero, dando cuenta que la relación entre el FMI observado y el estimado no pasa por el origen, lo que también se observa en los Anexos 2 y 3. Si bien teóricamente, la relación entre el FMI observado y el estimado se asume que pasa por el origen, en este caso en particular se observa que no sucede esto, llevando a la suposición de que es posible que este tipo de validación no sea el adecuado para interpretar o validar estos datos. Debido a que el estudio de la utilidad de estos estadísticos de validación trasciende el presente trabajo, no es posible hasta el momento tomar una decisión sobre la utilidad del R^2 obtenido a partir de una regresión sin intercepto entre valores observados y estimados.

9.3.2. Enfoque estadístico (GLM o GAM)

En el presente trabajo, las modelaciones MARS no se diferenciaron notoriamente respecto de GLM respecto a la predicción del FMI. Según Alvarado *et al.* (2010), la modelación MARS selecciona aquellas variables predictivas significativas y detecta posibles interacciones entre ellas, generando modelos más flexibles desde el punto de vista de la interpretación. Ya que las interacciones están siempre restringidas a alguna subregión, éstas quedan expresadas algebraicamente mediante las funciones base, logrando de esta forma establecer un modelo parsimonioso que representa sin ningún tipo de transformación adicional la naturaleza propia de las variables que se están trabajando. Crea nodos o puntos de corte que actúan como valores umbral para cada variable predictiva seleccionada, indicando el cambio que se genera en la contribución por parte de la función base a la respuesta en estudio.

Una vez seleccionado el modelo óptimo, MARS ajusta nuevamente un modelo para cada variable, de modo de determinar el impacto de la calidad del modelo al eliminar dicha variable; así se asigna un *ranking* relativo desde la variable más importante a la menos importante. De esta manera se definen variables de reemplazo o competidoras, con lo cual la metodología MARS permite tratar los

valores perdidos o faltantes generando una función base exclusiva para aquellas variables que no poseen información, es decir, se genera una función base de imputación cuya finalidad es imputar el valor promedio de la variable predictiva sin información.

En este sentido, se observó que la flexibilidad que teóricamente otorgan los modelos GAM produjo predicciones con relativamente pocos nodos, quizás indicando que no es necesaria tanta flexibilidad para predecir el FMI a partir de la ganancia de las diferentes mediciones de BMI. Es posible también que la elección del mejor modelo GAM en base al valor más bajo de *Generalized Criterion Value* (GCV) pueda estar seleccionando modelos muy poco flexibles, por lo que se necesitan revisar modelos con valores GCV más altos que los seleccionados, pero que se ajusten mejor a la variable respuesta. En este sentido, el criterio de selección de los modelos MARS fue más bien predictivo que descriptivo o exploratorio, al favorecer los valores GCV más bajos, a diferencia de haber utilizado un criterio más cualitativo, consistente en la revisión más detenida de cada modelo y de las interpretaciones que se pueden obtener al revisar los nodos que genera cada uno de ellos.

De todos modos, el aporte de los modelos GAM en el presente trabajo es que estos modelos muestran la aparición de nodos que marcan una relación diferencial entre FMI y las distintas mediciones de la ganancia de BMI, denotando un aumento de la relación, una vez superado un umbral de ganancia de BMI. Esta situación es evidente en todos los modelos, ya que tanto para hombres como para mujeres, el haber alcanzado cierto valor de ganancia del BMI genera un incremento de la relación entre esta ganancia FMI. Sin embargo lo anterior, ello no necesariamente se asocia a la aparición de un límite, tales como el sobrepeso o la obesidad, sino que marcan un cambio en la relación FMI-ganancia BMI. En este sentido, es interesante entender epidemiológicamente estos nodos y su asociación con FMI, lo que puede ayudar a detectar episodios tempranos de ganancia de peso, alertando sobre eventuales casos de sobrepeso u obesidad en el futuro.

Queda abierta la pregunta sobre si existen otros modelos estadísticos o variaciones de modelos GLM que puedan arrojar resultados más apropiados que los analizados en la presente investigación.

10. CONCLUSIONES

1. Las medidas de ganancia del BMI que obtuvieron mayor cantidad de estimadores de concordancia respecto a la variable respuesta FMI fue DBAZ y DPBMI, por lo que se sugieren como los mejores indicadores de cambio de BMI para la predicción de FMI a los 5 años. De todos modos, se acepta que las predicciones del FMI presentaron una baja concordancia en comparación con el FMI observado en la muestra de validación.
2. Los enfoques estadísticos GLM y GAM arrojaron buenos resultados en la construcción de los modelos. Se sugiere que GLM podría ser utilizado como un buen modelo para predecir las asociaciones entre la ganancia de las distintas mediciones del BMI sobre FMI, en tanto que GAM puede utilizarse para conocer qué valores específicos de ganancia de las mediciones de BMI pueden generar variaciones importantes en la estimación del FMI, por lo que pueden ser utilizados como puntos de corte para definir riesgo.
3. Es posible que los sesgos e incertezas observados entre el FMI observado y el estimado por los modelos propuestos se deba a que el rango etario utilizado para la modelación corresponde a una etapa de transición respecto a las medidas corporales y los estimadores del BMI, lo que puede afectar la precisión y certeza de los modelos predictivos estudiados.
4. Debido a que las variables predictivas explicaron una escasa fracción de la variabilidad de la variable respuesta, es posible que otras variables logren una mayor predictibilidad. También es posible que la brecha etaria entre las

variables predictivas y la respuesta pueda estar generando estas imprecisiones y sesgos, debido a la dinámica de cambio que viven los sujetos a esta edad.

11. BIBLIOGRAFÍA

Agras S, Hammer L, Mc Nicholas F, Kraemer H. 2004 Risk factors for childhood overweight: a prospective study from birth to 9.5 years. *Journal of Pediatrics*; 145:20-25.

Alvarado S, Silva C, Cáceres D. 2010 Modelación de episodios críticos de contaminación por material particulado (PM10) en Santiago de Chile. Comparación de la eficiencia predictiva de los modelos paramétricos y no paramétricos. *Gac Sanit*; 24(6):466-472.

Baird, J., Fisher, D., Lucas, P., Kleijnen, J., Roberts, H. y Law, C. 2005 Being big or growing fast: systematic review of size and growth in infancy and later obesity. *British Medical Journal*. 22; 331(7522): 929

Barlow SE, Trowbridge FL, Klish WS & Dietz WH (2002): Treatment of child and adolescent obesity: reports from pediatricians, pediatric nurse practitioners and registered dietitians. *Pediatrics* 110, 229–235.

Bell, C., Walley, A. y Froguel, P. 2005 The genetics of human obesity. *National Genetics Review*; 6: 221 – 235.

Burrows R, Gattás V, Leiva L, Barrera G, Burgueño M. 2001 Características biológicas, familiares y metabólicas de la obesidad infantil y juvenil. *Revista Médica de Chile*; 129: 1155-1162.

Cameron, N., Norgan, N. y Ellison, G. 2006 Introduction. En *Childhood Obesity: Contemporary Issues* Cap. 1 Introduction. CRC editores.

Cameron N, Pettifor J, De Wet T, Norris S. 2003 The relationship of rapid weight gain in infancy to obesity and skeletal maturity in childhood. *Obesity Research*; 11: 457-460.

Cole, T. 1991 Weight-stature indices to measure underweight, overweight and obesity. En Himes, J. *Anthropometric Assessment of Nutritional Status*. New York: Alan Liss p.83

Cole T. 2004 Children grow and horses race: is the adiposity rebound a critical period for later obesity. *BMC Pediatrics*; 4:6-11.

Cole, T. J., Bellizzi, M. C., Flegal, K. M., and Dietz, W. H. 2000 Establishing a standard definition for child overweight and obesity: international survey. *BMJ*, 320, 1240–1243.

Cole, T. y Rolland-Cachera, M. 2002 Measurement and definition. En Burniat W., Cole T., Lissau I. y Poskitt E. (eds.) *Child and Adolescent Obesity. Causes and Consequences, Prevention and Management*. Cambridge University Press p.1

Cole, T. Faith, M., Pietrobelli, A. y Heo, M. 2005 What is the best measure of adiposity change in growing children: BMI, BMI%, BMI z-score or BMI centile? *European Journal of Human Nutrition* 59, 419-425

Cole T. 2006 Childhood obesity: assessment and prevalence. En Cameron, N., Norgan, N. y Ellison, G. *Childhood Obesity: Contemporary Issues* Cap. 1. CRC editores. p.3-12.

Corvalán C, Uauy R, Stein AD, Kain J, Martorell R. 2009 Effect of growth on cardiometabolic status at 4 y of age. *American Journal of Clinical Nutrition*. 2009 Sep;90(3):547-55.

Crespo C, Smit E, Troiano R, Bartlett S, Macera C, Andersen R. 2001 TV watching, energy intake and obesity in US children: results from the III NHANES 1988-1994. *Arch Pediatr Adolesc Med*; 155: 360-365.

de Onis M, Blossner M. Prevalence and trends of overweight among preschool children in developing countries. *American Journal Clinical Nutrition* 2000; 72:1032–1039

Dietz W. 1997 Periods for risk in childhood and adult obesity – What do we need to learn? *Journal of Nutrition*; 127: 1884S – 1886S.

Dietz W. 2002 Forewords. En Burniat W., Cole T., Lissau I. y Poskitt E. (eds.) *Child and Adolescent Obesity. Causes and Consequences, Prevention and Management*. Cambridge University Press p. xv

Doherty, W. y Harkaway, J. 1990. Obesity and family systems: a family FIRO approach to assessment and treatment planning. *Journal of Marital and Family Therapy*, 16, 287–98.

DuBois, P. 1957 *Multivariate correlational analysis*. New York: Harper.

DuBois, P. 1970 Correlational analysis in training research. En P. Dubois y G. Mayo (Eds.) *Research strategies for evaluation training*. Chicago: Rand McNalley.

FAO 2004. *Human Energy Requirements*. Report of a Joint FAO/WHO/UNU Expert Consultation. Rome.

Franks, P., Hanson, R., Knowler, W., Sievers, M., Bennett, P. y Looker, H. 2010 Childhood obesity, other cardiovascular risk factors, and premature death. *The New England Journal of Medicine* 11;362(6):485-93.

Fogelholm, M. y Kukkonen-Harjula, K., 2000 Does physical activities prevent weight gain — a systematic review. *Obesity Revision*, 1, 95–111.

Frelut, M., Razakarivony, R., Cathelinau, L. y Navarro, J. 1995 Uneven occurrence of fatty liver in morbidly obese children: risk factors and impact of weight loss. *International Journal of Obesity*, 19 Suppl. 2, 122.

Frelut, M., Wiling, T., Navarro, J. y Debre, R. 1996 Fatty liver and hyperaminotransferasemia (ALT) in obese children: reversibility and correlation with fattening pattern but not overweight degree. *International Journal of Obesity*, 20 Suppl. 4, 147.

Friedman, J. 1991 Multivariate adaptative regression splines. *The Annals of Statistics* 19:1-141.

Garrow, J. y Webster, J. 1985: Quetelet's index (W/H²) as a measure of fatness. *International Journal of Obesity* 9:147-153.

Gasteyger C, Tremblay A. 2002 Metabolic impact of body fat distribution. *Journal of Endocrinological Investigation*. 25: 876–883.

Gavarry O, Giacomoni M, Bernard T, Seymat M, Falgairette G. 2003 Habitual physical activity in children and adolescents during school and free days. *Medical y Science in Sports and Exercise*; 35:525-531.

- Gennuso, J., Epstein, L., Paluch, R.A. y Cerny, F. 1998 The relationship between asthma and obesity in urban minority children and adolescents. *Archives of Paediatric and Adolescent Medicine*, 152, 1197–200.
- Gutiérrez-Gómez Y, Kain J, Uauy R, Galván M, Corvalán C 2009 Estado nutricional de preescolares asistentes a la Junta Nacional de Jardines Infantiles de Chile: concordancia entre indicadores de obesidad y obesidad central *Arch Latinoam Nutr*. 59(1):30-7
- Gunnel, D., Frankel, S., Nachahal, K., Peters, T. Y Smith D. 1998 Chilwood obesity and cardiovascular mortality: a 57–y follow-up study based on the Boyd Orr cohort. *American Journal of Clinical Nutrition*; 67: 1111-18.
- Hardin, J. y Hilbe, J. 2001 Generalized Linear Models and Extensions. Stata Press. 245pp.
- Hill, J., Wyatt, H., Reed, G. y Peters, J. 2003 Obesity and the environment: where do we go from here? *Science*; 299:853-5.
- Isnard-Mugnier, P., Vila, G., Nollet-Clemencon, C., Vera, L. y Rault, G. 1993 Etude controllee des conduites alimentaires et des manifestations emotionnelles dans une population d'adoloscentes obeses. *Archives Francaises de Pediatrie*, 50, 479–84.
- Kain, J., Lera, L., Rojas, J. y Uauy, R. 2007 Obesidad en preescolares de la Región Metropolitana de Chile. *Revista Médica de Chile*. 135:63-70.
- Kain J, Corvalán C, Lera L, Galván M, Uauy R. 2009 Accelerated growth in early life and obesity in preschool Chilean children. *Obesity* (Silver Spring). 2009 Aug;17(8):1603-8.
- Keijzer-Veen, M., Euser, A., van Montfoort, N., Dekker, F., Vandenbroucke, J. y van Houwelingen, H. 2005 A regression model with unexplained residuals was preferred in the analysis of the fetal origins of adult diseases hypothesis. *J Clin Epidemiol* 58:1320-1324.
- Kimm, S., Sweeney, C. y Janosky, J. 1991 Self-concept measures and childhood obesity: a descriptive analysis. *Journal of Developmental and Behavioral Pediatrics*, 12, 19–24.

Kinugasa, A., Tsunamoto, K., Furukawa, N., Sawada, T., Kusunoki, T. y Shimada, N. 1984 Fatty liver and its Wbrous changes found in simple obesity of children. *Journal of Pediatrics Gastroenterology and Nutrition*, 3, 408–14.

Kirchengast, S. 2010 Gender Differences in Body Composition from Childhood to Old Age: An Evolutionary Point of View. *Journal of Life Science* 2(1):1-10

Koffler, M., Raskin, R., Womble, D. y Helderman, J.H. 1991 Immunobiological consequences of regulation of insulin receptor on alloactivated lymphocytes in normal and obese subjects. *Diabetes*, 40, 364–70.

Kuczmarski R, Ogden C, Guo S, Grummer-Strawn L, Flegal K, Mei Z, et al. 2000 CDC growth charts for the United States: Methods and development. National Center for Health Statistics. *Vital Health Stat* 2002;11: 1-190.

Law, C., Barker, D., Osmond, C, Fall, C., Simmonds, S. 1992 Early growth and abdominal fatness in adult life. *Journal of Epidemiology and Community Health*,; 46:184-186.

Lamarche B. 1998 Abdominal obesity and its metabolic complications: implications for the risk of ischaemic heart disease. *Coronary Artery Diseases*. 9: 473–481.

Leunissen, R., Kerkhof, G., Stijnen, T. y Hokken-Koelega, Anita. 2009 Timing and Tempo of First-Year Rapid Growth in Relation to Cardiovascular and Metabolic Risk Profile in Early Adulthood. *JAMA* 301(21):2234-2242

Lewis, P., Stevens, J. 1991 Nonlinear modeling of time series using multivariate adaptive splines regression splines (MARS). *Journal of the American Statistic Association* 86:416.

Lin, L. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45: 255-268.

Lin, L. 2000. A note on the concordance correlation coefficient. *Biometrics* 56: 324-325.

Lissau, I., Burniat, W., Poskitt, E. y Cole, T. 2002 Prevention. En Burniat W., Cole T., Lissau I. y Poskitt E. (eds.) *Child and Adolescent Obesity. Causes and Consequences, Prevention and Management* Cambridge University Press Cap. 12: 243-269.

- Lobstein T, Baur L, Uauy R. 2004 Obesity in children and young people : a crisis in public health. Report of the IOTF Childhood Obesity Working Groups. *Obesity Revision*; 5:4-104.
- Lord, G.M., Matarrese, G., Howard, J.K., Baker, R.J., Bloom, S.R. y Lechler, R.I. 1998 Leptin modulates the T-cell immunoresponse and reverses starvation-induced immunosuppression. *Nature*, 394, 897–902.
- Luder, E., Melnik, T.A. y DiMaio, M. 1998. Association of being overweight with greater asthma symptoms in inner city black and Hispanic children. *Journal of Pediatrics*, 132, 699–703.
- Malecka-Tendera, E., Wrzesniewski, N., Kurkowska, M. y Kudla, M. 1998 Overweight in adolescent girls with menstrual irregularities is a risk factor for polycystic ovary syndrome. *International Journal of Obesity*, 22 Suppl. 4, S26.
- Mallory, G.B., Fiser, D.H. y Jackson, R. (1989). Sleep-associated breathing disorders in morbidly obese children and adolescents. *Journal of Pediatrics*, 115, 892–7.
- Martínez J. 2000 Body-weight regulation: causes of obesity. *Proceeding of the Nutrition Society*, 59:337-45.
- Maynard L, Wisemandle W, Roche A, Chumlea W, Guo S. y Siervogel, R. 2001 Childhood body composition in relation to body mass index. *Pediatrics* 107, 344–350.
- MINSAL, 2007 Norma para el manejo ambulatorio de la malnutrición por déficit y exceso en el niño(a) menor de seis años. Gobierno de Chile, Ministerio de Salud.
- Moran, J.R., Ghishan, F.K., Halter, S.A. y Green, H.L. 1983 Steatohepatitis in obese children: a cause of chronic liver dysfunction. *American Journal of Gastroenterology*, 78, 374–7.
- Must, A., Jacques, P., Dallal, G., Bajema, C. y Dietz, W. 1992 Long-term morbidity and mortality of overweight adolescents: a follow-up of the Harvard Growth Study of 1922 to 1935. *New England Journal of Medicine*, 327, 1350-5.
- Must, A. y Anderson, S. 2005 Childhood Obesity: Definition, classification and assessment. *Clinical Obesity in Adults and Children* Cap. 15 Blackwell Publishing.

Noguchi, H., Tazawa, Y., Nishinomiya, F. y Takada, G. 1995 The relationship between serum transaminase activities and fatty children with simple obesity. *Acta Paediatrica Japonica*, 37, 621–5.

Ong KK, Ahmed ML, Emmett PM, Preece MA, Dunger DB. 2000 The ALSPAC Study Team. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *British Medicine Journal*; 320: 967-971.

Pietrobelli, A., Faith, M., Allison, D., Gallagher, D., Chiumello, G. y Heymsfield, S. 1998 Body mass index as a measure of adiposity among children and adolescents: a validation study. *J. Pediatr.* 132, 204–210.

Reilly J., Armstrong J, Dorosty A. 2005 Early life risk factors for obesity in childhood: cohort study. *British Medicine Journal*; (advance online publication 29 March 2005).

Richardson, S., Hastor, V., Goodman, N. y Dornbusch, S. 1961 Cultural uniformity in reaction to physical disabilities. *American Sociological Review*, 26, 241–7.

Rojas J. y Corvalán C. 2010 Avance de la Junta Nacional de jardines Infantiles en la incorporación de los estándares de crecimiento OMS 2006. *Rev Chil Nutr.* 37(4): 408-417

Saelens B, Sallis J, Nader P, Broyles S, Berry C, y Taras, H. 2002 Home environment influences on children's TV watching from early to middle childhood. *Journal of Development Behavior Pediatrics*; 23:127-132.

Santos, M., Martínez H., Pérez, B., Albala. B. 2005 Epidemiología genética de la obesidad: estudios familiares. *Revista Médica de Chile*; 133: 349 – 361.

Schousboe K, Willemssen G, Kyvik K, Mortensen J, Boomsma D, Cornes B, 2003 Sex differences in heritability of BMI: a comparative study of results from twin studies in eight countries *Twins Research*; 6:409-21.

Schroeder D, Martorell R, Flores R. 1999 Infant, child growth and fatness and fat distribution in Guatemalan adults. *American Journal of Epidemiology*, 149: 177-185.

Shaffer, E. y Small, D. 1977 Biliary lipid secretion in cholesterol gallstone disease: the effect of cholecystectomy and obesity. *Journal of Clinical Investigation*, 59, 828–40.

Stein C. y Colditz G. 2004 The epidemic of obesity. *Journal of Clinical and Endocrinological Metabolism* 89:2522-2525.

- Tracy, V., De N. y Harper, J. 1971 Obesity and respiratory infection in infants and young children *British Medical Journal*, 1, 16–18.
- Van Itallie, T., Yang, M., Heymsfield, S., Funk, R., Boileau, R. 1990 Height-normalized indices of the body's fat-free and fat mass: potentially useful indicators of nutritional status. *American Journal of Clinical Nutrition*. 52: 953– 959.
- Velásquez, M., Salazar, G., Vio, F., Díaz, N. y Anziani, A. 2008 Validación de Ecuaciones antropométricas para evaluar composición corporal en niños preescolares chilenos. *Revista Médica de Chile* 136: 433-441.
- Wabitsch, M., Hauner, H., Heinze, E., Muche, R., Böckmann, A., Benz, R., Mayer, H. y Teller, W. 1995 Body fat distribution and steroid hormone concentrations in obese adolescent girls before and after weight reduction. *Journal of Clinical Endocrinology and Metabolism*, 80, 3469–75.
- Wang Y. y Lobstein T. 2006 Worldwide trends in childhood overweight and obesity. *International Journal of Pediatric Obesity* 1:11–25
- Wardle, J., Volz, C. y Golding, C. 1995 Social variation in attitudes to obesity in children. *International Journal of Obesity*, 19, 562–9.
- Wells J. 2000 A Hattori chart analysis of body mass index in infants and children. *International Journal of Obesity and Related Metabolic Disorders* 24: 325–329.
- Wells J., Fuller N., Dewit O., Fewtrell M., Elia M. y Cole, T. 1999 Four component model of body composition in children: density and hydration of fat free mass and comparison with simpler models. *American Journal of Clinical Nutrition*; 69: 904 – 912.
- Wells, J., Coward, W., Cole, T. y Davies, P. 2002 The contribution of fat and fat-free tissue to body mass index in contemporary children and the reference child. *International Journal of Obesity*. 26:1323-1328.
- Wells JC. Sexual dimorphism of body composition. *Best Practice of Research of Clinical Endocrinological Metabolism* 2007; 21 (3): 415-30.
- WHO 2006 Multicentre Growth Reference Study Group. WHO Child Growth Standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. Geneva: World Health Organization. Williamson, D.F., Dietary intake and physical activity as “predictors” of

weight gain in observational, prospective studies of adults, *Nutrition Review*, 54, S101–S109, 1996.

Zannolli, R. y Morgese, G. 1996 Distribution of BMI in children: prevalence of wasting and fattening conditions. *Ann. Hum. Biol.* 23, 63–69.

Zhang X, Inoue M, Matsuo K, He J, Gupta PC, Ramadas K, Tsugane S, Irie F, Tamakoshi A, Gao YT, Wang R, Shu XO, Tsuji I, Kuriyama S, Tanaka H, Satoh H, Chen CJ, Yuan JM, Yoo KY, Ahsan H, Pan WH, Gu D, Pednekar MS, Sauvaget C, Sasazuki S, Sairenchi T, Yang G, Xiang YB, Nagai M, Suzuki T, Nishino Y, You SL, Koh WP, Park SK, Chen Y, Shen CY, Thornquist M, Feng Z, Kang D, Boffetta P, Potter JD. 2011 Association between body-mass index and risk of death in more than 1 million Asians. *New England Journal of Medicine* 364(8):719-29.

Zwiauer, K., Caroli, M., Malecka-Tendera, E. y Poskitt, E. 2002 Clinical features, adverse effects and outcome. En Burniat W., Cole T., Lissau I. y Poskitt E. (eds.) *Child and Adolescent Obesity. Causes and Consequences, Prevention and Management* Cambridge University Press. Cap 7. 131-153.

ANEXO 1: CASOS FICTICIOS PARA EJEMPLIFICAR REGRESION GAM UTILIZANDO DBMI

ID	DBMI	SEXO	BF1	BF3	BF5	BF6	BF8	BF10	BF11	BF15	FMI
1	-2.27	0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.58
2	-1.00	0	1.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	3.58
3	0.00	0	1.00	0.92	0.00	0.00	1.05	0.00	0.00	0.00	4.12
4	2.26	0	1.00	3.18	0.00	0.00	3.31	0.00	0.00	0.00	5.43
5	2.27	0	1.00	3.19	0.00	0.01	3.32	0.00	0.00	0.00	5.44
6	.	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.58
7	-2.27	1	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	3.58
8	-1.00	1	1.00	0.00	1.00	0.00	0.05	0.00	0.00	0.05	3.56
9	0.00	1	1.00	0.92	1.00	0.00	1.05	0.00	0.00	1.05	3.66
10	2.26	1	1.00	3.18	1.00	0.00	3.31	0.00	1.19	3.31	6.43
11	2.27	1	1.00	3.19	1.00	0.01	3.32	0.01	1.20	3.32	6.40
12	.	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.58
13	.	.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.58

Tabla 1. Simulación de trece casos para entender la relación de funciones base y predictor de FMI.

ANEXO 2: GRAFICOS DE CONCORDANCIA ENTRE FMI OBSERVADO Y PREDICHO

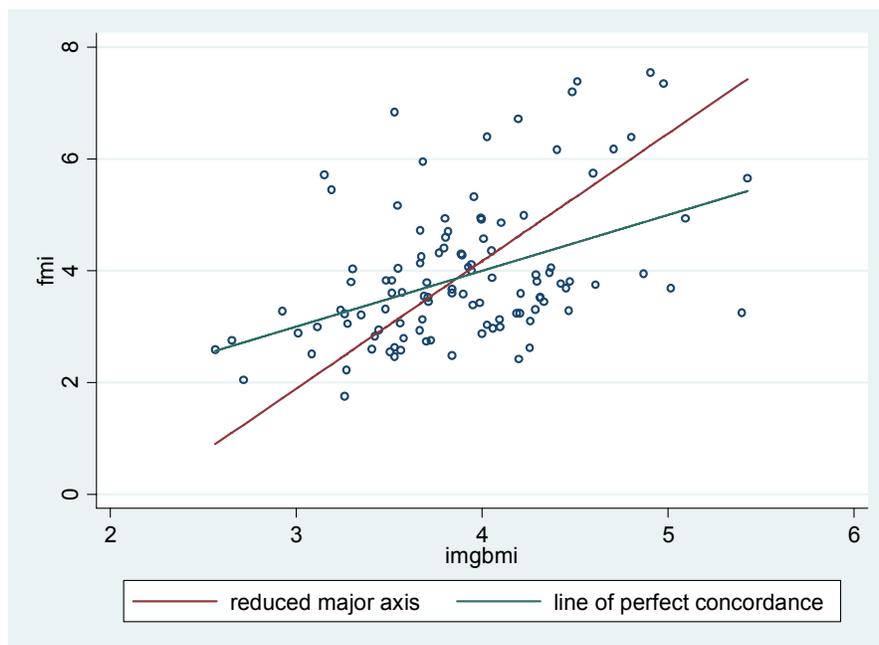


Figura 1. Gráfico de concordancia entre FMI observado y estimado a partir de un modelo GLM de DBMI y SEX.

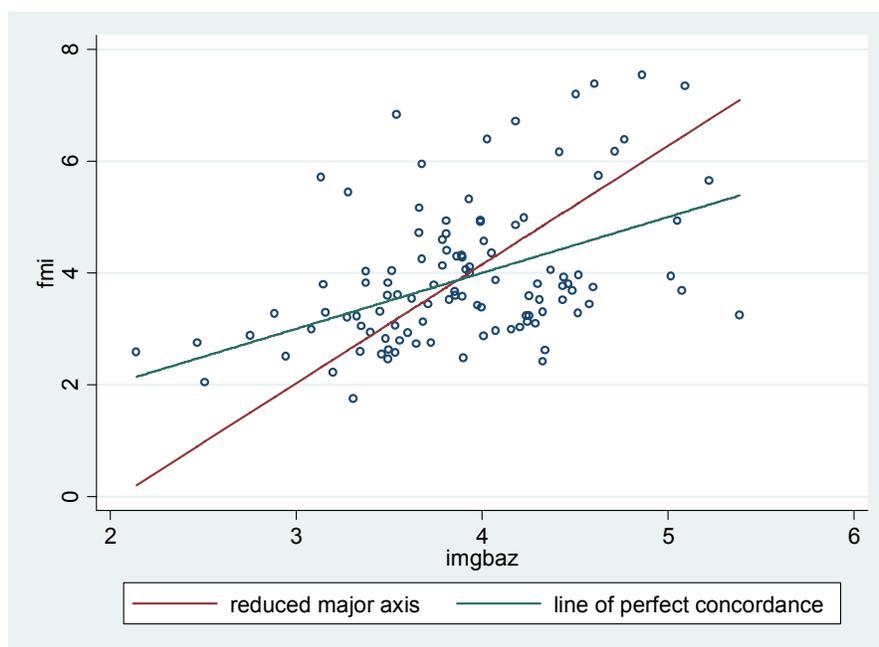


Figura 2. Gráfico de concordancia entre FMI observado y estimado a partir de un modelo GLM de DBAZ y SEX.

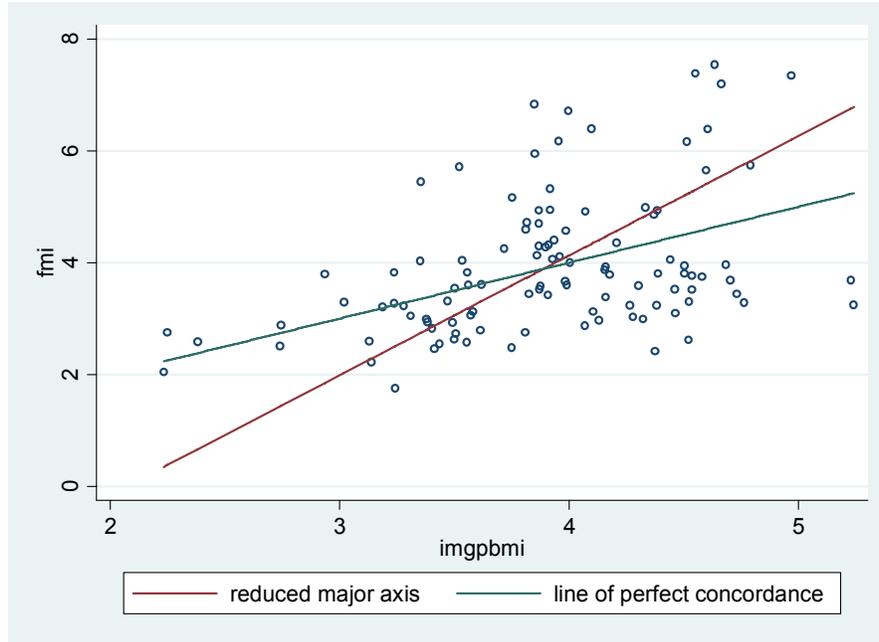


Figura 3. Gráfico de concordancia entre FMI observado y estimado a partir de un modelo GLM de DPBMI y SEX.

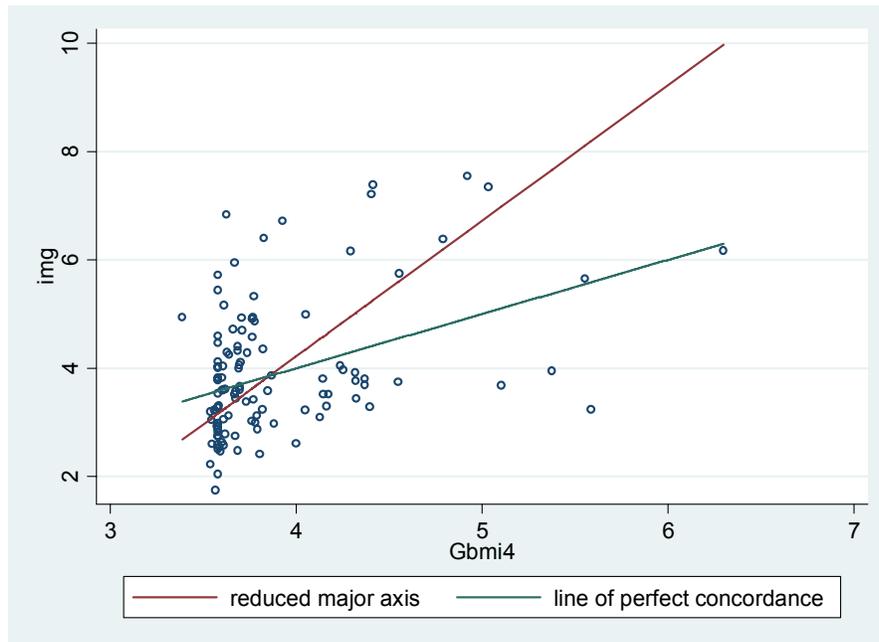


Figura 4. Gráfico de concordancia entre FMI observado y estimado a partir de un modelo GAM de DBMI y SEX.

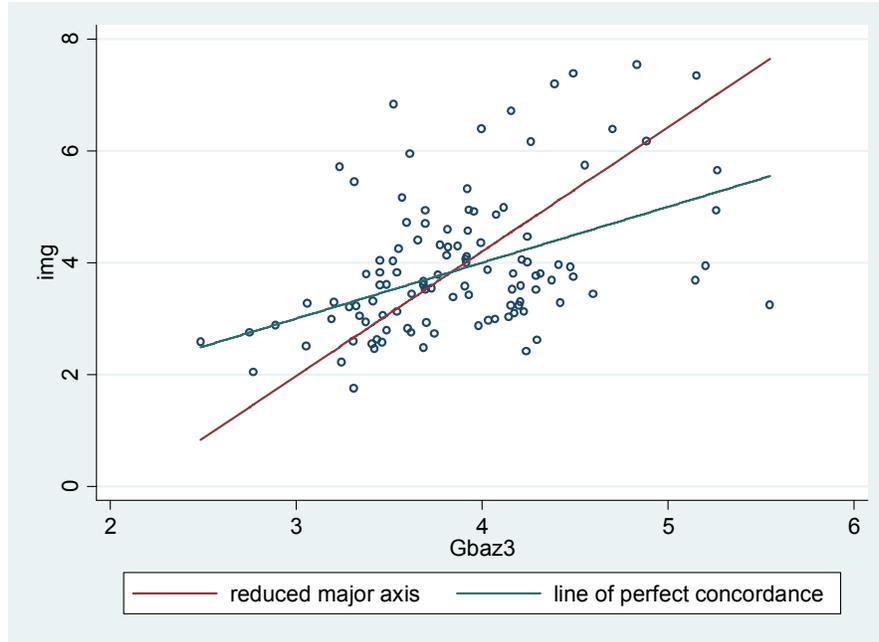


Figura 5. Gráfico de concordancia entre FMI observado y estimado a partir de un modelo GAM de DBAZ y SEX.

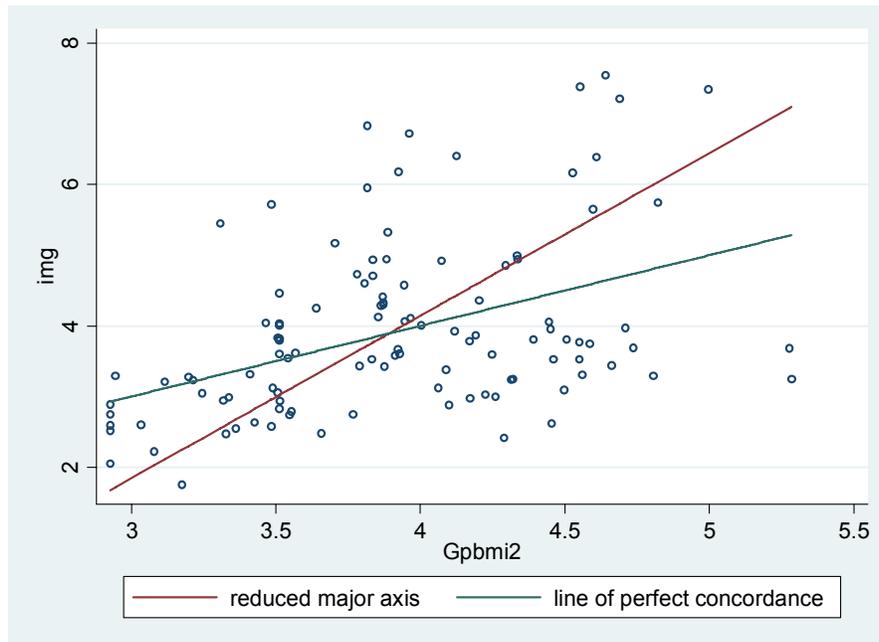


Figura 6. Gráfico de concordancia entre FMI observado y estimado a partir de un modelo GAM de DPBMI y SEX.

ANEXO 3: GRAFICOS DE DISPERSION CON RECTA DE AJUSTE ENTRE FMI OBSERVADO Y PREDICHO

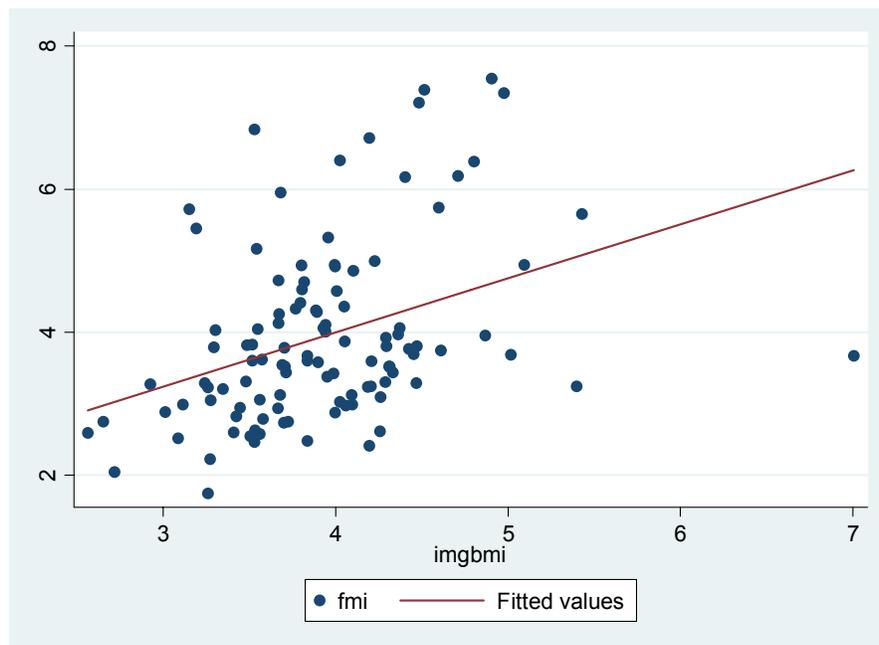


Figura 1. Gráfico de dispersión entre FMI observado y estimado a partir de un modelo GLM de DBMI y SEX.

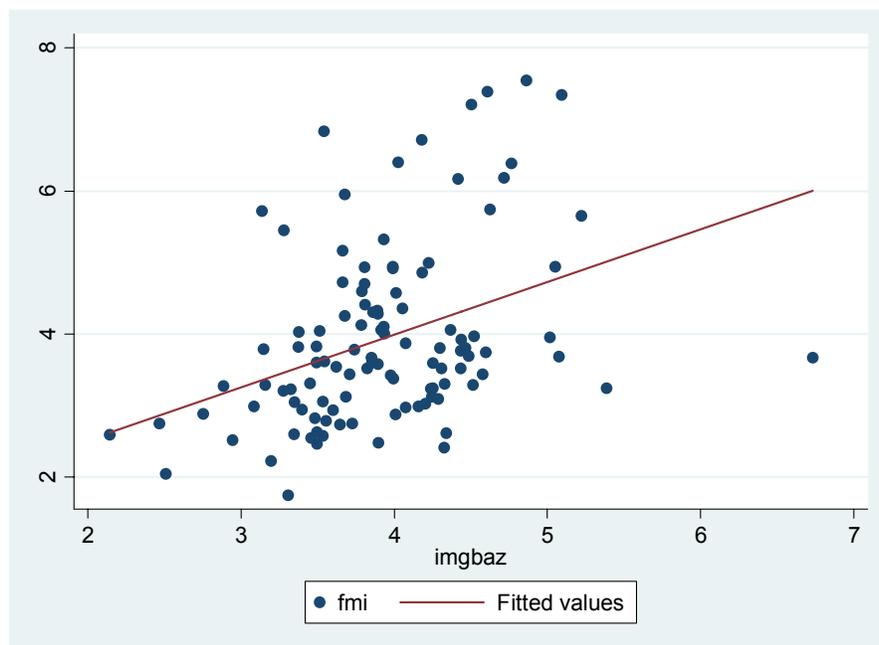


Figura 2. Gráfico de dispersión entre FMI observado y estimado a partir de un modelo GLM de DBAZ y SEX.

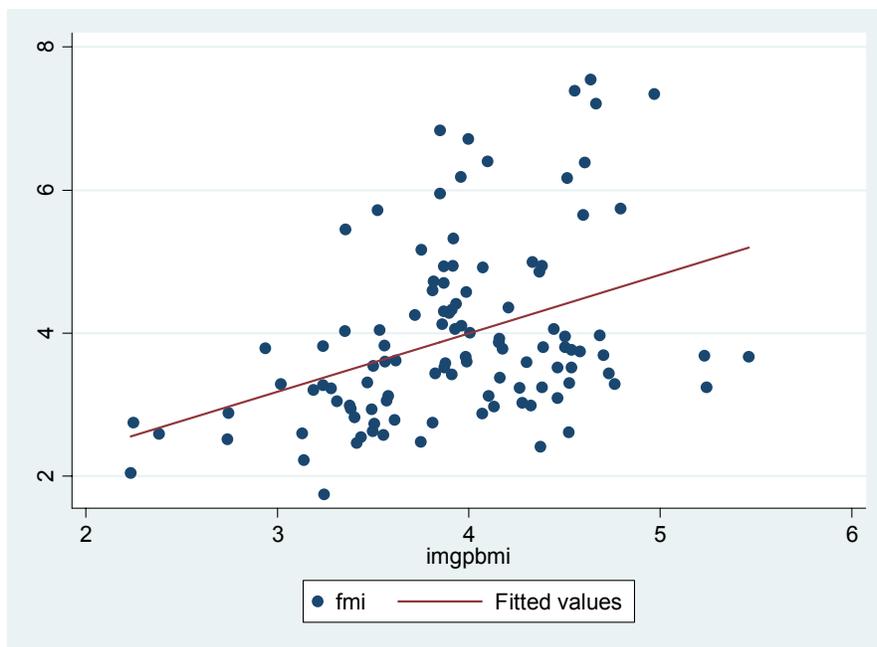


Figura 3. Gráfico de dispersión entre FMI observado y estimado a partir de un modelo GLM de DPBMI y SEX.

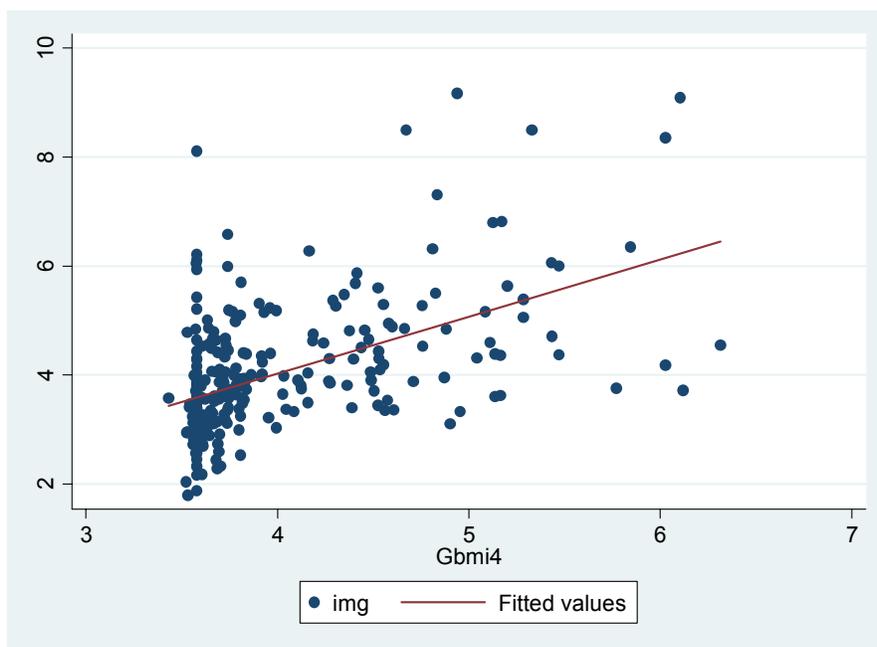


Figura 4. Gráfico de dispersión entre FMI observado y estimado a partir de un modelo GAM de DBMI y SEX.

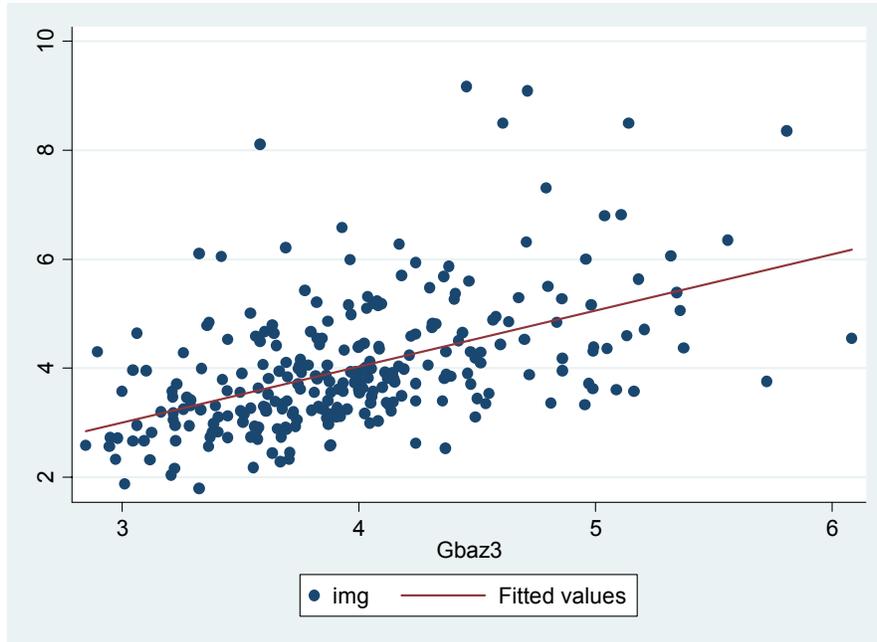


Figura 5. Gráfico de dispersión entre FMI observado y estimado a partir de un modelo GAM de DBAZ y SEX.

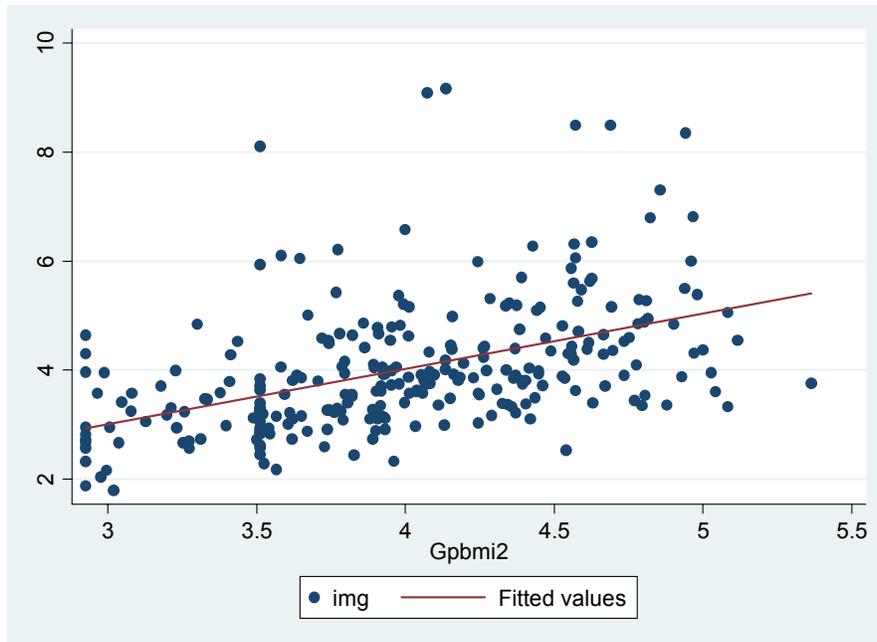


Figura 6. Gráfico de dispersión entre FMI observado y estimado a partir de un modelo GAM de DPBMI y SEX.