



Universidad de Chile
Facultad de Ciencias Sociales
Departamento de Antropología

Estimación de sexo en población chilena moderna a partir del fémur proximal

Memoria para optar al título de Antropólogo Físico



Autor: David Carvalho Sandoval
Profesor Guía: Rodrigo Retamal Yermani
Miércoles 14 de Noviembre del 2018

Agradecimientos

A Felipe Rodríguez, por brindarme invaluable ayuda durante las etapas más tempranas de este trabajo, y cuyos consejos fueron el germen de esta tesis.

A Yanis Valenzuela y Margarita Reyes, por toda el apoyo y compañía entregados durante las últimas instancias de mi formación de pregrado.

A Verónica Silva, por ser la formadora profesional que ha sido para mí y a quien le debo algunas de las oportunidades más importantes de mi corta carrera.

A Rodrigo Retamal, por ser irremplazable guía durante el proceso de realización de esta tesis.

A Nicole Barreux y especialmente a Ma. Cecilia Espinoza, por su asistencia con el material utilizado para esta investigación.

A Karlita Montero, por su comprensión, paciencia y voluntad para atender mis consultas y problemas de tesista.

A Alexandra Elbakyan, por su compromiso con la divulgación del conocimiento científico que resultó esencial para la realización de este trabajo.

A Franco Castellani, por su profesionalismo y enorme paciencia. El material fotográfico aquí exhibido es de su autoría.

A Jorge Carvallo y Ana Díaz, por su interés y acompañamiento durante todo este proceso.

A mi madre Nelly y mi padre Jorge, por darme la educación de que la esta tesis es reflejo.

Y a mis amigos Sebastián Vargas, Martín Ríos, Ignacio López, Pablo “Kiltro” Araya, Antonio Graneris, Nicolás Alarcón, Nicolás Aravena e Ignacio Tobar, por ser quienes son y acompañarme durante mi vida universitaria.

A todos ustedes, mis más sinceros agradecimientos.

Palabras Clave

Antropología Forense, Estimación de Sexo, Dimorfismo Sexual, Fémur Proximal, Regresión Logística, Chile

Resumen

La estimación de sexo es parte fundamental del proceso de identificación de restos humanos esqueletizados. Los métodos más utilizados se basan en la evaluación de la pelvis y el cráneo, pero es común que estos elementos no se preserven lo suficiente para ser evaluados. En estos casos se debe recurrir a otros elementos óseos, entre los cuales destaca el fémur dada su mejor preservación y marcado dimorfismo sexual. El fémur proximal, en particular, es la región mejor preservada y más dimórfica del hueso, por lo que su evaluación es especialmente útil para la estimación de sexo. En el presente estudio se elaboró una nueva metodología de estimación de sexo en población chilena moderna a partir del fémur proximal. Se evaluaron 8 variables métricas en una muestra de 270 individuos (200 individuos de entrenamiento, 70 individuos de validación) provenientes de la Colección Subactual de Santiago (Universidad de Chile) y se utilizó regresión logística para construir modelos predictivos del sexo de los individuos. Los resultados mostraron que los modelos basados en medidas del cuello femoral son los mejores estimadores de sexo, con hasta un 92.9% (modelos univariados) y un 95.7% (modelos multivariados) de acierto y un bajo sesgo por sexo. Otras medidas, particularmente las propuestas por Purkait (2005), reportaron bajos niveles de acierto, por lo que no fueron consideradas estimadores aceptables de sexo. Los hallazgos reportados pueden ser explicados por la naturaleza del dimorfismo sexual de las variables analizadas así como por la aproximación estadística empleada. Los modelos aquí propuestos son los más precisos reportados a la fecha para población chilena moderna y constituyen un aporte a la identificación de restos humanos en causas de violaciones a los Derechos Humanos cometidas durante la Dictadura Cívico-Militar en Chile.

Keywords

Forensic Anthropology, Sex Estimation, Sexual Dimorphism, Proximal Femur, Logistic Regression, Chile

Abstract

Sex estimation is a fundamental part of the process of identification of skeletonized human remains. The most used methods are based on the evaluation of the pelvis and skull, but it is common that these elements are not preserved well enough to be evaluated. In these cases, other bone elements should be used, among which the femur stands out given its better preservation and marked sexual dimorphism. The proximal femur, in particular, is the best preserved and most dimorphic region of the bone, so its evaluation is especially useful for sex estimation. In the present study a new methodology for sex estimation in modern Chilean population from the proximal femur was developed. 8 metric variables were evaluated in a sample of 270 individuals (200 individuals for training, 70 individuals for validation) from the Colección Subactual de Santiago (Universidad de Chile) and logistic regression was used to develop sex predictive models. The results showed that the models based on measurements of the femoral neck are the best sex estimators, with up to 92.9% (univariable models) and 95.7% (multivariable models) of accuracy and a low sex bias. Other measurements, particularly those proposed by Purkait (2005), reported low levels of accuracy so they were not considered acceptable sex estimators. The reported findings can be explained by the nature of the sexual dimorphism of the variables analyzed as well as by the statistical approach used. The models proposed here are the most accurate reported to date for modern Chilean population and represents a contribution to the identification of human remains in cases of human rights violations committed during the Civil-Military Dictatorship in Chile.

Índice de Contenidos

Portada	1
Agradecimientos	2
Palabras Clave y Resumen.....	3
Índice de Contenidos	5
Índice de Gráficos	6
Índice de Figuras	6
Índice de Tablas	7
Índice de Abreviaciones	7
Introducción	8
Antecedentes.....	9
Problematización	21
Pregunta, Hipótesis y Objetivos de Investigación.....	22
Materiales	23
Métodos.....	26
Resultados.....	32
Discusión	43
Conclusiones	51
Bibliografía.....	52
Anexos	59

Índice de Gráficos

Gráfico 1: Distribución de FNAL por sexo	34
Gráfico 2: Distribución de FNW por sexo	34
Gráfico 3: Distribución de AB por sexo	35
Gráfico 4: Distribución de AC por sexo	35
Gráfico 5: Distribución de BC por sexo	36
Gráfico 6: Distribución de GTFC por sexo.....	36
Gráfico 7: Distribución de GTLT por sexo	37
Gráfico 8: Distribución de LTFC por sexo	37

Índice de Figuras

Figura 1: Esquema de las variables FNAL y FNW en el fémur proximal	59
Figura 2: Esquema de las variables AB, AC y BC en el fémur proximal	60
Figura 3: Esquema de las variables GTFC, GTLT y LTFC en el fémur proximal	61
Figura 4: Espécimen de fémur característicamente femenino. Vista anterior	62
Figura 5: Espécimen de fémur característicamente femenino. Vista posterior	63
Figura 6: Espécimen de fémur característicamente masculino. Vista anterior.....	64
Figura 7: Espécimen de fémur característicamente masculino. Vista posterior	65

Índice de Tablas

Tabla 1: Distribución etaria de la muestra de entrenamiento por sexo (n=200)	23
Tabla 2: Variables analizadas	26
Tabla 3: ρ_c y p-valor para el coeficiente de correlación de concordancia de Lin (n=20).....	32
Tabla 4: Estadística descriptiva para cada lado y p-valor para pruebas t de Student pareadas (n=20).....	32
Tabla 5: Estadística descriptiva de la muestra de entrenamiento y p-valor para pruebas de Shapiro-Wilk, t de Student y Mann-Whitney-Wilcoxon (n=200)	33
Tabla 6: Regresiones logísticas (n=200)	38
Tabla 7: Validación de los modelos predictivos de sexo (n=70)	40
Tabla 8: Modelos propuestos para la estimación de sexo en población chilena moderna a partir del fémur proximal	42

Índice de Abreviaciones

FNAL	Distancia entre la base del gran trocánter y el ápex de la cabeza femoral
FNW	Diámetro mínimo del cuello femoral medido en sentido supero-inferior
AB	Distancia entre el punto más lateral del margen articular de la cabeza femoral (A) y el punto más medial del trocánter mayor (B)
AC	Distancia entre el punto más lateral del margen articular de la cabeza femoral (A) y el punto más pósteromedial del trocánter menor (C)
BC	Distancia entre el punto más medial del trocánter mayor (B) y el punto más pósteromedial del trocánter menor (C)
GTFC	Distancia entre el punto más lateral del gran trocánter (GT) y el punto más superior de la fovea capitis (FC)
GTLT	Distancia entre el punto más lateral del gran trocánter (GT) y el punto más superior de la carilla del trocánter menor (LT)
LTFC	Distancia entre el punto más superior de la carilla del trocánter menor (LT) y el punto más superior de la fovea capitis (FC)

Introducción

La presente Memoria de Título tuvo por objeto desarrollar una metodología de estimación de sexo a partir de mediciones osteométricas del fémur proximal aplicable a población chilena moderna, con el propósito de contribuir a la estimación del perfil biológico y la identificación de restos humanos en contextos forenses subactuales, particularmente en casos de Detenidos Desaparecidos y crímenes de lesa humanidad cometidos durante la Dictadura Cívico-Militar en Chile (1973-1989).

Esta Memoria se compone de las siguientes secciones:

- Antecedentes, donde se abordan los conceptos de Antropología Forense, identificación, sexo, dimorfismo sexual y estimación de sexo, y se exponen los estudios de estimación de sexo en fémur proximal y en población chilena
- Problematización, en que se plantea el vacío de conocimiento existente en la estimación de sexo a partir del fémur proximal en población chilena moderna y por qué es relevante subsanarlo
- Pregunta, Hipótesis y Objetivos de Investigación, donde se establecen las directrices del presente estudio
- Materiales, en que se especifican las características de la muestra y los criterios de inclusión/exclusión
- Métodos, donde se detallan las variables analizadas y los análisis estadísticos empleados
- Resultados, en que se resumen los resultados obtenidos para los análisis conducidos y se proponen modelos predictivos de sexo
- Discusión, donde se desglosan y analizan los resultados obtenidos a la luz de los antecedentes, se exponen las limitaciones del estudio y se sugieren líneas de investigación para futuros trabajos
- Conclusiones, en que se recapitulan los aspectos más importantes de la Memoria.

Antecedentes

Se denomina Ciencia Forense (en inglés *Forensic Science*) a la aplicación de la ciencia a cuestiones legales, sobre todo criminales (White, 2004; Tilstone, Savage y Clark, 2006; Pyrek, 2007; Fraser, 2010; Siegel y Mirakovits, 2010; Jackson y Jackson, 2017). Dado el vasto espectro de conocimiento científico aplicable al servicio de la ley, la Ciencia Forense constituye un campo multidisciplinario donde confluyen diversas áreas tales como la biología, patología, ingeniería, toxicología, medicina, entomología y genética, entre muchas otras más (Lerner y Lerner, 2006; Burns, 2013; Ubelaker, 2013; Houck, 2015; Katz y Halámek, 2016).

En el contexto legal en que se desenvuelve, las evidencias aportadas por la Ciencia Forense pueden ser cruciales para dar con la identidad de víctimas y victimarios, así como para declarar culpable a una persona o absolverla. Ante esta responsabilidad, los métodos empleados deben cumplir con altos estándares de rigurosidad, de acuerdo con el denominado estándar Daubert. Desde 1923, las cortes de los Estados Unidos decidían si el testimonio de un experto durante un juicio era válido o no basadas en el estándar Frye, llamado así por el caso *Frye v. United States* (Grivas y Komar, 2008). En él se establecía que el testimonio de un experto podía ser considerado válido en tanto presentara evidencia que tuviera aceptación generalizada por parte de la comunidad científica pertinente. Sin embargo, el año 1993 la Corte Suprema de los Estados Unidos determinó, en el contexto del caso *Daubert v. Merrell Dow Pharmaceuticals*, que la validez del testimonio de un experto no residía exclusivamente en la aprobación de la evidencia presentada sino en la pertinencia científica de las metodologías empleadas. Este nuevo criterio, denominado Daubert por el caso correspondiente, cambió de forma efectiva el foco de atención desde la experiencia y reputación de un experto hacia la calidad de sus métodos, mientras que se determinaron cinco puntos clave para que los jueces puedan juzgar la validez de un testimonio: (1) si los planteamientos del experto pueden o han sido testeados de forma objetiva, (2) si los métodos o teorías utilizados ha sido sujetos a revisión por parte de los pares científicos, (3) si se conoce el error potencial de los resultados, (4) si existen estándares que regulen la aplicación de las metodologías empleadas, y por último (5) si las teorías y métodos presentados poseen aceptación generalizada por parte de la comunidad científica. Posteriormente, los casos de *General Electric Co. v. Joiner* y *Kumho Tire Co. v. Carmichael* derivaron en la ampliación y flexibilización del estándar Daubert para la admisibilidad del testimonio de expertos en corte, siendo en conjunto denominados “trilogía Daubert” (Tilstone et al., 2006; Grivas y Komar, 2008; James y Nordby, 2009; Jamieson y Moenssens, 2009; Daeid, 2010; Siegel y Mirakovits, 2010; Shelton, 2011; Dirkmaat, 2012; Forrest y Kennedy en Ubelaker, 2013; Lesciotto, 2015). De las consideraciones establecidas por el estándar Daubert se desprende una obligación por parte de los peritos de emplear metodologías robustas, cuya confiabilidad pueda ser demostrada de forma objetiva ante la corte.

Dentro de la Ciencia Forense se inserta la Antropología Forense, que corresponde al subcampo de la Antropología Física en que las técnicas y conocimientos de la disciplina son aplicados a problemáticas de significancia médico-legal. Aunque los antropólogos forenses son capaces de trabajar con distintas materialidades, su importancia dentro de las ciencias forenses reside particularmente en el análisis de los restos óseos dada su experticia en osteología humana (Klepinger, 2006; Ubelaker en Schmitt, Cunha y Pinheiro, 2006; Adams, 2007; Houck, 2007; Embar-Seddon y Pass, 2009; Fraser y Williams, 2009; James y Nordby, 2009; Siegel y Mirakovits, 2010; Burns, 2013; Iscan y Stein, 2013; Steadman en Ubelaker, 2013; Tersigni-Tarrant y Shirley, 2013; Christensen, Passalacqua y Bartelink, 2014; Houck y Siegel, 2015; Okuda y Stephenson, 2015). Si bien la Antropología Forense puede servir a diversos propósitos, como determinar la pertinencia médico-legal de los restos o aportar evidencia sobre posibles crímenes, su objetivo principal, en el que culminan todas las líneas de evidencia, es asistir en la identificación de los restos óseos (Klepinger, 2006; Ubelaker en Schmitt et al., 2006; Adams, 2007; Houck, 2007; Embar-Seddon y Pass, 2009; Fraser y Williams, 2009; James y Nordby, 2009; Pickering y Bachman, 2009; Burns, 2013; Iscan y Stein, 2013; Steadman en Ubelaker, 2013; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014; Houck y Siegel, 2015).

Utilizado de forma distinta en otras ciencias forenses, el concepto de identificación en Antropología Forense se refiere al proceso mediante el cual se logran asociar de forma efectiva los restos analizados a una persona de identidad conocida (Lerner y Lerner, 2006; Tilstone et al., 2006; Adams, 2007; Jamieson y Moenssens, 2009; Fraser, 2010; Christensen et al., 2014). En este sentido, se habla de identificación positiva (o individualización) cuando la evidencia apunta a una sola persona y a la exclusión de cualquier otra. El análisis de ADN es la aproximación más confiable para identificar un individuo dada la relación directa que existe entre el genoma y la individualidad de una persona, por lo que es considerado actualmente como el “*gold standard*” para la identificación forense (Fraser, 2010; Shelton, 2011; Jackson y Jackson, 2017). No obstante, su aplicación posee varias limitaciones. La factibilidad de los análisis moleculares depende notablemente de condiciones de preservación y recuperación adecuadas, sobre todo frente a muestras degradadas; la contaminación es una complicación frecuente, particularmente en casos de restos mezclados, y la complejidad del manejo de la evidencia vuelve al proceso más propenso a errores producidos por una mala manipulación. Además, la obtención de las muestras necesarias para los análisis implica siempre algún grado de destrucción del material. Por otra parte, el análisis de ADN no puede conducirse *in situ*, requiriendo de un laboratorio especializado para su ejecución, y aún hoy las técnicas moleculares poseen un alto costo económico y demandan mucho tiempo, lo que supone una importante desventaja en casos forenses. Finalmente, el perfil genético obtenido como resultado del análisis sólo puede entregar una identificación positiva al ser contrastado con otras muestras de referencia (Pyrek, 2007; James y Nordby, 2009; Siegel y Mirakovits, 2010; Dirkmaat, 2012; Iscan y Stein, 2013; Tersigni-Tarrant y Shirley, 2013; Anastopoulou, Eliopoulos, Valakos y Manolis, 2014; Katz y Halámek, 2016; Barbaro, 2018).

Por otro lado, los métodos osteológicos propios de la Antropología Forense constituyen un acercamiento considerablemente mucho rápido, simple y barato a la identificación, generalmente conducible en el mismo lugar de recuperación de los restos y ajeno a los problemas de contaminación propios de los métodos moleculares. Sin embargo, los métodos antropológicos pueden entregar información valiosa sobre la identidad de los restos pero no una identificación positiva. Es por esto que actualmente ambas aproximaciones suelen conducirse en conjunto y de forma complementaria: la evaluación osteológica delimita significativamente la posible identidad de los restos, reduciendo el espectro de candidatos y por tanto la muestra de referencia necesaria para los análisis de ADN, contribuyendo así a un proceso notablemente más eficiente en tiempo y costo económico. Esta asistencia de la Antropología Forense a los métodos moleculares representa, sobre todo en casos complejos con presencia de restos mezclados como desastres naturales o fosas comunes, un complemento fundamental a la Genética Forense y constituye uno de sus mayores aportes dentro de la Ciencia Forense en la actualidad (Ubelaker Schmitt et al., 2006; Houck, 2007; Fraser y Williams, 2009; Dirkmaat, 2012; Steadman en Ubelaker, 2013; Christensen et al., 2014; Jackson y Jackson, 2017).

Para asistir en la identificación, el antropólogo forense se enfoca en la estimación de los parámetros biológicos básicos de los restos, lo que se conoce como perfil biológico. Los principales componentes del perfil biológico de un individuo son el sexo, la edad, la estatura y la ancestría, conjunto comúnmente llamado cuarteta básica o *big four* en inglés (Spradley, Jantz, Robinson y Peccerelli, 2008; Embar-Seddon y Pass, 2009; James y Nordby, 2009; Pickering y Bachman, 2009; Dirkmaat, 2012; Steadman en Ubelaker, 2013; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014; Houck y Siegel, 2015).

Dentro del perfil biológico, la estimación del sexo del individuo es fundamental: por una parte, al tratarse de un rasgo dicotómico – femenino o masculino -, se acota en aproximadamente un 50% la posible identidad del individuo y, por otro lado, la mayoría de los métodos de estimación de los demás componentes del perfil biológico exige conocer previamente el sexo del individuo, razón por la cual es siempre el primer componente en ser estimado (Dabbs y Moore-Jansen, 2010; Ross y Manneschi, 2010; Spradley y Jantz, 2011; Dawson, Ross y Mallett en Berthelot, 2014; Christensen et al., 2014; Garrido-Varas, Thompson y Campbell, 2014; Longe, 2015; Curate et al., 2016; Peckmann, Logar y Meek, 2016; Torimitsu et al., 2016). Considerando la importancia de la estimación de sexo en la Antropología Forense, el desarrollo de métodos certeros y confiables acordes al estándar Daubert resulta imperativo para la disciplina.

El sexo biológico de un individuo está determinado, en primera instancia, por una base genética. Los cromosomas sexuales determinan el sexo de un individuo como femenino si el par está constituido por dos cromosomas X, y como masculino si incluye un cromosoma X y uno Y. Esta divergencia genética se expresa en diferencias de diversa naturaleza entre ambos sexos, incluyendo los caracteres sexuales primarios, relativos a la anatomía genital y las gónadas, y los caracteres sexuales secundarios, que involucran diferencias de forma y tamaño. Estas diferencias morfológicas, en conjunto con otras fisiológicas y etológicas, constituyen el denominado dimorfismo sexual (Dirkmaat, 2012; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014).

Morfológicamente, los individuos masculinos poseen en promedio cuerpos más altos, grandes y con mayor masa muscular que los individuos femeninos, quienes tienden a ser más pequeños y tener menos musculatura. Esto se ve reflejado en el esqueleto, con huesos masculinos generalmente más largos, robustos y con inserciones musculares más marcadas, frente a los huesos femeninos, más pequeños y gráciles (Houck, 2007; du Jardin, Ponsaillé, Alunni-Perret, Quatrehomme, 2009; Black y Ferguson, 2011; Abarca, 2013; Tersigni-Tarrant y Shirley, 2013; Anastopoulou et al., 2014; Christensen et al., 2014; Dawson et al. en Berthelot, 2014; Houck y Siegel, 2015; Longe, 2015). Estas diferencias representan el dimorfismo sexual del esqueleto humano, y los métodos osteológicos de estimación de sexo se basan en su evaluación para establecer si un individuo es de sexo femenino o masculino.

No obstante, la estimación de sexo a partir de restos óseos posee algunas complejidades. Las diferencias morfológicas entre individuos masculinos y femeninos son causadas principalmente por la acción directa o indirecta de hormonas sexuales, por lo que su máxima expresión ocurre de forma posterior a la adolescencia. Como consecuencia, los individuos subadultos poseen un dimorfismo sexual mucho menos evidente. Es por esta razón que los métodos de estimación de sexo son más confiables en individuos adultos que en subadultos, a los que resulta mucho más difícil sexar y para los que se requieren métodos específicos (Klepinger, 2006; Adams, 2007; Houck, 2007; Pickering y Bachman, 2009; Dawson et al. en Black y Ferguson, 2011; Iscan y Stein, 2013; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014; Houck y Siegel, 2015; Okuda y Stephenson, 2015; Curate, Albuquerque, Ferreira y Cunha, 2017a).

Por otra parte, producto de la variabilidad humana, la expresión de un rasgo sexualmente dimórfico dentro de un sexo biológico generalmente se distribuye de forma normal en una población, desde los grados característicamente femeninos hasta los característicamente masculinos. Por lo tanto, cuando ambos sexos son incluidos en una secuencia pueden distinguirse dos curvas normales pero con un área de superposición entre las distribuciones de ambos sexos (Adams, 2007; Houck, 2007; Dawson et al. en Black y Ferguson, 2011; Dirkmaat, 2012; Iscan y Stein, 2013; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014; Houck y Siegel, 2015). Los individuos que se encuentran dentro de esta área son difíciles de sexar y suelen ser considerados como “indeterminados”. Con el fin de aumentar la confiabilidad y precisión de la estimación de sexo, particularmente para los individuos en esta zona de traslape, se recomienda evaluar tantas características como la preservación de los restos lo permita: a más rasgos evaluados, menor es la zona de superposición entre individuos masculinos y femeninos, lo que aumenta la probabilidad de acierto al sexo real del individuo (Albanese, 2003; Dirkmaat, 2012; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014).

Por otro lado, también existe una significativa variabilidad poblacional en la expresión del dimorfismo sexual. La interacción entre la carga genética específica y las condiciones ambientales particulares a las que se encuentra sometida cada población determinan su dimorfismo sexual. Dado que esta combinación es única para cada grupo humano, el dimorfismo sexual se expresa de forma específica para cada población (Frutos, 2003; Klepinger, 2006; Abarca 2013; Christensen et al., 2014; Clavero, Salicrú y Turbón, 2015;

Longe, 2015; Curate et al., 2016; Guyomarc'h et al., 2016). En consecuencia, distintas poblaciones pueden ubicarse en diferentes puntos del espectro del dimorfismo sexual humano, mientras que dentro de una población pueden presentarse diversos grados de dimorfismo sexual, convergiendo o divergiendo los individuos de ambos sexos. Esta variabilidad ha sido claramente constatada en la literatura (Klepinger, 2006; Ubelaker en Schmitt et al., 2006; Pickering y Bachman, 2009; Dirkmaat, 2012; Iscan y Stein, 2013). Adicionalmente, se ha observado una variación en el patrón intrapoblacional del dimorfismo sexual a lo largo del tiempo, explicado por los cambios genéticos y ambientales vividos por muchas poblaciones durante los últimos años, fenómeno conocido como cambio secular (Klepinger, 2006; Ubelaker en Schmitt et al., 2006; Abarca, 2013; Iscan y Stein, 2013; Tersigni-Tarrant y Shirley, 2013; Guyomarc'h et al., 2016). Dada esta variabilidad del dimorfismo esquelético, la aplicación indiscriminada de estándares osteológicos en poblaciones distintas y no-contemporáneas a aquellas en las que fueron elaborados resulta en la pérdida de confiabilidad y precisión. Para poder estimar sexo de forma certera es necesaria, entonces, la elaboración, ajuste y aplicación de métodos específicos para cada población (Klepinger, 2006; Ubelaker en Schmitt et al., 2006; Pickering y Bachman, 2009; Dirkmaat, 2012; Abarca, 2013; Tersigni-Tarrant y Shirley, 2013; Longe, 2015).

Con el fin de instrumentalizar el dimorfismo sexual del esqueleto humano para la estimación de sexo se han utilizado distintas aproximaciones. En este sentido, los métodos osteológicos pueden agruparse en dos grandes familias, los morfológicos y los métricos. Los métodos morfológicos (a veces llamados morfoscópicos o no-métricos) representan una evaluación visual y cualitativa de rasgos esqueléticos, buscando reconocer formas generales en el hueso, mientras que los métodos métricos involucran la medición cuantitativa de regiones específicas a partir de morfometría lineal, morfometría geométrica y datos angulares, entre otros (Dirkmaat, 2012; Iscan y Stein, 2013; Steadman en Ubelaker, 2013; Christensen et al., 2014; Houck y Siegel, 2015). Los métodos morfológicos son la aproximación más antigua, de uso más extendido y aplicación más rápida y sencilla, y comúnmente se considera que pueden aplicarse en poblaciones diferentes sin perder confiabilidad. Además, la mayoría de los métodos "clásicos" corresponden a enfoques morfológicos. No obstante, dependen fuertemente de la habilidad y experiencia del evaluador para interpretar los rasgos, teniendo una notable curva de aprendizaje y niveles altos de error inter-observador (Iscan y Stein, 2013; Christensen et al., 2014). Más aún, varios estudios han mostrado que los métodos morfológicos, contrario a la creencia general, no son aplicables de forma indiscriminada a poblaciones diversas, observándose niveles dispares de precisión (Albanese, Eklics y Tuck, 2008). Los métodos métricos, en cambio, son más objetivos, fáciles de evaluar e interpretar y menos dependientes de la experticia del investigador, reportando menores niveles de error intra e interobservador y una mayor reproductibilidad. Además, cuentan con un mayor potencial para análisis estadísticos, particularmente cuando se emplean muestras grandes, y son capaces de dar cuenta de diferencias sexuales leves pero sistemáticas que los métodos morfológicos pasan por alto (Albanese, 2003; Klepinger, 2006; Dabbs y Moore-Jansen, 2010; Black y Ferguson, 2011; Dirkmaat, 2012; Iscan y Stein, 2013; Christensen et al., 2014).

Diversos elementos del esqueleto han sido evaluados para estimar sexo. No obstante, las regiones más frecuentemente empleadas son la cintura pélvica y el cráneo, casi sin excepción mediante métodos morfológicos. Un protocolo muy extendido es la utilización de ambos elementos para sexar los restos.

La pelvis (i.e. los coxales y el sacro) se ha utilizado extensamente dado que es reconocida como el elemento más dimórfico del esqueleto humano. Esto se debe a que la anatomía de la cintura pélvica representa una adaptación no sólo a la bipedestación sino también a la concepción de bebés con grandes cerebros, característicamente humanos, en los individuos femeninos. De esta forma, la morfología de la pelvis evidencia no sólo el dimorfismo sexual de tamaño y robusticidad presente en todo el esqueleto, sino que además las fuerzas evolutivas combinadas de la bipedestación y la reproducción, dotándola de una forma distintiva en los individuos femeninos muy útil para diferenciar un sexo del otro (Klepinger, 2006; Ubelaker en Schmitt et al., 2006; Adams, 2007; Houck, 2007; Fraser y Williams, 2009; Pickering y Bachman, 2009; Dawson et al. en Black y Ferguson, 2011; Dirkmaat, 2012; Iscan y Stein, 2013; Steadman en Ubelaker, 2013; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014; Houck y Siegel, 2015). Así, la pelvis femenina tiene una forma corta y ancha, de disposición predominantemente horizontal, con superficies ilíacas extensas y planas, un ángulo subpúbico abierto, un diámetro interno amplio y redondo y un sacro de aspecto recto. En tanto, la pelvis masculina es más grande, alta y angosta, con una disposición general vertical, un espacio interno más estrecho, un sacro más curvado hacia anterior e inserciones musculares más prominentes (Klepinger, 2006; Ubelaker en Schmitt et al., 2006; Adams, 2007; Houck, 2007; Fraser y Williams, 2009; Pickering y Bachman, 2009; Steadman en Ubelaker, 2013; Tersigni-Tarrant y Shirley, 2013; Christensen et al., 2014; Houck y Siegel, 2015). Dentro de los métodos más extendidos y confiables basados en la evaluación de la pelvis se encuentran la metodología de Phenice (1969) que evalúa tres rasgos en los coxales - arco ventral, concavidad subpúbica y rama isquiopúbica -, y el método propuesto por Buikstra y Ubelaker (1994), que suma a los rasgos de Phenice la evaluación de dos características más – la escotadura ciática y el surco preauricular.

El cráneo, por otro lado, ha sido considerado tradicionalmente como el segundo mejor estimador de sexo. El dimorfismo en esta región es causado principalmente por la diferencia de tamaño y de impronta biomecánica en el cráneo y la mandíbula. De esta manera, el cráneo femenino suele ser más pequeño y tener un aspecto grácil y redondeado, mientras que el cráneo masculino es más grande, robusto y posee un aspecto más angulado, con inserciones musculares más pronunciadas (Adams, 2007; Houck, 2007; Pickering y Bachman, 2009; Dawson et al. en Black y Ferguson, 2011; Iscan y Stein, 2013; Steadman en Ubelaker, 2013; Christensen et al., 2014; Houck y Siegel, 2015). Destacan entre los métodos más utilizados de estimación de sexo en cráneo la evaluación de los rasgos compilados por Buikstra y Ubelaker (1994) a partir de cinco atributos craneales y mandibulares, y la metodología propuesta por Walrath, Turner y Bruzek (2004) basada en diez rasgos.

La evaluación de pelvis y cráneo para la estimación de sexo tiene, sin embargo, ciertas limitaciones. La pelvis es altamente susceptible al daño por factores tafonómicos, sobre todo la frágil región púbica, precisamente la que concentra mayor dimorfismo sexual. Por esta razón, es muy común tanto en contextos arqueológicos como forenses la incompletitud, fragmentación o ausencia de la pelvis que impide su evaluación. (Albanese et al., 2008; Soni, Dhall y Chhabra, 2010; Clavero et al., 2015; Longe, 2015; Curate et al., 2017b). Además, existe evidencia de que los métodos morfológicos basados en la cintura pélvica también son afectados por la variabilidad poblacional, observándose rendimientos inconsistentes, muy inferiores al aplicarlos en poblaciones distintas a las de referencia (Albanese 2003). Por otra parte, la confiabilidad de la estimación del sexo en cráneo ha sido puesta en duda ya que se han observado altos niveles de error inter-observador (Christensen et al., 2014). Así mismo, el cráneo también es frecuentemente afectado por procesos tafonómicos y traumas perimortem, siendo común la fragmentación o ausencia que impide la aplicación de métodos basados en este elemento (Albanese et al., 2008).

Es en este escenario que la estimación de sexo en otros elementos del esqueleto cobra importancia por dos razones cruciales. Primero, puesto que es común tanto en contextos arqueológicos como forenses recuperar restos fragmentarios o incompletos que impiden evaluar elementos tradicionalmente utilizados para estimar sexo, es necesario contar con métodos basados en otros elementos óseos sexualmente dimórficos que sí se encuentren disponibles (Soni et al., 2010; Spradley, Anderson y Tise 2014; Longe 2015; Curate et al. 2017b). Y segundo, porque pese a la alta valoración que regularmente tiene el cráneo en la estimación de sexo, varios componentes del post-cráneo, específicamente los huesos largos, poseen un mayor dimorfismo sexual y reportan mejores niveles de acierto como estimadores de sexo que este elemento (Dittrick y Suchey, 1986; Klepinger, 2006; Spradley y Jantz, 2011; Kim, Kwak y Han, 2013; Spradley et al., 2014).

Los métodos de estimación de sexo en elementos post-craneales no-pélvicos se basan en la evaluación del dimorfismo sexual de tamaño corporal y robusticidad entre ambos sexos donde, como se expuso anteriormente, los huesos masculinos tienden a ser más grandes y robustos mientras que los femeninos suelen ser más pequeños y gráciles. (Adams, 2007; Dawson et al. en Black y Ferguson, 2011; Berthelot, 2014; Christensen et al., 2014; Longe, 2015). Los métodos desarrollados en estos elementos se fundamentan casi sin excepción en aproximaciones métricas, mayoritariamente morfometría lineal, siendo fáciles de evaluar e interpretar así como menos subjetivos y dependientes de la familiaridad del observador con las variables evaluadas en comparación a los métodos morfológicos; además, los resultados obtenidos de su aplicación, dada su naturaleza numérica, permiten una mayor contrastación y estandarización de las metodologías (Houck, 2007; Dabbs y Moore-Jansen, 2010; Albanese, 2013; Houck y Siegel, 2015; Curate et al., 2016). Prácticamente todos los huesos del post-cráneo no-pélvico han sido estudiados para la estimación de sexo, incluyendo huesos largos, vértebras, escápula, esternón, costillas, patella, metacarpos, metatarsos y falanges de manos y pies. Una exposición detallada de la literatura atingente excede los objetivos del presente escrito, pero puede encontrarse en las publicaciones más importantes del tema dada su vigencia e importancia. Destacan dentro del esqueleto post-craneal los huesos largos, no solo por sus mejores niveles de acierto sino además porque

tienden a preservarse mejor que otros elementos óseos gracias a su robustez y estructura tubular (Soni et al., 2010), y porque su morfología permite medir fácilmente dimensiones particularmente idóneas para enfoques métricos, como largos y anchos (Kim et al., 2013).

Entre los elementos del post-cráneo, la evaluación del fémur reviste notables ventajas para la estimación de sexo. Corresponde a un elemento particularmente dimórfico: su morfología está influenciada no sólo por la diferencia de tamaño corporal y robusticidad entre ambos sexos, como en otros huesos largos, sino además por la directa relación biomecánica en que se encuentra con la pelvis que, como se expuso previamente, es la región más dimórfica del esqueleto humano (Albanese, 2003; Albanese et al., 2008; Christensen et al., 2014). Además, el fémur es el hueso más grande, pesado, robusto y resistente del esqueleto humano, lo que sumado a la protección que otorga la gruesa capa de tejido blando que se extiende sobre él (principalmente tejido muscular) lo hacen uno de los elementos óseos menos susceptible al daño, con mayor recuperación y mejor preservación en toda clase de contextos (Klepinger, 2006; Albanese et al., 2008; Kranioti, Vorniotakis, Galiatsou, Iscan y Michalodimitrakis, 2009; Kim, Lee y Kim, 2010; White, Black y Folkens, 2012; Kim et al., 2013; Anastopoulou et al., 2014; Mitra et al., 2014; Clavero et al., 2015; Gulhan, Harrison y Kiris, 2015; Curate et al., 2016; Djorojevic, Roldán, Botella y Alemán, 2016; Guyomarc'h et al., 2016; Curate et al., 2017a; Curate et al., 2017b). La buena conservación del fémur lo hacen particularmente útil para la identificación en contextos forenses, ya que en estos casos el cráneo y la pelvis se hallan frecuentemente fragmentados y mezclados, mientras que el fémur es comúnmente recuperado en mejor estado de preservación (Gulhan et al., 2015). Finalmente, el fémur es el hueso largo más estudiado del esqueleto humano, favoreciendo la comparación de resultados y la familiaridad del observador con el material analizado (Iscan y Stein, 2013). Estas características idóneas del fémur concuerdan con la vasta literatura existente que ha demostrado el excelente potencial de este elemento para estimar sexo (Curate et al., 2017b).

Los estudios recientes más importantes de estimación de sexo en fémur han considerado medidas a través de todo el hueso, incluyendo la cabeza, el cuello, los trocánteres, la diáfisis, los cóndilos, el largo, el peso, el volumen y la densidad (Mall, Graw, Gehring y Hubig, 2000; Asala, 2002; Albanese, 2003; Frutos, 2003; Asala, Bidmos y Dayal, 2004; Purkait, 2005; Brown, Ubelaker, Schanfield, 2007; Harma y Karakas, 2007; Albanese et al., 2008; Alunni-Perret, Staccini y Quatrehomme, 2008; Jantz, Kimmerle y Baraybar, 2008; du Jardin et al., 2009; Kranioti et al., 2009; Soni et al., 2010; Stull y Godde, 2012; Tise, Spradley y Anderson, 2012; Kim et al., 2013; Anastopoulou et al., 2014; Mitra et al., 2014; Sharma, Sharma, Jehan y Sastya, 2014; Spradley et al., 2014; Alunni, du Jardin, Nogueira, Buchet y Quatrehomme, 2015; Boldsen, Milner y Boldsen, 2015; Clavero et al., 2015; Gulhan et al., 2015; Hishmat et al., 2015; Longe, 2015; Meeusen, Christensen y Hefner, 2015; Curate et al., 2016; Djorojevic et al., 2016; Guyomarc'h et al., 2016; Curate et al., 2017a; Curate et al., 2017b). La mayoría de las investigaciones han utilizado morfometría lineal y muchas han considerado la elaboración de modelos predictivos basados en más de una variable.

Si bien se ha podido establecer que la totalidad del fémur es útil para la estimación de sexo, la epífisis proximal destaca dentro de la morfología femoral por ser la región que concentra mayor dimorfismo sexual. Esto se debe, primero, a que aloja una serie de inserciones musculares sexualmente dimórficas (Kim et al., 2013) y, segundo, a que es la región más influenciada por el dimorfismo de la pelvis, dada la estrecha relación biomecánica en que se encuentra con la cintura pélvica (Albanese, 2003; Asala et al., 2004; Albanese et al., 2008; Christensen et al., 2014), teniendo además un importante rol en la transmisión del peso corporal – sexualmente dimórfico – desde el esqueleto axial hacia las extremidades inferiores y en un amplio rango de movimientos propios de la articulación de la cadera (Asala et al., 2004, Djorojevic et al., 2016). Por otro lado, corresponde al área mejor protegida del fémur, dado el resguardo que ofrece tanto la voluminosa musculatura del extremo proximal de la pierna como los coxales (Netter, 2011; White et al., 2012), lo que favorece su preservación.

Dentro del fémur proximal, la cabeza femoral ha sido la estructura más estudiada y comúnmente es considerada como la mejor para la estimación de sexo (Curate et al., 2016). Las variables evaluadas han incluido el diámetro vertical, diámetro transversal, diámetro máximo y la circunferencia de la cabeza, tanto en modelos univariantes como multivariantes (Dittrick y Suchey, 1986; Steyn & Iscan, 1997; Mall et al., 2000; Asala, 2002; Albanese, 2003; Asala et al., 2004; Brown et al., 2007; Jantz et al., 2008; Kranjoti et al., 2009; Soni et al., 2010; Tise et al., 2012; Mitra et al., 2014; Spradley et al., 2014; Alunni et al., 2015; Boldsen et al., 2015; Clavero et al., 2015; Gulhan et al., 2015; Longe, 2015; Djorojevic et al., 2016; Guyomarc'h et al., 2016; Curate et al., 2017b). Dada su predominancia en la estimación de sexo, es común encontrar que en una población sólo se haya evaluado la cabeza femoral. No obstante, existen otras estructuras y variables sexualmente dimórficas en el fémur proximal, y varias medidas han sido propuestas para instrumentalizar este dimorfismo para la estimación de sexo (Center et al., 1998; Seidemann, Stojanowski y Doran, 1998; Frutos, 2003; Asala et al., 2004; Purkait, 2005; Brown et al., 2007; Albanese et al., 2008; Anastopoulou et al., 2014; Mitra et al., 2014; Clavero et al., 2015; Gulhan et al., 2015; Meeusen et al., 2015; Curate et al., 2016; Djorojevic et al., 2016; Curate et al., 2017b).

El cuello del fémur ha sido una de las estructuras mejor estudiadas en la epífisis proximal, con el ancho y el largo como las dos principales variables analizadas. El largo del cuello femoral ha sido abordado mediante más de una medida; no obstante, la aproximación métrica más estudiada para la estimación de sexo ha sido el largo del eje cuello femoral, comúnmente abreviado FNAL por sus siglas en inglés (Figura 1 en Anexos). FNAL fue originalmente propuesto por Center et al. (1998) como medición para evaluar el riesgo de fractura de cadera en hombres y mujeres adultos mayores. Al evidenciarse dimorfismo sexual, esta variable ha sido utilizada en los últimos años para estimar sexo. Gulhan et al. (2015) evaluaron la utilización de FNAL, entre otras medidas, en una muestra de población turca (Hospital Mehmet Akif Ersoy de Estambul) a través de imágenes radiológicas, obteniendo un acierto del 82.5% mediante un modelo calculado por análisis de función discriminante. Por su parte, Meeusen et al. (2015) utilizaron FNAL para estimar el sexo y la ancestría en una muestra de individuos estadounidenses afroamericanos y caucásicos (Robert J. Terry Anatomical Skeletal Collection y William M. Bass Donated Skeletal

Collection) y nativos americanos (Averbuch Site Skeletal Collection), reportando un acierto al sexo de los individuos del 83.51% mediante regresión logística; para la estimación de ancestría, sin embargo, no se obtuvieron resultados satisfactorios. Curate et al. (2016), por otro lado, emplearon una muestra de población portuguesa (Luís Lopes Collection para la muestra de entrenamiento, y Coimbra Identified Skeletal Collection para la muestra de validación) y regresión logística y algoritmo C4.5 como aproximaciones estadísticas, obteniendo entre un 80.1% y un 85.3% de acierto al sexo de los individuos. Finalmente, Curate et al. (2017b) reportaron entre un 76% y 79.5% de acierto para FNAL empleando REPTree para elaborar los modelos predictivos de sexo, también sobre una muestra portuguesa (Coimbra Identified Skeletal Collection para la muestra de entrenamiento; 21 Century Identified Skeletal Collection para la muestra de validación).

Por otra parte, el ancho del cuello femoral (FNW, por sus siglas en inglés) es una medida tradicional de la osteometría, ampliamente utilizada en estudios de estimación de sexo (Figura 1 en Anexos). Fue evaluada por primera vez por Seidemann et al. (1998) reportando un nivel de acierto en torno al 90% para una muestra de población estadounidense afrodescendiente y caucásica (Hamann-Todd Collection) mediante función discriminante. Posteriormente, Stojanowski y Seidemann (1999) reevaluarían el uso de FNW en una muestra de población actual de diversas ancestrías (Documented Collection at the University of New Mexico), reportando porcentajes de acierto entre un 82% y un 97% para ancestrías separadas y de un 85% para ancestrías combinadas utilizando modelos obtenidos por función discriminante. Frutos (2003) evaluó el potencial del ancho del cuello femoral para la estimación de sexo en una muestra de población guatemalteca (Fundación de Antropología Forense de Guatemala), obteniendo un 89.5% de acierto al sexar mediante función discriminante, mientras que Asala et al. (2004) alcanzaron un 78.2% de acierto para FNW con un modelo obtenido también por función discriminante aplicado sobre una muestra de tribus Sotho, Xhosa y Zulu (Raymond Darts Collection of Human Skeleton). Posteriormente, Mitra et al. (2014) evaluaron la estimación de sexo a través del ancho del cuello femoral en una muestra radiológica contemporánea iraní (Hospital Sina de Teherán), calculando puntos de corte mediante curva ROC, con lo que obtuvieron un 77% de acierto. Clavero et al. (2015) ocuparon una muestra española de tomografías computarizadas provenientes de hospitales de la Comunidad Autónoma de Castilla-La Mancha, reportando un 90.4% de acierto al sexo de los individuos al utilizar un modelo obtenido por regresión logística. Gulhan et al. (2015) también probaron esta medida en una muestra de tomografías computarizadas de origen turco provenientes del Departamento de Radiología del Hospital Mehmet Akif Ersoy de Estambul, obteniendo un 88% de acierto mediante función discriminante. Curate et al. (2016), además de FNAL, evaluaron también FNW en la misma muestra portuguesa y empleando las mismas aproximaciones estadísticas (regresión logística y algoritmo C4.5), alcanzando entre un 82.5% y un 85.3% de acierto. Djorojevic et al. (2016), por su parte, evaluaron FNW en una muestra española (San José Identified Skeletal Collection) mediante función discriminante, logrando un 91.9% de acierto. Finalmente, Curate et al. (2017b) además de FNAL también evaluaron FNW, obteniendo entre un 77% y un 79.5% de acierto al sexar los individuos de la muestra mediante REPTree.

Un conjunto de variables dentro del fémur proximal fue propuesto por Purkait (2005), en base a tres puntos ubicados sobre el margen articular de la cabeza y las apófisis de ambos trocánteres. Al trazar las distancias entre cada par de puntos se obtienen tres variables métricas que forman un triángulo (Figura 2 en Anexos). Purkait (2005) evaluó el potencial de estas variables para estimar sexo en una colección de India Central usando análisis de función discriminante para obtener los modelos predictivos de sexo. Los porcentajes de acierto reportados por el autor se ubicaron por sobre el 80%, llegando hasta un 86.4% para la combinación de las tres variables juntas, con la excepción de un modelo univariable (AB) que sólo reportó un 62.5% de acierto. Las variables propuestas por Purkait han sido posteriormente evaluadas para la estimación de sexo por otros autores en poblaciones diversas, obteniendo resultados variados. Así, Brown et al. (2007) evaluaron el triángulo de Purkait en una muestra de población estadounidense de origen mixto (Terry Collection) obteniendo desde un 69% hasta un 87% de acierto para variables individuales y combinadas mediante función discriminante. Anastopoulou et al. (2014), por su parte, pusieron a prueba las medidas del triángulo de Purkait en una muestra de población griega (Athens Collection) también mediante función discriminante, y obtuvieron rendimientos desde un 54.32% hasta un 74.04% para modelos univariados y multivariados. Djorojevic et al. (2016) además de FNW también evaluaron el triángulo de Purkait, tanto las fórmulas originalmente propuestas por el autor como otras nuevas específicas para la muestra obtenidas mediante función discriminante. Para los modelos originales, se obtuvieron rendimientos entre un 74.7% y un 81.7% de acierto general, en algunos casos con una gran diferencia en el acierto para cada sexo. En cambio, los nuevos modelos específicos obtuvieron aciertos desde un 76.9% hasta un 86.6% entre modelos univariados y multivariados, con sesgos por sexo significativamente menores que en los modelos originales.

Otro conjunto de variables dentro del fémur proximal fue también propuesto por Albanese et al. (2008), a partir de tres puntos asociados a regiones de la epífisis proximal – trocánter menor, trocánter mayor y fovea capitis. Las distancias entre cada par de puntos describen tres variables métricas que, como con las variables propuestas por Purkait (2005), forman un triángulo a través de la morfología proximal del fémur (Figura 3 en Anexos). Al comparar ambas propuestas, sin embargo, se puede caracterizar al esquema de Purkait como un triángulo “interno”, mientras que al de Albanese et al. (2008) como uno “externo”. Los autores reportaron porcentajes de acierto de hasta un 95% para modelos que incluían las variables señaladas. No obstante, pese a los buenos resultados reportados, las medidas propuestas por Albanese et al. (2008) apenas han sido reevaluadas por otros investigadores. La excepción la constituye el estudio antes citado de Clavero et al. (2015), en que los autores evaluaron GTFC – denominada DHT en su estudio – en una muestra española contemporánea mediante un modelo de regresión logística, con el que obtuvieron un porcentaje de acierto al sexo de un 89.5%.

Además de los modelos univariados señalados, se han evaluado algunas interacciones entre estas variables. Específicamente, Curate et al. (2016) probaron la combinación de FNAL y FNW para la estimación de sexo, reportando un acierto general desde un 82.9% hasta un 87.7%, mientras que Djorojevic et al. (2016) evaluaron la combinación de FNW con las variables del triángulo de Purkait (2005), obteniendo entre un 89.3% y un 94.6% de acierto al sexo de los individuos de la muestra.

Los antecedentes expuestos demuestran el potencial del fémur proximal para la estimación de sexo. En este sentido, es posible que la evaluación de estas variables sea un aporte en la estimación de sexo en población chilena moderna. No obstante, pocos estudios han sido llevados a cabo en esta línea. Dentro del post-cráneo no-pélvico, Ross y Manneschi (2010) alcanzaron un nivel de precisión sobre el 80% para medidas en el húmero y el fémur; Garrido-Varas et al. (2014), por su parte, entre un 70% y 90.5% para medidas en varios elementos del post-cráneo no-pélvico; por otro lado, Peckmann et al. (2016) reportaron un porcentaje del 80% - 86% para la escápula. Finalmente, O'Bright, Peckmann y Meek (2018) obtuvieron porcentajes de acierto para medidas en la tibia desde un 71.4 hasta un 89.2%.

En el caso específico del fémur, se han evaluado las variables de la circunferencia al centro de la diáfisis, largo máximo y ancho bicondilar, con un 82%, 76.9% y 84.1% de acierto, y, en la epífisis proximal, el diámetro vertical, transversal y máximo de la cabeza femoral, con un 84.8%, 87.7% y 86%, respectivamente (Ross y Manneschi, 2010; Garrido-Varas et al., 2014).

Problematización

En Chile, los escenarios más comunes en los que la Antropología Forense se desenvuelve son aquellos relacionados a violaciones a los Derechos Humanos cometidos durante la Dictadura Cívico-Militar (1973-1990) y, en menor medida, en casos de índole médico-legal actuales (e.g. homicidios) y desastres naturales (Ross y Manneschi, 2010; Garrido-Varas et al., 2014; Peckmann et al., 2016). Dado que el uso extenso de análisis de ADN para la identificación de restos es aún demasiado costoso para las instituciones chilenas, el proceso de individualización de las víctimas es abordado mediante un enfoque multidisciplinario, que incluye análisis de Antropología Forense (Garrido-Varas e Intriago, 2011).

En el contexto particular de los casos de violaciones a los Derechos Humanos, el cráneo y la cintura pélvica se encuentran comúnmente entre los elementos más afectados, el primero principalmente por traumas perimortem y el segundo por daño tafonómico, sobre todo la región púbica dada su desprotegida ubicación (Garrido-Varas et al., 2014). Además, es escenario común en estos contextos la presencia de restos mezclados y altamente fragmentados (Garrido-Varas e Intriago, 2011; Garrido-Varas et al., 2014). En este sentido, la elaboración de métodos de estimación de sexo en elementos mejor preservados y que puedan ser aplicados en restos fragmentados resulta esencial en el contexto chileno.

Si bien se han obtenido buenos resultados, los modelos publicados a la fecha, tanto para el fémur como para otros elementos post-craneales no-pélvicos, se han limitado a la inclusión de una sola variable, renunciando al potencial que tiene la combinación de más de una medida dentro de un mismo modelo (Ross y Manneschi, 2010; Garrido-Varas e Intriago, 2011; Garrido-Varas et al., 2014; Peckmann et al., 2016). Por otro lado, el fémur proximal que, como vimos, reviste importantes ventajas para la estimación de sexo, no ha sido evaluado en su totalidad. Por el contrario, los estudios se han limitado a medidas tradicionales de la cabeza femoral, agotando las posibles mediciones en esta estructura pero ignorando otras variables sexualmente dimórficas también presentes en la epífisis proximal del hueso. Por estas razones, esta insuficiente evaluación del fémur proximal representa un importante vacío de conocimiento, particularmente en el contexto forense chileno caracterizado por una deficiente recuperación del cráneo y la pelvis en el que contar con modelos de estimación de sexo en fémur proximal resultaría especialmente conveniente. El presente estudio tiene por objeto explorar la morfología del fémur proximal para desarrollar una metodología de estimación de sexo en población chilena moderna, en base a medidas inéditas e incluyendo modelos multivariados, aplicable en contextos forenses subactuales y de causas de violaciones a los Derechos Humanos cometidas durante la Dictadura Cívico-Militar en Chile.

Pregunta de Investigación

¿Cómo se relaciona la expresión de la morfología del fémur proximal con el sexo documentado en individuos de población chilena moderna?

Hipótesis

Existe una asociación estadísticamente significativa entre el sexo y la expresión de la morfología del fémur proximal, susceptible de ser utilizada para la estimación de sexo en población chilena moderna.

Objetivos

Objetivo General

Conocer e instrumentalizar la relación existente entre la expresión de la morfología del fémur proximal y el sexo documentado de los individuos para la estimación del sexo.

Objetivos Específicos

- a) Conocer la expresión de la morfología del fémur proximal en la muestra estudiada
- b) Analizar la relación entre la expresión de la morfología del fémur proximal en la muestra estudiada y el sexo documentado de los individuos
- c) Construir modelos predictivos del sexo documentado de los individuos a partir de las variables analizadas del fémur proximal
- d) Validar los modelos construidos, evaluando su certeza y precisión

Materiales

Se utilizó una muestra de 270 individuos, dividida aleatoriamente en dos submuestras: una de entrenamiento compuesta por 200 individuos (100 femeninos, 100 masculinos) empleada para la elaboración de los modelos predictivos de sexo, y otra de prueba compuesta por 70 individuos (35 femeninos, 35 masculinos) utilizada para la validación de los modelos construidos. El tamaño muestral considerado es similar o mayor a las muestras utilizadas comúnmente en estudios previos con objetivos similares (Tise et al., 2012; Kim et al., 2013; Anastopoulou et al., 2014; Mitra et al., 2014; Sharma et al., 2014; Spradley et al., 2014; Alunni et al., 2015; Boldsen et al., 2015; Clavero et al., 2015; Gulhan et al., 2015; Hishmat et al., 2015; Meeusen et al., 2015; Curate et al., 2016; Djorojevic et al., 2016; Guyomarc'h et al., 2016; Curate et al., 2017a; Curate et al., 2017b). La Tabla 1 muestra la distribución de la muestra por sexo y edad según la categorización etaria propuesta por Buikstra y Ubelaker (1994). Si bien los segmentos etarios no están equitativamente representados, varios estudios han descartado que exista una influencia significativa de la edad en la determinación de sexo a partir del fémur proximal (Center et al., 1998; Alunni-Perret et al., 2008; du Jardin et al., 2009; Djorojevic et al., 2016). Además, la muestra representa un amplio rango etario - entre 20 y 96 años.

Tabla 1: Distribución etaria de la muestra de entrenamiento por sexo(n=200)

Categoría	Rango Etario (años)	Sexo	Frecuencia
Adulto Joven	20-34	Femenino	18
		Masculino	15
Adulto Medio	35-49	Femenino	27
		Masculino	38
Adulto Mayor	50+	Femenino	55
		Masculino	47

Todos los individuos de la muestra fueron seleccionados de forma aleatoria y sin sesgo conocido desde la Colección Subactual de Santiago, tutelada por el Departamento de Antropología de la Universidad de Chile. Se trata de una colección de más de 1500 individuos rescatados del Cementerio General de Santiago, a través de diversas resacas realizadas durante la década de 1970 gracias a las gestiones del profesor Juan Munizaga, y durante 1993 como parte del proyecto Fondecyt N°1028-91, con el objetivo de contar con una colección osteológica para la actividad docente y la elaboración de estándares adecuados para la identificación forense aplicables a la población chilena. La colección representa un segmento de la población de Santiago fallecida entre 1950 y 1973, procedentes de tumbas transitorias pertenecientes a individuos sin deudos que los reclamasen y que, de no ser rescatados, estaban destinados a su destrucción (Lemp, Rodríguez, Retamal y Aspillaga, 2008; Abarca, 2011; Herrera, 2012; Morales et al., 2016; Peckmann et al., 2016).

Dada su preservación, tamaño, variabilidad y documentación, la Colección Subactual de Santiago posee un enorme valor científico para estudiar temas vinculados al área forense y constituye una de las colecciones osteológicas más importantes a nivel regional e incluso mundial (Lemp et al., 2008; Morales et al., 2016). Es por sus características que múltiples estudios y tesis se han llevado a cabo utilizando el material de la colección, incluyendo la evaluación y elaboración de métodos (Paredes, Hagn y Constantinescu, 1994; Díaz, 2010; Ross y Manneschi, 2010; Herrera, 2012; Abarca, 2013; Garrido-Varas et al., 2014; Espinoza; 2016; Peckmann et al., 2016; Herrera y Retamal, 2017) e investigaciones de otra índole (Aspillaga, Gonzáles y Paredes, 1984; Retamal, 2004; Urzúa, Huberman, Delgado, Pacheco y Retamal, 2009; Pacheco, 2010; Abarca, 2011; Pontigo, 2013). Para el presente estudio, las propiedades de la colección permitieron contar con una muestra relativamente grande, con equilibrio en la representación de ambos sexos y un amplio rango etario.

Una de las principales características de la Colección Subactual de Santiago es que cuenta con una extensa documentación, incluyendo información sobre el nombre, sexo, edad, causa de muerte y, en algunos casos, fecha de nacimiento y fallecimiento de los individuos (Lemp et al., 2008; Morales et al., 2016). Gracias a la contextualización histórica de la colección llevada a cabo por Abarca (2011), se sabe que los restos corresponden principalmente a personas de estrato socioeconómico bajo, habitantes de la periferia de la ciudad, que vivieron en precarias condiciones de subsistencia, con escaso acceso a servicios de salud, agua potable, electricidad y con una alimentación nutricionalmente deficiente. Estas condiciones ambientales desfavorables representan un importante factor biocultural en la colección que podría constituir una noxa al crecimiento y desarrollo de la población representada en la colección (Abarca, 2011).

No obstante, la Colección Subactual de Santiago es una fuente de material apropiada para la elaboración de metodologías aplicables a casos de violaciones a los Derechos Humanos cometidas durante la Dictadura Cívico-Militar en Chile puesto que, considerando las fechas de las defunciones de los individuos, la colección representa un sector de la población chilena contemporánea a las víctimas (Garrido-Varas et al., 2014; Peckmann et al., 2016). Además, existe una correspondencia entre la composición de la colección y los datos demográficos de las víctimas de violaciones a los DDHH, que se concentran en población de estratos sociales medios y bajos habitantes de la Región Metropolitana (Ballesteros, 1995). Considerando esto, la Colección Subactual de Santiago es la mejor fuente de material osteológico para el desarrollo de métodos antropológico-forenses disponible actualmente en el país, razón por la cual ha sido anteriormente empleada en estudios con aplicabilidad en estos contextos (Paredes et al., 1994; Díaz, 2010; Ross y Manneschi, 2010; Herrera, 2012; Abarca, 2013; Garrido-Varas et al., 2014; Espinoza; 2016; Peckmann et al., 2016; Herrera y Retamal, 2017).

Los criterios de inclusión/exclusión para los individuos fueron:

- a) Sexo documentado
- b) Adulthood de la región osteológica estudiada, evidenciada a través de la madurez completa de la epífisis proximal del fémur (i.e. completa fusión de la cabeza femoral, trocánter mayor y trocánter menor)
- c) Región proximal del fémur izquierdo libre de alteraciones tafonómicas o patológicas que pudieran afectar la medición de las variables evaluadas

Arbitrariamente, sólo se consideraron fémures izquierdos. Esta decisión metodológica se fundamenta en la convención común en investigaciones en huesos largos, y particularmente en fémur, de evaluar sólo el miembro izquierdo por defecto (Frutos, 2003; Asala et al., 2004; Harma y Karakas, 2007; Albanese et al., 2008; Alunni-Perret et al., 2008; du Jardin et al., 2009; Soni et al., 2010; Clavero et al., 2015; Longe, 2015; Meeusen et al., 2015; Curate et al., 2016; Guyomarc'h et al., 2016; Curate et al., 2017a). Además, la mayoría de los estudios que han evaluado la asimetría bilateral del fémur no han observado diferencias estadísticamente significativas entre ambos lados (MacLaughlin y Bruce, 1985; Mall et al., 2000; Purkait, 2005; Alunni-Perret et al., 2008; Kim et al., 2013; Anastopoulou et al., 2014; Mitra et al., 2014; Alunni et al., 2015; Boldsen et al., 2015; Hishmat et al., 2015; Djorojevic et al., 2016) y es práctica frecuente la utilización de ambos miembros indistintamente o en conjunto para la elaboración y validación de métodos de estimación de sexo (MacLaughlin y Bruce, 1985; Dittrick y Suchey, 1986; Steyn e Iscan, 1997; Asala, 2002; Albanese, 2003; Purkait, 2005; Brown et al., 2007; Albanese et al., 2008; Jantz et al., 2008; Tise et al., 2012; Kim et al., 2013; Anastopoulou et al., 2014; Spradley et al., 2014; Alunni et al., 2015; Boldsen et al., 2015; Clavero et al., 2015; Gulhan et al., 2015; Longe, 2015; Meeusen et al., 2015; Djorojevic et al., 2016; Guyomarc'h et al., 2016; Curate et al., 2017b). Más aún, para el caso chileno, no se han encontrado diferencias estadísticamente significativas entre ambos miembros a nivel del fémur proximal (Garrido-Varas et al., 2014). No obstante lo anterior, y ante la posibilidad de que existan tales diferencias entre ambos lados, se evaluó la asimetría bilateral para las variables analizadas para poder así asegurar la reproductibilidad del método en ambos miembros (véase Métodos).

Métodos

Variables

Las variables analizadas corresponden a medidas definidas y evaluadas por estudios previos en fémur proximal y que han demostrado potencial como predictores de sexo. En la Tabla 2 se detallan la definición y referencias para cada variable y en las Figuras 1, 2 y 3 (Anexos) se ilustran de forma gráfica.

Tabla 2: Variables analizadas

Variable	Definición	Referencias
FNAL	Distancia entre la base del gran trocánter y el ápex de la cabeza femoral	Center et al., 1998; Gulhan et al., 2015; Meeusen et al., 2015; Curate et al., 2016; Curate et al., 2017b
FNW	Diámetro mínimo del cuello femoral medido en sentido supero-inferior	Seidemann et al., 1998; Stojanowski y Seidemann, 1999; Frutos, 2003; Asala et al., 2004; Mitra et al., 2014; Clavero et al., 2015; Gulhan et al., 2015; Curate et al., 2016; Djorojevic et al., 2016; Curate et al., 2017b
AB	Distancia entre el punto más lateral del margen articular de la cabeza femoral (A) y el punto más medial del trocánter mayor (B)	
AC	Distancia entre el punto más lateral del margen articular de la cabeza femoral (A) y el punto más pósteromedial del trocánter menor (C)	Purkait, 2005; Brown et al., 2007; Anastopoulou et al., 2014; Djorojevic et al., 2016
BC	Distancia entre el punto más medial del trocánter mayor (B) y el punto más pósteromedial del trocánter menor (C)	
GTFC	Distancia entre el punto más lateral del gran trocánter (GT) y el punto más superior de la fovea capitis (FC)	
GTLT	Distancia entre el punto más lateral del gran trocánter (GT) y el punto más superior de la carilla del trocánter menor (LT)	Albanese et al., 2008; Clavero et al., 2015
LTFC	Distancia entre el punto más superior de la carilla del trocánter menor (LT) y el punto más superior de la fovea capitis (FC)	

Todas las mediciones fueron tomadas con un pie de metro digital (precisión de ± 0.02 mm para >100 mm; ± 0.03 mm para 100-200mm).

Error intraobservador

Para evaluar el error intraobservador, se midió una submuestra aleatoria de 20 fémures (10 femeninos, 10 masculinos) en dos oportunidades diferentes y se calculó el Coeficiente de Correlación de Concordancia de Lin, estadístico que mide el grado de concordancia entre pares de observaciones de una misma variable continua (Lin, 1989). Se consideró que existía una concordancia significativa entre ambas observaciones con $\rho_c \geq 0.9$ y significación estadística ($p < 0.05$)

Asimetría bilateral

Para estimar la asimetría bilateral, se empleó una submuestra aleatoria de 20 individuos (10 femeninos, 10 masculinos) en la que se midieron ambos miembros, se calculó la estadística descriptiva de cada lado y se realizaron pruebas t de Student pareadas para cada par de mediciones de distinto lado bajo la hipótesis nula de que no existen diferencias significativas entre ambas observaciones. Se consideró que existía asimetría bilateral significativa con $p < 0.05$.

Estadística descriptiva

La estadística descriptiva de los datos consideró el cálculo del promedio, desviación estándar, intervalo de confianza (95%), asimetría y curtosis para cada variable por sexo. Además, se obtuvieron gráficos de cajas y bigotes para ilustrar la distribución de cada variable por sexo. Se evaluó la normalidad de las variables para cada sexo mediante test de Shapiro-Wilk, rechazando normalidad con $p < 0.05$. Finalmente, se comprobó si existían diferencias estadísticamente significativas entre ambos sexos en cada variable con test t de Student y Mann-Whitney-Wilcoxon bajo la hipótesis nula de que los promedios de las variables son iguales en los individuos femeninos y masculinos, rechazándola con $p < 0.05$. Se utilizaron ambas pruebas dado que t de Student sólo permite trabajar con datos normalmente distribuidos, empleándose Mann-Whitney-Wilcoxon para las variables en que se descartó normalidad.

Modelos estadísticos

En el presente estudio se empleó regresión logística como aproximación estadística para crear los modelos predictivos de sexo en base a las variables analizadas.

La regresión logística es una forma de regresión que permite analizar y describir la relación existente entre una variable dependiente dicotómica, denominada variable respuesta, y una o más variables independientes, denominadas explicativas o predictoras, mediante un modelo matemático parsimonioso y de interpretación directa (Kleinbaum y Klein, 2010; Hosmer, Lemeshow y Sturdivant, 2013; Hyeoun-Ae, 2013). Para lograrlo, la regresión logística ajusta los datos a una distribución logística, que describe una curva característicamente sigmoidea (con forma de "S"), mediante la función enlace logit - que representa el logaritmo natural del *odd ratio* del evento -, y estima la probabilidad de ocurrencia del evento que representa la variable dependiente dicotómica, el denominado *log-odd*. En nuestro caso particular, el evento está representado por la ocurrencia de que el individuo sea de sexo masculino - codificado "1" -, y la no ocurrencia de que sea femenino - codificado "0".

El modelo resultante expresa la ocurrencia del evento como una función lineal de las variables independientes, de la forma:

$$L = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Donde L es el log-odd, α la constante (o *intercepto*), β_n el coeficiente para cada variable y X_n el valor de cada variable – en este caso, los valores de las medidas obtenidas para cada individuo. El coeficiente es un estimador numérico (calculado por máxima verosimilitud) del impacto que tiene cada variable explicativa sobre la variable dependiente. En los modelos aquí elaborados, un valor negativo (menor a 0) para L clasifica al individuo como femenino, mientras que uno positivo (superior a 0) lo clasifica como masculino.

Como parte de los resultados obtenidos de las regresiones logísticas, se obtienen los valores para los estadísticos de Wald y pseudo- R^2 . Wald es un estadístico utilizado para evaluar la significancia del coeficiente de una variable en términos de su contribución dentro de un modelo, mientras que pseudo- R^2 es un estimador del tamaño del efecto, entregando un aproximado de la proporción de la variabilidad de la variable dependiente que es explicada por el modelo (Hyeoun-Ae, 2013). Se consideró que un coeficiente era significativo con $p < 0.05$ para Wald. Para el estadístico pseudo- R^2 no existe una convención de niveles aceptables para el tamaño del efecto. No obstante, se espera que un mayor tamaño del efecto se relacione con un mejor rendimiento del modelo, aunque para aumentos discretos del pseudo- R^2 esto no siempre ocurre.

Se eligió la regresión logística como aproximación estadística para construir los modelos predictivos del sexo por su capacidad de generar resultados robustos exigiendo pocos requisitos. Específicamente, el cálculo de regresión logística no requiere normalidad ni homocedasticidad de los datos, ni asume linealidad en la relación entre la variable dependiente y las variables independientes, a diferencia de otros métodos de regresión como la función discriminante, que ha sido la aproximación estadística más ampliamente utilizada en la literatura (Norusis en Albanese, 2003; du Jardin et al., 2009; Kleinbaum y Klein, 2010; Hosmer et al., 2013; Hyeoun-Ae, 2013; Clavero et al., 2015). Esto es relevante puesto que tales condiciones pueden no darse, e incluso si se cumplen la regresión logística reporta mejores resultados que la función discriminante, con un menor porcentaje de clasificación errónea y un bajo sesgo por sexo (Albanese et al., 2008; Clavero et al., 2015). Por estas razones, ha podido demostrarse en la literatura que la regresión logística es una herramienta eficiente en el contexto antropológico (Guyomarc'h et al., 2016) y puede ser utilizada en lugar del análisis de función discriminante dentro de estudios forenses (Albanese et al., 2008).

Además de función discriminante y regresión logística, otros enfoques estadísticos han sido empleados en la literatura para la elaboración de modelos de estimación de sexo en fémur proximal, incluyendo Redes Neurales (du Jardin et al., 2009; Alunni et al., 2015), algoritmo C4.5 (Curate et al., 2016), CART (Curate et al., 2017a), Machine-learning (Curate et al., 2017b) y REPTree (Curate et al., 2017a; Curate et al., 2017b). Si bien estas aproximaciones alternativas han reportado resultados prometedores, a veces con mayores niveles de acierto que otros enfoques más tradicionales, el cálculo de modelos predictivos en base a estas herramientas puede resultar demasiado complejo, incluso limitante, y su aplicación

carece del potencial práctico que tienen enfoques más simples como el análisis de función discriminante o la regresión logística, cuyos modelos son económicos y fáciles de entender y usar (du Jardin et al., 2009). Además, no siempre se han obtenido mejores resultados para estos enfoques en comparación a la regresión logística (Alunni et al., 2015; Curate et al., 2016). Finalmente, estos métodos estadísticos sólo han sido explorados en la literatura, sin una extensa evaluación y validación por parte de otros investigadores. Siendo aún experimentales, quedan fuera del espectro de interés del presente estudio.

Para la predicción de una variable dependiente pueden construirse modelos univariados, que sólo incorporan una variable explicativa, y modelos multivariados, que incluyen más de una variable predictora de forma combinada. Dado que sólo incluyen una variable aislada, los modelos univariados pueden aplicarse en contextos de preservación parcial, frente a material incompleto o altamente fragmentado, cuando una preservación deficiente de los restos impide evaluar otras variables sexualmente dimórficas (Alunni-Perret et al., 2008; Alunni et al., 2015; Boldsen et al., 2015; Curate et al., 2016). No obstante, dado que se basan en diferencias absolutas de una sola dimensión, son más susceptibles a la composición de la muestra de entrenamiento utilizada para elaborar el modelo y, debido a que no alcanzan a rescatar el total de la variación entre ambos sexos, tienden a obtener un menor acierto y un mayor sesgo por sexo en comparación al uso de más de una variable combinada (Albanese, 2003; Alunni et al., 2015; Meeusen et al., 2015). Los modelos multivariados representan el caso opuesto. Al incluir más de una variable dentro de un mismo modelo, se logra recoger mayor dimorfismo sexual, lo que se traduce en una mayor certeza y precisión. Sin embargo, dado que requieren de la disponibilidad de más de una medida, requieren de una mejor preservación de los restos, lo que reduce su potencial de aplicación. Además, en la medida en que se incluyen variables altamente dimórficas dentro de un modelo, éste alcanzará un punto de saturación en que no se rescata más variabilidad asociada al sexo ni se obtiene un mayor acierto, por lo que la inclusión de nuevas variables pierde sentido al hacer al modelo menos eficiente – mismo rendimiento, pero más variables. Considerando las virtudes y desventajas de cada aproximación, en el presente estudio se construyeron modelos tanto univariados como multivariados para la estimación de sexo a partir de las variables analizadas.

Como criterios para la elección de las variables utilizadas en las regresiones logísticas se procuró que las combinaciones de medidas empleadas representasen escenarios realistas de recuperación de restos y además que incluyesen las interacciones de variables evaluadas previamente en la literatura. De esta forma, se realizaron ocho regresiones logísticas univariadas, una para cada variable individual (FNAL, FNW, AB, AC, BC, GTFC, GTLT y LTFC), y treinta regresiones logísticas multivariadas, incluyendo una para cada combinación posible de dos variables (veintiocho en total) y dos regresiones logísticas para combinaciones de tres variables - una para la combinación de las variables propuestas por Purkait (2005) y otra para las formuladas por Albanese et al. (2008). En consideración de los resultados obtenidos, se decidió a posteriori no elaborar más modelos de tres o más variables (véase Discusión).

Los modelos resultantes de las regresiones logísticas fueron nombrados en base a las variables incluidas, siguiendo el orden en que se han expuesto a lo largo del texto - FNAL, FNW, AB, AC, BC, GTFC, GTLT y LTFC. Para evitar cualquier confusión con las variables, los modelos figuran escritos en cursiva.

Validación

En la literatura, el método más comúnmente empleado para validar un modelo es la validación cruzada (*cross-validation*), en el cual un modelo es recalculado mediante iteraciones y luego reevaluado sobre una submuestra aleatoria de la muestra total de entrenamiento (e.g. *leave-one-out*). Si bien esta forma de validación tiene el beneficio de permitir evaluar un modelo utilizando la misma muestra de entrenamiento, tiene la desventaja de sólo estimar correctamente el error de predicción esperado, exagerando el acierto real del modelo y disimulando el sesgo de estimación (Curate et al., 2016; Curate et al., 2017a; Curate et al., 2017b). En consecuencia, para obtener una estimación más precisa del error de un modelo es necesario evaluarlo en una muestra diferente de la de entrenamiento. En consideración de esto, los modelos construidos fueron validados utilizando una muestra de prueba (n=70) constituida por individuos que no fueron considerados en la construcción de los modelos.

Para cada modelo, se calculó la sensibilidad, especificidad y precisión general. La sensibilidad evalúa la capacidad de un modelo para clasificar positivamente un registro, mientras que la especificidad la aptitud para clasificarlo negativamente (Larose y Larose, 2015). Arbitrariamente, se definió la sensibilidad como la proporción de individuos masculinos estimados correctamente, mientras que la especificidad como la proporción de femeninos estimados correctamente. La precisión general corresponde a la proporción de individuos correctamente estimados del total de casos, y permite medir el grado de acierto entre el sexo real y el estimado.

Modelos propuestos

Para definir los modelos propuestos para la estimación de sexo se siguió un procedimiento sustractivo en que del total de modelos obtenidos de las regresiones logísticas cada uno fue aprobado o descartado según el cumplimiento de los siguientes parámetros:

- a) Significancia estadística de todos los coeficientes del modelo (p -valor <0.05). Dado que una variable con un coeficiente no significativo no contribuye al poder explicativo del modelo, su incorporación carece de sentido. Los modelos multivariados en que una variable no resultó significativa pero igualmente el modelo reportó un mayor acierto que sus contrapartes univariadas fueron asimismo descartados pues al incluir una variable no significativa resultan poco confiables.
- b) Precisión general igual o superior al 80%. Dado que la probabilidad de acertar al sexo real de un individuo sólo por azar es del 50%, el acierto mínimo aceptable para un modelo debe ser superior a este porcentaje, y en la literatura un 80% es utilizado por convención (du Jardin et al., 2009; Christensen et al., 2014).
- c) Diferencia entre especificidad y sensibilidad menor al 10%. La existencia de una brecha importante en el acierto logrado para ambos sexos es indicativa de que el modelo está sesgado en favor de un sexo, incluso si posee una precisión general aceptable. Si bien no existe convención al respecto y sólo estudios excepcionales han establecido niveles máximos para esta diferencia (Albanese, 2003), se decidió imponer arbitrariamente un 10% de diferencia máxima aceptable como una forma de asegurar un adecuado nivel de confiabilidad de los modelos propuestos.
- d) Eficiencia de los modelos construidos. Si un modelo multivariable reporta un rendimiento idéntico al de un modelo univariable de una de las variables incluidas, pierde sentido su utilización frente a este último, que resulta más eficiente al requerir de menos medidas. Por otro lado, la incorporación de una medida más dentro de un modelo reduce de forma significativa su potencial de aplicación, por lo que debe justificarse en un aumento de la precisión del modelo frente a sus contrapartes univariadas. Por esta razón, los modelos que no resultaron eficientes fueron también descartados.

El manejo de los datos y los cálculos estadísticos fueron realizados con los softwares Microsoft Excel® 2016 y R® versión 3.4.4. Todos los valores numéricos obtenidos fueron redondeados a los tres decimales.

Resultados

Error intraobservador

La Tabla 3 resume los resultados obtenidos para el coeficiente de correlación de concordancia de Lin. Para todas las medidas se obtuvo un coeficiente superior a 0.9 con significancia estadística ($p < 0.05$).

Tabla 3: ρ_c y p-valor para el coeficiente de correlación de concordancia de Lin (n=20)

Variable	ρ_c	p-valor
FNAL	0.997	<0.05
FNW	0.993	<0.05
AB	0.963	<0.05
AC	0.927	<0.05
BC	0.925	<0.05
GTFC	0.999	<0.05
GTLT	0.907	<0.05
LTFC	0.991	<0.05

Asimetría bilateral

La Tabla 4 muestra la estadística descriptiva de cada lado y los resultados obtenidos para las pruebas t de Student pareadas. La prueba arrojó un p-valor mayor a 0.05 en todos los casos, a excepción de la variable GTLT.

Tabla 4: Estadística descriptiva para cada lado y p-valor para pruebas t de Student pareadas (n=20)

Variable	Lado	Promedio	Desviación Estándar	p-valor (t de Student)
FNAL	Izquierdo	89.408	6.509	0.823
	Derecho	89.272	6.484	
FNW	Izquierdo	29.555	3.145	0.362
	Derecho	29.768	3.295	
AB	Izquierdo	26.222	3.982	0.407
	Derecho	26.771	3.754	
AC	Izquierdo	50.174	6.489	0.265
	Derecho	51.711	4.927	
BC	Izquierdo	55.635	6.047	0.471
	Derecho	56.200	6.349	
GTFC	Izquierdo	90.409	6.483	0.143
	Derecho	89.656	6.425	
GTLT	Izquierdo	53.366	5.008	0.022
	Derecho	54.605	4.936	
LTFC	Izquierdo	69.078	6.389	0.538
	Derecho	68.733	6.184	

Estadística descriptiva

La Tabla 5 sintetiza la estadística descriptiva de la muestra de entrenamiento y los resultados obtenidos para las pruebas estadísticas utilizadas. Los promedios de todas las variables son distintos entre sexos, con la media de los individuos masculinos siendo mayor a la media de los individuos femeninos en todas las variables. A su vez, los intervalos de confianza del 95% para cada variable no se superponen entre sexos. Los resultados para la estimación de la asimetría muestran que todas las variables para cada sexo son aproximadamente simétricas, con la mayoría de ellas levemente inclinadas a la izquierda de la curva (asimetría mayor a 0), a excepción de GTLT, con una distribución inclinada a la derecha en ambos sexos (asimetría menor a 0). Los resultados obtenidos para la curtosis indican que todas las variables describen curvas aproximadamente mesocúrticas (valores cercanos a 3), a excepción de FNW para el sexo femenino, que presenta una curva leptocúrtica (superior a 3). Los resultados para la prueba de Shapiro-Wilk indican que todas las variables se distribuyen de forma normal en ambos sexos, con la excepción de FNAL en el sexo masculino y FNW en el sexo femenino ($p < 0.05$). Los resultados obtenidos para los test t de Student y Mann-Whitney-Wilcoxon señalan que en todas las variables analizadas existen diferencias estadísticamente significativas entre ambos sexos ($p < 0.05$). Los gráficos 1 al 8 muestran la distribución de cada variable por sexo, y las figuras 4 a 7 (Anexos) ilustran un fémur proximal característicamente masculino y uno característicamente femenino para las variables analizadas.

Tabla 5: Estadística descriptiva de la muestra de entrenamiento y p-valor para pruebas de Shapiro-Wilk, t de Student y Mann-Whitney-Wilcoxon (n=200)

Variable	Sexo	Promedio	D.E.	I.C. (95%)	Asimetría	Curtosis	S-W	Student/M-W-W
FNAL	Femenino	82.894	4.572	81.986 - 83.801	0.298	2.822	>0.05	<0.05
	Masculino	95.060	4.749	94.118 - 96.003	0.534	2.522	<0.05	
FNW	Femenino	27.311	1.862	26.942 - 27.681	0.286	5.174	<0.05	<0.05
	Masculino	31.789	2.152	31.362 - 32.216	0.055	2.793	>0.05	
AB	Femenino	25.153	3.981	24.363 - 25.943	0.146	2.735	>0.05	<0.05
	Masculino	28.275	4.903	27.302 - 29.248	0.264	3.056	>0.05	
AC	Femenino	46.606	4.339	45.745 - 47.467	0.319	3.609	>0.05	<0.05
	Masculino	51.869	5.671	50.744 - 52.994	0.259	2.730	>0.05	
BC	Femenino	51.462	3.880	50.692 - 52.232	0.494	3.047	>0.05	<0.05
	Masculino	57.330	4.337	56.469 - 58.190	0.236	3.111	>0.05	
GTFC	Femenino	83.588	4.700	82.655 - 84.520	0.162	2.676	>0.05	<0.05
	Masculino	96.268	4.830	95.309 - 97.226	0.345	2.511	>0.05	
GTLT	Femenino	48.592	4.017	47.794 - 49.389	-0.187	2.753	>0.05	<0.05
	Masculino	55.155	3.516	54.457 - 55.853	-0.233	2.777	>0.05	
LTFC	Femenino	63.051	4.627	62.132 - 63.969	0.389	3.505	>0.05	<0.05
	Masculino	70.746	5.456	69.664 - 71.829	0.028	2.902	>0.05	

D.E.: desviación estándar; I.C. (95%): intervalo de confianza del 95%; S-W: prueba de Shapiro-Wilk; Student/M-W-W: prueba t de Student y prueba de Mann-Whitney-Wilcoxon

Gráfico 1: Distribución de FNAL por sexo

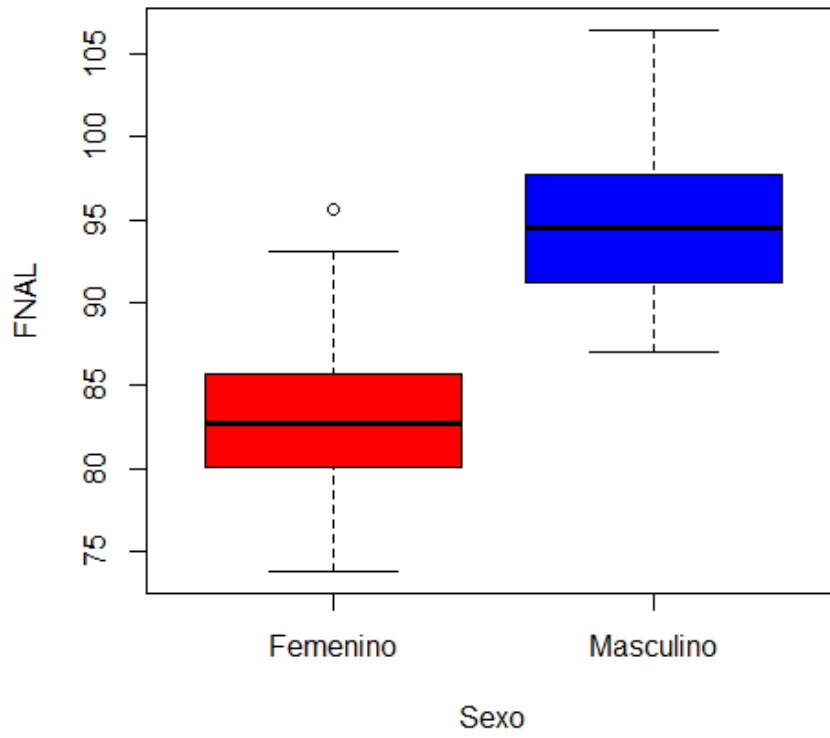


Gráfico 2: Distribución de FNW por sexo

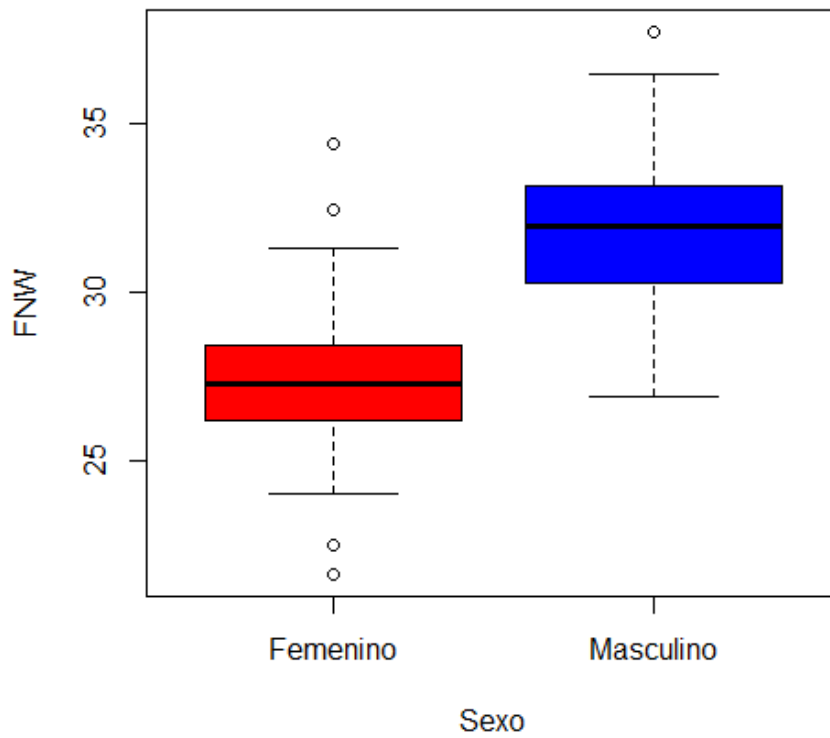


Gráfico 3: Distribución de AB por sexo

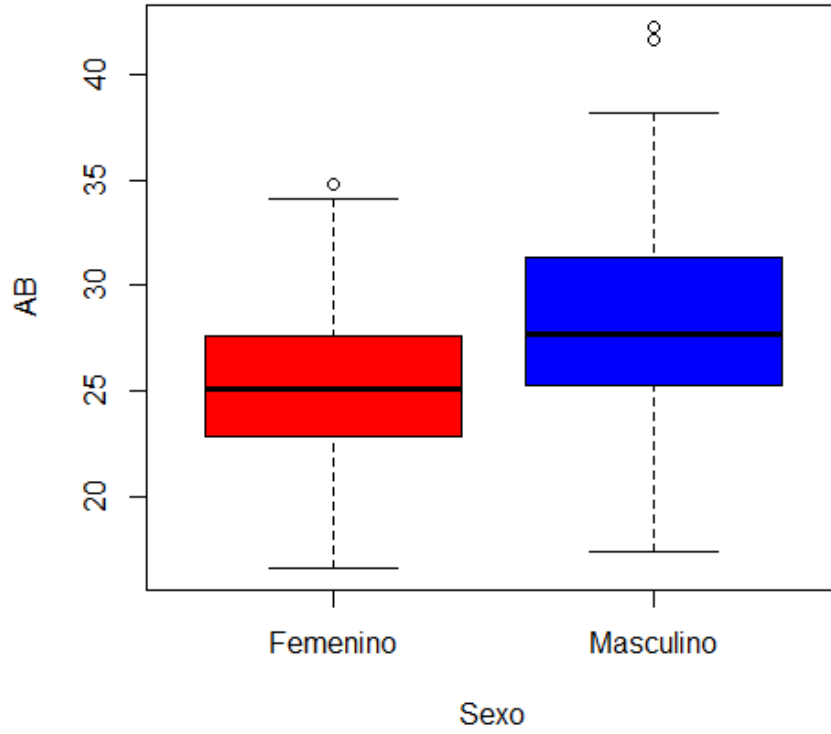


Gráfico 4: Distribución de AC por sexo

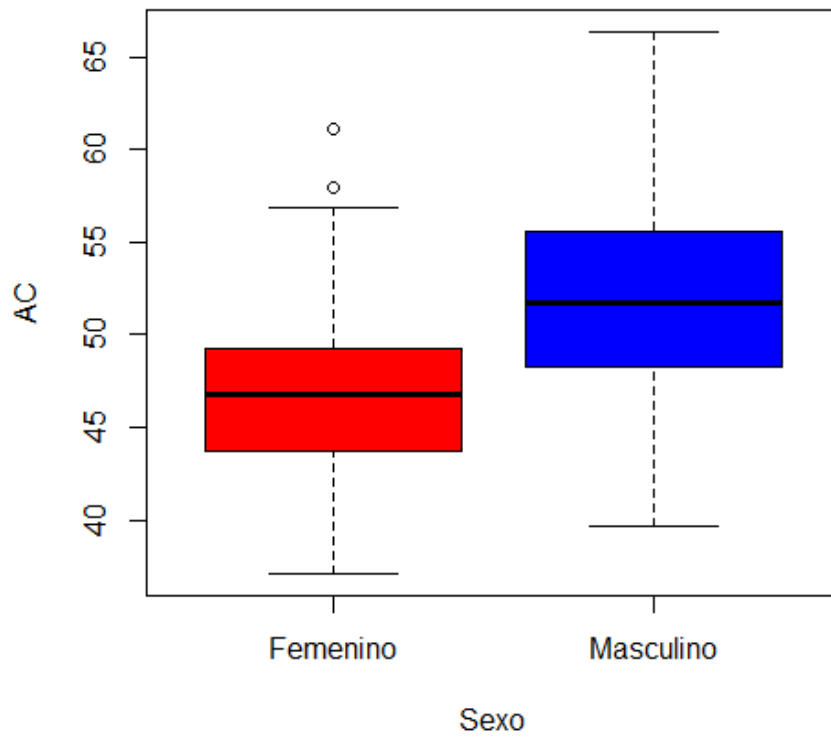


Gráfico 5: Distribución de BC por sexo

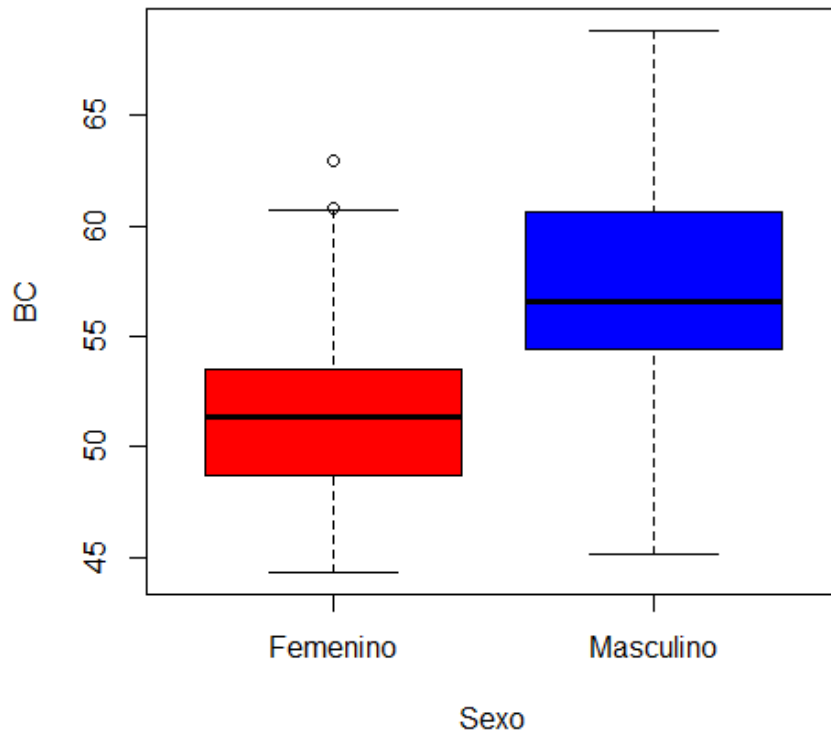


Gráfico 6: Distribución de GTFC por sexo

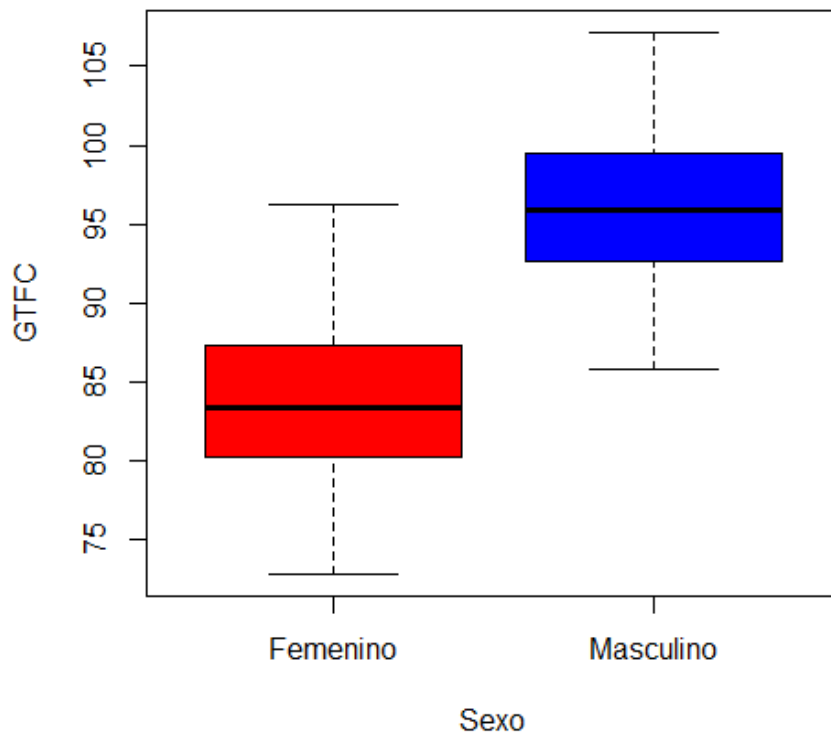


Gráfico 7: Distribución de GTLT por sexo

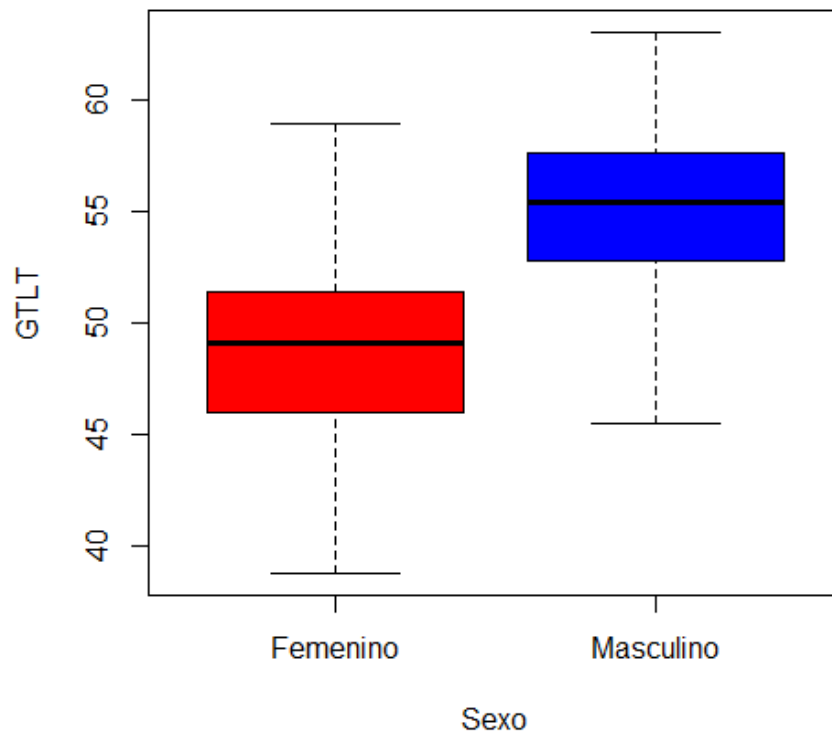
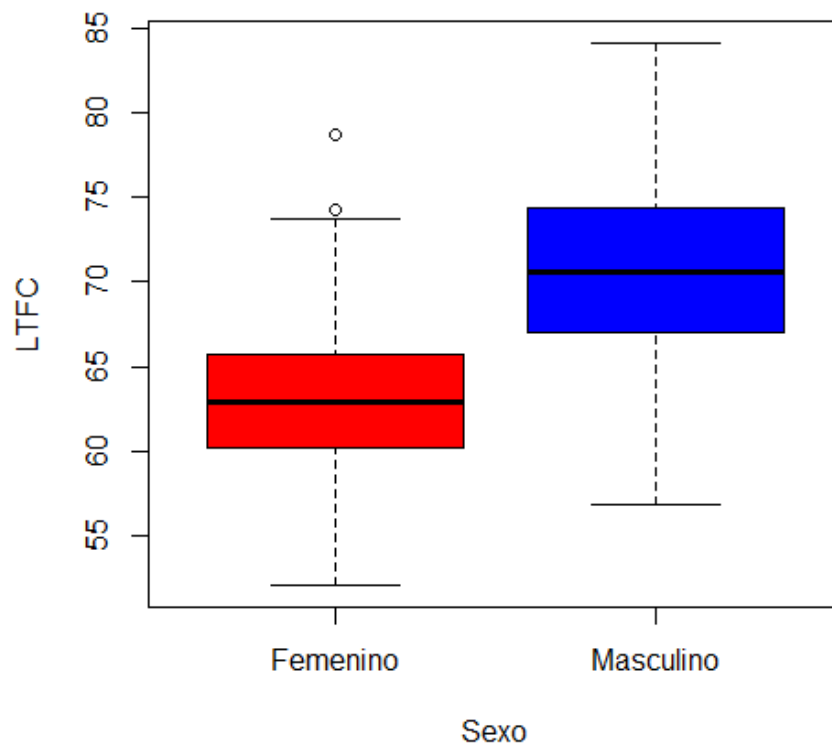


Gráfico 8: Distribución de LTFC por sexo



Modelos estadísticos

La Tabla 6 presenta los resultados obtenidos para las regresiones logísticas de cada modelo con el sexo como variable dependiente. La mayoría de los coeficientes de las variables incluidas en cada modelo resultaron significativos ($p < 0.05$), con la excepción de los coeficientes de AC en el modelo *FNAL+AC*, FNAL en el modelo *FNAL+GTFC*, LTFC en el modelo *FNAL+LTFC*, AB en el modelo *FNW+AB*, AB en el modelo *AB+LTFC*, AC en el modelo *AC+BC*, AC en *AC+GTFC*, AC en *AC+LTFC*, LTFC en *GTFC+LTFC*, AC en *Purkait* y LTFC en *Albanese* ($p > 0.05$). Los valores para pseudo- R^2 muestran una gran dispersión, pero para la mayoría de los modelos es superior a 0.6. Además, se evidencia un aumento general del tamaño del efecto en los modelos multivariantes frente a los univariantes.

Tabla 6: Regresiones logísticas (n=200)

Modelo	Variabes	Coefficientes	Error Estándar	Wald	p-valor	Pseudo-R2
<i>FNAL</i>	Constante	-57.513	9.084	-6.330	<0.0001	0.825
	FNAL	0.646	0.102	6.350	<0.0001	
<i>FNW</i>	Constante	-32.727	4.517	-7.240	<0.0001	0.721
	FNW	1.115	0.155	7.210	<0.0001	
<i>AB</i>	Constante	-4.242	0.965	-4.400	<0.0001	0.147
	AB	0.159	0.036	4.440	<0.0001	
<i>AC</i>	Constante	-10.493	1.798	-5.840	<0.0001	0.290
	AC	0.214	0.037	5.830	<0.0001	
<i>BC</i>	Constante	-19.127	2.804	-6.820	<0.0001	0.450
	BC	0.353	0.052	6.820	<0.0001	
<i>GTFC</i>	Constante	-57.525	9.211	-6.240	<0.0001	0.831
	GTFC	0.639	0.102	6.270	<0.0001	
<i>GTLT</i>	Constante	-24.376	3.385	-7.200	<0.0001	0.571
	GTLT	0.469	0.065	7.220	<0.0001	
<i>LTFC</i>	Constante	-20.255	2.891	-7.010	<0.0001	0.487
	LTFC	0.304	0.043	7.000	<0.0001	
<i>FNAL+FNW</i>	Constante	-61.222	10.118	-6.050	<0.0001	0.855
	FNAL	0.495	0.106	4.680	<0.0001	
	FNW	0.580	0.190	3.050	0.0023	
<i>FNAL+AB</i>	Constante	-59.939	9.650	-6.210	<0.0001	0.843
	FNAL	0.736	0.118	6.260	<0.0001	
	AB	-0.208	0.083	-2.510	0.0120	
<i>FNAL+AC</i>	Constante	-57.519	9.112	-6.310	<0.0001	0.825
	FNAL	0.646	0.104	6.190	<0.0001	
	AC	0.001	0.057	0.010	0.9933	
<i>FNAL+BC</i>	Constante	-62.143	10.127	-6.140	<0.0001	0.837
	FNAL	0.609	0.106	5.740	<0.0001	
	BC	0.144	0.069	2.090	0.0369	
<i>FNAL+GTFC</i>	Constante	-59.406	9.544	-6.220	<0.0001	0.836
	FNAL	0.259	0.205	1.270	0.2057	
	GTFC	0.404	0.202	1.990	0.0461	
<i>FNAL+GTLT</i>	Constante	-71.124	12.335	-5.770	<0.0001	0.860
	FNAL	0.594	0.116	5.120	<0.0001	
	GTLT	0.345	0.103	3.340	0.0008	
<i>FNAL+LTFC</i>	Constante	-57.627	9.189	-6.270	<0.0001	0.826
	FNAL	0.683	0.116	5.910	<0.0001	
	LTFC	-0.047	0.062	-0.750	0.4510	

Modelo	Variables	Coefficientes	Error Estándar	Wald	p-valor	Pseudo-R2
<i>FNW+AB</i>	Constante	-32.923	4.536	-7.260	<0.0001	0.722
	FNW	1.091	0.158	6.910	<0.0001	
	AB	0.034	0.055	0.610	0.5392	
<i>FNW+AC</i>	Constante	-35.963	5.074	-7.090	<0.0001	0.736
	FNW	1.064	0.160	6.670	<0.0001	
	AC	0.098	0.045	2.160	0.0308	
<i>FNW+BC</i>	Constante	-39.301	5.577	-7.050	<0.0001	0.751
	FNW	0.996	0.161	6.210	<0.0001	
	BC	0.186	0.062	3.020	0.0026	
<i>FNW+GTFC</i>	Constante	-64.519	11.100	-5.810	<0.0001	0.866
	FNW	0.606	0.192	3.160	0.0016	
	GTFC	0.518	0.110	4.710	<0.0001	
<i>FNW+GTLT</i>	Constante	-40.437	5.770	-7.010	<0.0001	0.763
	FNW	0.887	0.162	5.490	<0.0001	
	GTLT	0.277	0.081	3.420	0.0006	
<i>FNW+LTFC</i>	Constante	-38.981	5.469	-7.130	<0.0001	0.757
	FNW	0.962	0.161	5.980	<0.0001	
	LTFC	0.161	0.051	3.180	0.0015	
<i>AB+AC</i>	Constante	-13.995	2.230	-6.270	<0.0001	0.363
	AB	0.145	0.041	3.580	0.0003	
	AC	0.207	0.038	5.400	<0.0001	
<i>AB+BC</i>	Constante	-20.233	2.936	-6.890	<0.0001	0.471
	AB	0.089	0.042	2.130	0.0329	
	BC	0.329	0.053	6.240	<0.0001	
<i>AB+GTFC</i>	Constante	-58.458	9.359	-6.250	<0.0001	0.845
	AB	-0.173	0.080	-2.170	0.0302	
	GTFC	0.700	0.111	6.330	<0.0001	
<i>AB+GTLT</i>	Constante	-28.009	3.930	-7.130	<0.0001	0.605
	AB	0.136	0.048	2.860	0.0042	
	GTLT	0.467	0.067	6.930	<0.0001	
<i>AB+LTFC</i>	Constante	-20.395	2.926	-6.970	<0.0001	0.489
	AB	0.023	0.046	0.500	0.6196	
	LTFC	0.297	0.046	6.510	<0.0001	
<i>AC+BC</i>	Constante	-19.748	2.869	-6.880	<0.0001	0.460
	AC	0.065	0.045	1.430	0.1518	
	BC	0.306	0.060	5.140	<0.0001	
<i>AC+GTFC</i>	Constante	-58.804	9.545	-6.160	<0.0001	0.833
	AC	0.049	0.056	0.870	0.3844	
	GTFC	0.626	0.103	6.060	<0.0001	
<i>AC+GTLT</i>	Constante	-31.291	4.365	-7.170	<0.0001	0.635
	AC	0.165	0.044	3.770	0.0002	
	GTLT	0.445	0.069	6.430	<0.0001	
<i>AC+LTFC</i>	Constante	-20.370	2.916	-6.990	<0.0001	0.488
	AC	0.019	0.050	0.370	0.7102	
	LTFC	0.292	0.053	5.460	<0.0001	
<i>BC+GTFC</i>	Constante	-64.325	10.892	-5.910	<0.0001	0.846
	BC	0.157	0.069	2.280	0.0224	
	GTFC	0.618	0.110	5.630	<0.0001	
<i>BC+GTLT</i>	Constante	-28.712	3.908	-7.350	<0.0001	0.612
	BC	0.176	0.057	3.090	0.0020	
	GTLT	0.368	0.070	5.250	<0.0001	
<i>BC+LTFC</i>	Constante	-27.196	3.742	-7.270	<0.0001	0.572
	BC	0.230	0.056	4.140	<0.0001	
	LTFC	0.221	0.047	4.730	<0.0001	

Modelo	Variables	Coefficientes	Error Estándar	Wald	p-valor	Pseudo-R2
<i>GTFC+GTLT</i>	Constante	-66.268	10.903	-6.080	<0.0001	0.858
	GTFC	0.562	0.107	5.260	<0.0001	
	GTLT	0.297	0.100	2.970	0.0030	
<i>GTFC+LTFC</i>	Constante	-57.474	9.226	-6.230	<0.0001	0.832
	GTFC	0.649	0.112	5.810	<0.0001	
	LTFC	-0.014	0.062	-0.220	0.8239	
<i>GTLT+LTFC</i>	Constante	-38.731	5.392	-7.180	<0.0001	0.702
	GTLT	0.429	0.074	5.790	<0.0001	
	LTFC	0.244	0.048	5.140	<0.0001	
<i>Purkait</i>	Constante	-20.868	2.996	-6.970	<0.0001	0.480
	AB	0.091	0.042	2.150	0.0319	
	AC	0.068	0.047	1.460	0.1434	
	BC	0.278	0.062	4.510	<0.0001	
<i>Albanese</i>	Constante	-66.060	10.899	-6.060	<0.0001	0.858
	GTFC	0.571	0.115	4.960	<0.0001	
	GTLT	0.297	0.100	2.960	0.0031	
	LTFC	-0.015	0.069	-0.220	0.8233	

Validación

La Tabla 7 muestra los resultados obtenidos para la validación de cada uno de los modelos construidos. La mayoría de los modelos reportaron una precisión general mayor al 80%, a excepción de los modelos *AB*, *AC*, *BC*, *FNW+GTLT*, *AB+AC* y *AC+BC*, con niveles entre un 74.3% y un 78.6%. La sensibilidad resultó igual o mayor a la especificidad en casi todos los modelos, con la excepción de los modelos *AB*, *FNW+AB*, *FNW+BC* y *FNW+GTLT*, en donde la relación fue inversa. Destacan entre los resultados los rendimientos obtenidos para los modelos *FNAL*, *GTFC*, *FNAL+FNW*, *FNAL+AB*, *FNAL+AC*, *FNAL+BC*, *FNAL+GTFC*, *FNAL+GTLT*, *FNAL+LTFC*, *FNW+BC*, *FNW+GTFC*, *FNW+LTFC*, *AB+GTFC*, *AC+GTFC*, *BC+GTFC*, *BC+GTLT*, *GTFC+GTLT*, *GTFC+LTFC* y *Albanese*, con niveles de sensibilidad entre un 88.6% y un 100%, una especificidad entre un 77.1% y un 91.4% y porcentajes de precisión general desde un 90% hasta un 95.7%

Tabla 7: Validación de los modelos predictivos de sexo (n=70)

Modelo	Sensibilidad	Especificidad	Precisión General
<i>FNAL</i>	0.971	0.886	0.929
<i>FNW</i>	0.857	0.857	0.857
<i>AB</i>	0.657	0.857	0.757
<i>AC</i>	0.771	0.771	0.771
<i>BC</i>	0.800	0.743	0.771
<i>GTFC</i>	0.971	0.886	0.929
<i>GTLT</i>	0.829	0.800	0.814
<i>LTFC</i>	0.857	0.800	0.829
<i>FNAL+FNW</i>	0.943	0.914	0.929
<i>FNAL+AB</i>	0.971	0.886	0.929
<i>FNAL+AC</i>	0.971	0.886	0.929
<i>FNAL+BC</i>	0.943	0.914	0.929
<i>FNAL+GTFC</i>	0.971	0.886	0.929
<i>FNAL+GTLT</i>	0.943	0.914	0.929
<i>FNAL+LTFC</i>	0.971	0.886	0.929
<i>FNW+AB</i>	0.857	0.886	0.871

<i>Modelo</i>	Sensibilidad	Especificidad	Precisión General
<i>FNW+AC</i>	0.886	0.886	0.886
<i>FNW+BC</i>	0.886	0.914	0.900
<i>FNW+GTFC</i>	1.000	0.914	0.957
<i>FNW+GTLT</i>	0.514	0.971	0.743
<i>FNW+LTFC</i>	0.914	0.886	0.900
<i>AB+AC</i>	0.800	0.771	0.786
<i>AB+BC</i>	0.857	0.800	0.829
<i>AB+GTFC</i>	0.971	0.914	0.943
<i>AB+GTLT</i>	0.914	0.857	0.886
<i>AB+LTFC</i>	0.857	0.800	0.829
<i>AC+BC</i>	0.829	0.743	0.786
<i>AC+GTFC</i>	0.971	0.914	0.943
<i>AC+GTLT</i>	0.857	0.800	0.829
<i>AC+LTFC</i>	0.829	0.800	0.814
<i>BC+GTFC</i>	0.971	0.914	0.943
<i>BC+GTLT</i>	0.914	0.771	0.843
<i>BC+LTFC</i>	0.857	0.800	0.829
<i>GTFC+GTLT</i>	0.971	0.914	0.943
<i>GTFC+LTFC</i>	0.971	0.886	0.929
<i>GTLT+LTFC</i>	0.914	0.829	0.871
<i>Purkait</i>	0.857	0.800	0.829
<i>Albanese</i>	0.971	0.914	0.943

Modelos Propuestos

Siguiendo el proceso sustractivo detallado en la sección de Métodos, algunos modelos obtenidos como resultado de las regresiones logísticas fueron descartados por no cumplir con los parámetros establecidos como aceptables. En algunos casos, un modelo falló en cumplir con varios de estos requisitos. La Tabla 8 resume los 19 modelos propuestos para la estimación de sexo luego del proceso sustractivo de descarte.

En la mayoría de las regresiones logísticas las variables incluidas resultaron significativas (p -valor de Wald < 0.05). Sin embargo, esto no ocurrió para las variables AC en el modelo *FNAL+AC*, FNAL en *FNAL+GTFC*, LTFC en *FNAL+LTFC*, AB en *FNW+AB*, AB modelo *AB+LTFC*, AC en *AC+BC*, AC en *AC+GTFC*, AC en *AC+LTFC*, LTFC en *GTFC+LTFC*, AC en *Purkait* y LTFC en *Albanese*, que no aportan de forma estadísticamente significativa al poder explicativo de sus respectivos modelos, por lo que fueron descartados. Los modelos *FNW+AB* y *AC+GTFC* arrojaron resultados conflictivos: si bien incluyen una variable no significativa, sí reportaron una mejor precisión general que sus contrapartes univariadas gracias a una más alta especificidad. Aunque estos hallazgos pueden parecer contradictorios, los modelos mencionados fueron igualmente descartados en conformidad a lo establecido en la sección Métodos puesto que la sola presencia de una variable no significativa resta confiabilidad a los modelos.

En cuanto a la precisión general, se obtuvo un porcentaje de acierto mayor al 80% para la mayoría de los modelos calculados. Sin embargo, los modelos *AB*, *AC*, *BC*, *FNW+GTLT*, *AB+AC* y *AC+BC* reportaron una precisión general menor al 80% establecido como aceptable, por lo que fueron excluidos. Por otro lado, la diferencia entre el porcentaje de individuos femeninos y masculinos correctamente estimados no superó el 10% establecido salvo en los modelos *AB*, *FNW+GTLT* y *BC+GTLT*, siendo también causal de descarte.

Por otra parte, en los modelos *FNAL+AB*, *FNAL+AC*, *FNAL+GTFC*, *FNAL+LTFC*, *AB+LTFC*, *AC+LTFC*, *BC+LTFC*, *GTFC+LTFC*, *Purkait* y *Albanese* los porcentajes de acierto no mejoraron frente a sus contrapartes con menos variables, razón por la cual también fueron descartados. El modelo *FNW+GTLT*, en particular, no sólo no mejora el acierto de los modelos *FNW* y *GTLT* sino que reduce notablemente su rendimiento frente a estos últimos.

En consideración de que los dos modelos de tres variables calculados, *Purkait* y *Albanese*, no lograron combinar tres variables significativas dentro de un mismo modelo ni mejorar el acierto al sexo de los individuos en comparación a modelos con menos variables, se decidió a posteriori desestimar la elaboración de otros modelos de más de dos variables pues no se espera obtener una combinación de variables significativas ni mejores rendimientos frente a modelos con menos variables.

Tabla 8: Modelos propuestos para la estimación de sexo en población chilena moderna a partir del fémur proximal

Modelo	Fórmula	Sensibilidad	Especificidad	Precisión General
<i>FNAL</i>	$-57.513 + 0.646*FNAL$	97.1%	88.6%	92.9%
<i>FNW</i>	$-32.727 + 1.115*FNW$	85.7%	85.7%	85.7%
<i>GTFC</i>	$-57.525 + 0.639*GTFC$	97.1%	88.6%	92.9%
<i>GTLT</i>	$-24.376 + 0.469*GTLT$	82.9%	80.0%	81.4%
<i>LTFC</i>	$20.255 + 0.304*LTFC$	85.7%	80.0%	82.9%
<i>FNAL+FNW</i>	$-61.222 + 0.495*FNAL + 0.580*FNW$	94.3%	91.4%	92.9%
<i>FNAL+BC</i>	$-62.143 + 0.609*FNAL + 0.144*BC$	94.3%	91.4%	92.9%
<i>FNAL+GTLT</i>	$-71.124 + 0.594*FNAL + 0.345*GTLT$	94.3%	91.4%	92.9%
<i>FNW+AC</i>	$-35.963 + 1.064*FNW + 0.098*AC$	88.6%	88.6%	88.6%
<i>FNW+BC</i>	$-39.301 + 0.996*FNW + 0.186*BC$	88.6%	91.4%	90.0%
<i>FNW+GTFC</i>	$-64.519 + 0.606*FNW + 0.518*GTFC$	100%	91.4%	95.7%
<i>FNW+LTFC</i>	$-38.981 + 0.962*FNW + 0.161*LTFC$	91.4%	88.6%	90.0%
<i>AB+BC</i>	$-20.233 + 0.089*AB + 0.329*BC$	85.7%	80.0%	82.9%
<i>AB+GTFC</i>	$-58.458 + -0.173*AB + 0.700*GTFC$	97.1%	91.4%	94.3%
<i>AB+GTLT</i>	$-28.009 + 0.136*AB + 0.467*GTLT$	91.4%	85.7%	88.6%
<i>AC+GTLT</i>	$-31.291 + 0.165*AC + 0.445*GTLT$	85.7%	80.0%	82.9%
<i>BC+GTFC</i>	$-64.325 + 0.157*BC + 0.618*GTFC$	97.1%	91.4%	94.3%
<i>GTFC+GTLT</i>	$-66.268 + 0.562*GTFC + 0.297*GTLT$	97.1%	91.4%	94.3%
<i>GTLT+LTFC</i>	$-38.731 + 0.429*GTLT + 0.244*LTFC$	91.4%	82.9%	87.1%

Discusión

En el presente estudio se desarrollaron modelos de estimación de sexo en población chilena moderna mediante el uso de regresión logística a partir de medidas osteométricas ubicadas en el fémur proximal.

Error Intraobservador

La reproductibilidad de una metodología es importante, pues asegura que otros investigadores puedan replicar el método de forma precisa. En este sentido, los resultados obtenidos para la estimación del error intraobservador indican que las medidas analizadas no poseen un error de medición estadísticamente significativo, siendo por tanto reproducibles. Se puede concluir del bajo error obtenido que las medidas están claramente definidas y son fáciles de evaluar.

Asimetría Bilateral

La existencia de diferencias significativas entre ambos lados del cuerpo puede dificultar la aplicación de un método en contextos en que sólo un miembro está disponible. Los resultados obtenidos muestran que no existe una asimetría bilateral estadísticamente significativa para las variables evaluadas con la excepción de la variable GTLT. Se concluye entonces que ambos miembros son mayoritariamente simétricos.

En el caso de GTLT, en base a las medias aritméticas obtenidas para cada lado se infiere que el lado derecho es en promedio mayor que el izquierdo. Una explicación certera a este fenómeno resulta esquiva, particularmente considerando que es la única variable que evidenció asimetría bilateral, pero dado que no se encontraron niveles de error intraobservador significativos se puede concluir que esta diferencia es real y no representa un error de medición. Tomando en cuenta que los modelos predictivos fueron calculados en base a fémures izquierdos, la utilización de las fórmulas que incluyen la variable GTLT en el miembro derecho no es recomendable.

Estadística descriptiva e inferencial

Al comparar los promedios e intervalos de confianza para cada sexo en la muestra de entrenamiento se evidencia que todas las variables son mayores en los individuos masculinos en comparación con los femeninos. Esto concuerda con los resultados obtenidos para las pruebas estadísticas realizadas que confirman la existencia de diferencias estadísticamente significativas entre ambos sexos para cada variable. Se concluye entonces que todas las variables analizadas son sexualmente dimórficas, con mayores magnitudes en el sexo masculino respecto al femenino.

Modelos propuestos para la estimación de sexo en población chilena moderna

Se realizaron regresiones logísticas para todas las medidas - individualmente y en combinaciones - como variables independientes y el sexo como variable dependiente. De estas regresiones se obtuvieron modelos predictivos del sexo que luego fueron sometidos a validación en una muestra de prueba. Siguiendo un proceso sustractivo, varios modelos fueron descartados por no cumplir con los parámetros establecidos como aceptables. Habiendo excluido éstos, se propusieron 19 modelos para la estimación de sexo en población chilena moderna a partir del fémur proximal (Tabla 8, Resultados).

Para ilustrar la aplicación de estos modelos, se tomarán como ejemplos los modelos *FNAL* para los univariados y *GTFC* para los multivariados.

De la regresión logística para la variable *FNAL* se obtuvo una constante de -57.513 y un coeficiente para la variable de 0.646. Incorporando estos datos en la fórmula lineal se obtiene el modelo *FNAL*, con la forma:

$$\text{Modelo } FNAL = -57.513 + 0.646 * FNAL$$

De esta manera, en el caso de un individuo con un valor para *FNAL* de 100.34, al resolver la ecuación se obtiene 7.307, que clasifica al individuo como masculino. Esta clasificación, según los resultados obtenidos para la validación, será correcta el 92.9% de las veces.

En el caso del modelo *GTFC+FNW*, los resultados de la regresión logística para ambas variables combinadas arrojaron una constante de -64.519, un coeficiente de 0.518 para *GTFC* y un coeficiente de 0.606 para *FNW*. Reemplazando, se tiene:

$$\text{Modelo } GTFC+FNW = -64.519 + 0.518 * GTFC + 0.606 * FNW$$

Con este modelo, para un individuo con valores de 86.14 para *GTFC* y 29.73 para *FNW* se obtiene -1.882 al resolver la fórmula, valor que clasifica al individuo como femenino. Según lo reportado en la validación, esta estimación será correcta el 95.7% de las veces.

En concordancia con lo expuesto en la sección Métodos, se proponen modelos tanto univariados como multivariados. Los modelos multivariados permiten recoger más dimorfismo sexual y en consecuencia reportaron un mayor tamaño del efecto y un mejor acierto, mientras que los modelos univariados mostraron en general una menor precisión y un mayor sesgo por sexo. No obstante, estos últimos son particularmente útiles en contextos de alta fragmentación en que no puede obtenerse más de una medida (Kranioti et al., 2009; Boldsen et al., 2015).

Es importante que un modelo tenga una sensibilidad y especificidad similares pues porcentajes de acierto desiguales para individuos masculinos y femeninos es indicativo de que un modelo está sesgado en favor de uno u otro sexo. En este sentido, todos los modelos propuestos poseen un bajo sesgo por sexo, que no supera el 8.6% de diferencia en el acierto entre ambos sexos. Cabe mencionar que los modelos *FNAL+FNW*, *FNAL+BC* y *FNAL+GTLT* obtuvieron la misma precisión general que el modelo *FNAL*, pero la incorporación de las demás variables se tradujo en una reducción del sesgo por sexo al disminuir la sensibilidad pero aumentar la especificidad, razón por la cual se decidió conservarlos.

Salvo *FNW+BC*, todos los modelos propuestos alcanzaron mayores porcentajes de acierto en individuos masculinos frente a los femeninos, patrón evidenciado también en los modelos descartados y comúnmente reportado en la literatura (Curate et al., 2016). En algunos casos es posible que esta diferencia se deba a una representación desigual de los sexos en la muestra de entrenamiento empleada para elaborar el modelo (Kranioti et al., 2009), pero esto queda descartado en el presente estudio pues se procuró una representación equitativa de individuos femeninos y masculinos en la muestra. En cambio, la causa más probable a este fenómeno es la expresión misma del dimorfismo sexual evaluado: desde un punto de vista lineal, existen más individuos femeninos de apariencia “masculina” que individuos masculinos de apariencia “femenina” (du Jardin et al., 2009).

En cuanto a la precisión general, en los modelos univariantes se obtuvo un acierto superior al 80% para *FNW*, *GTLT* y *LTFC* llegando hasta un 85.7% para *FNW*, mientras que *FNAL* y *GTFC* tuvieron un mismo acierto del 92.9%. Comparados con el 90.5% reportado por Garrido-Varas et al. (2014) para la altura de la cavidad glenoidea de la escápula (EACG), las variables *FNAL* y *GTFC* son los mejores estimadores individuales de sexo descritos hasta el momento para población chilena moderna. Además, estos modelos, al incluir una sola variable, pueden aplicarse incluso en restos altamente fragmentados, algo especialmente cierto para la variable *FNW* que tiene una notable preservación (Curate et al., 2016). En los modelos multivariantes, por otro lado, la combinación de más de una medida dentro un mismo modelo significó un aumento generalizado del acierto, con la mayoría de los modelos obteniendo una precisión general superior al 90%. Destacan los modelos *AB+GTFC*, *BC+GTFC* y *GTFC+GTLT*, los tres con rendimientos idénticos que alcanzan un 94.3% de precisión general y una excelente sensibilidad y especificidad, y el modelo *FNW+GTFC*, con un 91.4% de precisión para los individuos femeninos, 100% para individuos masculinos y 95.7% de precisión general. Estos resultados convierten al modelo multivariable *FNW+GTFC* en el método de estimación de sexo en post-cráneo no-pélvico más preciso reportado a la fecha para población chilena moderna.

Los resultados obtenidos en el presente estudio concuerdan con la extensa bibliografía que establece que los huesos largos, y en particular el fémur, son muy buenos estimadores de sexo, y respaldan la postura de que este elemento debe ser considerado no como una alternativa a la pelvis y el cráneo si no como una evidencia de primera línea, de igual importancia que la pelvis y preferible por sobre el cráneo (Albanese, 2003; Spradley y Jantz, 2011). El acierto logrado para los modelos propuestos puede explicarse por dos factores principales: la naturaleza del dimorfismo sexual de las variables analizadas y el empleo de una aproximación estadística robusta como la regresión logística.

Las fuerzas selectivas combinadas de la bipedestación y la reproducción en la pelvis son visibles en el ángulo y el largo del cuello femoral (Albanese et al., 2008). Dado que FNAL y GTFC son ambas medidas del largo del cuello femoral, resulta lógico que los resultados obtenidos las establezcan como los mejores estimadores del sexo de los individuos, tanto individualmente como junto a otras variables. En este sentido, dado que ambas constituyen aproximaciones métricas similares para una misma dimensión anatómica, sus valores suelen ser muy cercanos. FNAL y GTFC entregan información redundante, como puede observarse en el modelo *FNAL+GTFC*, que reportó la misma precisión general, sensibilidad y especificidad que los modelos univariantes *FNAL* y *GTFC*. Un cálculo posterior del coeficiente de correlación de Spearman para ambas variables arrojó una alta correlación estadística ($\rho = 0.983$, $p\text{-valor} < 0.05$), confirmando la redundancia de los datos.

No obstante, ambas variables no son iguales. Como se señala en la Tabla 2 (Métodos) y se ilustra en las Figuras 1 y 3 (Anexos), cada medida representa la distancia lineal entre pares de puntos diferentes, localizados sobre la misma estructura – el eje del cuello femoral. FNAL representa la distancia entre la base del gran trocánter y el ápex de la cabeza, mientras que GTFC mide la distancia entre el punto más lateral del gran trocánter y el punto más superior de la fóvea capitis. Los puntos considerados en FNAL son más bien áreas que puntos específicos, mientras que los puntos incluidos en GTFC son hitos claramente delimitados en la anatomía del fémur proximal. Esto implica que para medir FNAL se requiere de una observación algo más cuidadosa versus GTFC, cuya evaluación suele ser inmediata. Si bien esta diferencia no afecta los niveles de reproductibilidad de las variables (como lo evidencian los resultados obtenidos para el error intraobservador), sí indica que la medición de FNAL posee una curva de aprendizaje un poco más pronunciada que GTFC, como fue evidenciado durante la realización de la presente investigación y resulta relevante de señalar para su evaluación por parte de otros investigadores. Por otro lado, precisamente debido a que GTFC considera dos puntos muy delimitados, es más frecuente su afectación tafonómica, mientras que los puntos a evaluar en FNAL presentan una mejor conservación, al tratarse de áreas más robustas y mejor protegidas. Esta diferencia tiene implicancias prácticas: FNAL es una variable mejor preservada que GTFC.

Finalmente, GTFC resulta ser una aproximación ligeramente más precisa a la estimación de sexo que FNAL, lo que puede evidenciarse al comparar el tamaño del efecto de ambos modelos univariantes, con un pseudo- R^2 levemente superior en *GTFC* que en *FNAL*. Si bien esta diferencia en términos del tamaño del efecto no se vio reflejada en los rendimientos (que para ambos modelos fue idéntico), es posible que al emplear una muestra significativamente más grande se observe una mayor precisión para el modelo *GTFC* frente

a *FNAL*. Esto concuerda también con los coeficientes obtenidos para cada variable dentro del modelo combinado *FNAL+GTFC*, donde *GTFC* resultó significativa pero *FNAL* no, sugiriendo que *GTFC* recoge más dimorfismo sexual del largo del cuello femoral que *FNAL*, que sólo entrega información redundante.

La comparación entre ambas variables puede sintetizarse de la siguiente manera: si bien ambas medidas son similares puesto que miden una misma dimensión, *FNAL* constituye una medición más factible de ser tomada, mientras que *GTFC* resulta ser levemente mejor estimador de sexo y además es más fácil de medir. Se recomienda, entonces, la utilización de *GTFC* por sobre *FNAL* cuando sea posible, aunque en ambos casos los resultados son muy buenos y la diferencia entre ambos rendimientos es mínima. Finalmente, y pese a que resulte en gran parte redundante, se aconseja la evaluación de ambas variables si están disponibles, en concordancia con las recomendaciones comunes para una mejor praxis en Ciencias Forenses (Christensen et al., 2014; Barbaro, 2018)

En cuanto a *FNW*, los resultados obtenidos se explican principalmente por la diferencia de tamaño corporal entre ambos sexos; la mayor masa corporal de los individuos masculinos requiere un mayor grosor del cuello femoral (Mitra et al., 2014). Esto podría explicar también por qué no se lograron niveles tan altos de acierto como con *FNAL* o *GTFC*, pues mientras estas variables reflejan de forma más directa el dimorfismo sexual pélvico, *FNW* está más relacionado con el peso corporal, parámetro en el que otros factores, y no sólo el sexo, tienen una importante influencia (Anastopoulou et al., 2014).

Los resultados obtenidos indican que las variables que estiman dimensiones del cuello femoral, como lo son el largo y el ancho, son los mejores estimadores individuales de sexo. En concordancia, el modelo que reportó mayor precisión fue *FNW+GTFC*, que combina las medidas de ancho y largo del cuello en una sola ecuación, recogiendo el máximo posible de dimorfismo sexual del fémur proximal.

AB, *AC*, *BC*, *GTLT* y *LTFC*, por otra parte, fueron los estimadores con menores rendimientos, con las variables del triángulo de Purkait (2005) no alcanzando siquiera niveles aceptables de acierto (80%) para los modelos univariados y la mayoría de sus combinaciones. *GTLT* y *LTFC*, a diferencia de *FNAL*, *GTFC* y *FNW*, no constituyen una aproximación métrica a un parámetro biomecánico específico del fémur como lo son el ancho y el largo del cuello femoral. En cambio, son medidas que simplemente buscan representar la morfología general de la epífisis proximal del hueso mediante un triángulo (Albanese et al., 2008). En consecuencia, solo se aproximan de forma indirecta al dimorfismo sexual de la región, lo que explica los menores porcentajes de acierto para la estimación de sexo. El caso quizá más complejo de explicar es el de las variables del triángulo de Purkait (2005). Los puntos B y C están posicionados sobre estructuras del fémur proximal, como lo son los trocánteres mayor y menor, que alojan inserciones musculares involucradas en la transmisión de la masa corporal, mientras que el punto A se ubica sobre el margen articular de la cabeza y cumple una función más bien referencial para el triángulo conformado por las distancias entre estos puntos. Dado que el peso corporal es sexualmente dimórfico y a que la distribución de éste hacia el fémur es también dimórfica dada una morfología pélvica diferencial entre ambos sexos, es esperable que las variables

AB, AC y BC sean a su vez dimórficas (Purkait, 2005; Anastopoulou et al., 2014; Djorojevic et al., 2016). Si bien se pudo evidenciar un dimorfismo sexual estadísticamente significativo para las tres variables, éste no resultó lo suficientemente marcado como para estimar el sexo de los individuos de la muestra con un acierto aceptable, ni en los modelos univariantes ni en la mayoría de las combinaciones posibles para las variables. Sólo se obtuvo un acierto sobre el 80% con los modelos *AB+AC* y *Purkait* - descartado en favor del primero, más eficiente - y con modelos que incorporaban las variables de este triángulo junto con otras variables más dimórficas.

Los bajos rendimientos obtenidos para las variables de Purkait (2005) pueden explicarse, por una parte, por una importante variabilidad poblacional en el dimorfismo de estas medidas, evidenciada previamente por otros autores que han reevaluado las variables. En este sentido, AB, AC y BC serían menos dimórficas en población chilena moderna que en población de la India Central, estadounidense, griega y española (Purkait, 2005; Brown et al., 2007; Anastopoulou et al., 2014; Djorojevic et al., 2016). Pero, por otra parte, a partir de la evaluación de estas variables en el material se pudo observar que la relación entre los valores obtenidos para las medidas y el sexo de los individuos es contraproducente: en teoría, puesto que el fémur proximal masculino es más grande que el femenino, las distancias entre los puntos aumentan en los individuos masculinos frente a los femeninos pero, por otro lado, los individuos masculinos suelen tener trocánteres con inserciones más prominentes, lo que tiende a acercar los puntos el uno hacia el otro, acortando las distancias entre ellos. Este fenómeno podría contribuir a los bajos niveles de acierto reportados por las variables propuestas por Purkait (2005).

Otro factor que puede explicar los resultados obtenidos para los modelos construidos es la aproximación estadística empleada. Como se expuso en la sección Métodos, en el presente estudio se utilizó regresión logística dado que permite calcular modelos robustos exigiendo pocos requisitos a los datos, algo particularmente cierto frente al análisis de función discriminante. En este sentido, comparaciones entre ambas aproximaciones en el fémur han confirmado que mejores resultados se pueden obtener mediante regresión logística (du Jardin et al., 2009; Meeusen et al., 2015; Curate et al., 2017a). La explicación a este hecho radica en que la regresión logística no asume que la relación entre las variables independientes y la variable dependiente es lineal, a diferencia del análisis de función discriminante que sí asume linealidad (Kleinbaum y Klein, 2010; Hosmer et al., 2013; Hyeoun-Ae, 2013). Esta diferencia es fundamental puesto que la evidencia sugiere que la relación entre las variables osteométricas y el sexo de un individuo no es lineal (du Jardin et al., 2009). Esto se traduce en que los modelos obtenidos mediante regresión logística estiman de forma más precisa la relación entre las variables consideradas, alcanzando de esta manera un mayor acierto al estimar sexo.

Pese a los beneficios que tiene el uso de la regresión logística por sobre el análisis de función discriminante, este último sigue siendo la aproximación más utilizada en la literatura, con una subrepresentación de la regresión logística (Albanese, 2003). Considerando lo expuesto anteriormente, los resultados aquí reportados abogan por una mayor utilización de esta aproximación estadística para la elaboración de modelos predictivos para la estimación de sexo en restos esqueléticos.

El presente estudio no está exento de limitaciones. La expresión del dimorfismo sexual es específica para cada población puesto que está determinada por la interacción entre la composición genética y las condiciones ambientales propias de cada grupo humano (Frutos, 2003; Klepinger, 2006; Ubelaker en Schmitt et al., 2006; Pickering y Bachman, 2009; Dirkmaat, 2012; Abarca 2013; Iscan y Stein, 2013; Christensen et al., 2014; Clavero et al., 2015; Longe, 2015; Curate et al., 2016; Guyomarc'h et al., 2016). Los modelos propuestos en este estudio fueron elaborados para tener aplicabilidad en casos forenses subactuales en población chilena, como causas de violaciones a los Derechos Humanos cometidas durante la Dictadura Cívico-Militar, en base a la contemporaneidad y representatividad de la muestra utilizada (Colección Subactual de Santiago) respecto a las víctimas de desaparición forzada. Por esta razón, no deben ser considerados válidos para otras poblaciones y su aplicación debe limitarse a población chilena moderna. Esto incluye no sólo a poblaciones distantes sino que también a otras poblaciones de Latinoamérica, puesto que se ha evidenciado que la expresión del dimorfismo sexual difiere significativamente entre poblaciones latinoamericanas pese a que tradicionalmente han sido agrupadas como si de un solo grupo homogéneo se tratase (Bright, Peckmann y Meek, 2018).

Los modelos aquí propuestos tampoco pueden ser considerados válidos para población chilena actual, puesto que la población representada por la Colección Subactual de Santiago – población de Santiago de estrato socioeconómico bajo fallecida entre 1950 y 1973 – no evidenciaría el proceso de cambio secular presente en la población chilena contemporánea, al menos a nivel del dimorfismo sexual de la estatura (Abarca, 2013). Si bien dentro de las variables aquí analizadas sólo se ha estudiado y evidenciado cambio secular para FNW (Stojanowski y Seidemann, 1999; Alunni-Perret, Staccini y Quatrehomme, 2003), es esperable que las transformaciones ambientales y genéticas experimentadas por la población chilena actual se reflejen en un comportamiento diferente de las variables analizadas. Por lo tanto, para contar con un método aplicable a contextos actuales se requiere la elaboración de modelos nuevos en base a muestras pertinentes.

Considerando los alcances del estudio aquí expuesto, algunas líneas de investigación quedan abiertas para estudios futuros. La mayoría de los métodos de estimación de componentes del perfil biológico en población chilena han sido elaborados en base a la Colección Subactual de Santiago, por lo que no deben ser considerados válidos para población chilena actual teniendo en cuenta la naturaleza población-específica del dimorfismo sexual esquelético anteriormente mencionada. Por esta razón, probablemente la línea de investigación más urgente sea la elaboración de estándares de estimación de sexo construidos en base a muestras contemporáneas, que permitan desarrollar metodologías aplicables a población chilena actual. Dado que no existen colecciones osteológicas representativas de este grupo, un enfoque factible es el trabajo con población viva mediante aproximaciones tales como la radiografía o la tomografía computarizada, herramientas que ya han sido utilizadas con éxito en otros contextos actuales para obtener los datos necesarios para construir modelos predictivos de sexo (Harma y Karakas, 2007; Kranioti et al., 2009; Stull y Godde, 2012; Kim et al., 2013; Mitra et al., 2014; Clavero et al., 2015; Gulhan et al., 2015; Hishmat et al., 2015).

Por otro lado, en esta tesis se elaboró una metodología a partir del fémur proximal dadas sus ventajas para la estimación de sexo. No obstante, otros elementos sexualmente dimórficos pueden ser explorados en estudios venideros para elaborar modelos predictivos de sexo en población chilena. Tomando en cuenta que en el caso nacional ya se han evaluado escápula, clavícula, húmero, radio, ulna, fémur, tibia y fíbula (Ross y Manneschi, 2010; Garrido-Varas et al., 2014; Peckmann et al., 2016; O’Bright et al., 2018), podrían investigarse otros elementos como costillas, carpos, tarsos y falanges. Además, los estudios conducidos hasta ahora no han explotado completamente el potencial de cada elemento, pudiendo analizarse otras medidas inéditas o bien emplear enfoques no explorados hasta ahora como la regresión logística, de la que el presente estudio es el único exponente para población chilena. También puede resultar interesante la evaluación de enfoques experimentales como *machine learning* o redes neurales como una forma de aportar nuevos antecedentes sobre estas novedosas herramientas técnicas.

Finalmente, otra línea de investigación posible, esta vez ligada a la Bioarqueología, que requiere de estudios no realizados a la fecha es la elaboración de estándares aplicables a poblaciones arqueológicas presentes en el territorio chileno. Como en el caso de la población chilena actual, no existen estándares específicos para este grupo, por lo que la reconstrucción del perfil biológico suele conducirse mediante metodologías elaboradas en base a poblaciones de referencia muy distintas, con una consecuente pérdida de confiabilidad significativa en los métodos empleados. La evaluación de colecciones como Pica 8, tutelada por el Departamento de Antropología de la Universidad de Chile y constituida por cerca de 150 individuos (Lemp et al., 2008; Pacheco, 2010), así como la puesta en valor de material sin sistematizar como el resguardado por el Museo Nacional de Historia Natural (proceso en el cual el autor participó como parte de su práctica profesional) permitirían la elaboración de nuevos estándares aptos para ser aplicados a restos de temporalidad arqueológica.

Conclusiones

En el presente trabajo se elaboraron modelos estadísticos para la estimación de sexo en población chilena moderna mediante regresiones logísticas en base a variables osteométricas del fémur proximal. Las variables que miden proporciones del cuello femoral – FNAL, GTFC y FNW – resultaron ser los mejores estimadores del sexo de los individuos pues aproximan de forma más directa el dimorfismo sexual de la epífisis proximal. Por otro lado, las variables del triángulo de Purkait (AB, AC y BC) así como las demás variables del triángulo de Albanese (GTLT y LTFC) reportaron menores aciertos dado que recogen de manera más indirecta el dimorfismo de la región. Los modelos propuestos alcanzan un 92.9% de acierto para las ecuaciones univariadas y hasta un 95.7% para las multivariadas, en ambos casos con bajos sesgos por sexo, siendo los modelos más precisos reportados a la fecha para población chilena moderna.

Las limitaciones del presente estudio incluyen el espectro de aplicación de la metodología propuesta, que sólo puede ser considerada válida para población chilena moderna. La estimación de sexo en otras poblaciones distintas tanto temporal como geográficamente exige la elaboración de estándares propios construidos en base a muestras pertinentes. Futuros estudios pueden buscar elaborar metodologías de estimación de perfil biológico aplicables a población chilena contemporánea, evaluar otros elementos esqueléticos y/u otras variables para la estimación de sexo en población chilena moderna, explorar la utilización de enfoques experimentales o construir métodos aptos para poblaciones arqueológicas.

El estudio aquí presentado podría ser un aporte a la identificación de restos humanos en contextos de causas de violaciones a los Derechos Humanos en Chile cometidas durante la Dictadura Cívico-Militar (1973-1989), al proveer de una nueva metodología de estimación de sexo en población chilena moderna mediante la evaluación osteométrica del fémur proximal.

Bibliografía

- Abarca, V. (2011). Efectos de la nutrición sobre el Dimorfismo Sexual expresado en la Estatura (SSD) de una muestra de población chilena Subadult. Memoria para optar al Título Profesional de Antropóloga Física. Santiago. Facso, Universidad de Chile.
- Abarca, V. (2013). Desarrollo de nueva fórmula para estimar estatura en población chilena adulta, a partir del largo del fémur. Tesis para optar al grado de Magíster en Ciencias de la Nutrición mención Promoción de la Salud y Prevención de Enfermedades Crónicas no Transmisibles. Santiago. INTA, Universidad de Chile
- Adams, B. J. (2007). Inside Forensic Science. Forensic Anthropology. New York: Infobase Publishing
- Albanese, J. (2003). A Metric Method for Sex Determination Using the Hipbone and the Femur. *Journal of Forensic Sciences*, 48 (2), 263-273
- Albanese, J. (2013). A Method for Estimating Sex Using the Clavicle, Humerus, Radius, and Ulna. *Journal of Forensic Sciences*, 1413-1419.
- Albanese, J., Eklics, G., y Tuck, A. (2008). A Metric Method for Sex Determination Using the Proximal Femur and Fragmentary Hipbone. *Journal of Forensic Sciences*, 53 (6), 1283-1288.
- Alunni-Perret, V., Staccini, P., y Quatrehomme, G. (2003). Reexamination of a measurement for sexual determination using the super-inferior femoral neck diameter in a modern European population. *Journal of Forensic Sciences*, 48 (3), 517-520.
- Alunni, V., du Jardin, P., Nogueira, L., Buchet, L. y Quatrehomme, G. (2015). Comparing discriminant analysis and neural network for the determination of sex using femur head measurements. *Forensic Science International*, 253, 81-87.
- Alunni-Perret, V., Staccini, P., y Quatrehomme, G. (2008). Sex determination from the distal part of the femur in a French contemporary population. *Forensic Science International*, 175, 113–117.
- Anastopoulou, I., Eliopoulos, C., Valakos, E. y Manolis, S. (2014). Application of Purkait's triangle method on a skeletal population from southern Europe. *Forensic Science International*, 245, 203.e1-203.e4.
- Asala, S. (2002). The efficiency of the demarking point of the femoral head as a sex determining parameter. *Forensic Science International*, 127, 114-118.
- Asala, S. A., Bidmos, M. A. y Dayal, M. R. (2004). Discriminant function sexing of fragmentary femur of South African blacks. *Forensic Science International*, 145, 25–29.
- Aspillaga, E., Gonzáles, M. y Paredes, C. (1984). Estimación de la capacidad torácica a través de los huesos. *Estudios Atacameños*, 7, 249-254.
- Ballesteros, E. (1995). La memoria y el olvido. *Detenidos Desaparecidos en Chile*. Santiago: CEME
- Barbaro, A. (2018). *Manual of Forensic Science*. Boca Raton: CRC Press.
- Berthelot, C. M. (2014). Metric, nonmetric, and geometric morphometric methods of sex estimation using the distal humerus. Tesis para optar al grado de Master of Arts. Boston University Open BU.

- Black, S., y Ferguson, E. (2011). *Forensic Anthropology. 2000 to 2010*. Boca Ratón: CRC Press.
- Boldsen, J., Milner, G. y Boldsen, S. (2015). Brief Communication: Sex Estimation from Modern American Humeri and Femora, Accounting for Sample Variance Structure. *American Journal of Physical Anthropology*, 158, 745-750.
- O’Bright, L., Peckmann, T. y Meek, S. (2018). Is “Latin American” population-specific? Testing sex discriminant functions from the Mexican tibia on a Chilean sample. *Forensic Science International*
- Brown, R., Ubelaker, D. y Schanfield, M. (2007). Evaluation of Purkait’s Triangle Method for Determining Sexual Dimorphism. *Journal of Forensic Sciences*, 52 (3), 553-556.
- Buikstra, J. E., y Ubelaker, D. H. (1994). *Standards for data collection from human skeletal remains*. Fayetteville: Arkansas Archeological Survey Research Series No. 44.
- Burns, K. R. (2013). *Forensic Anthropology Training Manual*. Upper Saddle River: Pearson.
- Center, J., Nguyen, T., Pocock, N., Noakes, K. Kelly, P., Eisman, J. y Sambrook, P. (1998). Femoral Neck Axis Length, Height Loss and Risk of Hip Fracture in Males and Females. *Osteoporosis International*, 8, 75-81
- Christensen, A. M., Passalacqua, N. V., y Bartelink, E. J. (2014). *Forensic Anthropology. Current Methods and Practice*. Oxford: Elsevier.
- Clavero, A., Salicrú, M. y Turbón, D. (2015). Sex prediction from the femur and hip bone using a sample of CT images from a Spanish population. *International Journal of Legal Medicine*, 129, 373-383.
- Curate, F., Coelho, J., Gonçalves, D., Coelho, C., Ferreira, M., Navega, D., y Cunha, E. (2016). A Method for Sex Estimation Using the Proximal Femur. *Forensic Science International*, 266, 579.e1–579.e7
- Curate, F., Albuquerque, A., Ferreira, I. y Cunha, E. (2017a). Sex estimation with the total area of the proximal femur: A densitometric approach. *Forensic Science International*, 275, 110-116.
- Curate, F., Umbelino, C., Perinha, A., Nogueira, C., Silva, A. y Cunha, E. (2017b). Sex determination from the femur in Portuguese populations with classical and machine-learning classifiers. *Journal of Forensic and Legal Medicine*, 52, 75-81
- Dabbs, G. R., y Moore-Jansen, P. H. (2010). A Method for Estimating Sex Using Metric Analysis of the Scapula. *Journal of Forensic Sciences*, 55(1), 149-152.
- Daeid, N. (2010). *Fifty Years of Forensic Science: A Commentary*. Chichester: Wiley-Blackwell.
- Díaz, P. (2010). *Valor Predictivo de los Rasgos Anatómicos del Cráneo utilizados en la Estimación Visual del Sexo en Población Chilena: un Análisis de Morfometría Geométrica*. Tesis de Grado para optar al Título profesional de Antropólogo Físico. Santiago. Facso, Universidad de Chile
- Dirkmaat, D. C. (2012). *A Companion to Forensic Anthropology*. Chichester: Wiley Blackwell.

- Dittrick, J. y Suchey, J. (1986). Sex Determination of Prehistoric Central California Skeletal Remains Using Discriminant Analysis of the Femur and Humerus. *American Journal of Physical Anthropology*, 70, 3-9.
- Djorojevic, M., Roldán, C., Botella, M. y Alemán, I. (2016). Estimation of Purkait's triangle method and alternative models for sex assessment from the proximal femur in the Spanish population. *International Journal of Legal Medicine*, 130, 245-251.
- du Jardin, P., Ponsaille, J., Alunni-Perret, V., y Quatrehomme, G. (2009). A comparison between neural network and other metric methods to determine sex from the upper femur in a modern French population. *Forensic Science International*, 192, 127.e1–127.e6.
- Embar-Seddon, A. y Pass, A. (2009). *Forensic Science*. Pasadena: Salem Press.
- Espinoza, Ma. C. (2016). Evaluación del método de estimación de edad al momento de la muerte en superficie auricular en una muestra de población chilena Subactual (Colección Subactual de Santiago). Memoria para optar al título de Antropóloga Física. Santiago. Facso, Universidad de Chile.
- Fraser, J. y Williams, R. (2009). *Handbook of Forensic Science*. Cullompton: Willan Publishing
- Fraser, J. (2010). *Forensic Science. A Very Short Introduction*. Oxford: Oxford University Press.
- Frutos, L. (2003). Brief Communication: Sex Determination Accuracy of the Minimum Supero-Inferior Femoral Neck Diameter in a Contemporary Rural Guatemalan Population. *American Journal of Physical Anthropology*, 122, 123-126.
- Garrido-Varas, C. e Intriago, M. (2011). Managing commingled remains from mass graves: Considerations, implications and recommendations from a human rights case in Chile. *Forensic Science International*, 219, e19-e24.
- Garrido-Varas, C., Thompson, T., y Campbell, A. (2014). Parámetros Métricos para la Determinación de Sexo en Restos Esqueléticos Chilenos Modernos. *Chungará*, 46 (2), 285-293.
- Grivas, C. y Komar, D. (2008). *Kumho, Daubert*, and the Nature of Scientific Inquiry: Implications for Forensic Anthropology. *Journal of Forensic Sciences*, 53 (4), 771-776.
- Gulhan, O., Harrison, K. y Kiris, A. (2015). A new computer-tomography-based method for sex estimation: Development of Turkish population-specific standards. *Forensic Science International*, 255, 2-8.
- Guyomarc'h, P., Velemínská, J., Sedlak, P., Dobisíková, M., Švenkrťová, I. y Brůžek, J. (2016). Impact of secular trends on sex assessment evaluated through femoral dimensions of the Czech population. *Forensic Science International*, 262, 284.e1-284.e6.
- Harma, A. y Karakas, H. (2007). Determination of sex from the femur in Anatolian Caucasians: A digital radiological study. *Journal of Forensic and Legal Medicine*, 14, 190-194.
- Herrera, Ma. J. y Retamal, R. (2017). Reliability of age estimation from iliac auricular Surface in a subactual Chilean sample. *Forensic Science International*.
- Herrera, Ma. J. (2012). Evaluación del método de estimación de edad a través de la superficie auricular del ilion en una muestra chilena Subactual (Cementerio

General). Memoria para optar al título profesional de Antropóloga Física. Santiago. Facso, Universidad de Chile.

- Hishmat, A., Michiue, T., Sogawa, N., Oritani, S., Ishikawa, T., Fawzy, I., Mohsen, M. y Maeda, H. (2015). Virtual CT morphometry of lower limb long bones for estimation of the sex and stature using postmortem Japanese adult data in Forensic identification. *International Journal of Legal Medicine*, 129, 1173-1182.
- Hosmer, D., Lemeshow, S. y Sturdivant, R. (2013). *Applied Logistic Regression*. Hoboken: Wiley.
- Houck, M. y Siegel, J. (2015). *Fundamentals of Forensic Science*. Oxford: Elsevier.
- Houck, M. (2007). *Forensic Science. Modern Methods of Solving Crime*. Westport: Praeger.
- Houck, M. (2015). *Professional Issues in Forensic Science*. San Diego: Elsevier.
- Hyeoun-Ae, P. (2013). An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain. *Journal of Korean Academy of Nursing*, 43 (2), 154-164.
- İşcan, M. Y., y Steyn, M. (2013). *The Human Skeleton in Forensic Medicine*. Springfield: Charles C Thomas Publisher Ltd.
- Jackson, A. y Jackson, J. (2017). *Forensic Science*. Harlow: Pearson.
- James, S. y Nordby, J. (2009). *Forensic Science. An Introduction to Scientific and Investigative Techniques*. Boca Raton: CRC Press.
- Jamieson, A. y Moenssens, A. (2009). *Wiley Encyclopedia of Forensic Science*. Chichester: Wiley.
- Jantz, R., Kimmerle, E. y Baraybar, J. P. (2008). Sexing and Stature Estimation Criteria for Balkan Populations. *Journal of Forensic Sciences*, 53 (3), 601-605.
- Katz, E. y Halámek, J. (2016). *Forensic Science. A Multidisciplinary Approach*. Wienheim: Wiley-VCH.
- Kim, D.I., Lee, S.S. y Kim, Y.S (2010). Statistical analysis of bone elements excavated from the forensic context, *Korean Journal of Physical Anthropology*, 23, 1-8.
- Kim, D., Kwak, D. y Han, S. (2013). Sex determination using discriminant analysis of the medial and lateral condyles of the femur in Koreans. *Forensic Science International*, 233, 121-125.
- Kleinbaum, D. y Klein, M. (2010). *Logistic Regression. A Self-Learning Text*. New York: Springer.
- Klepinger, L. L. (2006). *Fundamentals of Forensic Anthropology*. Urbana: Wiley-Liss.
- Kranioti, E. F., Vorniotakis, N., Galiatsou, C., Iscan, M. Y., y Michalodimitrakis, M. (2009). Sex identification and software development using digital femoral head radiographs. *Forensic Science International*, 189, 113.e1-113.e7
- Larose, D. T., y Larose, C. D. (2015). *Data Mining and Predictive Analytics*. Hoboken: John Wiley y Sons.
- Lemp, C., Rodríguez, M., Retamal, R. y Aspillaga, E. (2008). Arqueología del depósito: manejo integral de las colecciones bioantropológicas en el Departamento de Antropología de la Universidad de Chile. *Conserva*, 12, 69-96.
- Lerner, K. y Lerner, B. (2006). *World of Forensic Science*. Farmington Hills: Thompson Gale.

- Lesciotto, K. (2015). The Impacto of *Daubert* on the admissibility of Forensic Anthropology Expert Testimony. *Journal of Forensic Sciences*, 60 (3), 549-555.
- Lin, L. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, 45 (1), 255-268.
- Longe, S. (2015). Sex Estimation In Forensic Anthropology Using the Radius, Femur, And Scapula. Honors Thesis. Texas State University.
- MacLaughlin, S. y Bruce, M. (1985). A Simple Univariate Technique for Determining Sex From Fragmentary Femora: Its Application to a Scottish Short Cist Population. *American Journal of Physical Anthropology*, 67, 413-417.
- Mall, G., Graw, M., Gehring, K. y Hubig, M. (2000). Determination of sex from femora. *Forensic Science International*, 113, 315-321.
- Meeusen, R. A., Christensen, A. M., y Hefner, J. T. (2015). The Use of Femoral Neck Axis Length to Estimate Sex and Ancestry. *Journal of Forensic Sciences*, 60 (5), 1300-1304.
- Mitra, A., Khadijeh, B., Vida, A. P., Ali, R. N., Farzaneh, M., Maryam, V. F., y Vahid, Y. (2014). Sexing based on measurements of the femoral head parameters on pelvic radiographs. *Journal of Forensic and Legal Medicine*, 23, 70-75.
- Morales, H., Quiñones, E., Krapivka, S., Barreaux, N., Espinoza, Ma. C., Flores, S., Galimany, J., González, R., Jara, K. (2016). Puesta en Valor de la Colección Osteológica Subactual de Santiago. Informe no publicado. Proyecto Trayectorias de Patrimonialización, Iniciativa Bicentenario. Universidad de Chile.
- Netter, Frank H. (2011). *Atlas de Anatomía Humana*. Barcelona: Elsevier.
- Okuda, M. y Stephenson, F. (2015). *A Hands-On Introduction to Forensic Science. Cracking the Case*. Boca Raton: CRC Press.
- Pacheco, A. (2010). Bioarqueología del Conflicto. Lesiones esqueléticas en la Colección Pica 8 y su relación con situaciones de violencia interpersonal durante el Período Intermedio Tardío (Región de Tarapacá, Chile). Memoria para obtener el Título de Antropólogo con especialidad en Antropología Física. Santiago. Facso, Universidad de Chile.
- Paredes, C., Hagn, J. C. y Constantinescu, F. (1994). Identificación: determinación de la edad en sínfisis púbica. *Revista Chilena de Antropología*, 12, 163-178.
- Peckmann, T. R., Logar, C., y Meek, S. (2016). Sex estimation from the scapula in a contemporary Chilean population. *Science and Justice*. *****
- Phenice, T. W. (1969). A newly developed visual method of sexing the Os Pubis. *American Journal of Physical Anthropology*, 30 (2), 297-301.
- Pickering, R. y Bachman, D. (2009). *The Use of Forensic Anthropology*. Boca Raton: CRC Press.
- Pontigo, F. (2013). Evaluando nuevos enfoques metodológicos para el estudio de dietas en poblaciones prehistóricas: Microbioma oral y PIXE. Memoria para optar al Título Profesional de Antropólogo Físico. Santiago. Facso, Universidad de Chile.
- Purkait, R. (2005). Triangle identified at the proximal end of femur: a new sex determinant. *Forensic Science International*, 147, 135-139.
- Pyrek, K. (2007). *Forensic Science Under Siege. The Challenges of Forensic Laboratories and the Medico-Legal Investigation System*. Oxford: Elsevier.
- Retamal, R. (2004). Efectos de la Deformación Craneana Intencional sobre la expresión del Dimorfismo Sexual en cráneos adultos: un análisis morfométrico

geométrico. Memoria para optar al Título de Antropólogo Físico. Santiago. Facso, Universidad de Chile.

- Rissech, C., Schaefer, M. y Malgosa, A. (2008). Development of the femur – Implications for age and sex determination. *Forensic Science International*, 180, 1-9.
- Ross, A. H., y Manneschi, M. (2010). New identification criteria for the Chilean population: Estimation of sex and stature. *Forensic Science International*
- Schmitt, A., Cunha, E., y Pinheiro, J (2006). *Forensic Anthropology and Medicine. Complementary Sciences from Recovery to Cause of Death*. Totowa, Estados Unidos: Humana Press.
- Seidemann, R. M., Stojanowski, C. M., y Doran, G. H. (1998). The Use of the Supero-Inferior Femoral Neck Diameter as a Sex Assessor. *American Journal of Physical Anthropology*, 107, 305–313.
- Stojanowski, C. M., y Seidemann, R. M. (1999). A reevaluation of the sex prediction accuracy of the minimum supero-inferior femoral neck diameter for modern individuals. *Journal of Forensic Sciences*, 44 (6), 1215–1218.
- Sharma, R., Sharma, S., Jehan, M. y Sastya, A. (2014). Sexual dimorphism of the femoral bicondylar angle in Gwalior, Madhya pradesh region. *International Journal of Healthcare and Biomedical Research*, 2 (4), 38-45.
- Shelton, D. (2011). *Forensic Science in Court*. Lanham: Rowman & Littlefield Publishers.
- Siegel, J. y Mirakovits, K. (2010). *Forensic Science. The Basics*. Boca Raton: CRC Press.
- Soni, G., Dhall, U., y Chhabra, S. (2010). Determination of Sex from Femur: Discriminant Analysis. *Journal of the Anatomical Society of India*, 59 (2), 216-221.
- Spradley, K., y Jantz, R. L. (2011). Sex Estimation in Forensic Anthropology: Skull Versus Postcranial Elements. *Journal of Forensic Sciences*, 56 (2), 289-296
- Spradley, K., Anderson, B. E., y Tise, M. L. (2014). Postcranial Sex Estimation Criteria for Mexican Hispanics. *Journal of Forensic Sciences*, 60, 1-5. ****
- Spradley, K., Jantz, R. L., Robinson, A., y Peccerelli, F. (2008). Demographic Change and Forensic Identification: Problems in Metric Identification of Hispanic Skeletons. *Journal of Forensic Sciences*, 53 (1), 21-28.
- Steyn, M. e Iscan, M. (1997). Sex determination from the femur and tibia in South African whites. *Forensic Science International*, 90, 111-119.
- Stull, K. y Godde, K. (2012). Sex Estimation of Infants Between Birth and One Year Through Discriminant Analysis of the Humerus and Femur. *Journal Of Forensic Sciences*, 58 (1), 13-20.
- Tersigni-Tarrant, M. y Shirley, N. (2013). *Forensic Anthropology. An Introduction*. Boca Raton: CRC Press.
- Tilstone, W., Savage, K. y Clark, L. (2006). *Forensic Science. An Encyclopedia of History, Methods, and Techniques*. Santa Barbara: ABC-CLIO.
- Tise, M., Spradley, K. y Anderson, B. (2012). Postcranial Sex Estimation of Individuals Considered Hispanic. *Journal of Forensic Sciences*, 58 (S1), S9-S14.
- Torimitsu, S., Makino, Y., Saitoh, H., Sakuma, A., Ishii, N., Yajima, D., Inokuchi, G., Motomura, A., Chiba, F., Yamaguuchi, R., Hashimoto, M., Hoshioka, Y. e Iwase, H.

- (2016). Sex estimation based on scapula analysis in a Japanese population using multidetector computed tomography. *Forensic Science International*, 262, 285.e1-285.e5
- Ubelaker, D. (2013). *Forensic Science. Current Issues, Future Directions*. Chichester: Wiley-Blackwell.
 - Urzúa, I., Huberman, J., Delgado, I., Pacheco, A. y Retamal, R. (2009). Prevalencia de Caries y Pérdida de Dientes de una Población Adulta Chilena Nacida en el Siglo XIX. *Revista Clínica de Periodoncia, Implantología y Rehabilitación Oral*, 2 (3), 175-178.
 - Walrath, D., Turner, P. y Bruzek, J. (2004). Reliability Test of the Visual Assessment of Cranial Traits for Sex Determination. *American Journal of Physical Anthropology*, 125, 132-137.
 - White, P. (2004). *The Essentials of Forensic Science*. Cambridge: The Royal Society of Chemistry.
 - White, T., Black, M. y Folkens, P. (2012). *Human Osteology*. Oxford: Elsevier.

Anexos

Figura 1: Esquema de las variables FNAL y FNW en el fémur proximal



Figura 2: Esquema de las variables AB, AC y BC en el fémur proximal



Figura 3: Esquema de las variables GTFC, GTLT y LTFC en el fémur proximal



Figura 4: Espécimen de fémur característicamente femenino. Vista anterior



Figura 5: Espécimen de fémur caracteristicamente femenino. Vista posterior



Figura 6: Espécimen de fémur característicamente masculino. Vista anterior



Figura 7: Espécimen de fémur característicamente masculino. Vista posterior

