

**UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE CIENCIAS MÉDICAS
ESCUELA DE NUTRICIÓN**

**TRABAJO DE INVESTIGACIÓN DE LA LICENCIATURA EN NUTRICIÓN:
INFORME FINAL**

***“Aporte metodológico
para el estudio de la red
causal del
Cáncer de Mama:
El sesgo de información
y su impacto en la
estimación de los efectos
de componentes de la
dieta”***



Alumna: Paula Carreño

Directora: Dra. María Del Pilar Díaz

Co-directora: Dra. Julia Becaria Coquet

UNIVERSIDAD NACIONAL DE CÓRDOBA
FACULTAD DE CIENCIAS MÉDICAS
ESCUELA DE NUTRICIÓN

TRABAJO DE INVESTIGACIÓN DE LA
LICENCIATURA EN NUTRICIÓN

*“Aporte metodológico para el estudio de la red causal del cáncer de mama:
El sesgo de información y su impacto en la estimación de los
efectos de componentes de la dieta”*

Tribunal de evaluación:

Med. Bollati Maria Alicia Lic. Ruiz Bruner Mercedes Dra. Díaz María Del Pilar

Autora:

Carreño Paula

Nota: _____

Fecha: _____

N° de tesis: _____

*“Las opiniones expresadas por los autores de este Seminario Final no
representan necesariamente los criterios de la Escuela de Nutrición de la
Facultad de Ciencias Médicas”*

AGRADECIMIENTOS

A Pilar y Julia por guiarme en el camino de la investigación con gran dedicación, dedicándome tiempo y enriqueciéndome tanto en lo profesional como en lo personal.

A las integrantes del tribunal evaluador.

A la Escuela de Nutrición y al Proyecto PICT 2012-1019, FONCyT.

A mis papas y mi hermano por apoyarme, acompañarme y motivarme durante toda la carrera, gracias por compartir cada logro conmigo y sobre todo por confiar tanto en mi.

A toda mi familia que siempre me apoyó e incentivó a seguir adelante. A mi abuela por las tantas velitas prendidas antes de cada examen.

A mi prima Bela por estar presente en cada momento haciendo que la distancia no se un impedimento para estar juntas.

A mis incondicionales amigas que me dio la facultad que hicieron el camino más fácil y divertido, gracias por todos los momentos compartidos y por festejar cada pasito conmigo.

A mi compañera desde el primer grupo de estudio, Belu, gracias por calmar mi estrés y rendir cada final conmigo.

A mis amigos de siempre que me apoyaron desde el primer momento y estuvieron presentes, motivándome siempre a crecer y superarme.

A mi amiga Sori por ser creadora de la imagen de esta tesis y acompañarme en cada detalle, incluso en los momentos más difíciles.

A todo el equipo de Estadística, a Los Epipichis y a Unite al Mejor Abrazo por siempre estar dispuestos a ayudar y acompañarme.

“Aporte metodológico para el estudio de la red causal del cáncer de mama: el sesgo de información y su impacto en la estimación de los efectos de componentes de la dieta”

Trabajo de Investigación de Licenciatura en Nutrición

Área temática de investigación: Epidemiología y Salud Pública.

Autores: Carreño P, Dra. Becaria Coquet J, Dra. Diaz M, P.

Resumen

Introducción: La alimentación se asocia con la ocurrencia de cáncer de mama (CM), en especial la ingesta inadecuada de algunos nutrientes promueve su desarrollo. No obstante, su cuantificación es compleja y sensible a la presencia de sesgos, impactando en la estimación de las medidas de riesgo.

Objetivo: Valorar el impacto del sesgo de información presente en la valoración cuantitativa nutricional del Valor Energético Total (VET) y de la fibra, originadas en un estudio caso-control de CM en Córdoba, en el periodo 2008-2016.

Metodología: Estudio epidemiológico observacional de tipo caso-control (343 casos; 606 controles). Se valoraron medidas antropométricas, características bio-socio-culturales, actividad física e ingesta alimentaria, mediante encuesta cuali-cuantitativa. Se realizaron análisis descriptivos e inferenciales mediante modelos de regresión logística con datos completos para evaluar la asociación entre presencia de CM y las variables de exposición (VET y fibra), incluyéndose covariables de interés. Posteriormente, se aplicó imputación múltiple (IM) para datos faltantes.

Resultados: Se observó asociación positiva entre el VET elevado y el CM (OR: 2,4; IC 95%: 1,043; 2,305). Se estimó una relación inversa entre la ingesta de fibra y el desarrollo de CM (OR:0,64; IC 95%:0,505- 0,812). Con la aplicación de IM se obtuvo mayor significancia de esas variables, además los efectos del IMC y estado de menopáusico se evidenciaron luego del uso de IM.

Conclusiones: Existe mayor riesgo de presentar CM en mujeres con VET elevado, menor cociente fibra insoluble/soluble, obesidad y que se encuentran en la post-menopausia, condicionantes identificados luego de que la técnica de IM fuera aplicada.

Palabras claves: adenocarcinoma de mama- nutrición- datos faltantes- imputación

ÍNDICE

INTRODUCCIÓN	6
PLANTEAMIENTO DEL PROBLEMA	8
OBJETIVO GENERAL	9
OBJETIVOS ESPECÍFICOS	9
MARCO TEÓRICO	10
EPIDEMIOLOGÍA DEL CÁNCER	10
CÁNCER DE MAMA	12
ROL DE LA ALIMENTACIÓN EN EL DESARROLLO DE CÁNCER DE MAMA	14
SESGO POR DATOS FALTANTES: MÉTODO DE IMPUTACIÓN MÚLTIPLE	16
HIPÓTESIS	19
VARIABLES DE ANÁLISIS	19
DISEÑO METODOLÓGICO	20
TIPO DE ESTUDIO	20
MUESTRA	21
OPERACIONALIZACIÓN DE LAS VARIABLES	22
TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS	26
PLAN DE ANÁLISIS DE LOS DATOS	27
RESULTADOS	30
ANÁLISIS EXPLORATORIO:	30
SESGO DE INFORMACIÓN Y ROL DE LA DIETA EN LAS ESTIMACIONES DE RIESGO DE CÁNCER DE MAMA	41
DISCUSIÓN	45
CONCLUSIONES	50
BIBLIOGRAFIA	52
ANEXOS	59
ANEXO 1. Categorización de la variable nivel socioeconómico	59
GLOSARIO	61

INTRODUCCIÓN

Las transiciones demográficas y epidemiológicas, la urbanización y los cambios en los modelos de desarrollo, modificaron la relación de las poblaciones con sus alimentos propiciando la aparición en niveles epidémicos de diferentes cánceres y eventos cardiovasculares. (1-3)

A nivel mundial, los tumores malignos representan la segunda causa de muerte por enfermedad, con una incidencia creciente, especialmente en países de bajos y medianos ingresos. (4) Argentina no es la excepción y ha seguido la tendencia epidemiológica mundial hacia un rápido aumento de la carga relativa de las enfermedades no transmisibles siendo Córdoba una de las provincias con mayores tasas de incidencia y mortalidad. (5,6) Díaz y cols. describieron los cánceres más frecuentes en la provincia: mama, cérvix, colon y pulmón en el sexo femenino, y de próstata, pulmón, colon y vejiga en el masculino (5). Diversos factores de riesgo son potencialmente carcinógenos, entre ellos, la alimentación (4), asociada significativamente a la ocurrencia del cáncer. Se estima que el 35% de los tumores malignos estarían relacionados con hábitos alimentarios poblacionales (5).

El cáncer de mama (CM) es el de mayor incidencia y mortalidad en Córdoba, en mujeres y está espacialmente distribuido de manera agregada (no aleatoria) (6,7). Asimismo, se asocia a factores socio-económicos y patrones dietarios de las poblaciones urbanas y rurales (3,8). El Grupo de Epidemiología Ambiental del Cáncer en Córdoba (GEACC), de la Universidad Nacional de Córdoba, identificó patrones alimentarios y reportó que una dieta habitual con elevadas ingestas de calorías, grasas, especialmente saturadas, carnes procesadas, alimentos refinados y alcohol se asocia al aumento del riesgo de CM (1, 5). Esto evidencia el papel de la nutrición como factor determinante en el desarrollo de la carcinogénesis mamaria (5,9,10).

Trabajos del GEACC, basados en estudios tipo caso-control, estudios ecológicos y prospectivos, identificaron variables que disminuyen el riesgo de desarrollar CM. (11) Se reporta en la bibliografía que entre los factores protectores, se destaca una dieta con alto contenido de alimentos ricos en carotenos, micronutrientes y fibra (12). Respecto a la fibra dietética, revisiones sistemáticas asocian su mayor consumo con un menor riesgo

de presentar CM, reduciéndose un 5% por cada incremento adicional de 10 gramos de fibra (13).

Por otra parte, a pesar de la plausibilidad biológica entre el consumo de grasas y el riesgo de CM, todavía existe controversia sobre esta relación; reportándose una modesta (y directa) asociación entre ellos (5,14). Este plan propone estudiar, desarrollando metodología confiable, estas relaciones en la red causal del CM.

Para valorar el rol de la dieta en la etiología y prevención del CM y el aporte que determinados nutrientes ejercen en el crecimiento del riesgo de desarrollarlo es menester una correcta estimación de la ingesta alimentaria. Esto supone mejorar los métodos de evaluación de ingesta de alimentos y la metodología analítica-estadística para estimar confiablemente posibles asociaciones, aún cuando los estudios incluyan poblaciones heterogéneas o en presencia de factores de confusión.

Los estudios epidemiológicos, principalmente los caso-control, son susceptibles a errores sistemáticos (sesgos) en la valoración alimentaria, provocando inconsistencias en las estimaciones que reportan, y por ende, inferencias no confiables. Estos sesgos aparecen por errores en el diseño del trabajo, ya sea en la selección de individuos, en la información recogida o en su análisis, produciéndose diferencias entre grupos, no atribuible al factor estudiado. Si el sesgo surge durante la recolección de datos, cuando no se recaba toda la información (bases incompletas de datos), se denomina de información. Este es frecuente en las investigaciones biomédicas y afecta principalmente a las covariables originando estimaciones poco precisas. Por ende, requiere un tratamiento especial al estimar riesgos o efectos de las mismas (15-17).

Este plan estudia el rol que la fibra y el Valor Energético Total (VET), ejercen sobre CM, estimando de manera confiable los riesgos asociados desde un exhaustivo análisis de sensibilidad. Indaga mediante técnicas estadísticas novedosas, la presencia de sesgos de información citados como importantes en los estudios de tipo caso-control, y valora su impacto en las estimaciones de los efectos de esos nutrientes en el CM.

PLANTEAMIENTO DEL PROBLEMA

¿Existe un impacto significativo del sesgo de información en las estimaciones de los efectos atribuidos al VET y el consumo de fibra, originadas a partir del estudio caso-control de cáncer de mama en Córdoba?

OBJETIVO GENERAL

Valorar el impacto sesgo de la información presente en la valoración cuantitativa nutricional del VET y de la fibra, originadas a partir del estudio caso-control de cáncer de mama en Córdoba, en el periodo 2008-2016.

OBJETIVOS ESPECÍFICOS

- ✚ Caracterizar la ingesta de nutrientes, especialmente VET y consumo de fibra, como así también las características y hábitos de vida de las mujeres que forman parte del estudio caso-control de mama.
- ✚ Estimar el sesgo de información, provocado por la presencia de datos faltantes, que afecta la estimación de los riesgos de cáncer de mama, asociados a la ingesta de fibra y el VET.
- ✚ Estimar el rol que ejercen esos componentes dietarios en la ocurrencia del cáncer de mama, considerando el sesgo de información y controlando por algunas covariables reconocidas en la red causal del cáncer de mama.

MARCO TEÓRICO

EPIDEMIOLOGÍA DEL CÁNCER

La rápida y no planificada urbanización, sumada a la mundialización de modos de vida poco saludables, el envejecimiento de la población, las dietas poco saludables y la inactividad física son algunos de los factores que han favorecido la aparición de Enfermedades Crónicas No Transmisibles (ECNT). El cáncer, al igual que las enfermedades cardiovasculares, las enfermedades respiratorias y la diabetes, forma parte de este grupo de enfermedades, las cuales tienden a ser de larga duración, generalmente con una lenta progresión, y resultan de la combinación de factores genéticos, fisiológicos, ambientales y conductuales. Dichas patologías afectan a todos los grupos de edad, en todas las regiones y países, y no realiza distinciones en relación al nivel de ingresos (18,19).

El cáncer no es una sola enfermedad, sino que involucra a un grupo de más de cien enfermedades en las cuales las células en una parte del cuerpo comienzan a crecer sin control. Existen muchos tipos de cáncer y puede originarse en diferentes tejidos del cuerpo, incluso migrar y propagarse a otras partes del cuerpo (metástasis) (20,21). Se trata de una enfermedad de origen monoclonal, es decir, que comienza en una sola célula en la cual se producen una serie de mutaciones genéticas que ocasionan un cambio de su fenotipo normal y la convierten en una célula neoplásica. La carcinogénesis constituye un proceso multietápico, que se resume en las siguientes etapas: iniciación, promoción y progresión. Esta última, es exclusiva de la transformación maligna e implica la capacidad de invadir tejidos vecinos o a distancia. Cabe destacar que, para que una célula se convierta en neoplásica se requieren varias mutaciones en varios genes, lo cual ocurre luego de mucho tiempo de estar expuesto a un agente carcinogénico (22).

Actualmente, el cáncer es la segunda causa de muerte, precedida por las enfermedades cardiovasculares, ya sea a nivel mundial, nacional como provincial. (23-25). A nivel mundial, de acuerdo a lo declarado por la OMS, dicha enfermedad, ocasionó 8,8 millones de defunciones en el año 2015 (21). Asimismo, el cáncer constituye la segunda causa de muerte en América, donde se estima que 2,8 millones de personas son diagnosticadas y 1,3 millones mueren por cáncer cada año. Las últimas estimaciones realizadas por la Agencia Internacional de Investigación sobre Cáncer (IARC) de acuerdo

a los datos disponibles a nivel mundial para el año 2018, muestran que Argentina presenta una tasa de incidencia de 212 por cada 100.000 habitantes (considerando ambos sexos y todos los tumores a excepción de piel no melanoma), posicionándonos a nivel mundial en un país con incidencia de cáncer media- alta y en séptimo lugar en Latinoamérica. A partir de los datos provenientes de Registro de Cáncer de Base Poblacional (RCBP) de nuestro país, la IRAC ha estimado para Argentina una incidencia de 269 y 290 casos nuevos de cáncer por cada 100.000 varones y mujeres, respectivamente (26). Díaz y cols., estudiaron la incidencia de cáncer en la provincia, reportando tasas de incidencia totales estandarizadas y ajustadas por edad, iguales a 263,53 DE=138,34 y 200,45 DE=98,30 (x100000 hab/año) (27).

CÁNCER DE MAMA

El término "cáncer de mama" hace referencia a un tumor maligno que se ha desarrollado a partir de células mamarias que comienzan a crecer sin control. (28) Este crecimiento descontrolado se debe a mutaciones en el ADN que convierten a las células normales de los senos, en células cancerosas. Dichas células normalmente forman un tumor que a menudo se puede observar en una radiografía o se puede palpar como una masa o bulto. (29) El CM ocurre casi exclusivamente en las mujeres, y generalmente, se origina en las células de los lobulillos (carcinomas lobulillares), o en los conductos galactóforos (carcinomas ductuales). Con menos frecuencia, el cáncer de mama puede originarse en los tejidos estromales, que incluyen a los tejidos conjuntivos grasos y fibrosos de la mama (28,29). Asimismo, el CM se puede propagar cuando las células cancerosas alcanzan la sangre o el sistema linfático y llegan a otras partes del cuerpo (29). El cáncer de mama es el segundo cáncer más frecuente en el mundo y ocupa el quinto lugar como causa de muerte entre todos los cánceres (30). Esta patología constituye el cáncer más frecuente en las mujeres tanto en los países desarrollados como en los países en desarrollo (31). En este sentido, el Observatorio Global de Cáncer de la OMS, reveló en sus estimaciones que el CM corresponde al 24% de todos los cánceres a nivel mundial en mujeres para el año 2018 (30). En América Central y del Sur la mortalidad por CM ha aumentado en las últimas décadas; un estudio a nivel mundial acerca de las tasas de incidencia y mortalidad por cáncer de mama, estandarizadas por edad, revelan que durante el período 2011-2016, Argentina, Brasil y Uruguay registraron las tasas de incidencia más altas, incluso nuestro país presentó tasas de mortalidad por CM elevadas, al igual que Cuba y Uruguay (32). Según las estimaciones de incidencia del Observatorio Global de Cáncer de la OMS, en Argentina, el CM es el de mayor magnitud en cuanto a ocurrencia para el 2018, registrándose un volumen de más de 21.000 casos al año, lo que representa el 17% de todos los tumores malignos y casi un tercio de los cánceres femeninos. En mujeres, la tasa de incidencia es mayor, registrándose 73 casos por cada 100.000 mujeres, lo cual representa el 32% de todos los tumores malignos (26,30).

En la provincia de Córdoba, el CM es la neoplasia maligna más frecuente (14.968 casos, 17,5% del total), seguido del cáncer colorrectal (9.125 casos), pulmón (7.888 casos) y próstata (7.516 casos). A nivel provincial el CM representa el sitio tumoral más frecuente en mujeres (23). Asimismo, Díaz y col., en su trabajo acerca de la distribución

“Aporte metodológico para el estudio de la red causal del cáncer de mama: el sesgo de información y su impacto en la estimación de los efectos de componentes de la dieta”

espacial de la incidencia de cáncer revela, a partir de cocientes sitio-específicos estandarizados por edad, que el cáncer de mama junto con el de cuello de útero son los más incidentes en mujeres (27).

ROL DE LA ALIMENTACIÓN EN EL DESARROLLO DE CÁNCER DE MAMA

Las transformaciones socioeconómicas y demográficas actuales que ocurren a nivel mundial se encuentran acompañadas por cambios en el estilo de vida, en los patrones dietéticos y en el perfil epidemiológico de las enfermedades prevalentes (33). El cáncer forma parte de dicho perfil epidemiológico, y constituye una encrucijada compleja de determinantes relativos al orden genético, ambiental y a los diferentes estilos de vida, tanto a nivel individual como colectivo (34).

Como se mencionó anteriormente, las tasas de incidencia de CM en la provincia responden a patrones espaciales no aleatorios, por lo que sugieren que las distribuciones geográficas de esta enfermedad podrían estar asociadas, en parte, a factores epigenéticos como a las características ambientales, socioeconómicas y de estilo de vida (6,27). Asimismo, al analizar las tendencias de mortalidad por cáncer de mama en Córdoba, se observa que el comportamiento de la mortalidad por CM se encuentra asociada a factores de índole contextual (demográficos, históricos, económicos, políticos (35).

El cáncer comparte con las principales ECNT algunos factores de riesgo relacionados con el comportamiento como el bajo consumo de frutas y hortalizas, el sedentarismo, el elevado índice de masa corporal y hábitos tóxicos que incluyen el consumo de tabaco e ingesta excesiva de alcohol (34,36). En relación con la alimentación, existen estudios que revelan el papel de la misma como un importante factor vinculado a la ocurrencia del cáncer (6,37). Se reconoce que el 35% de los tumores malignos están asociados con factores alimentarios, y serían prevenibles mediante una alimentación y nutrición apropiadas. Así, dietas con alta densidad calórica, elevada proporción de carnes procesadas, alimentos refinados, grasas y alcohol, se asocian a un incremento del riesgo de desarrollar cáncer de mama, próstata y colon, mientras que una alimentación que incluya importantes cantidades de vegetales y frutas, con alto contenido de micronutrientes y fibra, se encuentra relacionada a una disminución del riesgo (5). Biológicamente, los nutrientes pueden tener efectos anticancerígenos y cancerígenos, dada su posible influencia en la oxidación, metilación, reparación de ADN y otras capacidades, posiblemente afectando el inicio, promoción y progresión del cáncer (38). Es así que la fibra alimentaria constituye uno de los nutrientes que cobra relevancia en la alimentación en relación con el CM, debido a su efecto protector. Por lo tanto, su consumo

en cantidades adecuadas se asocia significativamente con un menor riesgo de padecer CM, reduciéndose un 5% por cada incremento adicional de 10 gramos de fibra. (13, 39-41). Asimismo, cabe aclarar que el cáncer manifiesto puede estar separado por varias décadas de los factores dietéticos relacionados, ya que el efecto de la dieta puede ser acumulativo y solo el impacto a largo plazo puede traducirse en enfermedad. (6)

El estado nutricional también constituye un factor relevante, ya que estudios epidemiológicos indican que el sobrepeso y la obesidad, junto con la dieta y el estilo de vida, contribuyen al riesgo de algunos tipos de cáncer. Esta tendencia en los países en desarrollo está asociada con un mayor consumo de alimentos grasos y / o azucarados y una disminución de la actividad física (42). Más específicamente en el CM, hayamos que la obesidad se asocia con una mayor recurrencia y mortalidad por dicha patología (39). En este sentido, cobra gran importancia el valor energético total de la dieta, ya que la densidad energética alimentaria se ha asociado positivamente con el índice de masa corporal (IMC) y otros factores de riesgo para el cáncer de mama posmenopáusico, lo cual sugiere una asociación positiva estadísticamente significativa entre el valor energético total de la dieta y el riesgo de cáncer de mama posmenopáusico (43).

SESGO POR DATOS FALTANTES: MÉTODO DE IMPUTACIÓN MÚLTIPLE

La dieta es un complejo repertorio de exposiciones que están fuertemente correlacionadas. Los individuos se exponen a la dieta en diferentes grados, lo cual genera que la evaluación de la ingesta de alimentos sea difícil y esté sujeta a múltiples sesgos (15-17, 44).

Tanto en investigación como en vigilancia epidemiológica las encuestas son y continuarán siendo elementos básicos de recopilación de datos en salud y en especial en el área de las enfermedades crónicas. Por lo que el cuestionario es una herramienta ampliamente utilizada en el sector salud para recolectar información y principalmente para determinar el consumo de nutrientes de manera indirecta. Esto da lugar a sesgos u errores que afectan la validez de la información obtenida, convirtiendo la medida de exposición a factores dietéticos nutricionales en una importante limitación para la epidemiología nutricional (44-47).

Uno de estos sesgos se origina por la presencia de datos faltantes, es decir, a partir aquellos valores no disponibles que serían útiles o significativos para el análisis de los resultados (48). Estos valores ausentes representan falta de información en el contenido de una o varias variables en un conjunto de datos, y pueden deberse a factores como la no respuesta en una encuesta, la falta de alguna medición, la pérdida en el proceso de recolección, etc. (49). Dichos datos faltantes son inevitables en la investigación epidemiológica y clínica, pero su potencial para socavar la validez de los resultados de la investigación a menudo se ha pasado por alto en la literatura médica, en la medida que los elementos que participan en estos estudios pueden tener características distintas respecto de los que no participan, lo que puede ocasionar que la muestra deje de ser representativa de la población. Es así que, generalmente la falta de datos no es reportada y si se menciona el hecho, no se explicita la manera en que fueron abordados (50,51).

Generalmente los programas estadísticos asumen que se trabaja con datos completos e incorporan opciones, no siempre las más adecuadas, para imputar observaciones sin que el usuario se dé cuenta de ello (16,51). Por lo que, los investigadores generalmente abordan los datos faltantes al incluir en el análisis solo casos completos, es decir, aquellos individuos que no tienen datos faltantes en ninguna de las variables requeridas para ese análisis; por lo tanto, cuando un sujeto tenga solo un dato

faltante en una de las variables, los datos de dicha persona deben excluirse del análisis junto con todos los demás valores de las otras variables de la encuesta realizada a ese individuo. Sin embargo, los resultados de tales análisis, incluso en situaciones en que el porcentaje de valores ausentes sea bajo en cada variable, pueden estar sesgados y a su vez, puede producirse un efecto acumulativo de los datos faltantes en varias variables que a menudo conduce a la exclusión de una proporción sustancial de la muestra original, lo que a su vez provoca una pérdida sustancial de precisión y potencia (15,16,48,50). Asimismo, está ampliamente documentado que la aplicación de procedimientos inapropiados de sustitución de información introduce sesgos y reduce el poder explicativo de los métodos estadísticos, le resta eficiencia a la fase de inferencia y puede incluso invalidar las conclusiones del estudio (47,51).

Existen diferentes tipos de datos faltantes y muchas razones por las cuales pueden ocurrir, esto resulta fundamental para enfrentar la ausencia de datos en el momento de analizar los resultados, donde lo principal es decidir si la pérdida es aleatoria (46). Hayamos tres tipos principales de patrones de datos faltantes: *Missing Completely at Random (MCAR)*, *Missing At Random (MAR)* y *Noting At Random (NMAR)*, se afirma que un proceso de datos omitidos se genera en forma: *completamente al azar*, es decir, cuando no hay diferencias sistemáticas entre los valores faltantes y los valores observados; *al azar*, donde cualquier diferencia sistemática entre los valores faltantes y los valores observados puede explicarse por diferencias en los datos observados; o *no al azar*, cuando, incluso después de tener en cuenta los datos observados, las diferencias sistemáticas permanecen entre los valores faltantes y los valores observados (47). Asimismo, a pesar de que se reconoce que los criterios estadísticos son fundamentales para la elección del método de imputación, es necesario tener claridad sobre el uso que se hará de la información (51).

Comúnmente, una variedad de enfoques se utilizan para tratar el análisis con datos faltantes. La imputación múltiple es un enfoque general del problema de la falta de datos, cuyo objetivo consiste en reemplazar los datos perdidos por estimaciones que surgen de la creación de varios conjuntos de datos imputados diferentes y la combinación adecuada de los resultados obtenidos de cada uno de ellos. Este proceso dispone de una primera etapa en la que se crean múltiples copias del conjunto de datos, con los valores faltantes reemplazados por valores imputados, los cuales son muestreados a partir de su

distribución predictiva en función de los datos observados. Cabe aclarar, que en lugar de imputar un valor único para cada dato perdido, cada uno de ellos se sustituye por m datos simulados que representa la incertidumbre del valor a imputar. Posteriormente, se acude al uso de métodos estadísticos para ajustar el modelo de interés a cada uno de los conjuntos de datos imputados. Los errores estándar se calculan utilizando las reglas de Rubin (1987), que tienen en cuenta la variabilidad en los resultados entre los conjuntos de datos imputados, lo que refleja la incertidumbre asociada con los valores faltantes. Las asociaciones estimadas en cada uno de los conjuntos de datos imputados diferirán debido a la variación introducida en la imputación de los valores faltantes, y solo son útiles cuando se promedian juntas para dar asociaciones estimadas generales. Así, se obtienen inferencias válidas porque estamos promediando la distribución de los datos faltantes dados los datos observados (15-17,50).

En conclusión, se puede afirmar que disponer de un archivo de datos completos sería lo ideal, sin embargo cuando esto no ocurre, aplicar métodos de imputación tiene el potencial de mejorar la validez de la investigación en nutrición, siempre y cuando se haga de manera apropiada (44,50).

HIPÓTESIS

El diseño del análisis de sesgo de información y su cuantificación permite obtener estimaciones de riesgo del VET y el consumo de fibra diferentes a las estimaciones clásicas, en el estudio de la ocurrencia del cáncer de mama.

VARIABLES DE ANÁLISIS

A continuación se exponen las variables analizadas, definiendo como variable respuesta (*outcome*) a la presencia de CM y como exposición el consumo de fibra y el valor energético total de la dieta. También se incluyen las covariables y variables de ajuste en los modelos y metodologías utilizadas, tal como se puede observar en el siguiente cuadro:

Cuadro 1. Variables de análisis

VARIABLES PRINCIPALES	
Variable de respuesta	Variable de exposición
Presencia de cáncer de mama.	Valor Energético Total (VET) – Consumo de fibra.
COVARIABLES	
Edad - Nivel socioeconómico - Hábito de fumar - Actividad física - Índice de Masa Corporal (IMC) - Estado ginecológico - Hijos - Lactancia materna.	

DISEÑO METODOLÓGICO

El presente Trabajo de Investigación de Licenciatura (TIL) se enmarca en el proyecto titulado "Epidemiología Ambiental del Cáncer y otras Enfermedades No Transmisibles en Córdoba: Integrando conceptos y métodos para aportar al conocimiento de sus etiologías" (PICT 2012-1019, PICT 2008-1814, FonCyT –ANPCYT, INC 2014-2017, MinSal), que desarrolla el Grupo de Epidemiología Ambiental del Cáncer en el Instituto de Investigaciones en Ciencias de la Salud (CONICET, UNC). El mismo tiene registro en el Ministerio de Salud de Córdoba, fue evaluado por la Comisión de Bioética de la FCM-UNC (Res. N° 237), por el CIEIS del Hospital Nacional de Clínicas (08/2007) y asignado por el CoEIS del MinSal Córdoba, en el marco del Registro 742 (proyecto global del GEACC) del Registro Provincial de Investigaciones en Salud (08/2007)

TIPO DE ESTUDIO

Estudio epidemiológico observacional de tipo caso-control diseñado para los cánceres más incidentes en Córdoba. Particularmente, se indagó el cáncer de mama (C50 CIE-10) (52), en población de mujeres mayores a 18 años, de la provincia de Córdoba, durante el período 2008–2016.

MUESTRA

Se trabajó con una muestra aleatoria de $n=949$ mujeres, existiendo información de $n_1=343$ casos y $n_2=606$ controles, previo consentimiento informado (firmado) por parte de cada participante. Se consideraron casos a mujeres con diagnóstico histopatológicamente confirmado de adenocarcinoma de mama reciente, residentes de la provincia de Córdoba en el período 2008-2016. La muestra incluyó aproximadamente 2 controles por cada caso, para trabajar con una potencia de por lo menos el 80%. Los sujetos controles fueron mujeres que procedían del mismo departamento de residencia que los casos y con edad similar (± 5), que no padecían o hayan padecido alguna enfermedad oncológica, ni otras asociadas al tumor en cuestión, y que no presentaran hábitos alimentarios particulares (por costumbres, salud o creencias religiosas). Fueron criterios de exclusión, para toda la muestra ser menor de 18 años, presentar una neoplasia benigna, o alguna patología asociada al cáncer en estudio, o hábitos alimentarios particulares, o presentar como tumor primario otra localización tumoral diferente de la estudiada.

OPERACIONALIZACIÓN DE LAS VARIABLES

Presencia de cáncer de mama.

Definición conceptual: tumor que se origina cuando las células en el seno comienzan a crecer en forma descontrolada. El tumor es maligno (cáncer) si las células pueden crecer penetrando (invadiendo) los tejidos circundantes o propagándose (metástasis) a áreas distantes del cuerpo. El cáncer de mama ocurre casi exclusivamente en las mujeres (1,4% del cáncer total de mama, en Córdoba, es en hombres (53).

Clasificación: variable cualitativa nominal dicotómica.

Indicadores y/o categorías: ausencia/presencia, según diagnóstico histopatológicamente de adenocarcinoma mamario reciente (no más de tres años previos al momento de realización de la encuesta).

Consumo de fibra.

Definición conceptual: Son sustancias de origen vegetal, hidratos de carbono o derivados de los mismos, excepto la lignina, que resisten la hidrólisis de las enzimas digestivas humanas y llegan intactos al colon, donde algunos pueden ser hidrolizados y fermentados por la microbiota colónica (54). Según su solubilidad en el agua la fibra dietética puede clasificarse en: soluble e insoluble.

Clasificación: variable cuantitativa continua.

Indicadores y/o categorías: se utilizaron la cantidad de gramos (g) totales y discriminados por fibra soluble e insoluble. Asimismo, a partir de los gramos (g) diarios consumidos diferenciados en fibra soluble e insoluble, se construyó el cociente de fibra, correspondiente a la relación fibra insoluble sobre soluble.

Valor Energético Total.

Definición conceptual: energía que consume un organismo, está representado por la tasa metabólica basal (TMB), la actividad física (AF) y la termogénesis inducida por la dieta (TID) (55).

Clasificación: variable cuantitativa continua.

Indicadores y/o categorías: Se utilizó la cantidad de calorías totales por día. Por conveniencia del análisis, se adoptó también categorizar a las calorías totales diarias en tres clases (terciles).

Edad.

Definición conceptual: tiempo que ha vivido una persona (56).

Clasificación: variable cuantitativa continua.

Indicadores y/o categorías: en años.

Nivel socioeconómico.

Definición conceptual: el nivel socioeconómico se basa en la integración y ponderación de distintas variables medidas a través de los siguientes indicadores(57):

- Ocupación del principal sostén del hogar (PSH).
- Nivel educativo del PSH.
- Cantidad de aportantes del hogar.
- Posesión de algunos bienes y servicios (autos, computadora, internet, tarjeta de débito).
- Cobertura de salud

Clasificación: variable cualitativa ordinal.

Indicadores y/o categorías: se analizó como categórica, dividida en 3 niveles: bajo, medio y alto.

Hábito de fumar.

Definición conceptual: aspirar y despedir el humo del tabaco, de forma habitual o reiterada (58).

Clasificación: variable cualitativa nominal dicotómica.

Indicadores y/o categorías: fumadores (sujetos que reportaron ser o haber sido fumadores)/ no fumadores (sujetos que reportaron no haber fumado).

Actividad física.

Definición conceptual: se la define como cualquier movimiento corporal producido por los músculos esqueléticos que exija gasto de energía (59), calculada a partir de la sumatoria de los equivalentes metabólicos (METs) de acuerdo a cada tipo de actividad (60).

Clasificación: variable cuantitativa continua.

Indicadores y/o categorías: se evaluó la actividad física efectuada cinco años previos al diagnóstico de la enfermedad en los casos, y en los cinco anteriores a la realización de la entrevista en los controles. Para ello, se utilizó una versión adaptada del Cuestionario Internacional de Actividad Física, el cual se encuentra estandarizado y validado en diferentes partes del mundo (60). A partir del mismo se evaluó y categorizó el nivel de actividad física usando como criterios los equivalentes metabólicos semanales (METs) de cada tipo de actividad. La variable fue tratada como continua, como así también se la analizó categorizada en leve, moderado y alto.

Índice de Masa Corporal.

Definición conceptual: El índice de masa corporal (IMC) es un índice utilizado frecuentemente para clasificar el sobrepeso y la obesidad en adultos (61). Se calcula dividiendo la masa corporal (kg) por la talla elevada al cuadrado (m^2), a partir de las medidas de peso y talla declaradas.

Clasificación: variable cuantitativa continua.

Indicadores y/o categorías: se clasificó de acuerdo a las categorías determinadas por la OMS: bajo peso ($IMC < 18,5 \text{ kg/ m}^2$), normopeso ($IMC 18,5 - 24,9 \text{ kg/ m}^2$), pre-obesidad ($IMC 25 \text{ kg/ m}^2 - 29,9 \text{ kg/ m}^2$) y obesidad ($IMC > 29,9 \text{ m}^2$) (61).

Estado menopáusico.

Definición conceptual: es el cese permanente de los periodos menstruales de una mujer, diagnosticada de manera retrospectiva, una vez que han transcurrido 12 meses consecutivos de amenorrea, que sucede por atresia de los folículos ováricos y su capacidad de producir estrógenos ante el estímulo de las hormonas folículo-estimulante y luteinizante (62).

Clasificación: variable cualitativa nominal dicotómica.

Indicadores y/o categorías: Si/No. En el grupo de los casos se consideró si las mujeres menstruaban o no al momento del diagnóstico de la enfermedad, y en los controles al momento de la realización de la encuesta; esta diferenciación se debe a que en ocasiones el tratamiento oncológico recibido para hacer frente al cáncer de mama les induce la menopausia.

Presencia de hijos.

Definición conceptual: se refiere a si las mujeres tuvieron hijos o no (nulíparas).

Clasificación: variable cualitativa nominal dicotómica.

Indicadores y/o categorías: no (nulíparas)/ si (tuvo uno o más hijos).

Práctica de Lactancia Materna.

Definición conceptual: La lactancia natural es una forma sin parangón de proporcionar un alimento ideal para el crecimiento y el desarrollo sanos de los lactantes; también es parte integrante del proceso reproductivo, con repercusiones importantes en la salud de las madres (63).

Clasificación: variable cualitativa nominal dicotómica.

Indicadores y/o categorías: se categorizó de acuerdo a la cantidad de meses que las mujeres amamantaron, considerando si las mismas no amantaron a sus hijos o lo hicieron por menos de seis meses; y si lo hicieron durante seis meses o más. Dicha categorización se basa en evidencia científica que estima un efecto protector para el CM cuando la lactancia materna se practica durante seis meses o más (64-66).

TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Se trabajó con la información que dispone el GEACC sobre este estudio caso-control. Se comenzó con el acondicionamiento y puesta a punto de la base de datos, cuya información ya fue recolectada previamente a partir de un cuestionario de frecuencia alimentaria (FFQ). En cada entrevista, y como lo reportan los estudios del equipo citado, se utilizó un cuestionario de frecuencia cuali-cuantitativa de 127 alimentos, métodos de preparación y cocción, junto con un Atlas de Alimentos ambos validados (67). Cabe aclarar, que dicha información alimentaria indaga sobre 5 años antes del diagnóstico (en casos) o de la entrevista (en controles). Otros detalles se encuentran en Tumas et al (68).

La frecuencia de consumo (veces por semana, al mes y al día) y el tamaño de porción, fueron convertidas a cantidades de consumo (g/día y kcal/día) y posteriormente, a gramos de nutrientes consumidos/día. El cuestionario también registró datos personales (edad, sexo, estado civil, características físicas), antropométricos (talla, peso), socio-demográficos, sobre actividad física y consumo de tabaco.

PLAN DE ANÁLISIS DE LOS DATOS

Se usaron herramientas de estadística descriptiva, incluyendo la construcción de tablas, gráficos y medidas resumen, de acuerdo a la naturaleza de las variables. Asimismo, se utilizaron pruebas de hipótesis para la verificación de hipótesis de igualdad de medias en grupos poblacionales, así como tablas de contingencia para la estimación de las asociaciones con las covariables recogidas mediante la estimación del test de chi-cuadrado. Se adoptó un nivel de significación igual a 0,05.

Se construyeron modelos de regresión logística múltiple para analizar la asociación entre la variable respuesta, de naturaleza dicotómica (Presencia /Ausencia de cáncer de mama) y las variables de exposición, incluyéndose covariables de interés.

Posteriormente, y utilizando el enfoque metodológico desarrollado por la Dra. Becaria Coquet en su tesis doctoral (16), se siguieron sus lineamientos generales y se realizó un exhaustivo análisis del sesgo de información o error sistemático por datos faltantes en los métodos de cuantificación de la ingesta de nutrientes, recurriendo al enfoque de imputación múltiple (50) en las covariables identificando el mecanismo de imputación (MCAR, MAR, MNAR). Se ajustaron modelos de regresión logística múltiple (*outcome*, la presencia de CM) y la exposición el VET e ingesta de fibra, ajustando por covariables (edad, BMI, etc.). Se realizó un análisis comparativo entre los resultados derivados del análisis de la base solo con los datos completos (menor n) y el análisis de imputación múltiple (completando por simulación). Para este último, se generaron 20 bases de datos imputadas (50) y se realizó la imputación múltiple de datos en los casos donde exista más del 10% de datos faltantes (51). Se postularon como variables a imputar aquellas asociadas al CM en la bibliografía (lactancia, BMI, actividad física, hábito de fumar, estado ginecológico) y que presenten datos faltantes.

Las covariables imputadas fueron actividad física, estrato socioeconómico y lactancia materna debido a que presentaban importantes porcentajes de datos faltantes (25,61%, 43,10% y 15,38% respectivamente). En el caso de la Actividad Física, la variable fue tratada como continua, a partir del cálculo de los equivalentes metabólicos semanales (MET's).

En la tabla 1 se muestran las distribuciones de frecuencias de las variables de interés y la distribución de los datos faltantes. Se puede observar que, en relación con la

actividad física, el 28,35% de las mujeres fueron clasificadas como sedentarias, no obstante, existe un gran porcentaje de datos faltantes para esta variable.

Respecto al Estrato Socioeconómico, se construyó la variable a partir de 8 ítems planteados por la Asociación Argentina de Marketing (Anexo 1). Es por ello que, al momento de imputar, se imputaron cada una de estas variables que presentaron datos faltantes y luego se realizó el cálculo final del Estrato Socioeconómico. De las 8 variables que componen esa clasificación, fueron imputados los siguientes componentes: obra social, cantidad de aportantes del hogar, computadora, internet, tarjeta de débito, autos; por tener más de un 10% de datos faltantes.

En relación con Lactancia Materna, se categorizó la variable en dos categorías, tomando como punto de corte los 6 meses de lactancia materna. En la tabla 1, se puede observar que, si bien existió un porcentaje de datos faltantes elevado, un 64% de las mujeres reportó haber amamantado 6 meses o más, mientras que un 15% declaró no haberlo hecho.

Además de esta distribución, se realizaron análisis exploratorios adicionales, los cuales indicaron que las variables con datos faltantes (estrato socioeconómico, actividad física y lactancia materna) se encuentran asociadas a diferentes covariables (presencia de cáncer de mama, edad, VET, estrato socioeconómico e IMC), por lo que los patrones de pérdida de datos faltantes parecen asumir un mecanismo de tipo MAR.

Los análisis antes mencionados, se llevaron a cabo mediante el software Stata V14.0 (69).

Tabla 1. Sujetos y datos faltantes: Distribución de frecuencias absolutas y relativas de resultado, exposición y covariables. Córdoba 2008-2016.

	N	%	% de observaciones con datos faltantes							
Total	949	100	Actividad Física	Obra Social	Nº de Proveedores	Computadora	Internet	Tarjeta débito	Autos	Lactancia materna
Presencia CM										
No	606	63,86	28,88	22,94	9,90	10,40	10,40	10,56	11,06	14,19
Si	343	36,14	19,63	19,83	8,16	8,16	8,16	8,16	8,75	17,49
VET										
Categoría 1	278	29,29	34,53	14,39	5,76	6,47	6,47	6,83	7,91	16,19
Categoría 2	298	31,40	26,51	22,82	10,40	10,40	10,40	10,40	10,40	17,11
Categoría 3	373	39,30	18,23	26,54	10,99	11,26	11,26	11,26	11,80	13,40
Cociente fibra										
1	359	37,83	19,22	26,18	9,19	9,47	9,47	9,47	10,31	15,04
2	302	31,82	27,15	25,17	10,93	11,26	11,26	11,26	11,26	17,22
3	288	30,35	31,94	12,85	7,64	7,99	7,99	8,33	9,03	13,89
Edad										
<45 años	162	17,07	29,63	19,14	8,64	8,64	8,64	8,64	9,88	27,78
45-60 años	339	35,72	20,35	24,19	10,32	10,91	10,91	11,21	10,91	12,39
>60 años	448	47,21	28,13	20,98	8,71	8,93	8,93	8,93	9,82	13,17
IMC										
<25 kg/m ²	446	47,60	27,13	24,66	7,85	8,07	8,07	8,07	8,07	18,83
25-30 kg/m ²	303	32,34	25,08	19,47	8,25	8,58	8,58	8,91	9,24	12,87
>30 kg/m ²	188	20,06	22,87	20,21	13,83	13,83	13,83	13,83	15,96	11,70
Desconocido	12	1,26	25	0	16,67	25	25	25	25	8,33
Actividad Física										
Sedentario	269	28,35	-	44,98	8,18	7,81	7,81	7,81	8,18	15,24
Moderado	232	24,45	-	28,02	9,05	9,05	9,05	9,05	9,05	13,36
Vigorosa	205	21,60	-	9,76	21,95	22,44	22,44	22,44	22,93	18,54
Desconocido	243	25,61	-	0,41	0	1,23	1,23	1,65	2,88	14,81
Lactancia Materna										
No	147	15,49	23,81	20,41	4,76	6,12	6,12	6,80	6,80	-
Si	608	64,07	33,22	19,90	9,38	9,54	9,54	9,54	10,20	-
Desconocido	194	20,44	3,09	28,87	12,37	12,37	12,37	12,37	12,89	-
Estado ginecológico										
No	234	24,66	27,35	21,37	8,12	8,12	8,12	8,12	8,12	8,12
Si	664	69,97	26,66	23,64	10,24	10,69	10,69	10,84	10,69	10,84
Desconocido	51	5,37	3,92	0	1,96	1,96	1,96	1,96	1,96	2,04

RESULTADOS

En este apartado se exponen los resultados obtenidos, en una primera instancia, a partir del análisis exploratorio para lograr caracterizar la muestra constituida por $n_1=343$ casos y $n_2=606$ controles. En una segunda instancia, se presentan los resultados del análisis inferencial, mostrando lo obtenido para afrontar el sesgo de información por datos faltantes presentes en las estimaciones de riesgo de CM, asociadas a la ingesta de fibra y el VET, y estimando así, el rol que ejercen dichos componentes en la ocurrencia del CM.

ANÁLISIS EXPLORATORIO

En relación a la caracterización del valor energético total de la dieta se puede observar que el promedio diario obtenido fue igual a 2796,57 kcal con un desvío estándar (DE) igual a 1142,75, siendo que un 50% de la muestra consume habitualmente hasta 2594 kcal por día. Un 25% de las mujeres mostró un consumo menor o igual a 2065 kcal y un 75% un VET diario menor o igual a 3296 kcal. La Figura 1 ilustra la distribución de frecuencias de dicho consumo.

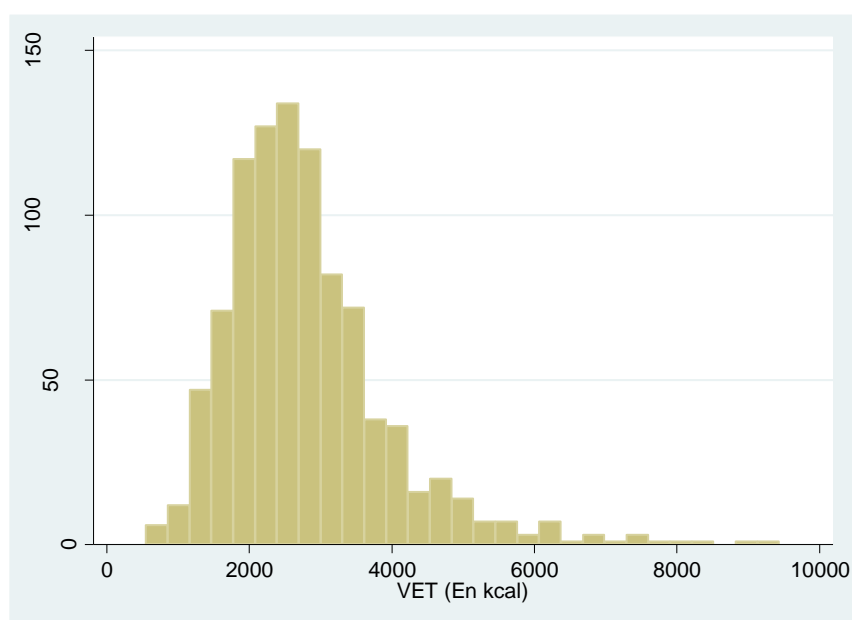


FIGURA 1. Distribución de frecuencias del VET

Se indagó sobre características socio-culturales, hábito de fumar y de desayunar, estudios, trabajo y actividad física en los sujetos bajo estudio. Cabe destacar que para el VET diario existieron diferencias significativas en relación al trabajo ($p=0,02$), donde se observó que quienes no trabajan presentaban un VET promedio por día igual a 3088,12 kcal, superior al consumo de quienes si trabajan (Tabla 2). Asimismo, se observaron diferencias significativas, con 90% de confianza, para el VET en relación al estrato socioeconómico.

Respecto a las variables reproductivas, entre ellas el estado ginecológico (definido por la presencia / ausencia de menopausia) y lactancia materna (teniendo en cuenta aquellas mujeres que lactaron durante 6 meses o más respecto de las que no), no se hallaron diferencias significativas entre aquellas mujeres que presentaban CM y quienes no (Tabla 2).

Tabla 2. Características socio-culturales, hábito de fumar y desayunar, estudios trabajo y actividad física en relación al VET diario, según casos y controles, Córdoba 2008-2016.

COVARIABLE	VET: Media (desvío estándar)	P-valor
Estrato Socioeconómico		
<i>Bajo</i>	2705,63 (1082,21)	0,06
<i>Medio</i>	2907,50 (1144,01)	
<i>Alto</i>	2654,77 (1038,14)	
Trabajo		
<i>No</i>	30,88,12 (1485,55)	0,02
<i>Si</i>	2756,81 (1086,51)	
<i>Jubilado o pensionado</i>	2732,18 (1104,86)	
<i>Jubilado y trabaja</i>	2767,95 (912,29)	
Estudios		
<i>Sin instrucción</i>	2681,94 (1197,05)	0,22
<i>Primaria incomplete</i>	2818,67 (1134,74)	
<i>Primaria complete</i>	2900,99 (1181,95)	
<i>Secundaria incomplete</i>	2599,29 (1088,51)	
<i>Secundaria complete</i>	2841,59 (1071,45)	
<i>Terciario o universitario</i>	2711,52 (1177,82)	
Nivel de actividad física		
<i>Leve</i>	2962,21 (1087,76)	0,21
<i>Moderado</i>	2797,01 (997,96)	
<i>Alto</i>	2947,69 (1322,82)	
Hábito de fumar		
<i>Si</i>	2814,49 (57,86)	0,68
<i>No</i>	2783,45 (48,35)	
Hábito de desayunar		
<i>Si</i>	2787,56 (38,76)	0,78
<i>No</i>	2820,39 (120,69)	
Estado menopáusico		
<i>Si</i>	2876,18 (84,27)	0,07
<i>No</i>	2724,31 (40,19)	
Lactancia materna		
<i>Si</i>	2842,05 (49,29)	0,39
<i>No</i>	2770,01 (70,29)	

Al estimar la correlación entre la presencia de CM y la edad, esta fue significativa tanto para los casos como para los controles ($p=0,011$), observándose una disminución del promedio de kilocalorías (kcal) consumidas por día a medida que aumenta la edad, tal como se puede observar a partir del diagrama de la dispersión en la Figura 2.

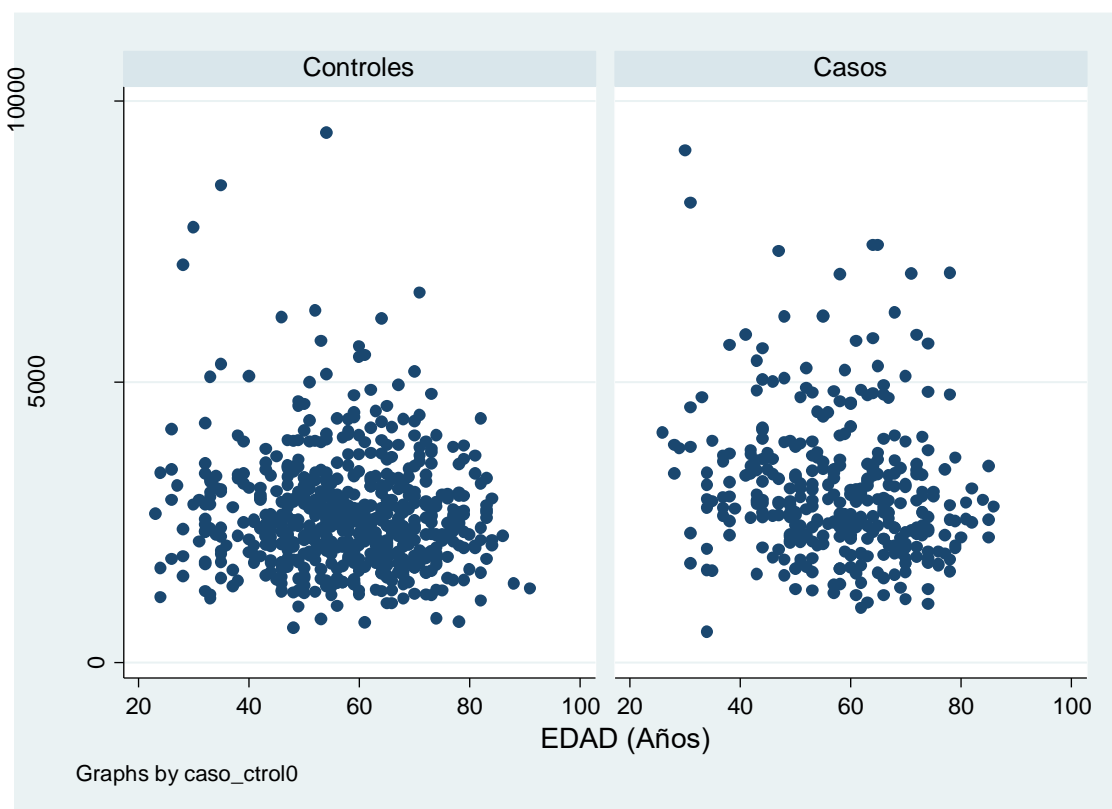


FIGURA 2. Diagrama de dispersión para edad y VET. Córdoba 2008-2016.

En cuanto a las variables antropométricas indagadas, que incluyeron peso, talla, circunferencia de cintura e IMC, estas no se correlacionaron con el VET habitual de la dieta de las mujeres estudiadas (Figura 3).

“Aporte metodológico para el estudio de la red causal del cáncer de mama: el sesgo de información y su impacto en la estimación de los efectos de componentes de la dieta”

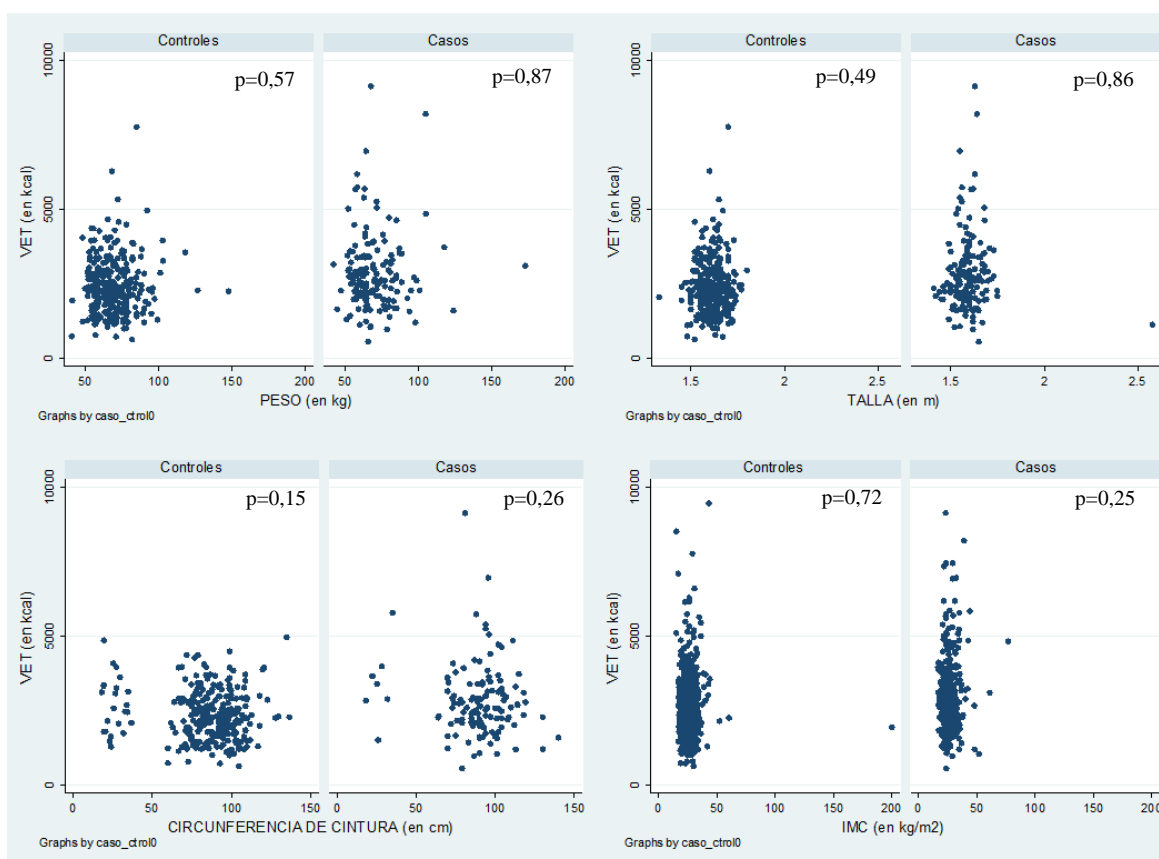


FIGURA 3. Diagrama de dispersión para las variables antropométricas y VET. Córdoba 2008-2016.

En la tabla 3 se muestran las medidas resúmenes y la distribución de frecuencias del VET entre los casos y los controles. Cabe aclarar que el 50% de los casos perteneció a la categoría de VET más alta y que la media entre los grupos de estudios fue significativamente diferente. Estos resultados estarían indicando que el promedio del VET para los casos es mayor al de los controles ($p=0,001$).

TABLA 3. Distribución de frecuencias y medidas resúmenes del VET según casos y los controles, Córdoba 2008-2016.

VARIABLE	CASO	CONTROLES	P-valor
	n (%)	n (%)	
VET Categorizado			
Tercil 1 (548,73; 2148,48 kcal)	74 (21,6)	204 (33,6)	0,001
Tercil 2 (2151,41; 2860,22 kcal)	96 (27,9)	202 (33,3)	
Tercil 3 (2862,76; 9438,47 kcal)	173 (50,5)	200 (33,1)	
VET	3082,93 (1270, 81)*	2634,48 (1029,85)*	0,001

*Media (desvío estándar)

Respecto a la fibra alimentaria, el consumo promedio diario de fibra obtenido fue igual a 21,53 gramos (DE igual a 8,89 gramos), siendo que un 50% de la muestra consume habitualmente hasta 20,06 g por día. El 25% de las mujeres presentó un consumo de fibra diario menor o igual 15,43 g y el 75% de la muestra consume habitualmente hasta 25,99 gramos por día. La Figura 4 ilustra dicho consumo. Cabe aclarar que, de acuerdo a su clasificación, se encuentra la fibra soluble, cuyo consumo promedio fue de 5,94 gramos (DE= 2,48); y la insoluble, con un consumo promedio igual a 15,59 gramos (DE=6,73).

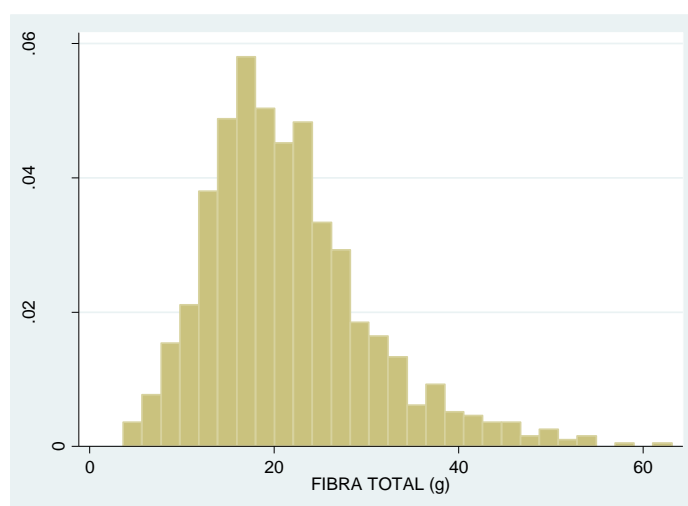


FIGURA 4.

*de frecuencias del consumo total de fibra.
Córdoba 2008-2016.*

Distribución

Respecto de las características socio-culturales, hábito de fumar y de desayunar, estudios, trabajo y actividad física en los sujetos bajo estudio, los resultados obtenidos indican que el consumo de fibra diario se asoció solamente al nivel de actividad física ($p=0,02$), siendo mayor el consumo de promedio de fibra en aquellos que realizan actividad física intensa (con una media de 22,77 g) en relación con las que mantenían un nivel leve de actividad física (con una media de 29,80 g), tal como se muestra en la tabla 3.

Por otra parte, las variables reproductivas como el estado ginecológico y lactancia materna), no se encontraron asociadas al consumo de fibra (Tabla 4).

Tabla 4. Características socio-culturales, hábito de fumar y desayunar, estudios trabajo y actividad física en relación al consumo de fibra, según casos y controles, Córdoba 2008-2016.

VARIABLE	Consumo de Fibra: Media (desvío estándar)	P-valor
Estrato Socioeconómico		
<i>Bajo</i>	22,33 (9,49)	0,18
<i>Medio</i>	22,91 (22,91)	
<i>Alto</i>	21,14 (8,34)	
Trabajo		
<i>No</i>	22,22 (9,36)	0,55
<i>Si</i>	21,61 (8,95)	
<i>Jubilado o pensionado</i>	21,08 (8,50)	
<i>Jubilado y trabaja</i>	22,58 (9,60)	
Estudios		
<i>Sin instrucción</i>	22,57 (6,86)	0,37
<i>Primaria incomplete</i>	20,86 (8,39)	
<i>Primaria complete</i>	20,91 (7,69)	
<i>Secundaria incomplete</i>	21,16 (9,65)	
<i>Secundaria complete</i>	22,66 (9,56)	
<i>Terciario o universitario</i>	21,92 (9,63)	
Actividad física		
<i>Leve</i>	20,80 (7,61)	0,02
<i>Moderado</i>	22,43 (8,78)	
<i>Alto</i>	22,77 (9,93)	
Hábito de fumar		
<i>Si</i>	21,27 (9,11)	0,43
<i>No</i>	21,72 (8,73)	
Hábito de desayunar		
<i>Si</i>	21,71 (0,31)	0,07
<i>No</i>	20,09 (0,85)	
Estado menopáusico		
<i>Si</i>	21,62 (0,36)	0,54
<i>No</i>	21,21 (0,53)	
Lactancia materna		
<i>Si</i>	21,37 (0,37)	0,11
<i>No</i>	22,41 (0,58)	

La caracterización también incluyó la estimación de diferencias por rango de edad, sin presentar asociación en los casos ($p=0,12$), ni en los controles ($p=0,52$), tal como se puede observar a partir del diagrama de la dispersión en la Figura 5.

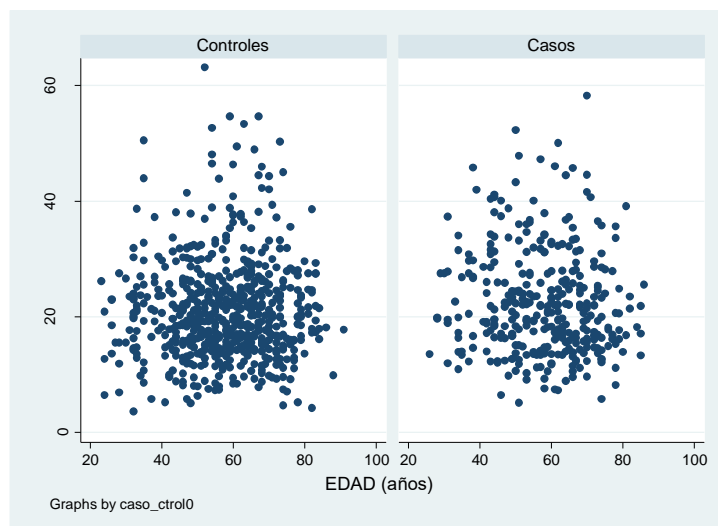


FIGURA 5. Diagrama de dispersión para edad y fibra total. Córdoba 2008-2016.

Respecto a las variables antropométricas analizadas, cabe mencionar que si bien la circunferencia de cintura no mostró correlación con el consumo habitual de fibra para los controles ($p=0,11$), en el grupo de casos fue diferente ($p=0,04$) observándose una disminución del consumo de fibra total a medida que aumenta la circunferencia de cintura; tal como ilustra la Figura 6. El resto de las variables antropométricas no presentaron asociación. (Figura 7).

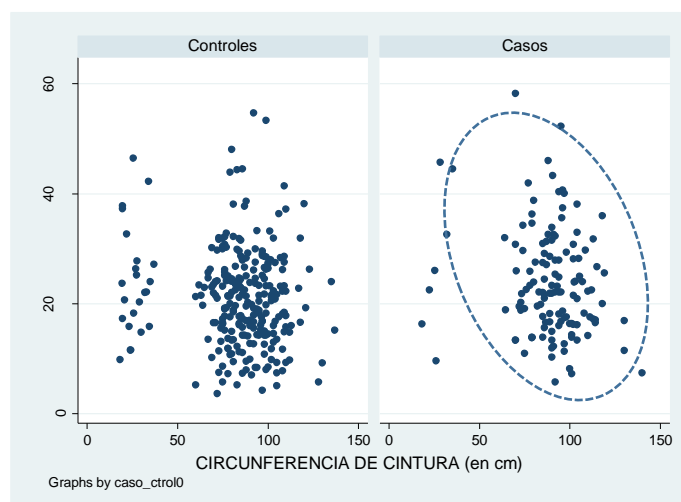


Figura 6. Diagrama de dispersión para la circunferencia de cintura y fibra total. Córdoba 2008-2016.

“Aporte metodológico para el estudio de la red causal del cáncer de mama: el sesgo de información y su impacto en la estimación de los efectos de componentes de la dieta”

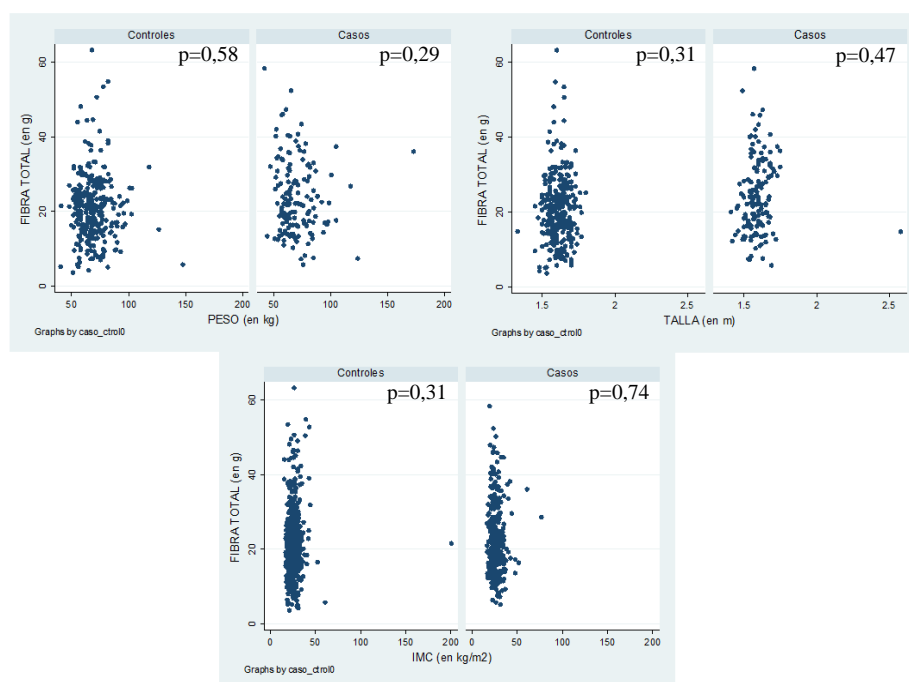


Figura 7. Diagrama de dispersión para las variables antropométricas y fibra total

Al analizar las variaciones entre casos y controles, se encontraron diferencias significativas respecto al consumo diario de fibra total ($p=0,02$), tal como se puede observar en la Tabla 5 y la distribución de frecuencias en la figura 8.

En cuanto, al consumo habitual de fibra insoluble, no se observaron diferencias significativas ($p=0,09$) entre los grupos en estudio. Por el contrario, se hallaron diferencias significativas en el consumo diario de fibra soluble ($p=0,001$) (Tabla 5), donde el promedio del consumo diario de fibra soluble para los casos es significativamente mayor al de los controles ($p=0,001$).

TABLA 5. Medidas resúmenes de fibra según casos y los controles, Córdoba 2008-2016

VARIABLE	CASOS	CONTROLES	P-valor
	Media (desvío estándar)	Media (desvío estándar)	
Fibra total	22,37 (9,11)	21,06 (8,74)	0,0287
Fibra insoluble	16,07 (6,98)	15,32 (6,58)	0,0983
Fibra soluble	6,29 (2,51)	5,73 (2,44)	0,0008
Cociente fibra	2,59 (0,67)	2,74 (0,66)	0,0019

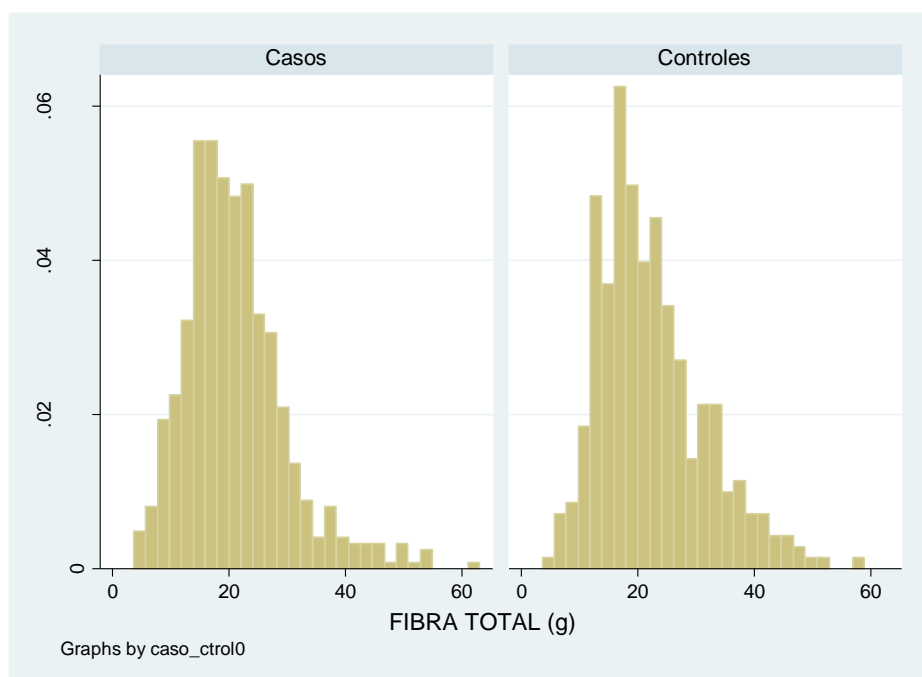


FIGURA 8. Distribución de frecuencias del consumo de fibra total por casos y controles. . Córdoba 2008-2016.

Por otra parte, se calculó el cociente entre el consumo correspondiente a la fibra insoluble y el de la fibra soluble, obteniéndose un promedio igual a 2,69 con un desvío estándar de 0,67; siendo que, para el 75% de la muestra, dicho cociente fue igual 3,06. Asimismo, se hallaron diferencias significativas respecto al grupo de los casos y los controles ($p=0,001$), siendo el coeficiente de estos mayor al de los casos, tal como se puede observar en la Tabla 5.

SESGO DE INFORMACIÓN Y ROL DE LA DIETA EN LAS ESTIMACIONES DE RIESGO DE CÁNCER DE MAMA

A continuación se exponen los resultados de la estimación de los modelos de regresión logística, ajustados con el fin de evaluar la relación entre la presencia de CM y la ingesta de fibra, como así también con el VET de la dieta, incluyendo como covariables a la edad, el estado ginecológico, la actividad física, el estrato socioeconómico, el IMC y la lactancia materna. Se generó un modelo para cada una de las variables de exposición (fibra y VET) y se analizó la base incluyendo la información “completa” de los sujetos, representando, aproximadamente, un tercio de las mujeres estudiadas, ya que de la base total, solo 289 de estas presentaban completitud (datos completos) en todas las variables incluidas en el análisis e indagadas en las encuestas. Los resultados indicaron que las mujeres con un VET habitual entre 2151 y 2860 (tercil 2) presentaron dos veces más chances de presentar CM que aquellas con valores de VET habitual pertenecientes al primer tercil, es decir, promedios de VET comprendidos entre 549 y 2148 kcal (IC 95%: 1,51-3,72; OR: 0,64). El análisis del consumo de fibra dietética demostró un efecto protector, por lo cual su consumo se asocia con una disminución en un 40% del riesgo de padecer CM (IC 95%: 0,041-0,91 OR: 2,4), tal como se puede observar en la Tabla 6. Respecto a las covariables de ambos modelos de riesgo, no se obtuvieron valores significativos.

Posteriormente, la metodología aplicada en este estudio caso-control consistió en la aplicación de esquemas de Imputación Múltiple de datos faltantes en covariables, asumiendo que los mismos responden a un patrón de datos aleatorios, es decir, MAR (*Missing At Random*). A partir de esto, y como se expresó en la sección anterior, siguiendo la metodología propuesta por Becaria Coquet (2017) se construyeron modelos de imputación para cada una de las variables imputadas, teniendo en cuenta las variables incluidas en los análisis anteriores y variables predictivas de los valores faltantes correspondientes. Dicho análisis, incluye la regresión logística donde la presencia de la patología constituye el *outcome* y el VET o el cociente de fibra la exposición, en los cuales la muestra alcanzó los 848 sujetos estudiados, es decir, 89% de las mujeres de la base de datos completa.

Al comparar los resultados derivados del análisis de datos completos y el análisis luego de la aplicación de esquemas de Imputación Múltiple, tal como se muestra en la

Tabla 6, se puede observar que, en ambos, existió un efecto promotor de consumo del valor energético elevado y protector de la variable de la relación fibra soluble/ fibra insoluble. Sin embargo, con la aplicación de Imputación Múltiple, el efecto de otras variables como IMC y el estado ginecológico se evidenció solo cuando se llevó a cabo dicha metodología, observándose un mayor riesgo de padecer CM en aquellas mujeres con obesidad (IC 95%: 1,04–2,31; $p=0,02$) y postmenopáusicas (IC 95% 1,12–2,84; $p=0,01$). Cabe señalar que mediante el proceso de imputación, se obtuvieron intervalos de confianza más precisos para las variables de exposición (cociente fibra y VET), tal como se puede observar en la tabla 6.

Cuando se consideró la medida diagnóstica de imputación (IVR, incremento de la varianza relativo) del modelo final de riesgo, la misma fue igual 0,03, para el VET, y de 0,04 para el cociente fibra; indicando que la variabilidad de la muestra estimada para este conjunto de variables fue sólo un 3% y un 4% respectivamente mayor de aquella que se hubiese observado en el caso que las covariables hubiesen presentado datos completos. Por lo tanto, se puede concluir que se observó un mayor riesgo de presentar cáncer de mama en mujeres que presentaron mayor VET, un menor cociente fibra insoluble/soluble, obesidad y que se encuentran en la menopausia. Estos resultados ponen en evidencia que la imputación múltiple es relevante en campo de la epidemiología nutricional del cáncer, para mejorar la precisión de las estimaciones, y aportar resultados más confiables para direccionar acciones en salud pública.

Tabla 6. Estimación de medidas de asociación (ORs) para el consumo de fibra y VET con el cáncer de mama. Análisis de casos completos y de imputación múltiple de datos. Estudio caso-control. . Córdoba 2008-2016.

	Análisis de casos completos (n=289 ; 30,4 %)			Imputación múltiple (n=848; 89,35%)			% Incremento del error estándar
	Odds Ratio	95% CI	p valor	Odds Ratio	95% CI	p valor	
MODELO A							
Cociente de fibra	0,607	[0,407; 0,905]	0,014	0,641	[0,505; 0,812]	0,001	0,07
Edad	1,005	[0,977; 1,035]	0,718	0,984	[0,969 1,001]	0,053	0,33
IMC							
<i>Bajo y normopeso</i>	REF	-	-	-	-	-	-
<i>Sobrepeso</i>	0,909	[0,508; 1,629]	0,750	1,186	[0,842; 1,669]	0,329	0,05
<i>Obesidad</i>	1,594	[0,812; 3,127]	0,175	1,576	[1,064; 2,336]	0,023	0,11
Actividad física	0,999	[0,999; 1,001]	0,965	1,001	[0,999; 1,001]	0,543	9,25
Estrato socioeconómico							
<i>Alto</i>	REF	-	-	-	-	-	-
<i>Medio</i>	1,497	[0,826; 2,713]	0,183	1,251	[0,871; 1,795]	0,226	4,42
<i>Bajo</i>	1,852	[0,901; 3,811]	0,094	0,984871	[0,646; 1,501]	0,943	2,85
Lactancia materna							
<i><6 meses</i>	REF	-	-	-	-	-	-
<i>≥6 meses</i>	0,781	[0,473; 1,291]	0,335	0,818	[0,606; 1,104]	0,189	2,19
Estado menopáusico							
<i>No</i>	REF	-	-	-	-	-	-
<i>Si</i>	1,096	[0,531; 2,261]	0,804	1,803	[1,139; 2,852]	0,012	0,04

MODELO B							
VET							
<i>Categoría 1</i> (548,7; 2148,5 kcal)	REF	-	-	-	-	-	-
<i>Categoría 2</i> (2151,4; 2860,2 kcal)	1,088	[0,575; 2,061]	0,795	1,301	[0,882; 1,918]	0,184	0,06
<i>Categoría 3</i> (2862,76; 11992 kcal)	2,069	[1,151; 3,719]	0,015	2,433	[1,043 ; 2,305]	0,030	0,06
Edad	1,008	[0,979;1,0376]	0,605	0,9868042	[0,971 ; 1,003]	0,107	0,34
IMC							
<i>Bajo y normopeso</i>	REF	-	-	-	-	-	-
<i>Sobrepeso</i>	0,961	[0,536; 1,721]	0,892	1,169	[0,828; 1,65]	0,374	0,05
<i>Obesidad</i>	1,743	[0,888; 3,421]	0,106	1,551	[1,043; 2,305]	0,030	0,10
Actividad física	0,999	[0,999; 1,001]	0,924	1,001	[0,999; 1,001]	0,591	8,73
Estrato socioeconómico							
<i>Bajo</i>	REF	-	-	-	-	-	-
<i>Medio</i>	1,356	[0,753; 2,442]	0,310	1,205	[0,839 ; 1,731]	0,311	4,10
<i>Alto</i>	1,764	[0,857; 3,634]	0,124	0,963	[0,632; 1,467]	0,860	2,83
Lactancia materna							
<i>No</i>	REF	-	-	-	-	-	-
<i>Si</i>	0,803	[0,486; 1,327]	0,392	0,800126	[0,591 ; 1,084]	0,150	2,54
Estado menopáusico							
<i>No</i>	REF	-	-	-	-	-	-
<i>Si</i>	1,138	[0,551; 2,355]	0,726	1,786	[1,124 ; 2,839]	0,014	0,04

DISCUSIÓN

El cáncer es uno de los principales problemas de salud pública a nivel mundial, ya que constituye la segunda causa de muerte tanto en el mundo, como en América (4,70). Específicamente, en América Central y América Latina la mortalidad por CM ha ido aumentando en las últimas décadas (32). En este sentido las transiciones demográfica, epidemiológica, sanitaria y nutricional, constituyen procesos sociales subyacentes, que pueden condicionar la salud, enfermedad y atención de las poblaciones, derivando en cambios en los estilos de vida (16,32,71). Entre los cánceres más frecuentes se halla el CM (26) y en estrecha relación con estos procesos, dicha patología es cada vez más frecuente en las mujeres de Argentina como así también en la provincia de Córdoba; representando el 25% del total de todos los tumores, y la primera causa de muerte por cáncer en población femenina (6,71).

Reportes epidemiológicos demuestran que la alimentación es un importante factor vinculado a la ocurrencia del cáncer, reconociéndose que el 35% de los tumores malignos están asociados con factores alimentarios, y serían prevenibles mediante una alimentación y nutrición apropiadas (5,9,10,15,16). Es por ello, que a través del presente trabajo se caracterizó la ingesta de fibra y el valor energético total de la dieta, como así también las características y hábitos de vida de las mujeres que forman parte del estudio caso control de CM en Córdoba. Asimismo, se estimó el rol de dichos componentes dietarios en la ocurrencia del CM. Los resultados obtenidos demostraron un mayor riesgo de presentar CM en mujeres con un VET elevado, un menor cociente fibra insoluble/soluble, obesidad y que se encuentran en la post-menopausia. Se ha reportado que en Argentina, y más específicamente en Córdoba, dietas con alta densidad calórica, elevada proporción de carnes procesadas, alimentos refinados, grasas y alcohol, se asocian a un incremento del riesgo de desarrollar CM; mientras que una alimentación que incluya importantes cantidades de vegetales y frutas, con alto contenido de micronutrientes y fibra, se encuentra relacionada a una disminución del riesgo (5,12). Datos epidemiológicos afirman esta asociación de dietas ricas en fibra con una reducción del riesgo de CM, principalmente en mujeres posmenopáusicas. Dichos resultados son concordantes con los obtenidos en el presente estudio, donde se evidenció que la fibra ejerce un papel protector sobre la ocurrencia CM, principalmente en mujeres postmenopáusicas (40,72,73). Al analizar la variable VET, estudios sugieren que

poblaciones con dietas de elevada densidad calórica se asocian a una morbi-mortalidad por CM (43), coincidiendo con lo observado en el presente trabajo. En este sentido, es importante destacar el rol del IMC, ya que se ha reportado que la densidad calórica de la dieta se asocia positivamente con el mismo (43), lo cual establece que la obesidad se relaciona con un riesgo mayor de ocurrencia y mortalidad por CM (42,74,75).

En el presente estudio se pudo identificar un mayor riesgo de padecer CM en aquellas mujeres menopáusicas, lo cual tiene concordancia con estudios que describen el papel relevante que juegan los estrógenos en la aparición del CM, en donde la menopausia después de los 54 años de edad incrementa el riesgo de padecerlo (75,76). Estos hallazgos constituyen un aporte al conocimiento de la red causal del CM en la provincia de Córdoba.

La existencia de valores ausentes es frecuente en trabajos de investigación en salud pública, e ignorarlos conlleva a la pérdida de potencia del estudio y a la obtención de estimadores ineficientes y posiblemente sesgados (46,77,78). Por lo tanto, la presencia de datos faltantes, es una situación a la que permanentemente se enfrentan investigadores y tomadores de decisiones, ante la cual, habitualmente, dichos datos faltantes no son reportados y, si se mencionan, no se explicita la manera en que fueron enfrentados (46,47). En los estudios epidemiológicos caso-control retrospectivos, tal como es el caso de este trabajo, el sesgo de información se presenta muy frecuentemente (15,16). Por lo que, el presente estudio permitió llevar a cabo un estudio exhaustivo del sesgo de información derivado de datos faltantes presente en este estudio epidemiológico caso-control de cáncer de mama en Córdoba, construyéndose un marco metodológico para su valoración. Si bien, no hay ninguna forma totalmente satisfactoria para el manejo de los datos faltantes (46), la imputación múltiple (IM), mediante la creación de varios conjuntos de datos imputados plausibles diferentes y la combinación adecuada de los resultados obtenidos de cada uno de ellos (47), constituye el método más adecuado para reponer información omitida, (79) ya que dichos procedimientos de IM generan buenos resultados, aún con porcentajes de omisión del 30, 40 o 50% (47,49). No obstante, utilizar algún procedimiento inapropiado puede introducir sesgos en el valor de los estimadores y en su error estándar, al tiempo que podría distorsionar la potencia de las pruebas de hipótesis (80), lo que sugiere reflexionar acerca de la mejor manera de obtener estimadores que generen inferencia válida a partir de datos imputados (47). Aquí resulta, relevante un adecuado tratamiento de los datos ya que, en ocasiones no se aplican las metodologías adecuadas para enfrentar dicho sesgo (15,16), lo cual conduce a una pérdida

de precisión y, en consecuencia, a una incorrecta interpretación de los resultados (77,83). Es por ello, que en este caso antes de proceder a la imputación de datos se evaluó el patrón de datos faltantes, es decir, si correspondía a *Missing Completely At Random* (MCAR), *Missing At Random* (MAR) y *Noting At Random* (NMAR), entendiendo que el tipo de datos faltantes en el conjunto de datos determina el método apropiado para usar en el manejo de los mismos antes de que comience un análisis estadístico formal (16).

La metodología consta de varias etapas, y en cada simulación se analizan la matriz de datos completos a partir de métodos estadísticos convencionales y posteriormente se combinan los resultados para generar estimadores robustos, su error estándar e intervalos de confianza (47). Posteriormente, al comparar los resultados obtenidos de un análisis de datos completos con los obtenidos luego de aplicar Imputación Múltiple de datos, se obtienen diferencias en los p-valores (15,16,82). El presente trabajo, en términos generales, siguió los lineamientos desarrollados por Becaria Coquet (16); para ello se construyeron, en un primer lugar, modelos de regresión logística múltiple para analizar la asociación entre la variable respuesta, de naturaleza dicotómica (Presencia /Ausencia de cáncer de mama) y las variables de exposición, incluyéndose covariables de interés. Y luego, se llevó a cabo el análisis de imputación múltiple en donde los datos faltantes asumían un patrón de datos faltantes de tipo MAR. Dicha imputación se aplicó sobre aquellas variables cuya falta de datos era mayor al 10%, considerando que la Imputación Múltiple puede también ser beneficiosa para la estimación de los coeficientes de covariables relativamente completas en presencia de otras covariables incompletas (83). Entre estas covariables imputadas se identificaron: Actividad Física, Estrato Socioeconómico y Lactancia Materna debido a que presentaban importantes porcentajes de datos faltantes (25,61%, 43,10% y 15,38% respectivamente).

Al comparar los modelos de regresión logrados a partir de los datos completos con aquellos que surgen de la Imputación Múltiple, se observó que ambos enfoques mostraron el efecto promotor del valor energético total elevado y el efecto protector de la fibra, respecto a la ocurrencia de cáncer de mama, imputación los p-valores fue más significativo. No obstante, los efectos de otras variables, como el Índice de Masa Corporal y el estado menopáusico solo se evidenciaron luego del uso del esquema de Imputación Múltiple propuesto. Esto se torna valioso porque identifica covariables importantes en la red causal del cáncer las cuales no se hubieran interpretado sin la modificación del método estadístico a utilizar en la etapa de modelación.

Es por todo lo antes expuesto, que la metodología de Imputación Múltiple de datos aplicada contribuyó a mejorar la validez de las estimaciones de riesgo derivadas de estudios epidemiológicos sobre enfermedades crónicas, como el cáncer, mostrando beneficios al considerar mayor información en los análisis. Dicha metodología apuntó hacia la obtención de estimaciones confiables derivadas de los estudios caso-control, y observacionales en general, con el fin de obtener recomendaciones alimentarias válidas y específicas, en especial para diferentes grupos poblacionales en riesgo.

Probablemente, el proceso de imputación sigue siendo desconocido para muchos profesionales no estadísticos, que le atribuyen una alta complejidad, y quizás un objetivo que no es exactamente el que persigue. Sin embargo, resulta oportuno que los investigadores realicen esfuerzos para llevar adelante metodologías estadísticas adecuadas para el tratamiento del sesgo de información, de manera tal que se obtengan mejoras en la validez de los resultados de la investigación médica y una reducción en el desperdicio de recursos causado por la falta de datos, mejorando así la calidad de las producciones científicas. (15,16,50)

Por otra parte, es necesario hacer mención a algunas limitaciones de este trabajo, las cuales se detallan a continuación. Es posible que resulte oportuno aplicar un enfoque de patrones alimentarios, considerando que los nutrientes, en este caso la fibra, no se consumen de forma individual o aislada, sino como parte de un determinado patrón alimentario, dicho enfoque ha ganado interés durante últimos años en el campo de la Epidemiología Nutricional, dado que permite la caracterización de la dieta de manera integral y extrapolable a recomendaciones alimentarias (5,15,16). Otra debilidad de este estudio se relaciona con las dificultades operacionales que supone el mantenimiento de conjuntos de datos múltiples, tal como afirman Rao y Shao (1992), quienes plantean que las agencias estadísticas prefieren, en general, el uso de la imputación simple. Este problema se hace aún más grave en el caso de muestras con un tamaño muestral muy elevado, como por ejemplo en el caso de muestras nacionales (78).

Debido a que el CM constituye uno de los tumores más prevalentes en nuestro país y más específicamente en la provincia de Córdoba, dicha investigación constituye un aporte significativo para la red causal del CM, colaborando en su prevención, como así también, en la promoción de la salud, al traducir dichas contribuciones en recomendaciones alimentarias poblacionales. En este contexto, cabe destacar también la necesidad de continuar con el trabajo de investigación en el escenario de la epidemiología

“Aporte metodológico para el estudio de la red causal del cáncer de mama: el sesgo de información y su impacto en la estimación de los efectos de componentes de la dieta”

nutricional del cáncer, en donde puede resultar provechoso aplicar el método de la imputación múltiple en aquellos estudios que consideren el componente alimentario, de manera tal que se obtengan estimaciones de calidad y que por lo tanto, den lugar a un accionar oportuno en el ámbito de la salud de las poblaciones.

CONCLUSIONES

Considerando que el CM constituye uno de los cánceres más relevantes debido a su elevada incidencia y mortalidad, tanto en el mundo como en nuestro país e incluso en nuestra provincia, fue de interés arribar resultados significativos tanto desde el punto de vista nutricional como metodológico, logrando combinar positivamente ambas cuestiones y obteniendo aportes útiles e interesantes para ser aplicados. Puntualmente, se construyó un marco metodológico para el análisis de datos faltantes sustentado en métodos de inferencia estadística; en el cual se estimó la manera en que el sesgo de información, provocado por la presencia de datos faltantes, afecta a las estimaciones de riesgo para el CM, asociadas a la alimentación, más específicamente, en relación con la ingesta de fibra y el valor energético total de la dieta.

En relación con la hipótesis planteada, se logró verificar que el diseño del análisis de sesgo de información y su cuantificación permite obtener estimaciones de riesgo más precisas del VET y el consumo de fibra con respecto a las estimaciones clásicas para la ocurrencia de riesgo de CM, encontrándose además, valoraciones diferentes a las estimaciones clásicas al analizar las covariables del modelo de riesgo.

Por lo tanto, cabe mencionar que entre los hallazgos más relevantes se incluyen:

- ✚ Se comprobó que dietas con VET elevado constituyen un escenario de exposición que favorece la presencia de CM.
- ✚ Se verificó que la ingesta de fibra alimentaria cumple un rol protector en la ocurrencia de CM.
- ✚ Resultó relevante el análisis de covariables del modelo de riesgo a saber: se evidenció que tanto la obesidad como la etapa post-menopáusica en las mujeres se encuentran asociadas a un mayor riesgo de padecer CM.
- ✚ La falta de datos es frecuente en estudios epidemiológicos nutricionales, en donde se pudo demostrar que la imputación múltiple de datos constituye una herramienta adecuada para el tratamiento del sesgo de información, derivando en estimaciones de mayor validez e identificación de covariables.

Es decir que, los resultados obtenidos demostraron que el riesgo de presentar CM es mayor en aquellas mujeres con un VET elevado, un menor cociente fibra insoluble/soluble, obesidad y que se hallan en la etapa post-menopausia. Estas

estimaciones constituyen un aporte valioso a la red causal de CM en relación con aspectos nutricionales. A la vez, pone en evidencia el rol protagónico de la alimentación en la ocurrencia de dicho cáncer, resultando oportuno y necesario continuar con investigaciones en el campo de la epidemiología nutricional, con un adecuado marco metodológico, de manera tal que se logre arribar a conclusiones sustentadas en evidencia empírica sólida, para fundamentar políticas en materia de salud pública que incidan en la calidad de vida de las comunidades y den lugar a la promoción de la salud.

Resulta evidente el rol protagónico que adquiere la alimentación en los eventos de salud, como futura Licenciada en Nutrición, considero sumamente importante trabajar en la producción científica de recomendaciones nutricionales de calidad que logren valorar la alimentación a partir de mediciones confiables. Frente a este escenario, cobra vital importancia la necesidad del desarrollo de metodologías que sean de utilidad y se encuentren al alcance de los profesionales de la salud. Y a partir de ello, implementar dichas recomendaciones en el ámbito de la salud pública, desde la promoción y prevención para el abordaje de las ECNT; poniendo énfasis en la promoción de la salud de las mujeres, y mejorando su calidad de vida.

BIBLIOGRAFÍA

1. Herrán F, Oscar Fernando, Prada G, Gloria Esperanza, & Ardila L, Maria F. Ingesta usual de macronutrientes y energía en Bucaramanga, Colombia: Análisis de registros de consumo 1998-2003. *Revista chilena de nutrición*. 2007; 34(4), 307-319.
2. Martínez-Montañez OG, Uribe-Zúñiga P, Hernández-Ávila M. Políticas públicas para la detección del cáncer de mama en México. *Salud Pública de México*. 2009; 51: 350-60.
3. Tumas N, Pou SA, Díaz MP. Inequidades en salud: análisis sociodemográfico y espacial del cáncer de mama en mujeres de Córdoba, Argentina. *Gaceta Sanitaria*. 2017; 31(5):396-403.
4. Organización Mundial de la Salud (OMS). Cáncer. 2018 [Internet]. [Citado 1 de abril de 2019]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/cancer>.
5. Pou SA, Niclis C, Aballay LR, Tumas N, Román MD, Muñoz SE, et al. Cáncer y su asociación con patrones alimentarios en Córdoba (Argentina). *Nutrición Hospitalaria*. 2014; 29(3):618-28.
6. Diaz MP, Corrente JE, Osella AR, Muñoz SE, Aballay LR. Modelling spatial distribution of cancer incidence in Cordoba, Argentina. *Applied Cancer Research*, 2010; 30(1):12-25.
7. Pou S, Tumas N, Becaria Coquet J, Niclis C, Román MD, Díaz, MP. (2017). Burden of cancer mortality and differences attributable to demographic aging and risk factors in Argentina, 1986-2011. *Cadernos de Saúde Pública*. Epub March 09, 2017; 33(2).
8. Díaz MP, Osella AR, Aballay LR, Muñoz SE, Lantieri MJ, Butinof M, Meyer Paz R, Pou S, Eynard AR, La Vecchia C. Cancer incidence pattern in Cordoba, Argentina. *European Journal of Cancer Prevention*, 2009; 18: 259-266.
9. Ronco ÁL, De Stéfani E, Mendoza B, Abbona E, Deneo-Pellegrini H. Patrones dietarios y riesgo de cáncer de mama: un análisis factorial de alimentos y nutrientes. *Revista Médica del Uruguay*, 2016; 32(4):242-53.
10. Pierart Z C, Rozowsky N J. Papel de la nutrición en la prevención del cáncer gastrointestinal. *Revista chilena de nutrición*. 2006; 33(1):8-13.
11. De La Quintana AG, Pou Sonia, Forte CA, Aballay LR. Ingesta alimentaria, estado nutricional y estilos de vida pro-inflamatorios en una cohorte de mujeres con cáncer de mama de la ciudad de Córdoba. *Revista de Nutrición de la CONFELANYD*. 2014; 4:49-63.
12. Anderson C, Harrigan M, George SM, Ferrucci LM, Sanft T, Irwin ML, et al. Changes in diet quality in a randomized weight loss trial in breast cancer survivors: the lifestyle, exercise, and nutrition (LEAN) study. *NPJ Breast Cancer*. 2016; 2:16026.
13. Harvie M, Howell A, Evans D. Can diet and lifestyle prevent breast cancer: what is the evidence? *ASCO*. 2015; 35:66–73

14. Mourouti N, Kontogianni MD, Papavagelis C, et al. Diet and breast cancer: a systematic review. *Int J Food Sci Nutr*. 2014;1–42.
15. Becaria Coquet J, Tumas N, Osella AR, Tanzi M, Franco I, Diaz MP. Cancer and modifiable lifestyle factors in Argentinean women: addressing missing data in a case–control study. *Asian Pac J Cancer Prev*. 2016; 17 (10):4567– 75.
16. Becaria Coquet J. Sesgos en el contexto de la Epidemiología Nutricional del Cáncer: Enfoque Integral de su cuantificación en el marco de los estudios caso-control en Córdoba [Tesis doctoral]. Córdoba, Argentina; 2017.
17. Becaria Coquet J, Muñoz S, Diaz MP. A framework to address potential bias in colorectal cancer: its implementation on a nutritional epidemiologic study in Argentina. *Indian Journal of Applied Research*. 2017; 7 (12).
18. Organización Mundial de la Salud (OMS). Enfermedades no transmisibles. 2018 [Internet]. [Citado 18 de octubre de 2019]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/noncommunicable-diseases>
19. Organización Mundial de la Salud (OMS). 10 datos sobre las enfermedades no transmisibles. 2013. [Internet] [Citado 19 de octubre de 2019]. Disponible en: https://www.who.int/features/factfiles/noncommunicable_diseases/es/
20. American Cancer Society. ¿Qué es el cáncer? 2016. [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <https://www.cancer.org/es/cancer/aspectos-basicos-sobre-el-cancer/que-es-el-cancer.html>
21. Organización Mundial de la Salud (OMS) Cáncer. 2019 [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <http://www.who.int/topics/cancer/es/>
22. Civetta M.T, Civetta J. Carcinogénesis. *Salud pública Méx* [Revista en Internet]. 2011 citado 19 de octubre de 2019]; 53(5): 405-414.
23. Registro Provincial de Tumores de Córdoba (RPTC). Informe sobre cáncer en la Provincia de Córdoba 2004-2009. Ministerio de Salud de la Provincia de Córdoba. Editorial de la Provincia de Córdoba. Córdoba, 2013.
24. Ministerio de Salud y Desarrollo Social. Mortalidad. 2013 [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <http://www.msal.gov.ar/ent/index.php/vigilancia/areas-de-vigilancia/mortalidad>
25. Dirección de Estadística e Información en Salud. Estadísticas Vitales. 2017. [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <http://www.deis.msal.gov.ar/wp-content/uploads/2019/04/Sintesis-nro-5-natalidad-y-mortalidad-2017.pdf>
26. Instituto Nacional de Cáncer. Estadísticas – Incidencia. 2018 [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <https://www.argentina.gob.ar/salud/instituto-nacional-del-cancer/estadisticas/incidencia>

27. Diaz MP, Corrente JE, Osella AR, Muñoz SE, Aballay LR. Modelling spatial distribution of cancer incidence in Cordoba, Argentina. *Applied Cancer Research* 2010; 30(2)245-252.
28. Breast Cancer Org. ¿Qué es el cáncer de mama? 2019 [Internet] [Citado 19 de octubre de 2019]. Disponible en: https://www.breastcancer.org/es/sintomas/cancer_de_mama/que_es_cancer_mama
29. American Cancer Society. ¿Qué es el cáncer de seno? 2019. [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <https://www.cancer.org/es/cancer/cancer-de-seno/acerca/que-es-el-cancer-de-seno.html>
30. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 11. [Internet]. Lyon, France: Centro Internacional de Investigaciones sobre el Cáncer; 2013. Disponible en: <http://globocan.iarc.fr>
31. Organización Mundial de la Salud (OMS). Cáncer de mama: prevención y control. 2019 [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <https://www.who.int/topics/cancer/breastcancer/es/>
32. Di Sibio A, Abriata G, Forman D, Sierra MS. Female breast cancer in Central and South America. *Cancer Epidemiol.* 2016; 44 (1):110-S120.
33. Aballay LR, Eynard A.R, Díaz MP, Navarro A., Muñoz S.E .Overweight and obesity: a review of their relationship to metabolic syndrome, cardiovascular disease, and cancer in South America. *NUTRITION REVIEWS*, Volume 71, Issue 3. 2013. 168–179,
34. Instituto Nacional de Cáncer. Análisis de Situación de Salud por Cáncer, Argentina. 2016. [Internet]. [Citado 19 de octubre de 2019]. Disponible en: <http://www.msal.gov.ar/images/stories/bes/graficos/0000000925cnt-2016-12-21-boletin-epidemiologia.pdf>
35. Tumas N, Niclis C, Osella A, Díaz MP, Carbonetti A. Tendencias de mortalidad por cáncer de mama en Córdoba, Argentina, 1986–2011: algunas interpretaciones sociohistóricas. *Rev Panam Salud Pública.* 2015; 37(4/5):330–6.
36. Ministerio de Salud y Desarrollo Social. Hábitos y estilos de vida [Internet] 2019 [citado 19 de octubre de 2019]. Disponible en: <https://www.argentina.gob.ar/salud/Cancer/factores-de-riesgo/habitos-y-estilos-de-vida>
37. Instituto Nacional de Cáncer. Factores de riesgo: Dieta. 2015 [Internet] [Citado 19 de octubre de 2019]. Disponible en: <https://www.cancer.gov/espanol/cancer/causas-prevencion/riesgo/dieta>
38. Michels KB. Nutritional epidemiology—past, present, future. *Int J Epidemiol.* 2003; 32 (4): 486-488.
39. Anderson C, Harrigan M, George SM, Ferrucci LM, Sanft T, Irwin ML, et al. Changes in diet quality in a randomized weight loss trial in breast cancer survivors: the lifestyle, exercise, and nutrition (LEAN) study. *NPJ Breast Cancer.* 2016;2:16026.

40. Chen S, Chen Y, Ma S, Zheng R, Zhao P, Zhang L, et al. Dietary fibre intake and risk of breast cancer: A systematic review and meta-analysis of epidemiological studies. *Oncotarget*. 2016;7(49):80980–9.
41. Pierce JP, Natarajan L, Caan BJ, Flatt SW, Kealey S, Gold EB, et al. Dietary change and reduced breast cancer events among women without hot flashes after treatment of early-stage breast cancer: subgroup analysis of the Women’s Healthy Eating and Living Study. *Am J Clin Nutr*. 2009;89(5):1565S-1571S.
42. Aballay LR, Osella AR, De La Quintana AG, Diaz MP. Nutritional profile and obesity: results from a random-sample population-based study in Córdoba, Argentina. *Eur J Nutr*. 2016; 55(2):675-685.
43. Hartman TJ, Gapstur SM, Gaudet MM, Shah R, Flanders WD, Wang Y, et al. Dietary Energy Density and Postmenopausal Breast Cancer Incidence in the Cancer Prevention Study II Nutrition Cohort. *J Nutr*. 2016; 146(10):2045–50.
44. Sangra A.R, Codina F.A. Identificación, impacto y tratamiento de datos perdidos y atípicos en epidemiología nutricional. *Rev Esp Nutr Comunitaria* 2015; 21(Supl. 1):188-194 ISSN 1135-3074.
45. Choi B, Granero R, Pak A. Catálogo de sesgos o errores en cuestionarios sobre salud. *Rev Costarr Salud Pública* 2010; 19: 106-118.
46. Dagnino S. J. Datos faltantes (missing values). *Rev Chil Anest* 2014; 43: 332-334.
47. Medina F, Galván M. Imputación de datos: teoría y práctica. 2007 [Internet] [Citado 19 de octubre de 2019]. Disponible en: <https://www.cancer.gov/espanol/cancer/causas-prevencion/riesgo/dieta>
48. Carracedo Martínez E, Figueiras A. Tratamiento estadístico de la falta de respuesta en estudios epidemiológicos transversales. *Salud Pública Méx*. 2006;48(4):341-7. }
49. Acock AC. Working with missing values-multiple imputation. In *A Gentle Introduction to Stata* (Fourth edition.). College Station, Texas, USA: Stata Press, 2014.
50. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009; 338: b2393.
51. Bennett D. How can I deal with missing data in my study? *Aust N Z J Public Health*. 2001; 25(5):464-469.
52. OPS (Organización Panamericana de la Salud) /OMS (Organización Mundial de la Salud). Clasificación Internacional de Enfermedades (CIE) [Internet]. [Citado 8 de abril de 2019]. Disponible en: https://www.paho.org/hq/index.php?option=com_content&view=article&id=3561:2010-clasificacion-internacional-enfermedades-cie&Itemid=2560&lang=es
53. American Cancer Society. Acerca del cáncer de seno. 2017. [Internet]. [Citado 1 de abril de 2019]. Disponible en: <https://www.cancer.org/es/cancer/cancer-de-seno/acerca/que-es-el-cancer-de-seno.html>

54. Escudero Álvarez E. y González Sánchez P. La fibra dietética. *Nutr. Hosp.* 2006; 21 (2) 61-72
55. Vargas M, Lancheros L, Barrera MP. Energy expenditure in repose related to body composition in adults. *Rev Fac Med Univ. Nac. Colomb.* 2011; 59(1):43-58.
56. Real Academia Española. 2018. Edad. Real Academia Española. Disponible en <http://lema.rae.es/drae/srv/search?id=HLafKWLkRDXX2hFUevue>
57. Asociación Argentina de Marketing (AAM) y Cámara de Control de Audiencia (CCMA). Índice de Nivel Socioeconómico 2002. [Internet]. [Citado 1 de abril de 2019]. Disponible en: <http://www.iprofesional.com/adjuntos/documentos/18/0001836.pdf>.
58. Real Academia Española. (2018). Fumar. Real Academia Española. Disponible en <http://lema.rae.es/drae/srv/search?id=HLafKWLkRDXX2hFUevue>
59. Organización Mundial de la Salud (OMS). Actividad física. 2017 [Internet]. [Citado 1 de abril de 2019]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/cancer>.
60. The international Physical Activity Questionnaire. International Physical Activity Questionnaire-Short Format. 2005. [Internet]. [Citado 1 de abril de 2019]. Disponible en: <https://sites.google.com/site/theipaq>.
61. Organización Mundial de la Salud (OMS). 10 datos sobre la obesidad. 2008. [Internet]. [consultado 3 de abril de 2019]. Disponible en: <https://www.who.int/features/factfiles/obesity/facts/es/>
62. Torres Jiménez AP, Torres Rincón JM. Climaterio y menopausia. *Revista de la Facultad de Medicina (México)*. 2018; 61(2):51-8.
63. Organización Mundial de la Salud (OMS). Recomendación de la OMS sobre la alimentación del lactante. 2001. [Internet]. WHO. [Citado 3 de abril de 2019]. Disponible en: https://www.who.int/nutrition/topics/infantfeeding_recommendation/es/
64. Aguilar Cordero M. J., González Jiménez E., Álvarez Ferre J., Padilla López C. A., Mur Villar N., García López P. A. et al. Lactancia materna: un método eficaz en la prevención del cáncer de mama. *Nutr. Hosp.* [Internet]. 2010 [Citado 13 de agosto de 2020]; 25 (6): 954-958. Disponible en: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0212-16112010000600010&lng=es.1.
65. Pérez EGG, Flores CC, Amezcua LD, Ramírez VDA. Factores de riesgo para el cáncer de mama en México: revisión de estudios en poblaciones mexicanas y México-americanas. *Revista CES Salud Pública*. 2014;5 (1):50–8.
66. Morey RJR, González LMB, Morey JR, Navarro VL. Caracterización de mujeres con riesgo de cáncer de mama. *Universidad Médica Pinareña*. 2019; 15(3):338–48.
67. Navarro A, Osella AR, Muñoz SE, Guerra V, Eynard AR, Lantieri MJ. Reproducibility and validity of a food frequency questionnaire in assessing dietary intakes

and food habits in epidemiological cancer studies in Argentina. *J Exp Clin Cancer Res.* 2001; 20:203-207 (sin número)

68. Tumas N, Niclis C, Aballay LR, Osella AR, Díaz MP. Traditional dietary pattern of South America is linked to breast cancer: an ongoing case-control study in Argentina. *Eur J Nutr.* 2014; 53 (2):557–66.

69. StataCorp. 2015. *Stata Statistical Software: Release 14.* College Station, TX: StataCorp LP.

70. Organización Panamericana de la Salud (OPS)/ Organización Mundial de la Salud (OMS) Argentina - Primer Foro Multidisciplinario sobre Cáncer con el apoyo de la OPS | OPS/OMS [Internet]. 2016 [Citado 26 de mayo de 2020]. Disponible en: https://www.paho.org/arg/index.php?option=com_content&view=article&id=10052:pimer-foro-multidisciplinario-cancer-apoyo-ops&Itemid=234

71. Tumas N. “Determinantes sociales y ambientales del cáncer de mama. Estudio socioecológico y de caso-control en la provincia de Córdoba (Argentina), últimas décadas [Tesis doctoral]. Córdoba, Argentina; 2016.

72. Sangaramoorthy M, Koo J, John EM. Intake of bean fiber, beans, and grains and reduced risk of hormone receptor-negative breast cancer: the San Francisco Bay Area Breast Cancer Study. *Cancer Med.* 2018; 7 (5):2131-44.

73. Majid HA, Keow LP, Islam T, Su TT, Cantwell M, Taib NA, et al. Nutritional Status of Breast Cancer Survivors 1 Year after Diagnosis: A Preliminary Analysis from the Malaysian Breast Cancer Survivorship Cohort Study. *J Acad Nutr Diet.* 2018; 118(4):705-13.

74. Aguilar Cordero MJ, Neri Sánchez M , Padilla López CA, Pimentel Ramírez ML, García Rillo A, Mur Villar N. Sobrepeso/obesidad en mujeres y su implicación en el cáncer de mama: edad de diagnóstico. *Nutr. Hosp.* [Internet]. 2012 [Citado 26 de mayo de 2020]; 27(5): 1643-1647. Disponible en: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0212-16112012000500040&lng=es.

75. Pérez Zumano SE, Sandoval Alonso L, Tapia Martínez H. Factores de riesgo para cáncer de mama. Revisión de la literatura: Rol potencial de Enfermería. *Enferm univ* [Internet]. 16 de abril de 2018 [Citado 26 de mayo de 2020]; 6 (3). Disponible en: <http://revista-enfermeria.unam.mx:80/ojs/index.php/enfermeriauniversitaria/article/view/330>

76. Pacheco J. Estrés oxidativo en el climaterio y menopausia y cáncer ginecológico. *Rev Per Ginecol Obstet.* [Internet] 2010 [Citado 26 de mayo de 2020]; 56: 108-119. Disponible en: http://sisbib.unmsm.edu.pe/BVRevistas/ginecologia/vol56_n2/pdf/a06v56n2.pdf

77. Hernandez G, Moriña D, Navarro A. Imputación de valores ausentes en salud pública: conceptos generales y aplicación en variables dicotómicas. *Gaceta Sanitaria.* 1 de marzo de 2017; 31.

78. Muñoz Rosas J F, Alvarez Verdejo E. Métodos de imputación para el tratamiento de datos faltantes: aplicación mediante R/Splus. Revista de métodos cuantitativos para la economía y la empresa. [Internet] 2009 [Citado 26 de mayo de 2020]; (7). 3–30. Disponible en: <http://www.upo.es/RevMetCuant/art25.pdf>
79. Rubin, DB. Multiple Imputation for Nonresponse in Surveys. J. Wiley & Sons, New York.1987.
80. Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2da ed. Estados Unidos: Wiley. 2002. p. 389.
81. National Research Council (NRC). The Prevention and Treatment of Missing Data in Clinical Trials. Panel on Handling Missing Data in Clinical Trials.2010. [Internet]. [Citado 26 de mayo de 2020]. Disponible en: <http://www.cytel.com/hs-fs/hub/1670/file-2411099288-pdf/Pdf/MissingDataNationalAcademyof-Medicine.2010.pdf>
82. Ibrahim, JG, Chu H, Chen MH. Missing Data in Clinical Studies: Issues and Methods. J Clin Oncol. 2012; 30, 3297–303.
83. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Stat Med. 2010; 29, 2920–31.

ANEXOS

ANEXO 1. Categorización de la variable nivel socioeconómico

Índice del nivel socioeconómico 2002. Asociación Argentina de Marketing – Cámara de Control y Medición de Audiencias. 2003. Argentina.

Se calcula por sumatoria de puntos de las variables seleccionadas de interés para la construcción del mismo. Para construir los estratos sociales se tuvieron en cuenta las siguientes variables: ocupación del principal sostén del hogar, cantidad de personas que aportan en el hogar, posesión de bienes y servicios, tipo de atención médica y nivel educativo del principal sostén del hogar.

Las categorías son las siguientes:

1. Nivel económico social alto.
2. Nivel económico social medio alto.
3. Nivel económico social medio típico.
4. Nivel económico social bajo superior.
5. Nivel económico social bajo inferior.
6. Marginal.

OCUPACIÓN DEL PRINCIPAL SOSTÉN DEL HOGAR	PUNTAJE (Puntos)
Dueño o alta dirección de empresas (más de 50 empleados)	32
Dueño o alta dirección de empresas (6 a 50 empleados)	28
Dueño o alta dirección de empresas (1 a 5 empleados)	22
Profesionales independientes	16
Técnico independientes y en relación de dependencia	12
Comerciantes sin personal, artesanos, encargados y supervisores	10
Empleados sin jerarquía	7
Supervisores	10
Obreros calificados o especializados	6
Trabajos autónomos u obreros no calificados	4
Ocupación sin inserción en act. económica genuina (changas)	2
Pasivos e inactivos (jubilados)	4
Desocupados	2

Cantidad de aportantes en el hogar:

- 4 ó + 9 puntos.
- 2 a 3..... 7 puntos.
- 1..... 1 punto.

Nivel educativo del PSH:

- Univ. completo o Postgrado.....13 puntos.
- Univ. incompleto o Terciario o Secundario completo.....4 puntos.
- Secundario inc., primario completo o incompleto.....0 puntos.

Posesión de Bienes y Servicios:

- Conexión a internet en el hogar.....8 puntos.
- Computadora en el hogar.....6 puntos.
- Tarjeta de débito.....5 puntos.
- Cantidad de autos (de menos de 15 años de antigüedad)
 - ➔ 2 ó +.....22 puntos.
 - ➔ 1.....11 puntos.

Tipo de atención médica:

- Privada o con obra social o medicina prepaga....5 puntos.
- Hospital público0 puntos.

Puntuaciones:

CÓDIGO	CLASE	PUNTAJE (Puntos)
1	Alta	55 a 100
2	Media alta	37 a 54
3	Media típica	29 a 36
4	Baja superior	17 a 28
5	Baja inferior	12 a 16
6	Marginal	0 a 11

GLOSARIO

Enfermedades Crónicas No Transmisibles (ECNT): Son afecciones de larga duración con una progresión generalmente lenta. Entre ellas destacan: las enfermedades cardiovasculares, el cáncer, las enfermedades respiratorias crónicas y la diabetes.

Tumor: Masa anormal de tejido que aparece cuando las células se multiplican más de lo debido o no se mueren cuando deberían. Los tumores son benignos (no cancerosos) o malignos (cancerosos). Las masas benignas a veces crecen mucho pero no se diseminan y tampoco invaden los tejidos cercanos ni otras partes del cuerpo. Las masas malignas suelen diseminarse o invadir los tejidos cercanos, y también es posible que se diseminen a otras partes del cuerpo a través de la sangre y el sistema linfático.

Cáncer: Es un proceso de crecimiento y diseminación incontrolados de células. Puede aparecer prácticamente en cualquier lugar del cuerpo. El tumor suele invadir el tejido circundante y puede provocar metástasis en puntos distantes del organismo.

Cáncer de mama (CM): tumor maligno que se ha desarrollado a partir de células mamarias, debido a un crecimiento descontrolado de las mismas.

Prevalencia: se define como el número de casos existentes de una enfermedad u otro evento de salud dividido por el número de personas de una población en un período específico. Cada individuo es observado en una única oportunidad, cuando se constata su situación en cuanto al evento de interés.

Incidencia: se define como el número de casos nuevos de una enfermedad u otra condición de salud dividido por la población en riesgo de la enfermedad (población expuesta) en un lugar específico y durante un período específico.

Valor energético total (VET): es la cantidad de energía que se le provee al organismo.

Fibra: grupo de diferentes sustancias de origen vegetal, que son resistentes a la digestión y absorción en el intestino delgado, pero que pueden ser hidrolizados y fermentados en el colon.

Media: Es el promedio aritmético de todos los valores de una distribución.

Desvío estándar: medida que representa la dispersión de los valores observados con respecto a la media.

Regresión logística: modelo estadístico multivariado que describe la probabilidad de que un proceso ocurra en función de un número determinado de factores cualitativos y/o cuantitativos.

Imputación Múltiple (IM): metodología en la que los datos faltantes, son sustituidos por $m > 1$ valores simulados para hacer un uso eficiente de los datos, obteniendo estimadores no sesgados y que reflejen adecuadamente la incertidumbre que la no-respuesta parcial.

Odds ratio (OR): medida estadística que refleja la posibilidad de que un evento se presente en un grupo de población frente a la probabilidad de que no ocurra. Es utilizado en estudios epidemiológicos transversales y de casos y controles, así como en los meta-análisis.