



Francisco Martín Pérez

UNA REVISIÓN SISTEMÁTICA PARA DETERMINAR LA CALIDAD DE LOS MODELOS PREDICTIVOS DE INCIDENCIA DE FIBRILACIÓN AURICULAR EN LA COMUNIDAD

UNIVERSITAT



Miguel Hernández

UNA REVISIÓN SISTEMÁTICA PARA DETERMINAR LA CALIDAD DE LOS MODELOS PREDICTIVOS DE INCIDENCIA DE FIBRILACIÓN AURICULAR EN LA COMUNIDAD

Tesis doctoral

Francisco Martín Pérez

Universidad Miguel Hernández de Elche

Departamento de Medicina Clínica

2017

UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

FACULTAD DE MEDICINA

Departamento de Medicina Clínica



**UNA REVISIÓN SISTEMÁTICA PARA DETERMINAR LA CALIDAD DE LOS
MODELOS PREDICTIVOS DE INCIDENCIA DE FIBRILACIÓN AURICULAR EN
LA COMUNIDAD**

TESIS DOCTORAL REALIZADA POR:

Francisco Martín Pérez

DIRIGIDA POR EL PROFESOR DOCTOR:

D. Vicente Francisco Gil Guillén

San Juan de Alicante, 2017

DOCTORADO EN INVESTIGACIÓN CLÍNICA (RD 1393/2007)



D. Francisco Javier Fernández Sánchez, Director del Departamento de Medicina Clínica de la Universidad Miguel Hernández

AUTORIZA:

La presentación y defensa como Tesis Doctoral del trabajo “UNA REVISIÓN SISTEMÁTICA PARA DETERMINAR LA CALIDAD DE LOS MODELOS PREDICTIVOS DE INCIDENCIA DE FIBRILACIÓN AURICULAR EN LA COMUNIDAD” realizado por Francisco Martín Pérez bajo la dirección del Prof. Dr. D. Vicente Francisco Gil Guillén.

Lo que firmo en San Juan de Alicante, a Once de Septiembre de Dos Mil Diecisiete.

Prof. D. Francisco Javier Fernández Sánchez

Director del Departamento de Medicina Clínica



D. Vicente Francisco Gil Guillén como Director de Tesis Doctoral

CERTIFICA:

Que el trabajo “UNA REVISIÓN SISTEMÁTICA PARA DETERMINAR LA CALIDAD DE LOS MODELOS PREDICTIVOS DE INCIDENCIA DE FIBRILACIÓN AURICULAR EN LA COMUNIDAD” realizado por Francisco Martín Pérez ha sido llevado a cabo bajo mi dirección y se encuentra en condiciones de ser leído y defendido como Tesis Doctoral en la Universidad Miguel Hernández.

Lo que firmo para los oportunos efectos en San Juan de Alicante, a Once de Septiembre de Dos Mil Diecisiete.

Prof. D. Vicente Francisco Gil Guillén

Director de Tesis Doctoral

En primer lugar me gustaría agradecer,

Al Dr. Vicente Francisco Gil Guillén, mi director de tesis, por ser un excelente investigador y una gran persona. Gracias por hacer posible la realización de este trabajo y confiar en mí.

Al Dr. Antonio Palazón Bru por su paciencia, optimismo, perseverancia y dedicación para orientarme y animarme a realizar este trabajo. Ha sido un placer trabajar contigo.

Por otro lado quiero dedicar,

Con todo mi cariño a mis padres por darme la oportunidad de estudiar y apoyarme incondicionalmente en todas las metas que me propongo.

A mis hermanos, José Guillermo y Antonio, por confiar siempre en mí, darme los consejos necesarios en el momento oportuno y por compartir risas y lloros a lo largo de mi vida.

A mi pareja, Camelia, que siempre está en los momentos difíciles mostrándome su comprensión y cariño, creyendo en mí y motivándome para lograr todas mis metas.

Por último a toda mi familia y amigos, aunque especialmente a mi amigo Crespo por el esfuerzo y apoyo que ha mostrado hacia mí desde siempre.

Muchas gracias a todos.

ÍNDICE



ÍNDICE

ABREVIATURAS	13
RESUMEN	17
1. INTRODUCCIÓN	25
1.1. La fibrilación auricular	27
1.1.1. Definición	28
1.1.2. Epidemiología (prevalencia, incidencia y factores de riesgos)	31
1.1.3. Tratamiento.....	34
1.1.4. Consecuencias en los pacientes e impacto económico de la FA	36
1.2. Modelos de riesgo	37
1.2.1. Evaluación del rendimiento y validación de los modelos de riesgo	41
1.3. Directrices para la comunicación de estudios en ciencias de la salud	45
1.3.1. Revisiones sistemáticas y metaanálisis: declaración PRISMA.....	48
1.3.2. Modelos predictivos: declaración TRIPOD	53
1.3.2.1. Ejemplos de revisiones sistemáticas de modelos predictivos anteriores a la declaración TRIPOD	57
2. JUSTIFICACIÓN	61
3. HIPÓTESIS	65
4. OBJETIVOS	69
5. MATERIAL Y MÉTODOS	73
5.1. Diseño del estudio	75
5.2. Criterios de inclusión	77
5.3. Extracción de los datos, análisis y presentación de resultados	78
6. RESULTADOS	81
6.1. Esquema del estudio	83
6.2. Análisis de los modelos predictivos encontrados	85
7. DISCUSIÓN	89
7.1. Interpretación de los resultados de la discusión.....	91
7.2. Respuesta a las preguntas en una revisión sistemática de la Escuela de Epidemiología Clínica Canadiense	95
7.3. Limitaciones del estudio	96
8. CONCLUSIONES	99
REFERENCIAS	103

ABREVIATURAS



ABREVIATURAS

ACRÓNIMOS/SIGLAS	DESCRIPCIÓN
AAS	ÁCIDO ACETILSALICÍLICO
ACO	ANTICOAGULANTES ORALES
ACV	ACCIDENTE CEREBROVASCULAR
AOS	APNEA OBSTRUCTIVA DEL SUEÑO
ARIC	ATHEROSCLEROSIS RISK IN COMMUNITIES
AUC	AREA UNDER CURVE
AVK	ANTAGONISTAS DE LA VITAMINA K
BIC	CRITERIO DE INFORMACIÓN BAYESIANO
CHARGE-AF	COHORTS FOR HEART AND AGING RESEARCH IN GENOMIC EPIDEMIOLOGY
CV	CARDIOVASCULAR
CONSORT	CONSOLIDATED STANDARDS OF REPORTING TRIALS
DOI	DIGITAL OBJECT IDENTIFIER
ECV	ENFERMEDADES CARDIOVASCULARES
ECG	ELECTROCARDIOGRAMA
EHRA	EUROPEAN HEART RHYTHM ASSOCIATION
EPOC	ENFERMEDAD PULMONAR OBSTRUCTIVA CRÓNICA
EPV	EVENTOS POR VARIABLE
ERC	ENFERMEDAD RENAL CRÓNICA
ERT	ENFERMEDAD RENAL TERMINAL
FA	FIBRILACIÓN AURICULAR
FHS	FRAMINGHAM HEART STUDY
GRIPS	GENETIC RISK PREDICTION STUDIES
HTA	HIPERTENSIÓN ARTERIAL
IC	INSUFICIENCIA CARDIACA
IMC	ÍNDICE DE MASA CORPORAL
INR	INTERNATIONAL NORMALIZED RATIO
lpm	LATIDOS POR MINUTO
NACO	NUEVOS ANTICOAGULANTES ORALES NO DEPENDIENTES DE LA VITAMINA K
PRISMA	PREFERRED REPORTING ITEMS FOR SYSTEMATIC REVIEWS AND META-ANALYSES
QUOROM	QUALITY OF REPORTING OF METAANALYSIS
REMARK	REPORTING RECOMMENDATIONS FOR TUMOR MARKER PROGNOSTIC STUDIES
ROC	RECEIVER OPERATING CHARACTERISTIC CURVE
STARD	STANDARDS FOR REPORTING OF DIAGNOSTIC ACCURACY STUDIES
STROBE	STRENGTHENING THE REPORTING OF OBSERVATIONAL STUDIES IN EPIDEMIOLOGY
TRIPOD	TRANSPARENT REPORTING OF A MULTIVARIABLE PREDICTION MODEL FOR INDIVIDUAL PROGNOSIS OR DIAGNOSIS
VI	VENTRÍCULO IZQUIERDO
WHS	WOMEN'S HEALTH STUDY

RESUMEN



RESUMEN

Introducción

La fibrilación auricular (FA) es la arritmia sostenida más común, y una de las enfermedades cardiovasculares (ECV) más frecuentes. Se trata de una de las principales causas de accidente cerebrovascular (ACV), insuficiencia cardiaca (IC), muerte súbita y morbilidad cardiovascular (CV) en todo el mundo.

El número estimado de personas en todo el mundo con FA en 2010 era superior a los 33 millones (20.9 millones de hombres y 12.6 millones de mujeres), siendo mayor su incidencia y prevalencia en los países desarrollados, donde se ha convertido en uno de los problemas de salud pública más importantes y una causa significativa del aumento de los costes de atención de la salud.

Los modelos predictivos son herramientas útiles en la práctica clínica habitual para determinar la ocurrencia de un suceso en base a una serie de parámetros ayudando a determinar, por ejemplo, a qué pacientes hay que realizar un seguimiento más estricto. En el caso concreto de la FA son conocidos una serie de factores de riesgo, los cuales pueden clasificarse como modificables y no modificables. Si logramos que el conjunto de todos estos factores ofrezca una predicción precisa de qué paciente desarrollará FA, dispondremos de una herramienta realmente útil para disminuir su incidencia, actuando en los diferentes factores de riesgo modificables, y reduciendo con ello la incidencia de ACV, IC, muerte súbita y morbilidad CV.

Justificación e hipótesis

No obstante, si estos modelos predictivos no están contruidos o validados internamente de forma correcta en su metodología estadística, presentarán limitaciones serias a la hora de ser empleados en la práctica clínica. Dicho de otro modo, si estos modelos no han sido implementados con la metodología estadística adecuada, la validez de sus predicciones puede ser puesta en entredicho, y su utilidad en la práctica clínica podría quedar limitada o anulada.

Objetivo

El objetivo de este trabajo ha sido la realización de una revisión sistemática de publicaciones científicas en las que se describa el desarrollo y validación interna de modelos predictivos multivariantes utilizando como variable principal el desarrollo de FA en pacientes de población general y sin diagnóstico previo de esta patología.

Material y métodos

Para este propósito, primero se realizó una revisión sistemática de acuerdo a la declaración PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses* -elementos preferenciales para los informes y publicación sobre protocolos de revisiones sistemáticas y metaanálisis-), utilizando las bases de datos MEDLINE (a través de Pubmed), Scopus y Google Scholar para efectuar una búsqueda bibliográfica exhaustiva y determinar artículos científicos que construyan y validen internamente modelos predictivos de FA que pueden aplicarse en población general. La búsqueda bibliográfica se realizó incluyendo todos los artículos publicados antes del 26 de septiembre de 2016.

Posteriormente, se utilizó la declaración TRIPOD (*Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis* -presentación transparente de un modelo de predicción multivariable para el pronóstico o diagnóstico individual-) para analizar si la metodología estadística de dichos modelos se correspondía con aquella recomendada en la literatura científica, y si el modo de informar de los resultados fue adecuado.

Resultados

Tras la búsqueda bibliográfica se obtuvo un total de 2784 publicaciones. Tras eliminar los duplicados, este número se redujo a 2380, de los cuales 2367 fueron excluidos, dejando finalmente 13 publicaciones para ser analizadas a texto completo. Tras este análisis, 9 más fueron excluidas, con lo únicamente 4 publicaciones fueron finalmente incluidos en la revisión.

Una vez analizados los resultados de los cuatro modelos, fue posible observar que existían algunas deficiencias en la metodología estadística y en la forma de detallarla en la publicación asociada, si bien en general los aspectos clave del

desarrollo y validación del modelo predictivo fueron realizados con exactitud y rigurosidad en todos ellos.

Conclusiones

La falta de rigor a la hora de presentar los cuatro modelos predictivos encontrados pudo ser debida a que todos estos fueron publicados con anterioridad a la declaración TRIPOD. Por lo que de ahora en adelante todos los modelos predictivos de la incidencia de FA en la comunidad deberían estar basados en la metodología estadística más adecuada según se ha analizado en esta revisión, y utilizar la declaración TRIPOD como guía a la hora de publicar los resultados en una revista científica.



ABSTRACT

Introduction

Atrial fibrillation (AF) is the most common sustained arrhythmia and one of the most common cardiovascular diseases. It is associated with an increased risk of stroke, congestive heart failure, sudden death, or cardiovascular morbidity.

The estimated number of people affected by AF throughout the world in 2010 was superior to 33 million -20.9 million male and 12.6 million female-. Its prevalence is higher in developed countries, where it has become an important cause of the increase in public healthcare system costs.

Risk prediction models are very useful tools that are widely used in medical practice to estimate the probability of a given event in terms of a given set of parameters, helping to determine, for instance, which patients need to be continuously tracked. AF, in particular, is associated with a set of well-defined risk factors which can be classified as adjustable or non-adjustable. If one can accurately determine which patients will be affected by AF by appropriately combining these risk factors, we will have a powerful tool to reduce its incidence. By acting upon the adjustable risk factors, the risk of congestive heart failure, stroke, or sudden death can be also accordingly reduced.

Justification and hypotheses

However, if risk prediction models are not adequately built or internally validated according to statistical methodology, its application in daily medical practice will be questionable. In other words, if the risk models have not been implemented using the proper statistical methodology, their predictions cannot be trusted and its application in medical practice will be considerably restricted.

Objective

The main objective of this work has been the realization of a systematic review of scientific publications dealing with the development and internal validation of multivariable risk prediction models to determine the incidence of AF in the general population without previous diagnosis.

Material and methods

For this purpose, a systematic review following the guidelines of the PRISMA - Preferred Reporting Items for Systematic reviews and Meta-Analyses- statement was realized. To do so, we first used the MEDLINE (through PubMed), Scopus, and Google data bases to perform a through search of scientific works published before September 2016 dealing with the construction and internal validation of AF risk prediction models applicable to the general population.

Afterwards, the TRIPOD -Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis- statement was employed to analyze whether the statistical methodology employed in the selected models was appropriated according to scientific literature, and the whether the reporting of the results was accurate.

Results

After the bibliographic search, a total of 2784 publications were obtained. After removing the duplicated results, this number was reduced to 2380 publications, 2367 of which were excluded. Therefore, only 13 publications were finally thoroughly analyzed. After this analysis, 9 more publications were also excluded. Therefore, only 4 publications were eventually included in this systematic revision. Although the key aspects in the description of the development and internal validation of the risk models were accurately described, after analyzing the results of the four models it was determined that there were several deficiencies in the statistical methodology and its reporting.

Conclusions

These deficiencies may be explained because of the fact that these four models were published before the TRIPOD statement had been developed and published. For this reason, AF risk prediction models in the general population developed from now on should employ the most suitable statistical methodology, as we analyzed in this review, as well as closely follow the TRIPOD statement upon publishing the results in a scientific journal.

1. INTRODUCCIÓN

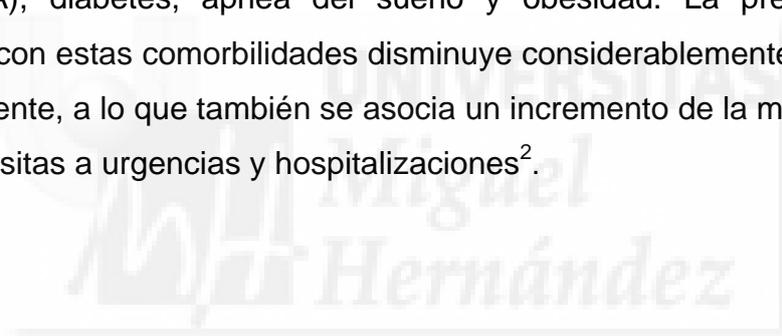


1. INTRODUCCIÓN

1.1. La fibrilación auricular

La fibrilación auricular (FA) es la arritmia sostenida más común, y una de las enfermedades cardiovasculares (ECV) más frecuentes. Se trata de una de las principales causas de accidente cerebrovascular (ACV), insuficiencia cardiaca (IC), muerte súbita y morbilidad cardiovascular (CV) en todo el mundo^{1,2}, constituyendo un importante problema de salud pública.

Esta arritmia es asintomática en al menos el 20% de los pacientes³. Los síntomas más comunes son dificultad respiratoria, palpitaciones, fatiga, dolor torácico y vértigo. Por ello, la mayor parte de los diagnósticos se realizan en exploraciones físicas rutinarias, o en un preoperatorio. La FA puede producirse de forma aislada, aunque es más común su aparición junto con otras ECV, hipertensión arterial (HTA), diabetes, apnea del sueño y obesidad. La presencia de la enfermedad con estas comorbilidades disminuye considerablemente la calidad de vida del paciente, a lo que también se asocia un incremento de la mortalidad y del número de visitas a urgencias y hospitalizaciones².



1.1.1. Definición

La FA es un tipo de taquiarritmia supraventricular, caracterizada por una frecuencia auricular muy rápida, por lo general superior a 300 latidos por minuto (lpm) (aunque puede oscilar entre 300 y 600 lpm), desorganizada y mecánicamente ineficaz¹. El diagnóstico de la FA se realiza mediante un electrocardiograma (ECG), el cual presenta patrones típicos caracterizados por intervalos R-R totalmente irregulares y ondas P indistinguibles o no definidas^{1,2}. Puede clasificarse en las siguientes cinco categorías en función de la presentación, duración y terminación espontánea de los episodios²:

1. **FA diagnosticada por primera vez**: la FA no ha sido diagnosticada previamente, con independencia de la duración de la arritmia o la presencia y la gravedad de los síntomas relacionados con ella.
2. **FA paroxística**: FA recurrente (más de dos episodios) limitada en el tiempo, con una duración típica de 24-48 horas (aunque puede llegar hasta 7 días), y que termina espontáneamente. Los episodios de FA que se revierten en los primeros siete días se consideran paroxísticos.
3. **FA persistente**: la FA se mantiene durante más de siete días. Los episodios que duran más de 7 días requieren terminación por cardioversión eléctrica directa o intervención farmacológica.
4. **FA persistente de larga duración**: los episodios de FA duran más de 1 año tras adoptar una estrategia de control del ritmo cardiaco.
5. **FA permanente**: el paciente y su médico asumen la FA, y no se adoptan intervenciones para el control del ritmo cardiaco. En caso de aplicarse medidas para el control del ritmo, la FA pasaría a clasificarse como FA persistente de larga duración.

Los mecanismos que ocasionan la FA pueden variar de un individuo a otro, y permiten identificar diferentes tipos clínicos^{2,4}:

1. **FA secundaria a enfermedad cardiaca estructural**: la FA se presenta en pacientes con disfunción sistólica o diastólica del ventrículo izquierdo (VI), HTA de larga duración u otra enfermedad cardiaca estructural. La

aparición de FA en estos pacientes es una causa frecuente de hospitalización.

2. **FA focal:** Pacientes con rachas auriculares y episodios cortos de FA paroxística. A menudo en pacientes jóvenes muy sintomáticos con ondas auriculares diferenciadas (FA gruesa), ectopia auricular y/o taquicardia auricular que progresa a FA.
3. **FA poligénica:** La FA en pacientes portadores de variantes genéticas comunes que se asocian a su aparición precoz.
4. **FA monogénica:** FA en pacientes con miocardiopatías hereditarias, incluidas las canalopatías.
5. **FA postoperatoria:** La cirugía cardiotorácica, incluida la cirugía de revascularización coronaria, es un factor de riesgo importante para el desarrollo de FA en el período postoperatorio. Se trata de una FA de nueva aparición (generalmente autolimitada en el tiempo) tras cirugía mayor (normalmente cardíaca) en pacientes sin antecedentes de FA y que estaban en ritmo sinusal antes de la cirugía.
6. **FA en pacientes con estenosis mitral y válvulas cardíacas protésicas:** FA en pacientes con estenosis mitral, tras cirugía de válvula mitral y, en algunos casos, otra valvulopatía.
7. **FA en atletas:** Generalmente FA paroxística, relacionada con la duración y la intensidad del ejercicio.

Los pacientes con FA presentan habitualmente una serie de síntomas tales como letargo, palpitaciones, disnea, opresión torácica, trastornos del sueño o estrés psicosocial, que se traducen en una calidad de vida significativamente peor que otros individuos sanos^{2,5,6}. La *European Heart Rhythm Association* (EHRA) ha propuesto una escala para evaluar los síntomas de los pacientes con FA, los cuales están diferenciados en cinco clases de menor a mayor gravedad⁷:

Clase 1: La FA no causa síntoma alguno.

Clase 2a: Los síntomas de la FA son leves. La actividad física diaria normal del paciente no está afectada por los síntomas de la FA.

Clase 2b: Los síntomas de la FA son moderados. La actividad física diaria normal del paciente no está afectada por los síntomas de la FA, pero los síntomas suponen un problema para el paciente.

Clase 3: Los síntomas de la FA son graves. La actividad física diaria normal del paciente está afectada por los síntomas de la FA.

Clase 4: Los síntomas de la FA son discapacitantes, y la actividad física diaria normal del paciente se ve interrumpida.



1.1.2. Epidemiología (prevalencia, incidencia y factores de riesgos)

El número estimado de personas en todo el mundo con FA en 2010 era superior a los 33 millones (20.9 millones de hombres y 12.6 millones de mujeres), siendo mayor su incidencia y prevalencia en los países desarrollados⁸, donde se ha convertido en uno de los problemas de salud pública más importantes y una causa significativa del aumento de los costes de atención de la salud. La Fig. 1(a) muestra la prevalencia mundial de FA ajustada por edad por 100.000 habitantes en 2010, donde es posible observar que la prevalencia de la enfermedad es menor en Asia y la región del Pacífico (<500), mientras que es máxima (>700) en Norteamérica⁸. La prevalencia y la incidencia de la FA, así como su incremento medio, fueron menores en países en vías de desarrollo, como los ubicados en el África subsahariana, comparadas con las regiones desarrolladas tales como Norteamérica. En España, cerca de un 8.5% de la población mayor de 60 años está afectada por FA, siendo el porcentaje de afectados mayor del 16% para edades superiores a 85 años. Así, la incidencia de la enfermedad en España se estima en unos 96.000 nuevos casos cada año³. La Fig. 1(b) muestra la prevalencia de la FA en España en función de la edad⁹. Es posible observar que el porcentaje de la prevalencia de la enfermedad aumenta notablemente con la edad de los pacientes, siendo del 11% para edades comprendidas entre 68-72 años, y cercano al 15% para edades comprendidas entre 73-79 años. La FA está asociada con un aumento del riesgo a largo plazo de ACV, IC y mortalidad por todas las causas, especialmente en las mujeres^{10,11}. La muerte por ACV puede reducirse en gran medida mediante la anticoagulación, mientras que las muertes debidas a otros problemas CV, tales como IC y muerte súbita, continúan siendo comunes incluso en pacientes con FA tratados. Las hospitalizaciones resultantes, las consultas ambulatorias y de medicina general y los tratamientos con fármacos (incluido el coste de la atención médica) suponen una carga considerable para el sistema público de salud.

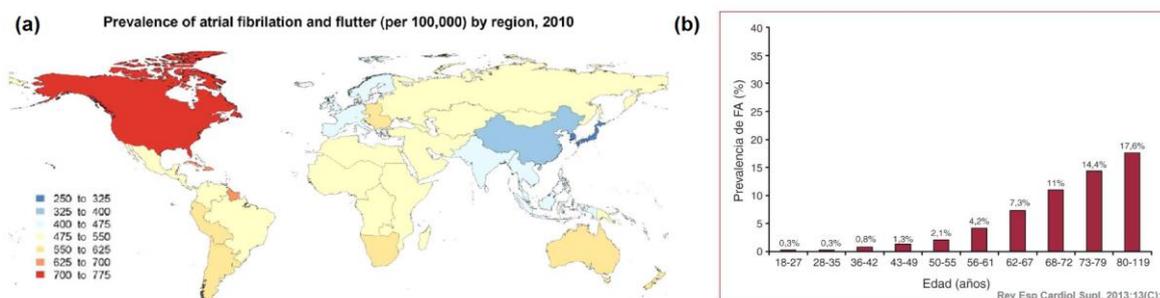


Figura 1. (a) Mapa mostrando la prevalencia de FA ajustada por edad por 100.000 habitantes en 2010⁸. (b) Prevalencia de FA en España según edad⁹.

El número estimado de pacientes con FA en la Unión Europea hacia 2030 es de 14-17 millones, lo que indica una prevalencia aproximada del 3% para adultos mayores de 20 años, con una incidencia de 120.000-215.000 nuevos casos cada año¹². El incremento en la prevalencia de la FA puede ser atribuido al aumento de la esperanza de vida (mayor prevalencia en personas mayores⁸), la mejor detección de la FA silente y otras condiciones predisponentes de la FA¹³. Algunas ECV y entidades concomitantes que aumentan el riesgo de padecer FA y complicaciones asociadas a esta, y cuyo conocimiento y posterior tratamiento son fundamentales para el adecuado tratamiento de los pacientes, son²:

- 1. IC:** Está presente en muchos pacientes afectados de FA, compartiendo con esta última factores de riesgo similares y la misma fisiopatología. La IC y la FA pueden ser causa una de otra, y agravar mutuamente los síntomas a través de diferentes mecanismos, tales como el remodelado cardiaco estructural, la activación neurohormonal o la afección del VI relacionada con la frecuencia cardiaca. El tratamiento de los pacientes afectados de FA es el mismo independientemente de si sufren también IC o no, si bien es conveniente tener en cuenta algunas consideraciones.
- 2. HTA:** Se trata de un factor de riesgo importante en pacientes afectados de FA. La HTA no controlada aumenta el riesgo de ACV y complicaciones hemorrágicas, y puede conducir a la FA recurrente. El control de la presión arterial es, por tanto, indispensable y debe estar incluida en el tratamiento integral de pacientes con FA.
- 3. Valvulopatías:** Aproximadamente el 30% de los pacientes afectados de FA sufren alguna valvulopatía, que únicamente puede ser detectada

mediante ecocardiografía. Están asociadas con un aumento del riesgo tromboembólico, incrementando el riesgo de ACV en los pacientes con FA. Al igual que la IC, las valvulopatías y la FA interactúan y se exacerban y perpetúan mutuamente por la sobrecarga de volumen y presión, taquimiocardiopatía y factores neurohumorales.

4. **Diabetes mellitus:** La diabetes mellitus y la FA coexisten con frecuencia debido a su asociación con otros factores de riesgo. Se trata de un factor de riesgo importante que incrementa el riesgo de ACV y otras complicaciones de la FA.
5. **Obesidad:** La obesidad aumenta el riesgo de padecer FA, con un incremento progresivo directamente proporcional al incremento del índice de masa corporal (IMC). Los pacientes obesos con FA pueden tener un mayor riesgo de sufrir ACV isquémico, tromboembolia o muerte.
6. **Enfermedad pulmonar obstructiva crónica (EPOC), apnea del sueño u otras enfermedades respiratorias:** La FA está asociada con la apnea obstructiva del sueño (AOS). Esta amplifica los cambios de las presiones intratorácicas que, a través de la activación vagal, pueden provocar el acortamiento del potencial de acción e inducir la aparición de la FA.
7. **Enfermedad renal crónica (ERC):** La FA está presente en un 15-20% de los pacientes con ERC, ya que comparten diversos factores de riesgo tales como la edad, la HTA o la diabetes mellitus.

Hábitos perjudiciales para la salud incluidos en el estilo de vida tales como el tabaquismo o la ingesta significativa de alcohol son también factores determinantes en el incremento de la prevalencia de la FA².

1.1.3. Tratamiento

El tratamiento de la FA ha evolucionado sustancialmente durante la pasada década como resultado de nuevas opciones y estrategias terapéuticas, permitiendo minimizar la morbilidad y mortalidad asociadas a la FA. Las guías de práctica clínica proporcionan recomendaciones basadas en estudios clínicos y la opinión de expertos para el tratamiento de los pacientes con FA, y que han sido desarrolladas para proporcionar un cuidado óptimo a los pacientes. Estas guías para el tratamiento de la FA están enfocadas en tres áreas principales: prevención de embolismos sistémicos mediante el uso de anticoagulantes, control de la frecuencia y control del ritmo cardiaco²:

1. **Anticoagulantes:** La prevención de embolismos sistémicos ha demostrado ser un modo eficiente de reducir la morbilidad y mortalidad asociadas a la FA. Los antagonistas de la vitamina K (AVK) (warfarina y acenocumarol) fueron los primeros anticoagulantes orales (ACO) utilizados en pacientes con FA. Han sido ampliamente utilizados con excelentes resultados en todo el mundo, y a día de hoy han demostrado ser el único tratamiento seguro en pacientes con FA y valvulopatía mitral reumática o prótesis valvular mecánica. Según estudios clínicos¹⁴, el tratamiento con AVK reduce hasta en un 66% el riesgo de ACV y en un 25% el riesgo de muerte respecto de paciente sin tratamiento o tratados únicamente con ácido acetilsalicílico (AAS). Sin embargo, los AVK muestran una estrecha ventana terapéutica ya que su eficacia desciende considerablemente con INR (*International Normalized Ratio*) inferior a 1.8, y su toxicidad se incrementa con INR superior a 3.5, una gran variabilidad intra e interindividual en su efecto terapéutico, necesidad de frecuentes ajustes de dosis, gran variedad de interacciones con otros fármacos o alimentos, así como farmacodinámica y farmacocinética impredecibles. Los nuevos anticoagulantes orales no dependientes de la vitamina K (NACO), tales como el inhibidor directo de la trombina (dabigatrán) y los inhibidores del factor Xa (rivaroxabán, edoxabán y apixabán), han surgido como alternativa a la warfarina y el acenocumarol para la prevención de complicaciones trombóticas en pacientes con FA. Los NACO presentan un efecto anticoagulante predecible sin necesidad de una monitorización

regular, un inicio de acción rápido y escasas interacciones farmacológicas o alimentarias.

- 2. Control de la frecuencia cardiaca:** Esta técnica consiste en controlar el número de veces por minuto que se contraen los ventrículos del corazón del paciente, de manera que estos tienen tiempo suficiente para llenarse completamente de sangre. Este tratamiento es una parte integral del tratamiento de pacientes con FA. A pesar de que no permite recuperar un ritmo normal y el corazón no funciona adecuadamente, en general es suficiente para mejorar algunos de los síntomas de la FA. El control de la frecuencia cardiaca a corto y largo plazo puede lograrse mediante tratamiento farmacológico, que incluye bloqueadores beta (bisoprolol, carvedilol, metoprolol, nebivolol, esmolol), bloqueadores de los canales de calcio (diltiazem, verapamilo) o un tratamiento combinado.
- 3. Control del ritmo cardiaco:** La restauración y mantenimiento del ritmo sinusal está recomendado para mejorar los síntomas en pacientes con FA que continúan sintomáticos pese a recibir el tratamiento adecuado para controlar la frecuencia cardiaca. El control del ritmo cardiaco puede lograrse mediante cardioversión farmacológica con antiarrítmicos (flecainida, amiodarona, propafenona, ibutilida, vernakalant) o cardioversión eléctrica. La cardioversión farmacológica está indicada en FA persistente o en FA de menos de 48 h de evolución con buena tolerancia hemodinámica y en ausencia de cardiopatía estructural. La cardioversión eléctrica está generalmente indicada en FA con mala tolerancia hemodinámica, como tratamiento en la FA paroxística y persistente, o cuando la cardioversión farmacológica no es eficaz. En caso de que ninguna de las dos técnicas anteriores sea efectiva, la ablación con catéter, que permite eliminar el tejido cardiaco con actividad eléctrica anormal, es la técnica recomendada. En cualquier caso, los beneficios a largo plazo de la restauración del ritmo sinusal y la frecuencia cardiaca frente a la restauración de únicamente la frecuencia cardiaca no han sido probados de manera concluyente^{2,15}.

1.1.4. Consecuencias en los pacientes e impacto económico de la FA

A pesar de los avances en el tratamiento de la FA, esta enfermedad continúa estando asociada a un incremento de ACV, IC, muerte súbita y morbilidad CV en todo el mundo. De hecho, diversos estudios han demostrado que 20-30% de los pacientes con ACV isquémico ya habían sido previamente diagnosticados de FA, o lo fueron inmediatamente después. La FA está también asociada con lesiones de la sustancia blanca y trastornos cognitivos (incluso en pacientes anticoagulados), disfunción del VI (en un 20-30% de los pacientes) e IC, disminución de la actividad física, efectos sobre la salud mental (depresión) y bienestar del paciente y sus familiares, y un deterioro en general de la calidad de vida de los afectados. Todo ello lleva asociado un incremento considerable de las hospitalizaciones, con un 10-40% de los pacientes afectados de FA hospitalizados anualmente, no solo debido a la FA, sino también a IC, infarto de miocardio o complicaciones asociadas al tratamiento².

Durante los últimos 20 años ha habido un incremento de cerca del 66% en las hospitalizaciones por FA en Estados Unidos y Europa, debido a una combinación de factores que incluye el envejecimiento de la población, el aumento de la prevalencia de enfermedades cardíacas crónicas y un mayor número de diagnósticos. Todo ello implica tremendos costes para los sistemas de salud, los cuales se derivan fundamentalmente de las hospitalizaciones (52%) debidas a la FA o complicaciones relacionadas con ella (como los ACV), seguido de los tratamientos farmacológicos (23%) y consultas médicas (9%)¹⁶. Globalmente, el coste anual por paciente está próximo a los 3.000€¹⁷. Considerando la prevalencia de la FA en, por ejemplo, la Unión Europea (se estima que cerca de 4.5 millones de personas en Europa sufren FA paroxística o persistente), el coste total para los sistemas sanitarios europeos ascendería a más de 13.500 millones de euros anuales¹⁶. Según un estudio reciente¹⁸, el coste medio para el sistema nacional de salud por paciente con FA en España es de 2.118€/año, de los cuales cerca del 48% del total (esto es, 1.008€) representa los costes hospitalarios.

1.2. Modelos de riesgo

Los modelos de riesgo son modelos estadísticos utilizados para estimar el riesgo o probabilidad de que un paciente sufra una determinada enfermedad (esto es, modelos diagnósticos), o que un evento específico ocurra en el futuro (es decir, modelos pronósticos), y que son especialmente útiles para la práctica médica y la investigación¹⁹. En salud pública, los modelos predictivos pueden ser útiles para orientar intervenciones preventivas en pacientes con riesgo relativamente alto de tener o desarrollar una enfermedad²⁰. En la práctica clínica, los modelos de riesgo pueden ser, a grandes rasgos, útiles de dos modos²¹: Primero, pueden permitir una clasificación fiable de los pacientes en dos o más grupos de riesgo con diferente pronóstico. Dado que los tratamientos disponibles actualmente para una determinada enfermedad son limitados, no es aconsejable una clasificación demasiado exhaustiva de los pacientes en un número elevado de grupos de riesgo diferentes. Una clasificación apropiada puede permitir la determinación de la terapia más adecuada para un determinado paciente, ahorrar derivaciones o pruebas innecesarias y facilitar la comunicación entre médicos. Segundo, un modelo de riesgo puede ser utilizado para estimar el pronóstico de un paciente determinado. Desde el punto de vista del diagnóstico, las predicciones pueden ser útiles para estimar la probabilidad de que una enfermedad esté presente, indicando de manera fiable si es recomendable iniciar un tratamiento contra la enfermedad o si es necesario realizar más pruebas diagnósticas. Desde el punto de vista terapéutico, los modelos de riesgo pueden ser indicadores fiables a la hora de establecer un umbral a partir del cual el tratamiento está indicado, y donde los potenciales beneficios individuales superan a los riesgos potenciales del tratamiento, ayudando al médico y paciente a tomar la mejor decisión posible^{19,21}.

Una condición clave para la implementación de un modelo de riesgo en la práctica clínica es que las predicciones que este realiza sean fiables. Esto significa que si, a modo de ejemplo, el modelo predice un 10% de riesgo de una determinada enfermedad en pacientes con unas determinadas características, se espera que en promedio un 10% de los pacientes reales que reúnan estas condiciones padezcan la enfermedad. Por este motivo, si bien los modelos de

riesgo pueden ser muy útiles en la práctica clínica, sólo un reducido número de ellos acaba siendo utilizado rutinariamente²⁰. Algunos prerrequisitos que debe cumplir un modelo de riesgo para ser creíble desde un punto de vista clínico son²¹:

1. Todos los datos clínicamente relevantes de los pacientes deberían haber sido comprobados antes de su inclusión en el modelo.
2. La obtención de los datos necesarios de un paciente ha de ser sencilla, fiable, realizarse sin emplear más recursos de los estrictamente necesarios, y rápida, de modo que la predicción pueda ser realizada a tiempo y ser útil para orientar las decisiones clínicas.
3. La estructura del modelo debe ser obvia, y sus predicciones deben ser fácilmente interpretables por los médicos. Todo ello facilita la implementación del modelo de riesgo en la práctica clínica, minimizando posibles negligencias derivadas de su mala interpretación o utilización.
4. El cálculo del pronóstico para un determinado paciente ha de ser sencilla.

La potencia de un modelo predictivo depende del número total de eventos observados, el cual es a menudo mucho menor que el tamaño de la muestra. Así, si un determinado evento ocurre con una incidencia del 0.1%, un estudio con 10.000 pacientes contendrá de media tan sólo 10 eventos, los cuales constituirán el tamaño efectivo de la muestra bajo estudio. Si la muestra bajo estudio es muy pequeña, la incertidumbre asociada a los resultados del modelo predictivo puede ser muy grande, con lo que éste dejaría de ser capaz de generar resultados completamente confiables. En este caso, el modelo predictivo necesitará partir de suposiciones más sólidas, y únicamente será capaz de generar predicciones válidas a cuestiones sencillas, dejando las cuestiones más complejas para modelos en los que el tamaño efectivo de la muestra sea mayor. Dado que en la práctica el tamaño efectivo de las muestras suele ser bastante pequeño, el desarrollo de modelos predictivos capaces de generar predicciones confiables con un tamaño de muestra efectiva reducido está siendo objeto de intensas investigaciones²⁰.

Tanto en los modelos diagnósticos como en los pronósticos, las estimaciones de las probabilidades de una enfermedad o evento no están basadas únicamente en un único predictor, sino que son el resultado de integrar en un único modelo diferentes características y síntomas del paciente. Por ello, los modelos utilizados en la práctica clínica son inherentemente multivariantes, combinando estadísticamente múltiples predictores a los que se asigna un determinado peso específico, para obtener un riesgo o probabilidad¹⁹.

Las variables utilizadas para la construcción de los modelos predictivos son generalmente continuas y no discretas, ya que la utilización de estas últimas es en muchos casos poco realista, puede suponer la exclusión de información relevante para el modelo, y considerablemente reduce el potencial del modelo para detectar una posible asociación entre las diferentes variables y el resultado²¹.

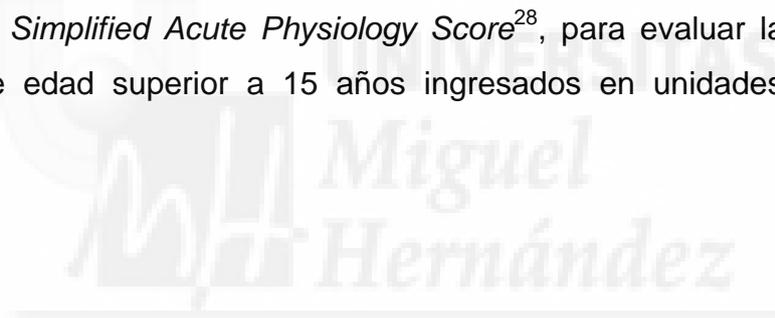
La técnica analítica más común para combinar estadísticamente los predictores en los modelos y obtener resultados a corto y largo plazo es la regresión de Cox. El modelo de riesgo de Cox es un modelo de regresión para el análisis de la supervivencia que describe la relación entre la incidencia de un evento y un conjunto de predictores. La función de riesgo describe la probabilidad instantánea de un evento en un instante de tiempo concreto t , o la probabilidad de que un sujeto bajo observación experimente el evento en una ventana de tiempo determinada alrededor de un instante concreto t . Matemáticamente, el modelo de Cox puede expresarse como:

$$h(t, x_1, \dots, x_N) = h_0(t) \times \exp\{b_1x_1 + \dots + b_Nx_N\}$$

donde $h(t, x_1, \dots, x_N)$ es la función de riesgo, la cual es dependiente del tiempo y del conjunto de N predictores (x_1, \dots, x_N) , los coeficientes (b_1, \dots, b_N) son pesos estadísticos que determinan el impacto de los N predictores sobre el resultado final del modelo, y $h_0(t)$ es la función de riesgo basal, que expresa el riesgo en el caso que todos los predictores son cero (esto es, $x_i=0$ para $i=1, \dots, N$). La función de riesgo basal $h_0(t)$ es una función arbitraria y no especificada, la cual se estima de manera no paramétrica (ya que únicamente depende del tiempo), una vez los coeficientes (b_1, \dots, b_N) han sido estimados. Por este motivo, el modelo de

regresión de Cox se define en ocasiones como un modelo semiparamétrico, lo que lo hace especialmente adecuado para estudios de supervivencia en los que el modelo paramétrico exacto es desconocido. Dado que $h_0(t)$ no está especificada, es posible estimar los coeficientes (b_1, \dots, b_N) , calcular las razones de riesgo y ajustar las curvas de supervivencia correspondientes a un gran número de situaciones, de modo que los resultados finales serán muy similares a los que se obtendrían con el modelo paramétrico, mucho más complejo de obtener^{22,23}.

Algunos ejemplos de modelos predictivos muy conocidos incluyen *Framingham Risk Score*²⁴, para estimar el riesgo cardiovascular de un individuo, *Ottawa Ankle Rules*²⁵, para ayudar al médico a decidir si un paciente con dolor de pie o talón debería ser irradiado con rayos X para diagnosticar una posible fractura, *EuroSCORE*²⁶, para evaluar el riesgo de fallecimiento después de una operación de corazón, *Nottingham Prognostic Index*²⁷, para después de la cirugía de cáncer de mama, o *Simplified Acute Physiology Score*²⁸, para evaluar la gravedad de pacientes de edad superior a 15 años ingresados en unidades de cuidados intensivos.



1.2.1. Evaluación del rendimiento y validación de los modelos de riesgo

Los modelos de riesgo son una herramienta de suma importancia en la práctica clínica. Para ello, es de suma importancia poder evaluar con exactitud la capacidad predictiva de un modelo, para así verificar que los resultados obtenidos son creíbles y confiables cuando el modelo es aplicado en nuevos sujetos fuera de la muestra utilizada para su construcción. Si se intenta evaluar la capacidad predictiva del modelo utilizando exactamente los mismos datos empleados en su construcción, el resultado obtenido tenderá a ser optimista y alejado de la realidad. Esto significa que el rendimiento aparente del modelo (es decir, el rendimiento estimado en la muestra) y el rendimiento real (referido a la población subyacente de la que se obtuvo la muestra) son diferentes. Este problema metodológico se denomina sobreajuste u *overfitting*¹⁹⁻²¹, y es muy común al intentar ajustar modelos que poseen un número elevado de grados de libertad (esto es, predictores). El resultado es que los datos utilizados en el desarrollo del modelo están excelentemente descritos, pero los resultados serán difícilmente aplicables a sujetos no contenidos en la muestra original.

El desarrollo de todo modelo de riesgo debe, por tanto, incluir un proceso de validación interna, utilizando los mismos datos que se utilizaron en el desarrollo del modelo, para determinar con exactitud cualquier optimismo en la capacidad predictiva del modelo. El rendimiento puede ser evaluado en términos de la calibración y la discriminación del modelo de riesgo^{29,30}:

1. La **calibración** hace referencia a la concordancia entre lo pronosticado por el modelo y lo que se observa efectivamente en la práctica. Por ejemplo, si el modelo predice un riesgo del 10% para una cierta enfermedad en un grupo de pacientes con unas características determinadas, dicha enfermedad debería poder ser observada en la práctica en aproximadamente 10 de cada 100 pacientes con esas características. Para cuantificar la capacidad predictiva del modelo en términos de la calibración, suele utilizarse el Test de Hosmer-Lemeshow. De acuerdo a esta técnica, los pacientes se dividen de acuerdo a su probabilidad de sufrir un determinado evento en diez grupos de riesgo (deciles), cada uno

(deciles), cada uno de ellos con idéntico número de pacientes, dentro de los cuales se mide la desviación entre lo predicho por el modelo y lo observado en la práctica. Las desviaciones se suman sobre los diez deciles de riesgo y se comparan utilizando el test de chi cuadrado (χ^2). Si existen diferencias estadísticamente significativas entre lo observado y lo predicho por el modelo (es decir, para $p < 0.05$, siendo p el nivel de significación), se considera que el modelo es incapaz de predecir adecuadamente los resultados en el grupo de pacientes de la muestra. La suma final de los deciles y sus χ^2 se conoce como bondad del ajuste (*goodness of fit*).

2. La **discriminación** evalúa la capacidad del modelo de riesgo para diferenciar entre los pacientes que sufrieron el evento bajo estudio, de aquellos que no. Si el modelo predice, por ejemplo, que 4 pacientes de una muestra de 10 individuos deben sufrir el evento bajo estudio, la discriminación tiene en consideración no sólo el número total de aciertos, sino que los individuos que sufren el evento sean exactamente los que el modelo predice. Así, si el modelo predijo que los pacientes 2, 3, 4 y 6 sufrirían el evento bajo estudio, pero en realidad lo sufrieron los pacientes 1, 3, 4 y 6, la tasa de predicciones correctas fue sólo para los pacientes 3, 4 y 6. Para ello, se utiliza el área bajo la curva (*Area under curve*, -AUC-) de eficacia diagnóstica (*Receiver operating characteristic curve*, -ROC-), la cual representa los verdaderos positivos en el eje de ordenadas, y los falsos positivos en el eje de abscisas. Si los resultados del modelo son perfectos (esto es, no hay falsos positivos), el área bajo la curva vale 1. Para que un modelo discrimine de modo correcto y sus predicciones tengan validez clínica, es preciso que el área bajo la curva sea igual o superior a 0.8.

Para corregir el *overfitting* y cuantificar una posible sobreestimación de la capacidad predictiva del modelo, la validación interna incluye métodos como el remuestreo o *bootstrapping*³¹. En esta técnica, la muestra original es replicada un número N de veces (infinitas veces en teoría). Posteriormente, las copias son unidas consecutivamente con el fin de crear una población mucho mayor que la

reales obtenidos en la práctica, y establecer con mejor precisión cuál es la capacidad predictiva real del modelo de riesgo¹⁹.



1.3. Directrices para la comunicación de estudios en ciencias de la salud

Los ensayos clínicos, los estudios observacionales, los estudios de pruebas diagnósticas y las revisiones sistemáticas son algunos de los estudios más frecuentemente publicados en el área de la investigación sanitaria. Estas publicaciones deben incluir toda la información sobre la metodología y los resultados que sean esenciales para juzgar la validez y la relevancia del estudio, de modo que sean útiles para otros investigadores y estos puedan utilizar los resultados y ser capaz de reproducirlos.

La evaluación sistemática de los artículos publicados ha demostrado que, sin embargo, un número elevado de ellos presenta habitualmente limitaciones y deficiencias graves. Estas incluyen la omisión de aspectos cruciales de la metodología, errores estadísticos, la presentación selectiva de sólo algunos resultados obtenidos, la presentación confusa de los datos y las gráficas del estudio, la malinterpretación de los resultados obtenidos o la citación selectiva e inapropiada de otros estudios^{32,33}. Si los resultados no están adecuadamente descritos, sino que están presentados de manera incompleta, ambigua o selectiva, otros investigadores serán incapaces de juzgar si el estudio y los datos presentados son fiables y pueden ser utilizados en otras investigaciones. Esto compromete la financiación de futuros proyectos de investigación, y conlleva serias consecuencias para la práctica clínica, además de suponer un evidente desperdicio de recursos.

Con el fin de describir adecuadamente y publicar los resultados de una investigación de modo que otros investigadores puedan evaluar adecuadamente el trabajo y utilizar sus resultados, existen guías en las que se especifica la información mínima que debe ser incluida en cada tipo de publicación. Entre estas guías se encuentran:

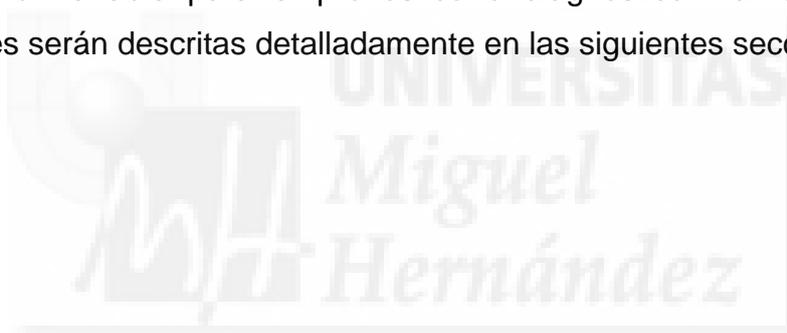
1. Declaración CONSORT³⁴ (*CONsolidated Standards Of Reporting Trials* - normas consolidadas para informar y publicar ensayos-). Es una guía formada por una lista de comprobación que contiene 25 elementos básicos orientados a la validación tanto interna como externa de ensayos clínicos controlados aleatorizados, y que deben ser incluidos en la posterior

descripción y publicación de los resultados, además de un diagrama de flujo que muestra un registro del flujo de participantes a lo largo de todo el ensayo. Toda la información sobre esta declaración está disponible en la página web de CONSORT (www.consort-statement.org).

2. Declaración STROBE³⁵ (*Strengthening the Reporting of Observational studies in Epidemiology* -fortalecimiento de la redacción y publicación de estudios de observación en epidemiología-). El principal objetivo de esta guía es el de informar a los investigadores sobre el modo adecuado de comunicar los resultados de un estudio observacional, para lo que consta de una lista de verificación de 22 puntos relacionados el título, el resumen, la introducción, y las secciones de metodología, resultados y discusión de las publicaciones. Toda la información sobre esta declaración está disponible en la página web de STROBE (www.strobe-statement.org).
3. Declaración STARD³⁶ (*Standards for Reporting of Diagnostic Accuracy Studies* -normas para informar y publicar sobre estudios de precisión diagnóstica-). Esta guía fue desarrollada para mejorar la calidad de los informes y publicaciones sobre estudios de precisión diagnóstica. Esta declaración consta de una lista de comprobación de 25 elementos, y recomienda la utilización de un diagrama de flujo con una descripción del diseño del estudio y el flujo de pacientes a lo largo de este.
4. Declaración REMARK³⁷ (*REporting recommendations for tumor MARKer prognostic studies* -recomendaciones de presentación de informes para estudios pronósticos de marcadores tumorales-). Esta guía contiene 20 directrices para la correcta publicación de estudios pronósticos con marcadores tumorales.
5. Declaración GRIPS³⁸ (*Genetic Risk Prediction Studies* -estudios de predicción de riesgo genético-). Esta guía contiene una lista de comprobación de 25 elementos agrupados en 6 dominios (título y resumen, introducción, metodología, resultados, discusión y otros), con el fin de mejorar la transparencia y calidad de la presentación de estudios de predicción de riesgo genético.

El objetivo de este trabajo es la realización de una revisión sistemática de publicaciones en las que se describa la construcción y validación interna de modelos predictivos de FA, para lo que ninguna de las declaraciones descritas anteriormente es completamente adecuada. Para este propósito, primero realizaremos una revisión sistemática de las publicaciones existentes utilizando la declaración PRISMA³⁹ (*Preferred Reporting Items for Systematic reviews and Meta-Analyses* -elementos preferenciales para los informes y publicación sobre protocolos de revisiones sistemáticas y metaanálisis-).

Posteriormente, analizaremos si la metodología estadística empleada en las publicaciones seleccionadas así como el modo en que los resultados son comunicados son las recomendados, para lo que utilizaremos la declaración TRIPOD¹⁹ (*Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis* -presentación transparente de un modelo de predicción multivariable para el pronóstico o diagnóstico individual-). Ambas declaraciones serán descritas detalladamente en las siguientes secciones.



1.3.1.Revisiones sistemáticas y metaanálisis: declaración

PRISMA

La declaración QUOROM (*Quality Of Reporting Of Metaanalysis* “calidad de presentación de metaanálisis”), publicada en 1999, fue la primera guía cuyo objetivo era establecer directrices para mejorar la calidad y transparencia de la presentación de metaanálisis de ensayos clínicos aleatorizados. Contiene una lista de comprobación de 18 puntos para ayudar a investigadores a publicar adecuadamente sus resultados en revistas médicas proporcionando toda la información que consideren relevante para su interpretación y utilización, recomendando además la utilización de un diagrama de flujo para describir el proceso. Sin embargo, los problemas identificados en los metaanálisis publicados, y las consiguientes mejoras introducidas en la metodología de elaboración de revisiones sistemáticas desde la publicación de la declaración QUOROM, condujeron a una actualización de las directrices originales en 2009. Esta nueva actualización y ampliación de la declaración QUOROM recibió el nombre de PRISMA.

La declaración PRISMA³⁹ consta de un total de 27 directrices y un diagrama de flujo más detallado y pormenorizado para describir el procedimiento, incorporando diversos nuevos aspectos conceptuales y metodológicos relacionados con los progresos realizados en la metodología de las revisiones sistemáticas.

Una mejora importante de la nueva lista está relacionada con la revisión de la terminología utilizada en revisiones sistemáticas y metaanálisis, así como la distinción clara entre estos dos conceptos: el metaanálisis es el procedimiento estadístico por el que se combinan los resultados de varios estudios independientes, y es solamente una parte (altamente recomendada pero no siempre realizable) de un proceso más amplio consistente en diversos pasos sucesivos, perfectamente determinados y reproducibles conocido como revisión sistemática. Además, si bien PRISMA se centra principalmente en las revisiones sistemáticas de ensayos aleatorizados, su aplicabilidad es mayor que la de QUOROM, pues también es útil para las revisiones de otro tipo de estudios, y presta especial atención a la evaluación del riesgo de sesgo al nivel de los

estudios incluidos, el cual es un componente esencial de cualquier revisión sistemática y debe ser considerado en los análisis, la interpretación de los resultados y las conclusiones de esta. No obstante, la declaración PRISMA no es un instrumento orientado a evaluar la calidad de una revisión sistemática.

La Tabla 1 muestra la lista de 27 directrices (en inglés) de la declaración PRISMA que deberían ser incluidos para la correcta publicación de una revisión sistemática, con o sin metaanálisis. La Fig. 2 muestra el diagrama de flujo de la información a través de las diferentes etapas de una revisión sistemática según la declaración PRISMA. Este diagrama, que consta de cuatro fases, parte al inicio de una búsqueda intensiva de literatura científica en bases bibliográficas y de datos, y que debe quedar reflejada en los registros. Posteriormente, este registro es revisado, las citas duplicadas son eliminadas, y algunas de las publicaciones son descartadas en base a un criterio de elegibilidad, con lo que se reduce el número total de publicaciones escogidas en la fase inicial. Finalmente, termina con la recapitulación de los estudios individuales incluidos en la revisión sistemática (síntesis cualitativa) y en el metaanálisis (síntesis cuantitativa). En este punto, PRISMA también diferencia claramente entre el número de estudios y de publicaciones, los cuales pueden no coincidir necesariamente: una publicación puede presentar varios estudios y a su vez, los resultados de un estudio pueden estar publicados en varias publicaciones simultáneamente. Por ello, el diagrama de flujo de PRISMA tiene en cuenta esta particularidad durante el proceso de revisión.

Todos los documentos de la declaración PRISMA, incluyendo la lista de directrices, el diagrama de flujo y su explicación correspondiente, están disponibles en www.prisma-statement.org.

Tabla 1. Lista de comprobación de la declaración PRISMA de los ítems para incluir en la publicación de una revisión sistemática, con o sin metaanálisis.

SECTION/TOPIC	NUMBER	CHECKLIST ITEM
TITLE		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
ABSTRACT		
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of what is already known.
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).
METHODS		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I ²) for each meta-analysis.
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.

SECTION/TOPIC	NUMBER	CHECKLIST ITEM
RESULTS		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).
DISCUSSION		
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

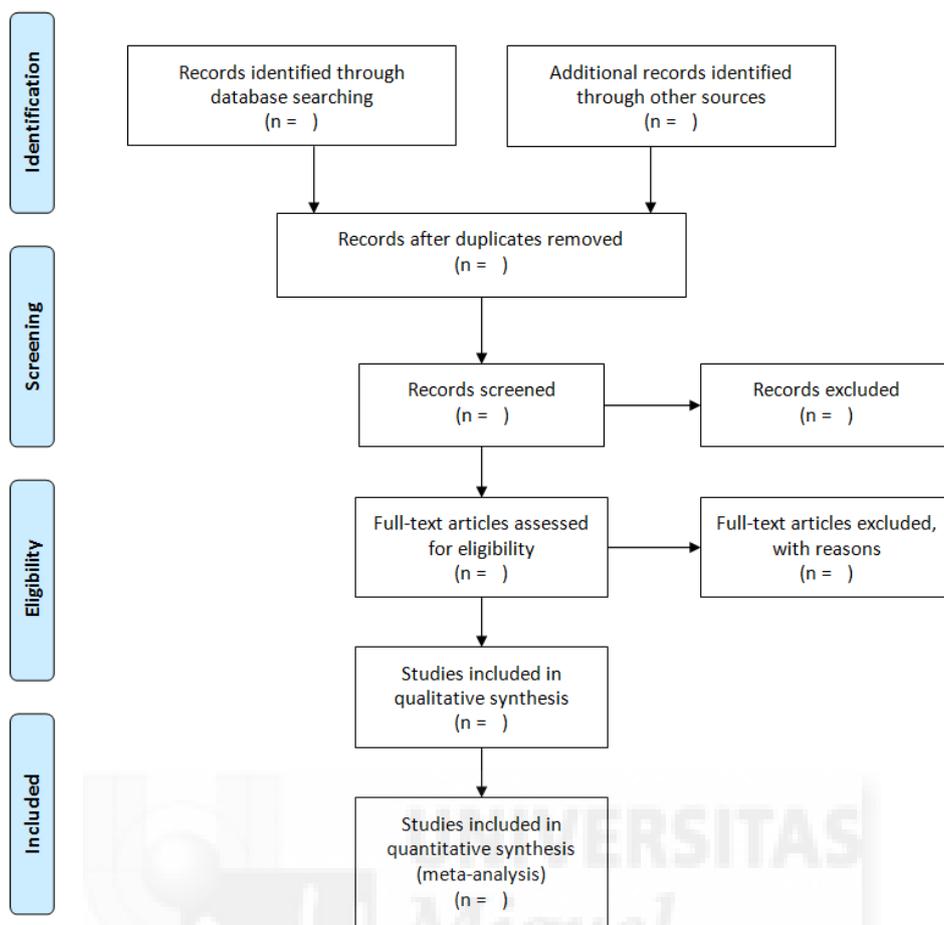


Figura 2. Diagrama de flujo de la información a través de las diferentes etapas del proceso de una revisión sistemática según la declaración PRISMA. El diagrama consta de cuatro fases bien diferenciadas: (1) Búsqueda bibliográfica y registro de la misma. (2) Eliminación de las citas duplicadas y revisión de las publicaciones restantes. (3) Elección de parte de las publicaciones en base a un criterio de elegibilidad. (4) Recapitulación de los estudios individuales incluidos en la revisión sistemática (síntesis cualitativa) y en el metaanálisis (síntesis cuantitativa).

1.3.2. Modelos predictivos: declaración TRIPOD

Muchas revisiones realizadas durante los últimos años han demostrado que los estudios publicados describiendo el desarrollo y validación de modelos predictivos multivariados en medicina son deficientes y de baja calidad, debido fundamentalmente a la ausencia de información detallada y transparente sobre los puntos clave de los estudios. Este hecho, lamentablemente muy recurrente, impide en último término que el resto de la comunidad científica pueda juzgar objetivamente los puntos fuertes y débiles de un determinado estudio publicado.

La declaración TRIPOD¹⁹ se publicó en 2015 precisamente con el fin de mejorar la calidad de los estudios publicados sobre modelos predictivos, ayudar a los autores trabajando en este campo a describir adecuadamente en sus publicaciones el desarrollo, validación y actualización de los modelos, y permitir a los editores de las revistas científicas, a los revisores durante el proceso de publicación y finalmente, a otros investigadores interesados, valorar adecuadamente el estudio. La declaración TRIPOD cubre explícitamente el desarrollo y validación de modelos predictivos tanto para la realización de diagnósticos como de pronósticos, y es aplicable en todos los ámbitos médicos independientemente del tipo de predictores utilizados.

Consta de 22 directrices (véase Tabla 2 -en inglés-) específicamente diseñadas para poner especial énfasis en los estudios de validación así como en los requisitos necesarios para la correcta comunicación de dichos estudios. Las diferentes directrices están relacionadas con el título y el resumen del trabajo (directrices 1 y 2), antecedentes y objetivos (directriz 3), métodos (directrices 4 a 12), resultados (directrices 13 a 17), discusión (directrices 18 a 20), así como con otros aspectos importantes del estudio que los autores consideren (directrices 21 y 22). Estas 22 directrices son aplicables tanto a estudios que abarquen únicamente el desarrollo de un modelo predictivo, como a los que abarquen su desarrollo y validación o únicamente su validación externa. Por ello, los investigadores que utilicen la declaración TRIPOD no necesariamente encontrarán relevantes las 22 directrices: mientras que los puntos 10a, 10b, 14 y 15 serán importantes para describir adecuadamente el desarrollo de un modelo predictivo, los puntos 10c, 10e, 12, 13c, 17 y 19a sólo tendrán relevancia en estudios que describan su validación externa. El resto de puntos de la

declaración TRIPOD son comunes y relevantes para todos los estudios referidos al desarrollo y validación de un modelo predictivo. Esta diferenciación está explícitamente indicada en la Tabla 2: las directrices relevantes únicamente para describir adecuadamente el desarrollo de un modelo están indicados con la letra D, las que están relacionadas únicamente con la validación están indicadas con la letra V, mientras que las relacionadas con ambos tipo de estudio están marcadas con D; V. No obstante, es importante recalcar que el objetivo de la declaración TRIPOD es únicamente proporcionar una guía para la comunicación adecuada de estudios de investigación en modelos predictivos, y no proporcionar directrices para desarrollar y validar dichos modelos.

Todos los documentos de la declaración TRIPOD, incluyendo la lista de directrices y su explicación correspondiente, están disponibles en www.tripod-statement.org.



Tabla 2. Lista de comprobación de la declaración TRIPOD para comunicar adecuadamente estudios de desarrollo o validación de modelos predictivos multivariantes.

SECTION/TOPIC	ITEM	CHECKLIST ITEM
TITLE AND ABSTRACT		
Title	1	D;V Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.
Abstract	2	D;V Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.
INTRODUCTION		
Background and objectives	3a	D;V Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.
	3b	D;V Specify the objectives, including whether the study describes the development or validation of the model or both.
METHODS		
Source of data	4a	D;V Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.
	4b	D;V Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.
Participants	5a	D;V Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.
	5b	D;V Describe eligibility criteria for participants.
	5c	D;V Give details of treatments received, if relevant.
Outcome	6a	D;V Clearly define the outcome that is predicted by the prediction model, including how and when assessed.
	6b	D;V Report any actions to blind assessment of the outcome to be predicted.
Predictors	7a	D;V Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.
	7b	D;V Report any actions to blind assessment of predictors for the outcome and other predictors.
Sample size	8	D;V Explain how the study size was arrived at.
Missing data	9	D;V Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.
Statistical analysis methods	10a	D Describe how predictors were handled in the analyses.
	10b	D Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.
	10c	V For validation, describe how the predictions were calculated.
	10d	D;V Specify all measures used to assess model performance and, if relevant, to compare multiple models.
	10e	V Describe any model updating (e.g., recalibration) arising from the validation, if done.

SECTION/TOPIC	ITEM	CHECKLIST ITEM
Risk groups	11	D;V Provide details on how risk groups were created, if done.
Development vs. Validation	12	V For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.
RESULTS		
Participants	13a	D;V Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.
	13b	D;V Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.
	13c	V For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).
Model development	14a	D Specify the number of participants and outcome events in each analysis.
	14b	D If done, report the unadjusted association between each candidate predictor and outcome.
Model specification	15a	D Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).
	15b	D Explain how to use the prediction model.
Model performance	16	D;V Report performance measures (with CIs) for the prediction model.
Model-updating	17	V If done, report the results from any model updating (i.e., model specification, model performance).
DISCUSSION		
Limitations	18	D;V Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).
Interpretation	19a	V For validation, discuss the results with reference to performance in the development data, and any other validation data.
	19b	D;V Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.
Implications	20	D;V Discuss the potential clinical use of the model and implications for future research.
OTHER INFORMATION		
Supplementary information	21	D;V Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.
Funding	22	D;V Give the source of funding and the role of the funders for the present study.

1.3.2.1. Ejemplos de revisiones sistemáticas de modelos predictivos anteriores a la declaración TRIPOD

La aparición de la declaración TRIPOD en 2015 aportó una nueva metodología para ayudar a los investigadores a comunicar adecuadamente los resultados de sus investigaciones, y contribuir de esta manera a mejorar la calidad y fiabilidad de las publicaciones, como ya se explicó en las secciones anteriores. El objetivo de esta sección es precisamente enfatizar la necesidad de unas directrices para la correcta comunicación de los resultados de estudios sobre modelos predictivos antes de la publicación de la declaración TRIPOD. Para ello, se expondrán brevemente los resultados correspondientes a tres revisiones sistemáticas de estudios sobre modelos de riesgo aplicados a la predicción de diabetes tipo 2 (fecha de la revisión: 2011), ERC (fecha de la revisión: 2013) y ECV (fecha de la revisión: 2016), todos ellos anteriores a la aparición de la declaración TRIPOD. En todas estas revisiones sistemáticas se encontró que un elevado porcentaje de los estudios sobre modelos de riesgo publicados no estaban adecuadamente descritos y contenían graves errores en el tratamiento estadístico de los datos.

Revisión 1: *Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting*⁴⁰

- **Descripción:** Búsqueda sistemática de estudios anteriores a mayo de 2011 describiendo el desarrollo de modelos predictivos que combinan dos o más variables para predecir el riesgo de prevalencia (casos no diagnosticados) o incidencia (casos diagnosticados) de diabetes tipo 2, y posterior extracción de la información clave de ellos: diseño del estudio, tamaño de la muestra y número de eventos, selección de los predictores, tratamiento de datos perdidos, estrategias para el desarrollo del modelo y rendimiento, entre otros. La revisión sistemática abarca 39 estudios, con la descripción de un total de 43 modelos de riesgo.
- **Resultados:** La revisión sistemática determinó que en nueve de los estudios analizados (23%) el número de eventos por variable (EPV) era reducido (inferior a 10), mientras que en catorce de ellos la información disponible no era suficiente para su determinación. El número de

predictores de riesgo en la mayor parte de los estudios oscilaba entre 4 y 64, y en siete de ellos su número no estaba adecuadamente descrito. Además, el método para su selección era inadecuado o no estaba adecuadamente descrito. Finalmente, aproximadamente la mitad de los estudios revisados (41%) no consideraba el tratamiento y manipulación de los datos perdidos. En conclusión, la revisión sistemática puso de manifiesto una descripción incompleta de los estudios publicados, así como una metodología en muchos casos errónea, lo que en último término comprometía seriamente la fiabilidad de las predicciones de los modelos y la calidad de las publicaciones, dificultando una valoración crítica de las mismas.

Revisión 2: A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods⁴¹

- **Descripción:** Búsqueda sistemática de estudios publicados hasta septiembre de 2011 describiendo el desarrollo de modelos predictivos que combinan dos o más variables para predecir el riesgo de prevalencia o incidencia de ERC o de evolución a enfermedad renal terminal (ERT), y posterior extracción de la información clave de ellos: diseño del estudio, tamaño de la muestra y número de eventos, selección de los predictores, tratamiento de datos perdidos, estrategias para el desarrollo del modelo y rendimiento, entre otros. La revisión sistemática abarca 11 estudios, con la descripción de un total de 14 modelos de riesgo. Más precisamente, 8 de los estudios describían 11 modelos de riesgo orientados a la predicción de ERC y ERT incidentes, mientras que 3 de ellos describían modelos para la predicción de ERC prevalente.
- **Resultados:** Los estudios analizados consideraron 97 predictores de riesgo, de los cuales 43 fueron finalmente empleados en total. La revisión sistemática determinó además que el método para su selección en 6 de los estudios considerados era inadecuado. Finalmente, prácticamente la totalidad de los estudios considerados no consideraba el tratamiento y

manipulación de los datos perdidos, o lo hacía de manera errónea. Al igual que en el ejemplo anterior, la revisión sistemática concluyó que la descripción de los estudios publicados era a grandes rasgos incompleta, y la metodología en muchos casos errónea, arrojando severas dudas sobre la fiabilidad de las predicciones de prevalencia o incidencia de ERC y ERT de los modelos estudiados.

Revisión 3: Prediction models for cardiovascular disease risk in the general population: systematic review^{A2}

- **Descripción:** Búsqueda sistemática de estudios publicados hasta junio de 2013 describiendo el desarrollo o validación externa de modelos predictivos multivariados para predecir el riesgo de ECV en la población general. Durante la posterior extracción de la información clave de cada uno de ellos, la revisión diferenció claramente entre los estudios referidos al desarrollo y a la validación externa de los modelos, respectivamente. Tras seleccionar 9965 referencias en la búsqueda bibliográfica, únicamente 212 de ellas cumplían los requisitos de elegibilidad impuestas por la revisión, describiendo en conjunto un total de 363 modelos predictivos. De las 212 publicaciones seleccionadas, 76 describían el desarrollo de uno o más modelos predictivos para predecir el riesgo de ECV, 49 describían tanto el desarrollo como la validación externa de dichos modelos, y 87 describían únicamente su validación externa.
- **Resultados:** La revisión sistemática encontró una gran heterogeneidad en los predictores utilizados en los diferentes estudios. En muchos casos, las publicaciones carecían además de información clínica o metodológica crucial para la correcta valoración de los estudios. Más concretamente, en 49 modelos (13%) el horizonte predictivo (esto es, como de lejos en el futuro es el modelo capaz de realizar predicciones) no estaba especificado, y en 92 de ellos (25%) faltaban datos cruciales para su correcta utilización en la predicción de riesgos individuales. Además, únicamente 132 de los modelos desarrollados (36%) habían sido externamente validados, de los cuales tan sólo 70 (19%) lo habían sido por

investigadores independientes. La conclusión general de la revisión sistemática fue que la utilidad de los modelos predictivos descritos en las publicaciones bajo consideración no era evidente, debido fundamentalmente a errores metodológicos, a la presentación incompleta de los resultados, y a la ausencia de validación externa y de estudios de impacto de dichos modelos.



2. JUSTIFICACIÓN



2. JUSTIFICACIÓN

Tal y como se explicó en las secciones anteriores, los modelos predictivos son herramientas útiles en la práctica clínica habitual para determinar la ocurrencia de un suceso en base a una serie de parámetros, ya que podemos disminuir el riesgo del suceso o saber en qué pacientes tenemos que mostrar un seguimiento más estricto.

En el caso concreto de la FA son conocidos una serie de factores de riesgo, los cuales pueden clasificarse como modificables y no modificables.

Un ejemplo de los factores de riesgo no modificables sería la edad, ya que un paciente siempre va a cumplir años y en consecuencia no podemos reducir su riesgo en base a este parámetro. A los factores de riesgo no modificables también se les denomina marcadores de riesgo.

Por otro lado, la obesidad es un factor de riesgo modificable para el desarrollo de FA que sí que podemos modificar, ya que siempre es posible indicar al paciente medidas higiénico-dietéticas que logren que este pierda peso corporal y en consecuencia, lograr una disminución de su IMC.

Si logramos que el conjunto de todos estos factores ofrezca una predicción precisa de qué paciente desarrollará FA, dispondremos de una herramienta realmente útil para disminuir su incidencia, actuando en los diferentes factores de riesgo modificables, y reduciendo con ello la incidencia de ACV, IC, muerte súbita y morbilidad CV. No obstante, si estos modelos predictivos no están contruidos o validados internamente de forma correcta en su metodología estadística, presentarán limitaciones serias a la hora de ser empleados en la práctica clínica. Dicho de otro modo, si estos modelos no han sido implementados con la metodología estadística adecuada, la validez de sus predicciones puede ser puesta en entredicho, y su utilidad en la práctica clínica podría quedar limitada o anulada.

Responder a través de una revisión sistemática a la cuestión de si estos modelos frutos de la investigación clínica tienen los parabienes necesarios para ser utilizados en la práctica clínica sería muy importante a la hora de la prevención primaria en la FA.

La búsqueda de las evidencias expuestas a través de una revisión sistemática constituye uno de los ejercicios de epidemiología clínica más importante a la hora de tomar decisiones con las mejores evidencias científicas basadas en las declaraciones PRISMA y TRIPOD, y que tiene como principal resultado el beneficio de los pacientes a la hora de aplicar los modelos predictivos de incidencias de FA.

Para ello, planteamos una revisión sistemática basada en los postulados de la medicina basada en la evidencia. Siguiendo a la Escuela Canadiense, en donde tras la definición clara y precisa de una pregunta clínica, planteamos y respondemos a las siguientes preguntas con respecto a la validez, metodología y utilidad de los resultados. Estas son:

1. ¿Se explicitaron los criterios de selección de los artículos para la revisión?
2. ¿Se evaluó la validez de los estudios fundamentales?
3. ¿La evaluación de los estudios fundamentales fue reproducible?
4. ¿Se analizó la variabilidad en los resultados de los estudios?
5. ¿Se combinaron apropiadamente los hallazgos de los estudios fundamentales?

3. HIPÓTESIS



3. HIPÓTESIS

Los modelos predictivos de incidencia de FA podrían presentar limitaciones en su metodología estadística que dificultaran su implementación en la práctica clínica habitual, al igual que se ha observado en revisiones similares anteriores realizadas para otras patologías, como la diabetes mellitus⁴⁰, la ERC⁴¹ o la ECV⁴².



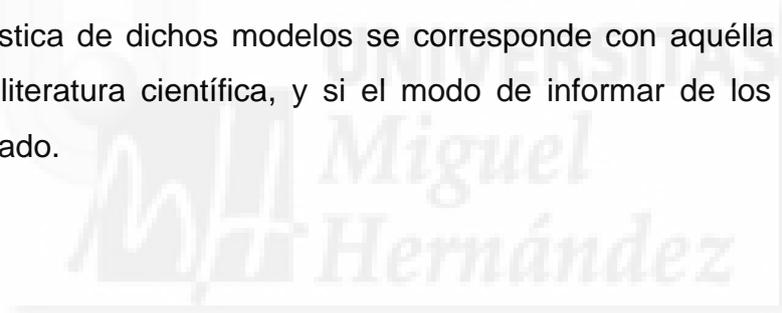
4. OBJETIVOS



4. OBJETIVOS

El objetivo de este trabajo es la realización de una revisión sistemática de publicaciones científicas en las que se describa el desarrollo y validación interna de modelos predictivos de FA aplicables en población general, y analizar si la metodología estadística y el modo en que los resultados son comunicados son los más adecuados. Más precisamente, el trabajo será dividido en las siguientes dos fases bien diferenciadas:

1. Realización de una revisión sistemática de acuerdo a la declaración PRISMA, utilizando las bases de datos MEDLINE, Scopus y Google Scholar para efectuar una búsqueda bibliográfica exhaustiva y determinar artículos científicos que construyan y validen internamente modelos predictivos de FA que pueden aplicarse en población general.
2. Utilización de la declaración TRIPOD para analizar si la metodología estadística de dichos modelos se corresponde con aquella recomendada en la literatura científica, y si el modo de informar de los resultados es adecuado.



5. MATERIAL Y MÉTODOS



5. MATERIAL Y MÉTODOS

5.1. Diseño del estudio

Se realizó una revisión sistemática de los artículos científicos que describían el desarrollo y validación interna de modelos predictivos multivariantes, utilizando como variable principal el desarrollo de FA en pacientes de población general y sin diagnóstico previo de esta patología. Se analizaron las bases de datos MEDLINE (a través de Pubmed), Scopus y Google Scholar utilizando las siguientes palabras clave:

1. MEDLINE: ("Atrial Fibrillation"[Mesh] OR "atrial fibrillation"[title/abstract]) AND ("Models, Statistical"[Mesh] OR "Nomograms"[Mesh] OR "predictive model"[title/abstract] OR "nomogram"[title/abstract] OR "scoring system"[title/abstract] OR "points system"[title/abstract] OR "risk score"[title/abstract] OR "prediction model"[title/abstract]) AND ("cohort"[title/abstract] OR "follow-up"[title/abstract] OR "longitudinal"[title/abstract] OR "development"[title/abstract] OR "incidence"[title/abstract]).
2. Scopus: TITLE-ABS-KEY ("atrial fibrillation") AND (TITLE-ABS-KEY ("predictive model") OR TITLE-ABS-KEY ("scoring system") OR TITLE-ABS-KEY ("nomogram") OR TITLE-ABS-KEY ("risk score") OR TITLE-ABS-KEY ("prediction model")) AND (TITLE-ABS-KEY ("incidence") OR TITLE-ABS-KEY ("cohort") OR TITLE-ABS-KEY ("longitudinal") OR TITLE-ABS-KEY ("follow-up") OR TITLE-ABS-KEY ("development")).
3. Google Scholar: ("predictive model" OR "scoring system" OR nomogram OR "risk score" OR "prediction model") AND "atrial fibrillation".

Se restringieron los artículos a aquellos publicados en inglés o en castellano que tenían resumen. La búsqueda bibliográfica se realizó incluyendo todos los artículos publicados antes del 26 de septiembre de 2016. Además, se analizaron las referencias de todos los trabajos que podían ser relevantes para cumplir nuestro objetivo primario.

No se analizaron las bases de datos EMBASE, Web of Knowledge, PAHO, WHOLIS, LILACS. Las causas de no utilizarlas para esta tesis doctoral quedan reflejadas en el punto 7.3. Limitaciones del estudio.



5.2. Criterios de inclusión

Se seleccionaron aquellos trabajos científicos que describieran el desarrollo y validación interna de un modelo predictivo para determinar la incidencia de FA en población general. Por otra parte, se excluyeron de la selección los siguientes tipos de artículos científicos:

1. Artículos de revisión.
2. Artículos describiendo el desarrollo de un modelo predictivo sin utilizar datos propios.
3. Artículos describiendo la validación externa de un modelo predictivo ya existente.



5.3. Extracción de los datos, análisis y presentación de resultados

Dado que la búsqueda bibliográfica se realizó en tres bases de datos diferentes, primeramente se identificaron las referencias duplicadas para ser eliminadas mediante la transformación de las búsquedas obtenidas en una única base de datos. Para ello, los registros obtenidos de las tres bases de datos fueron ordenados alfabéticamente en una única base de datos, y con este orden se determinó qué registros estaban duplicados. Tras este proceso, se analizaron los títulos y los resúmenes para excluir aquellos artículos científicos que no describieran modelos predictivos de FA en la comunidad. Este análisis fue realizado de forma paralela por otro revisor. Tras llegar a un consenso, todos aquellos trabajos científicos potenciales para el objetivo principal fueron examinados a texto completo.

Posteriormente, las referencias seleccionadas fueron analizadas a texto completo para determinar si estas cumplían todos los criterios de inclusión y ninguno de exclusión, lo cual fue realizado también de forma paralela e independiente por el otro revisor. Los resultados obtenidos fueron consensuados, no encontrándose ninguna discrepancia. Tras este análisis se procedió a la extracción de resultados de interés de las publicaciones seleccionadas.

Los datos extraídos de cada publicación fueron:

1. Diseño del estudio.
2. Tamaño de la muestra (número de pacientes totales analizados) y número de eventos (pacientes que desarrollaron FA en el estudio).
3. Definición de la variable principal.
4. Selección y categorización de las variables explicativas para predecir FA (factores de riesgo o predictores).
5. Manejo de datos perdidos.
6. Forma de construir el modelo predictivo.
7. Aplicabilidad del modelo en la práctica clínica habitual.

8. Aspectos de la validación interna: calibración y discriminación.

La elección de estos ítems está basada en la declaración TRIPOD y en revisiones sistemáticas realizadas para la diabetes mellitus⁴⁰, la ERC⁴¹ o la ECV⁴². Tras la extracción de todos los ítems de acuerdo a la declaración TRIPOD, se analizó cada modelo seleccionado para ver cuáles de ellos se satisfacían, y con ello realizar un análisis global de todos los modelos encontrados en nuestra búsqueda bibliográfica. La forma de redactar la revisión sistemática se basó en la declaración PRISMA, la cual indica los estándares internacionales para este tipo de estudio.



6. RESULTADOS



6. RESULTADOS

6.1. Esquema del estudio

El número total de referencias al inicio del estudio obtenidas en cada una de las bases de datos analizadas era:

1. MEDLINE: 1649.
2. Scopus: 1010, de los cuales hubo 74 sin resumen, lo que hizo un total de 936 trabajos para valorar.
3. Google Scholar: 199.

Juntando las publicaciones de estas tres bases de datos se obtuvo un total de 2784 registros (Fig. 3). Estos registros fueron ordenados alfabéticamente para así eliminar los duplicados, quedando un total de 2380 artículos. De estos, 2367 fueron excluidos ya que ni el título ni el resumen indicaban que la publicación describiera un modelo predictivo de las características planteadas, dejando un total de 13 artículos⁴³⁻⁵⁵ que fueron analizados a texto completo. El análisis de los mismos hizo que 9 publicaciones⁴³⁻⁵¹ más fueran excluidas: 2 por ser revisiones^{43,44}, 1 por no contar con datos propios⁴⁵ y 6 que realizaron únicamente validaciones externas⁴⁶⁻⁵¹ de modelos predictivos ya publicados previamente. En consecuencia, finalmente fueron analizados 4 artículos científicos⁵²⁻⁵⁵ determinando las características mencionadas en nuestra metodología.

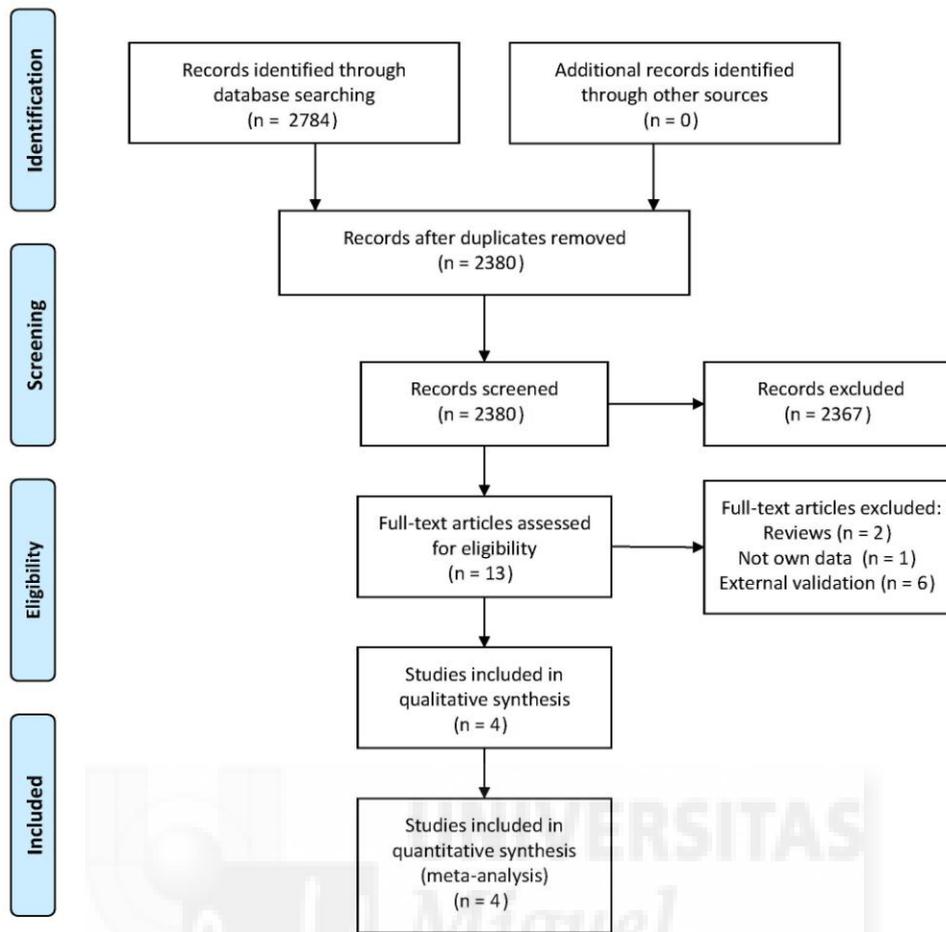


Figura 3. Esquema del estudio siguiendo la declaración PRISMA. Tras la búsqueda bibliográfica se obtuvo un total de 2784 publicaciones. Tras eliminar los duplicados, este número se redujo a 2380, de los cuales 2367 fueron excluidos, dejando finalmente 13 publicaciones para ser analizadas a texto completo. Tras este análisis, 9 más fueron excluidas, con lo únicamente 4 publicaciones fueron finalmente incluidos en la revisión.

6.2. Análisis de los modelos predictivos encontrados

En la Tabla 3 se observan las características extraídas de los 4 artículos finales⁵¹⁻⁵⁴ incluidos en nuestra revisión sistemática. Todos los modelos han recibido un acrónimo en la publicación correspondiente, lo que ayuda a la hora de ser comentados: *Atherosclerosis Risk In Communities*⁵² (ARIC), *Cohorts for Heart and Aging Research in Genomic Epidemiology*⁵³ (CHARGE-AF), *Framingham Heart Study*⁵⁴ (FHS) y *Women's Health Study*⁵⁵ (WHS). A continuación, comentaremos con más detalle cada una de las características:

1. Variable principal: se destaca que todos los modelos de predicción determinan la incidencia de FA a los 10 años, salvo el CHARGE-AF que lo hace hasta la mitad del tiempo, es decir, a los 5 años.
2. Tamaño de la muestra del estudio y número de eventos: el tamaño de la muestra osciló entre 4,764 y 18,556 pacientes, y los eventos entre 404 y 1,186.
3. El número de predictores tuvo un rango entre 6 y 26.
4. Razón de EPV: todos los estudios tuvieron un valor de esta razón superior a 10, siendo el mínimo valor de 19.8 y el máximo de 107.8.
5. Todos los estudios realizaron un tratamiento de los predictores continuos, como potencias, categorizaciones y logaritmos, y análisis de interacciones.
6. Todos los estudios analizaron sólo los casos completos y el FHS no indicó cómo procedió con los pacientes que no tenían todas las variables del estudio recogidas.
7. En lo referente a los criterios de inclusión de las variables, todos los modelos se basaron en modelos por pasos, salvo el FHS que no indica dicho procedimiento.
8. Dos estudios utilizaron sistemas de puntos y los restantes aplicaban la fórmula de Cox directamente, aunque el WHS no permite su aplicación por falta de conocimiento de la supervivencia basal.
9. En la discriminación, el ARIC utilizó el área bajo la curva ROC (AUC-ROC), mientras que el resto utilizó el estadístico C. La calibración fue realizada

mediante gráficos y pruebas basadas en el test χ^2 . La técnica de *bootstrapping* fue empleada por todos los estudios, salvo en el CHARGE-AF.



Tabla 3. Características principales de los modelos predictivos de fibrilación auricular seleccionados en la revisión sistemática.

CARACTERÍSTICAS	ARIC	CHARGE-AF	FHS	WHS
Variable principal	FA incidente en 10 años	FA incidente en 5 años	FA incidente en 10 años	FA incidente en 10 años
#pacientes/eventos	14,546 y 515	18,556 y 1,186	4,764 y 457	13,061 y 404
#predictores	26	11-14	12	6
EPV	19.8	84.7-107.8	38.1	67.3
Tratamiento de los predictores continuos	Categorizaciones e interacciones	Categorizaciones e interacciones	Potencias e interacciones	Logaritmos
Datos perdidos	Sólo casos completos	Sólo casos completos	No indicado	Sólo casos completos
Construcción del modelo (elección de variables)	Modelo de Cox por pasos hacia atrás	Modelo de Cox por pasos hacia atrás	Modelo de Cox: sin detalle sobre el descarte de variables	Modelo de Cox por pasos usando BIC
Presentación del modelo	Sistema de puntos	Aplicación directa de la fórmula del modelo de Cox	Sistema de puntos	Aplicación directa de la fórmula del modelo de Cox (faltan datos para aplicarla)
Validación	AUC-ROC y gráfico de calibración: <i>bootstrapping</i>	Estadístico C, pendiente de discriminación y estadísticos χ^2 : no <i>bootstrapping</i>	Estadístico C y gráfico de calibración: <i>bootstrapping</i>	Estadístico C, y estadísticos χ^2 : <i>bootstrapping</i>

7. DISCUSIÓN



7. DISCUSIÓN

7.1. Interpretación de los resultados de la discusión

Tras analizar alrededor de 2400 artículos obtenidos de las bases de datos MEDLINE, Scopus y Google Scholar, se han encontrado 4 publicaciones⁵²⁻⁵⁵ relevantes que describen el desarrollo y validación interna de un modelo predictivo para determinar la incidencia de FA en la comunidad. Tras analizar los resultados, es posible observar que existen algunas deficiencias en la metodología estadística y en la forma de detallarla en la publicación asociada, si bien en general los aspectos clave del desarrollo y validación del modelo predictivo han sido realizados con exactitud y rigurosidad en todos ellos.

Uno de los componentes clave a la hora de construir un modelo de las características planteadas es la razón de EPV, el cual tiene que ser mayor o igual a 10 para evitar el *overfitting*^{56,57}. Como podemos observar en la Tabla 3 y en el análisis de los resultados, todos los modelos analizados cumplieron este criterio de forma satisfactoria, ya que el mínimo valor de esta razón fue de 19.8. De haber existido *overfitting*, los resultados obtenidos al validar externamente el modelo en otro conjunto de datos no serían satisfactorios, esto es, el modelo no podría ser utilizado con fiabilidad en la práctica clínica habitual⁵⁸⁻⁶¹. El porcentaje de estudios referidos a otras enfermedades con una razón de EPV valor inferior a 10 ha oscilado entre 19.8 y 107.8, por lo que podemos decir que los modelos de predicción de FA en la comunidad son más rigurosos en este sentido que otros modelos predictivos.

A continuación, analizaremos el tratamiento de los predictores continuos, es decir, aquellas variables que son cuantitativas y continuas, tales como la edad o las cifras de tensión arterial. Estos predictores pueden tratarse utilizando la variable en su forma original, realizando transformaciones o bien categorizando en subgrupos⁶². La última opción es una práctica muy común en la literatura científica^{63,64}, estableciendo grupos según criterios clínicos basados en la media o en cuartiles, según cierto valor de p , etc. Si bien estas categorizaciones son muy frecuentes en la elaboración de modelos predictivos, no obstante es necesario tener en cuenta que su utilización no es en absoluto recomendable⁶². La realización de categorizaciones de sólo dos categorías de variables continuas

(dicotomía) suele resultar en una pérdida de información y de potencia estadística⁶⁵. Esto es debido a que los datos pueden tener una relación no lineal con la variable principal, por lo que es recomendable la utilización de *splines* o de funciones polinómicas⁶⁶. En este estudio se observó que el 50% de los modelos presentaron categorizaciones a la hora de manejar los predictores continuos. Estas cifras no difieren significativamente en revisiones sistemáticas de modelos predictivos desarrollados para otras patologías⁴⁰⁻⁴².

Respecto al tratamiento de los datos perdidos, fue posible observar que la mayoría de los estudios seleccionados empleó únicamente los casos completos, y que uno de ellos (el FHS) no ha indicado la metodología empleada para su análisis. Aunque esta es una práctica común, se ha comprobado que puede sesgar los resultados de la revisión⁶⁷. Por ello, es más recomendable aplicar la técnica de imputación múltiple^{68,69} sobre los datos perdidos ya que permite obtener resultados mucho más precisos. La técnica de la imputación consiste en realizar una estimación de los valores perdidos en base a los valores completos del resto de pacientes, esto es, en encontrar un patrón para poder “rellenar” los huecos correspondientes a los datos perdidos para con ello aumentar el tamaño muestral y no sesgar los resultados obtenidos. Por ello, sería deseable conocer qué procedimiento se ha seguido con los datos perdidos en esta publicación, para así poder validar los resultados publicados.

La siguiente cuestión que se aborda en la Tabla 3 tras los datos perdidos se refiere a la forma de construir el modelo multivariable. Generalmente, los modelos predictivos multivariables han sido desarrollados utilizando únicamente tests estadísticos como criterio a la hora de introducir o desechar una variable, esto es, algunos investigadores añaden una variable si su adicción al modelo es estadísticamente significativa o la eliminan si no muestra diferencias estadísticamente significativas con respecto al modelo sin ella. Otro modo de establecer qué variables deben ser introducidas en el modelo multivariable podría consistir en seleccionar solamente aquellas variables que sí fueron significativas en el análisis bivalente. Esta cuestión es clave, ya que aunque sean predictores no significativos, a la hora de realizar una predicción es posible que globalmente actúen de forma diferente al ajustarse por el resto de factores^{70,71} y que, en consecuencia, el modelo tenga mayor precisión. Con respecto a la FA, todos los

modelos utilizaron pasos basados en tests estadísticos, salvo el FHS, en el que no se indica cómo las variables del modelo final fueron seleccionadas.

Las principales diferencias entre los cuatro modelos surgen a la hora de estudiar su aplicabilidad en la práctica clínica. Dos de ellos han elaborado un sistema de puntos (ARIC y FHS), mientras que el resto (CHARGE-AF y WHS) son de aplicación directa de la fórmula de Cox. Para poder aplicar un sistema de puntos, el clínico tiene que conocer los valores de los diferentes predictores, los cuales darán una puntuación para cada predictor. Dichas puntuaciones son sumadas y dan un riesgo de desarrollo de FA. Este procedimiento es muy sencillo, y no necesita de dispositivos electrónicos para su cálculo, facilitando considerablemente su implementación de forma sistemática en las consultas médicas. Por otro lado, los dos modelos restantes (CHARGE-AF y WHS) necesitan un dispositivo electrónico para poder ejecutar los cálculos. Además, el último modelo analizado (WHS) no incluye todos los parámetros para su cálculo, lo que en última instancia impide su aplicación real en la práctica clínica.

El último punto de la Tabla 3 detalla los métodos para validar el modelo predictivo desarrollado en las cuatro publicaciones analizadas. Este paso es muy importante, ya que una vez el modelo ha sido desarrollado los autores deben asegurarse que este funciona a la hora de determinar la incidencia de FA. Existen diversos métodos de comprobar si el modelo “funciona” internamente: 1) utilización del mismo conjunto de datos que se utilizó en el desarrollo del modelo; 2) partir de este mismo conjunto de datos pero dividido en dos trozos, uno para construir y otro para validar el modelo; y 3) aplicar técnicas de *bootstrapping*^{72,73}. De todas ellas, la más recomendable es el *bootstrapping*, el cual ha sido aplicado en el 75% de los modelos, siendo el CHARGE-AF el único que no lo ha utilizado. A continuación, es necesario comprobar la discriminación y la calibración en la muestra de validación (o en las muestras *bootstrap*). La discriminación en los modelos de Cox debería ser analizada mediante el cálculo del estadístico C, y no mediante el AUC-ROC como en el modelo ARIC, ya que su valor no tiene en cuenta la existencia de datos censurados. Por otro lado, la calibración (comparación de eventos observados y esperados) se ha realizado en todos los casos comparando las tasas de eventos observados y esperados por grupos de

riesgo. Esta aproximación es correcta, pero sería mucho más recomendable la utilización de curvas suaves mediante *splines* lineales o *loess*⁷⁴.



7.2. Respuesta a las preguntas en una revisión sistemática de la Escuela de Epidemiología Clínica Canadiense

La revisión sistemática realizada es una revisión crítica muy relevante para la práctica clínica de los profesionales sanitarios, y además obtener una respuesta afirmativa a las 6 preguntas realizadas por la Escuela de Epidemiología Clínica Canadiense⁷⁵ a la hora de establecer su utilidad. Estas son:

- Se han establecido claramente las preguntas formuladas en la justificación del estudio, la hipótesis y los objetivos. Esta formulación se basa fundamentalmente en el método científico.
- Se ha explicado en la metodología de la revisión sistemática de forma muy exhaustiva los criterios de selección de los artículos para la revisión.
- En el departamento de medicina clínica y en la unidad de investigación del Hospital Universitario de Elda, a través de un trabajo sistemático y meticuloso se ha evaluado la validez de los estudios fundamentales.
- Se ha analizado de forma muy detallada que la evaluación de los estudios fundamentales fuera reproducible y tras el análisis exhaustivo del apartado material y métodos de los trabajos la respuesta fue afirmativa.
- Se ha analizado la variabilidad en los resultados de los estudios analizados.
- En los estudios seleccionados se combinaron apropiadamente los hallazgos de estos estudios.

También en su metodología se ha respetado las diferentes fases de una revisión sistemática a través de responder a un pregunta clínica concreta mediante la definición del objetivo, identificación de los estudios disponibles utilizando cuatro bases de datos, explicitando los criterios de selección, realizando una evaluación crítica de los estudios seleccionados y valorando la extracción de los datos, realizando estrategias de análisis. Por último, tal y como se comenta en el siguiente punto, se valoran las limitaciones del estudio identificando los sesgos que pueden amenazar a la validez de las conclusiones.

7.3. Limitaciones del estudio

Esta revisión sistemática se ha limitado a artículos publicados en inglés/castellano y no se tuvo en consideración la literatura gris (esto es, literatura que no sigue los canales convencionales de publicación). De este modo, sólo se valoraron publicaciones indexadas en las tres bases de datos analizadas (MEDLINE, Scopus y Google Scholar), por lo que es posible que algún otro modelo predictivo de las características analizadas no haya sido incluido.

Se han analizado artículos que valoran modelos de predicción en la literatura científica, restringiéndonos a aquellas publicaciones con alto factor de impacto, es decir, aquéllas que pertenecen al primer cuartil dentro de su categoría en el Journal Citation Reports. En estos artículos se ha valorado qué bases de datos han consultado para realizar su revisión sistemática y un análisis descriptivo de las mismas. Los datos crudos quedan reflejados en la Tabla 4. En ella observamos el número total de bases de datos utilizadas, teniendo una mediana de 2 con un rango intercuartílico de 1. En ella también destacamos la no utilización de las bases de datos Web of Knowledge, PAHO, WHOLIS, LILACS.

Tabla 4. Bases de datos empleados por los artículos científicos que han valorado modelos de predicción publicados en revistas de alto factor de impacto (cuartil1 del Journal Citation Reports).

DOI	Revista	MEDLINE	EMBASE	COCHRANE	CINAHL	Total
10.1186/1471-2288-14-40	BMC Med Res Methodol	Sí	No	No	No	1
10.1136/bmj.i2416	BMJ	Sí	Sí	No	No	2
10.1016/j.ajog.2015.06.013	Am J Obstet Gynecol	Sí	No	No	No	1
10.1186/1741-7015-9-103	BMC Med	Sí	Sí	No	No	2
10.1016/j.jclinepi.2012.06.020	J Clin Epidemiol	Sí	No	No	No	1
10.1002/bjs.10571	Br J Surg	Sí	Sí	Sí	No	3
10.1136/bmj.h4395	BMJ	Sí	Sí	No	Sí	3
10.1093/bja/aew476	Br J Anaesth	Sí	Sí	No	No	2
10.1002/14651858.CD012022	Cochrane Database Syst Rev	Sí	Sí	No	No	2

En lo referente a la distribución de las bases de datos utilizadas, observamos que la mayoría de estudios emplearon MEDLINE (100%) y EMBASE (66.67%). Además, en lo referente al número de bases utilizadas sólo un 22.2% empleó 3 bases de datos, cuando nosotros utilizamos un total de 3, lo que da mayor precisión a los resultados, incluyendo la más empleada en la literatura científica para la cuestión analizada (valoración de modelos de predicción). Esto se representa visualmente en la Fig. 4.

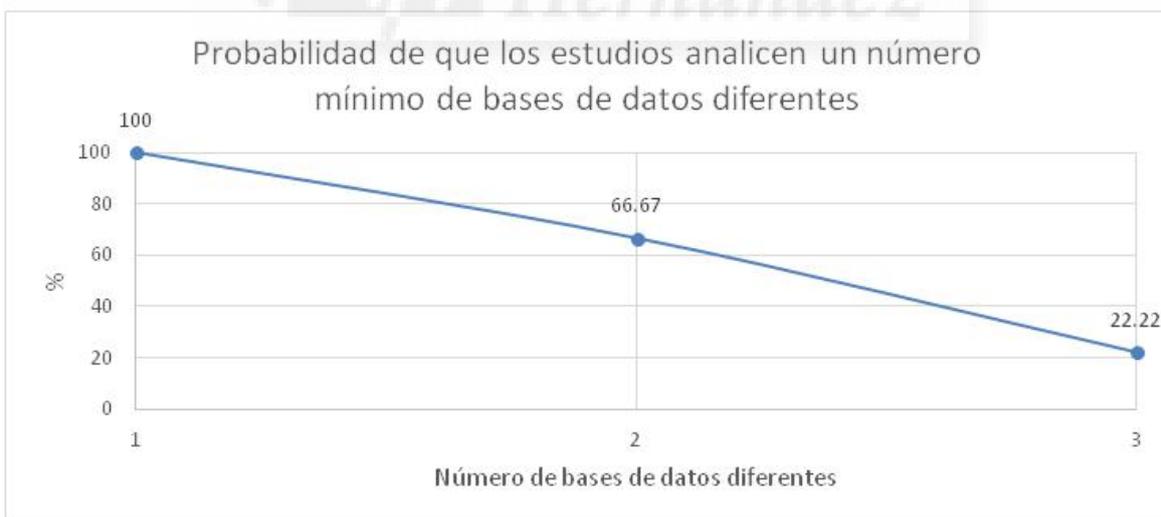
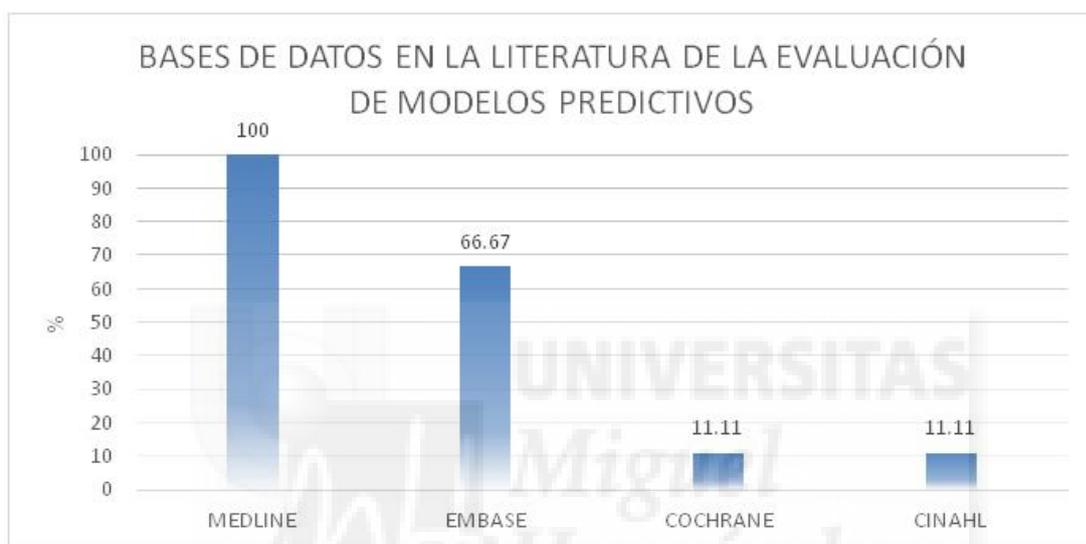


Figura 4. Distribución de las frecuencias de las bases de datos utilizadas por los artículos científicos que han valorado modelos de predicción publicados en revistas de alto factor impacto (cuartil 1 del Journal Citation Reports).

En consecuencia, estas bases de datos son de las más utilizadas en la literatura científica de la evaluación de modelos de predicción, siendo el número total de bases de datos consultadas superior al percentil 75 del número de bases de datos totales utilizadas en las revisiones sistemáticas publicadas en revistas científicas de alto impacto valorando modelos de predicción.

Los sesgos aceptados en una revisión sistemática pueden ser el de publicación, el de localización de estudios, el del idioma inglés, el de la base de datos, el de citación, el de publicación múltiple, el de datos ausentes o de mala calidad, y el de la definición de los criterios de inclusión.

Es posible introducir un sesgo cuando se establece los criterios de selección ya que estos se consensuan por unos investigadores que tienen mucha experiencia en la pregunta clínica realizada. Para minimizar este sesgo hemos consultado a varios expertos en FA y en metodología estadística con experiencia en revisiones sistemáticas.

Habitualmente, la búsqueda de las evidencias se realiza en una o varias bases de datos. El protocolo de revisión sistemática debe utilizar las bases de datos consultadas. La base de datos más utilizada es MEDLINE⁷⁶. Tal como se observa en la Fig. 4, en el 100% de los casos se utiliza MEDLINE. Además, las revisiones con tres bases de datos constituyen un criterio de calidad, ya que solo una de cada 5 revisión utilizan este número de bases de datos (la mayoría de las revisiones solo utilizan una o dos bases de datos).

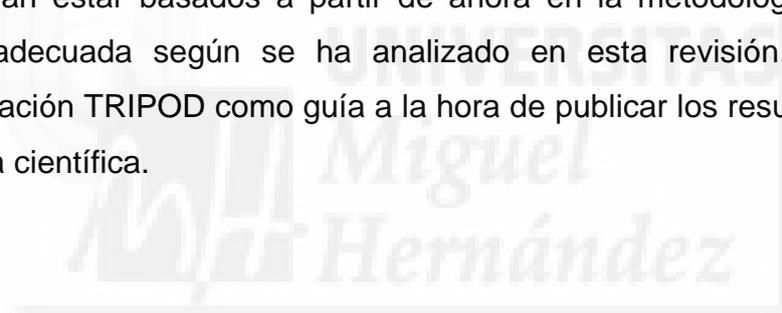
La base de datos EMBASE no fue utilizada en este estudio ya que ningún centro sanitario o universidad de la Comunidad Valenciana están suscritos a esta base de datos.

8. CONCLUSIONES



8. CONCLUSIONES

1. En esta revisión sistemática de las bases de datos MEDLINE, Scopus y Google Scholar fueron seleccionados 4 estudios publicados que describían el desarrollo y validación interna de un modelo predictivo para determinar incidencia de FA en población general.
2. El análisis de los cuatro modelos encontrados resaltó varias deficiencias metodológicas y carencias a la hora de presentar la forma de construcción y de validación interna del modelo predictivo.
3. La falta de rigor a la hora de presentar los cuatro modelos predictivos encontrados podría ser debida a que todos estos fueron publicados con anterioridad a la declaración TRIPOD (2015).
4. Todos los modelos predictivos de la incidencia de FA en la comunidad deberían estar basados a partir de ahora en la metodología estadística más adecuada según se ha analizado en esta revisión, y utilizar la declaración TRIPOD como guía a la hora de publicar los resultados en una revista científica.



REFERENCIAS



REFERENCIAS

1. Shenasa M, Camm AJ, Curtis AB, Mitchell LB, Mulder BA, Ozcan C et al. Management of Atrial Fibrillation: A Practical Approach. Oxford: Oxford University Press; 2015.
2. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B et al. 2016 ESC Guidelines for the Management of Atrial Fibrillation Developed in Collaboration With EACTS. *Eur Heart J*. 2016; **37**: 2893–2962.
3. López Gude MJ, Rodríguez Bezos D, Rodríguez Barrios JM. Análisis de coste-utilidad del manejo de la fibrilación auricular concomitante en España. *Gac Sanit*. 2010; **24**: 59-65.
4. Andrade J, Khairy P, Dobrev D, Nattel S. The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. *Circ Res*. 2014; **114**: 1453–1468.
5. Thrall G, Lane D, Carroll D, Lip GY. Quality of life in patients with atrial fibrillation: a systematic review. *Am J Med*. 2006; **119**: 448.e1–448.e19.
6. Dorian P, Jung W, Newman D, Paquette M, Wood K, Ayers GM et al. The impairment of health-related quality of life in patients with intermittent atrial fibrillation: implications for the assessment of investigational therapy. *J Am Coll Cardiol*. 2000; **36**: 1303–1309.
7. Wynn GJ, Todd DM, Webber M, Bonnett L, McShane J, Kirchhof P et al. The European Heart Rhythm Association symptom classification for atrial fibrillation: validation and improvement through a simple modification. *Europace*. 2014; **16**: 965–972.
8. Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation*. 2014; **129**: 837–847.
9. Martínez-Rubio A, Pujol Iglesias E, Bonastre Thio M, Guilera Mas E, Guindo Soldevila J. Epidemiología de la fibrilación auricular en España. *Rev Esp Cardiol Supl*. 2013; **13**:3-8.

10. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation*. 1998; **98**: 946–952.
11. Stewart S, Hart CL, Hole DJ, McMurray JJ. A population-based study of the longterm risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/ Paisley study. *Am J Med*. 2002; **113**: 359–364.
12. Zoni-Berisso M, Lercari F, Carazza T, Domenicucci S. Epidemiology of atrial fibrillation: European perspective. *Clin Epidemiol*. 2014; **6**: 213–220.
13. Schnabel RB, Yin X, Gona P, Larson MG, Beiser AS, McManus DD et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet*. 2015; **386**: 154–162.
14. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med*. 2007; **146**: 857–867.
15. Natale A, Jalife J, Jais P, Hocini M, Derval N, Ernst et al. *Atrial Fibrillation: From Bench to Bedside*. Totowa: Humana Press; 2008.
16. Fuster V, Rydén LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA et al. 2011 ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 Guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines developed in partnership with the European Society of Cardiology and in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *J Am Coll Cardiol*. 2011; **57**: e101-98.
17. Pérez-Villacastín J, Pérez Castellano N, Moreno Planas J. Epidemiología de la fibrilación auricular en España en los últimos 20 años. *Rev Esp Cardiol*. 2013; **66**:561-565.
18. Montes-Santiago J, Rodil V, Formiga F, Cepeda JM, Urrutia A. Características y costes de los pacientes ingresados por arritmias cardíacas en España. *Rev Clin Esp*. 2013; **213**: 235-239.

19. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Circulation*. 2015; **131**: 211–219.
20. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2009.
21. Williams C, Brunskill S, Altman D, Briggs A, Campbell H, Clarke M et al. Cost-effectiveness of using prognostic information to select women with breast cancer for adjuvant systemic therapy. *HTA*. 2006; **10**.
22. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part II: Multivariate data analysis-an introduction to concepts and methods. *BJC*. 2003; **89**: 431-436.
23. Bang H, Baraniuk S, Bentler P, Davis BR, Edwards LJ, Miller JP et al. *Epidemiology and Medical Statistics*. Amsterdam: Elsevier; 2008.
24. Anderson KM, Odell PM, Wilson PWF, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J*. 1991; **121**: 293–298.
25. Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med*. 1992; **21**: 384–390.
26. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg*. 1999; **16**: 9–13.
27. Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC et al. A prognostic index in primary breast cancer. *Br J Cancer*. 1982; **45**: 361–366.
28. Le Gall JR, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, Mercier P, Thomas R, Villers D. A simplified acute physiology score for ICU patients. *Crit Care Med*. 1984; **12**: 975–977.
29. Steyberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*. 2010; **21**: 128-138.

30. Sociedad Argentina de Terapia Intensiva. Terapia intensiva. 4^a ed. Buenos Aires: Panamerica; 2007.
31. Cramer D, Howitt D. The SAGE Dictionary of Statistics: A practical resource for students in the social sciences. Trowbridge: SAGE Publications; 2004.
32. Moher D, Altman DG, Kenney FS, Simera I, Wager E, Clarke M et al. Guidelines for Reporting Health Research: A User's Manual. Oxford: Wiley; 2014.
33. www.equator-network.org. The EQUATOR Network website and database is provided by the UK EQUATOR Centre, Centre for Statistics in Medicine, NDORMS, University of Oxford.
34. Cobos-Carbó A, Augustovski F. Declaración CONSORT 2010: actualización de la lista de comprobación para informar ensayos clínicos aleatorizados de grupos paralelos. Med Clin. 2011; **137**: 213-215.
35. Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Declaración de la Iniciativa STROBE (Strengthening the Reporting of Observational studies in Epidemiology): directrices para la comunicación de estudios observacionales. Gac Sanit. 2008; **22**: 144-150.
36. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L. STARD 2015 – An updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015; **351**: h5527.
37. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): Explanation and Elaboration. PLOS Med. 2012; **9**: e1001216.
38. Manterola C, Otzen T, Lorenzini N, Díaz A, Torres-Quevedo R, Claros N. Iniciativas Disponibles para el Reporte de Resultados en Investigación Biomédica con Diferentes Tipos de Diseño. Int J Morphol. 2013; **31**: 945-956.
39. Urrútia G, Bonfill X. Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. Med Clin. 2010; **135**: 507-511.

40. Collins GS, Mallett S, Omar O, Yu L. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Medicine*. 2011; **9**: 1741-7015.
41. Collins GS, Omar O, Shanyinde M, Yu LM. A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *J Clin Epidemiol*. 2013; **66**: 268-277.
42. Damen JAAG, Hooft L, Schuit E, Debray TPA, Collins GS, Tzoulaki I et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*. 2016; **353**: i2416.
43. Kaur A. Hybrid approach: Predictive data mining model for Atrial Fibrillation. *World Congress on Information and Communication Technologies*. 2011; 126-131.
44. Alvaro A, Norby FL. Predicting atrial fibrillation and its complications. *Circ J*. 2016; **80**: 1061–1066.
45. Brunner KJ, Bunch TJ, Mullin CM, May HT, Bair TL, Elliot DW et al. Clinical Predictors of Risk for Atrial Fibrillation: Implications for Diagnosis and Monitoring. *Mayo Clin Proc*. 2014; **89**:1498-1505.
46. Alonso A, Roetker NS, Soliman EZ, Chen LY, Greenland P, Heckbert SR. Prediction of Atrial Fibrillation in a Racially Diverse Cohort: The Multi-Ethnic Study of Atherosclerosis (MESA). *J Am Heart Assoc*. 2016; **5**: e003077.
47. Schnabel RB, Aspelund T, Li G, Sullivan LM, Suchy-Dicey A, Harris TB et al. Validation of an Atrial Fibrillation Risk Algorithm in Whites and African Americans. *Arch Intern Med*. 2010; **170**: 1909-1917.
48. Shulman E, Kargoli F, Aagaard P, Hoch E, Di Biase L, Fisher J. Validation of the Framingham Heart Study and CHARGE-AF Risk Scores for Atrial Fibrillation in Hispanics, African-Americans, and Non-Hispanic Whites. *Am J Cardiol*. 2016; **117**: 76-83.
49. Rodriguez F, Stefanick ML, Greenland P, Soliman EZ, Manson JE, Parikh N et al. Racial and ethnic differences in atrial fibrillation risk factors and

- predictors in women: Findings from the Women's Health Initiative. *Am Heart J.* 2016; **176**:70-77.
50. Christophersen IE, Yin X, Larson MG, Lubitz SA, Magnani JW, McManus DD. A comparison of the CHARGE–AF and the CHA2DS2–VASc risk scores for prediction of atrial fibrillation in the Framingham Heart Study. *Am Heart J.* 2016; **178**: 45-54.
51. Pfister R, Brägelmann J, Michels G, Wareham NJ, Luben R, Khaw KT. Performance of the CHARGE-AF risk model for incident atrial fibrillation in the EPIC Norfolk cohort. *Eur J Prev Cardiol.* 2015; **22**:932-939.
52. Chamberlain AM, Agarwal SK, Folsom AR, Soliman EZ, Chambless LE, Crow R et al. A Clinical Risk Score for Atrial Fibrillation in a Biracial Prospective Cohort (From the Atherosclerosis Risk in Communities (ARIC) Study). *Am J Cardiol.* 2011; **107**: 85–91.
53. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB et al. Simple Risk Model Predicts Incidence of Atrial Fibrillation in a Racially and Geographically Diverse Population: the CHARGE-AF Consortium. *J Am Heart Assoc.* 2013; **2**: e000102.
54. Schnabel RB, Sullivan LM, Levy D, Pencina MJ, Massaro JM, D'Agostino RB et al. Development of a Risk Score for Atrial Fibrillation in the Community; The Framingham Heart Study. *Lancet.* 2009; **373**: 739–745.
55. Everett BM, Cook NR, Conen D, Chasman DI, Ridker PM, Albert CM. Novel genetic markers improve measures of atrial fibrillation risk prediction. *Eur Heart J.* 2013; **34**: 2243-2251.
56. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR: A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* 1996; **49**: 1373-1379.
57. Feinstein AR. *Multivariable Analysis: An Introduction.* New Haven: Yale University Press; 1996.
58. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom. Med.* 2004; **66**: 411-421.

59. Altman DG. Prognostic models: a methodological framework and review of models for breast cancer. *Cancer Invest.* 2009; **27**: 235-243.
60. Mallett S, Royston P, Dutton S, Waters R, Altman DG. Reporting methods in studies developing prognostic models in cancer: a review. *BMC Med.* 2010; **8**: 20.
61. Concato J, Peduzzi P, Holford TR, Feinstein AR. Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. *J Clin Epidemiol.* 1995; **48**: 1495-1501.
62. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med.* 2006; **25**: 127-141.
63. Hukkelhoven CWPM, Rampen AJJ, Maas AIR, Farace E, Habbema JDF, Marmarou A et al. Some prognostic models for traumatic brain injury were not valid. *J Clin Epidemiol.* 2006; **59**: 132-143.
64. Leushuis E, van der Steeg JW, Steures P, Bossuyt PMM, Eijkemans MJC, van der Veen F et al. Prediction models in reproductive medicine. *Hum Reprod Update.* 2009; **15**: 537-552.
65. Lagakos SW. Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable. *Stat Med.* 1988; **7**: 257-274.
66. Royston P, Sauerbrei W. *Multivariable Model-Building: A Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Modelling Continuous Variables.* Chichester. John Wiley & Sons; 2008.
67. Little RA. Regression with missing X's: a review. *J Am Stat Assoc.* 1992; **87**: 1227-1237.
68. Marshall A, Altman DG, Royston P, Holder RL. Comparison of techniques for handling missing covariate data within prognostic modeling studies: a simulation study. *BMC Med Res Meth.* 2010; **10**: 7.
69. Vergouwe Y, Royston P, Moons KGM, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. *J Clin Epidemiol.* 2010; **63**: 205-214.

- 70.**Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996; **15**: 361-387.
- 71.**Sun GW, Shook TL, Kay GL: Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol.* 1996; **49**: 907-916.
- 72.**Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med.* 2000; **19**: 453-473.
- 73.**Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. *BMJ.* 2009; **338**: b605.
- 74.**Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. *J Clin Epidemiol.* 2016; **74**: 167-176.
- 75.**Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Epidemiología clínica. Ciencia básica para la medicina clínica.* 2ª ed. Buenos Aires: Panamericana; 1994.
- 76.**Argimon Pallas JM, Jiménez Villa J. *Métodos de investigación clínica y epidemiología.* 4ª ed. Barcelona: Elsevier; 2013.

