



**Escuela de
Ingeniería y Arquitectura
Universidad Zaragoza**

TRABAJO FIN DE MASTER

Un Sistema de Vigilancia del estado de la práctica de actividad física y los hábitos saludables en la mujer y su incidencia en el riesgo de cáncer de mama

Autor: Laura Muñoz Garatachea

Directoras: Natalia Ayuso Escuer y Raquel Trillo Lado

MÁSTER EN INGENIERÍA INFORMÁTICA
UNIVERSIDAD DE ZARAGOZA

Junio 2024

*A mis directoras Natalia y Raquel
por ser artífices de un proyecto tan ilusionante
y por ser colaboradoras necesarias para alcanzar este reto.*

*A mi hija Bárbara
mi motor y mi apoyo incondicional
Un detallazo su colaboración difundiendo por Whatsapp
A todos los que de un modo u otro me han apoyado.*

➡ Reenviado

Hola a todos! 🌞 os robo sólo un segundo...🕒 Dentro de su proyecto de TFM mi madre ha diseñado y desarrollado un cuestionario sobre actividad física🏃, y salud en la mujer👩🏫 con perspectiva de género👤 aparte de estudiar los hábitos deportivos, el proyecto se va a centrar en como estos afectan a las mujeres diagnosticadas con cáncer de mama👩🏫 El cuestionario lo podéis rellenar todas! 👩🏫! Os lo dejo por aquí para que le deis mucho amor❤️ y se agradece también difusión entre las mujeres de vuestro entorno! Muchas gracias💕💕😊 <https://cuantos01.unizar.es/tfm/cuestionario.php>

Resumen/Abstract

En este proyecto se ha implementado una plataforma integrada para la recopilación, el almacenamiento y la visualización y analítica de datos, aplicada al estudio del estado de la práctica de Actividad Física (AF) en la mujer, conformando un Sistema de Vigilancia para ésta. La plataforma desarrollada facilita alcanzar los objetivos [Objetivos de Desarrollo Sostenible \(ODS\) 3: Salud y bienestar](#) y [5: Igualdad de género](#) y se alinea con las recomendaciones de la [Organización Mundial de Salud \(OMS\)](#) a los estados y sistemas de salud de establecer sistemas de vigilancia que permitan evaluar la eficacia de las medidas de promoción de la [Actividad Física \(AF\)](#).

Además, se ha llevado a cabo el tratamiento de datos sobre hábitos saludables y variables físico-ginecológicas en la mujer que se habían evidenciado relevantes en el riesgo de desarrollar un cáncer de mama, permitiendo a nuestro sistema mejorar la precisión en la estimación de la magnitud del efecto de un estilo de vida saludable, que incluye la práctica regular de [AF](#). El Sistema de Información desarrollado que integra al Sistema de Vigilancia, facilita la generación de informes automatizados y conformes a diversas características socio-demográficas y situación personal de la mujer, proporcionando un enfoque de perspectiva de género. A diferencia del estado del arte, en el proyecto se incide en un enfoque inferencial del análisis estadístico que contribuye a superar las limitaciones de alcance de otras operaciones estadísticas sobre la práctica de [AF](#) en España; vinculando los resultados con las recomendaciones de la [OMS](#) y las pautas sobre [AF](#) requeridas para lograr un efecto protector en el riesgo de cáncer de mama (BC, por sus siglas en inglés -*Breast Cancer*-).

This project has implemented an integrated platform for data collection, storage, visualisation and analytics, applied to the study of the state of and the visualisation and analysis of data, applied to the study of the state of the practice of Physical Activity (PA) in women, forming a Surveillance System for women. The platform developed facilitates the achievement of Sustainable Development Goals (SDG) *3: Health and Well-being* and *5: Gender Equality* and is aligned with the recommendations of the World Health Organisation (WHO) to states and systems to establish monitoring systems to assess the effectiveness of the measures adopted in the promotion of [Physical Activity \(PA\)](#).

In addition, data has been processed on healthy habits and physical-gynecological variables in women that had been shown to be relevant to the risk of developing [Breast Cancer \(BC\)](#), allowing our system to improve the accuracy of the estimation of the magnitude of the effect of a healthy lifestyle, which includes the regular [PA](#). The Information System developed, which integrates the Surveillance System, facilitates the generation of automated reports tailored to different socio-demographic and personal situation providing a gender-sensitive approach. In contrast to the state of the art, the project focuses on an inferential approach to statistical analysis that contributes to overcoming the limitations of the scope of other statistical operations on the practice of [PA](#) in Spain; linking the results with the [World Health Organization \(WHO\)](#) recommendations and the [PA](#) guidelines required to achieve a protective effect on the risk of [BC](#).

Índice

	Pág.
1 Conociendo el dominio	1
1.1 La AF orientada a la salud	1
1.1.1 Tipos y medición de la AF	3
1.1.2 Dosis de AF recomendadas	5
1.1.3 Beneficios de la dosis de AF recomendada en la salud de la mujer	7
1.2 La AF y el cáncer de mama	10
1.2.1 Dosis y beneficios en la prevención	11
1.2.2 Dosis y beneficios durante el tratamiento	11
1.3 Estado de la AF en la mujer española	13
1.3.1 Medición de la AF en la población	13
1.3.2 Resultados arrojados por la Encuesta de Hábitos Deportivos (EHD)-2022	15
1.4 Conclusiones y formulación de objetivos	19
2 Adquiriendo datos relevantes	20
2.1 Recolección de datos: El Cuestionario	21
2.1.1 Definición del contenido	21
2.1.2 Diseño técnico	22
2.2 Almacenamiento	23
2.3 Tecnologías utilizadas y alojamiento Web	25
3 Desarrollando una aplicación Web dinámica e interactiva para la Visualización y Analítica	29
3.1 Contexto, motivación y objetivos	29
3.1.1 Elección de tecnologías	29
3.2 Descripción de tecnologías	31
3.2.1 Shiny App	31
3.2.2 Flexdashboard con R-Markdown	33
3.2.3 Otros paquetes y tecnologías complementarias	34
3.3 Estructura de la Aplicación	35
3.3.1 Componentes del cuadro de mandos	35
3.3.2 Secciones y elementos interactivos	35
3.3.3 Organización del código	36
3.4 Desarrollo de la aplicación	38
3.4.1 Preparación del entorno de trabajo	38
3.4.2 Preprocesamiento, acceso y carga de datos	39
3.5 Diseño de la Interfaz de Usuario (IU)	41
3.5.1 Lienzo del cuadro de mandos	41
3.5.2 Diseño de elementos interactivos	43

3.6	Implementación de la lógica de la aplicación	44
3.6.1	Manejo de entradas, salidas y funciones reactivas	44
3.6.2	Generación y renderizado de elementos visuales	45
3.6.3	Robustez de los elementos visuales	46
3.7	Despliegue y ejemplos de implementación	46
4	Conclusiones y futuras líneas de trabajo	51
	<i>Bibliografía</i>	54
	<i>Acrónimos</i>	55
	<i>Glosario</i>	57
	<i>Lista de Figuras</i>	58
	<i>Lista de Tablas</i>	59
	<i>ANEXO A: Cronología de actuaciones de los Physical Activiyy Guideliness Advisor Comittee (PAGAC) y los comités de expertos de la OMS</i>	60
	<i>ANEXO B: Metodología de los PAGACs para la extracción de evidencias científicas</i>	63
	<i>ANEXO C: Perspectiva de género: Análisis de resultados de la EHD en la población masculina</i>	66
	<i>ANEXO D: Solicitudes de Conformidad de Protección de Datos en Universidad de Zaragoza y de Evaluación del proyecto por el Comité de Ética de la Investigación en la Comunidad autónoma de Aragón (CEICA)</i>	70
	<i>ANEXO E: El Cuestionario</i>	72
	<i>ANEXO F: Scripts Python</i>	75
	<i>ANEXO G: Cronograma del proyecto</i>	78

Capítulo 1

Conociendo el dominio

1.1. La AF orientada a la salud

En esta sección se realiza el estudio sobre el estado del arte de la práctica de AF y la salud enfatizando en la evaluación del efecto y la contribución a la mejora de la salud.

La evidencia científica de los efectos beneficiosos de la práctica de AF en la salud se ha ido acumulando a lo largo de varias décadas. Los estudios que investigan la relación entre la AF y la salud datan de mediados del siglo XX y han continuado hasta el presente fortaleciendo la comprensión de los beneficios de la AF, a través de las conclusiones de numerosos trabajos científicos en diferentes ámbitos de la salud.

No obstante, el enfoque de la AF orientada a la salud comenzó a consolidarse en el siglo XXI por una creciente comprensión de la evidencia científica por parte de la comunidad médica, a la que contribuyeron los avances en la ciencia del deporte y de la medicina deportiva. Estos factores, unidos a la creciente preocupación de los sistemas sanitarios por los efectos negativos en la salud de un estilo de vida sedentario, favorecieron el surgimiento de iniciativas por parte de los departamentos de salud gubernamentales (fundamentalmente en Estados Unidos y, posteriormente desde la OMS) dirigidas a fomentar la práctica de AF a través de ambiciosos programas de promoción y la elaboración de recomendaciones para las personas sobre la práctica de AF orientada a la salud.

Los programas de promoción de AF, cuyo objetivo es el de dar a conocer y concienciar sobre los beneficios de la práctica regular de AF, brindando orientación y recursos para incorporarla de manera segura y efectiva en las rutinas diarias, se han convertido en una herramienta esencial para fomentar estilos de vida activos en la sociedad. Estos programas suman los esfuerzos de los departamentos e instituciones de salud, la comunidad médica y científica persiguiendo, además de la mejora de la condición física de las personas, una mejora de la salud mental, la calidad de vida y la reducción en el riesgo de padecer enfermedades crónicas, haciendo hincapié en el papel de la AF como pilar central de la salud y el bienestar.

Los programas y las pautas sobre AF que, en lo que va de siglo, se han acometido fundamentalmente desde el *Health and Human Services Department (HHS)* estadounidense a nivel nacional, y la OMS a nivel mundial, han estado avalados por los informes científicos elaborados por los comités asesores de expertos en los ámbitos propios de la salud y la AF. Los miembros de los PAGACs estadounidenses son los expertos en el campo de la AF y la salud que conforman el nexo de transferencia de los resultados de la investigación científica a la sociedad, en tanto que son los encargados de revisar y evaluar la evidencia científica disponible acerca del impacto de la AF en la salud. Se trata de expertos de reconocido prestigio, de carácter multidisciplinar en los campos de la medicina, la fisiología del ejercicio, la salud pública o la epidemiología, cuyo nombramiento utiliza rigurosos criterios de conocimiento de la investigación científica en AF, experiencia en el propósito de la investigación y la comunicación científica, y un interés manifiesto en la salud y el bienestar. Sus áreas de especialización incluyen la promoción de la salud y prevención de enfermedades crónicas, salud y rendimiento óseo, articular y muscular, obesidad y control del peso, AF y riesgo de lesiones musculoesqueléticas, revisiones sistemáticas y especialización en poblaciones específicas.

Los PAGACs colaboran estrechamente con los comités de expertos de la OMS cuyos informes sustentan tanto las pautas sobre AF recomendadas por la organización, como la puesta en marcha

de los programas mundiales de promoción de AF. A modo de aproximación en el conocimiento de las actuaciones más relevantes acometidas, iniciadas en 2008, en la **Figura 1.1** se sintetiza su cronología en el transcurso de los últimos 15 años. Una descripción más detallada de los contextos, los sucesivos programas de promoción de la AF de la OMS y del departamento de salud estadounidense, así como las actualizaciones de evidencias científicas realizadas periódicamente por los PAGAC y comités asesores de la OMS, y las respectivas pautas de AF puede consultarse el **Anexo A**.

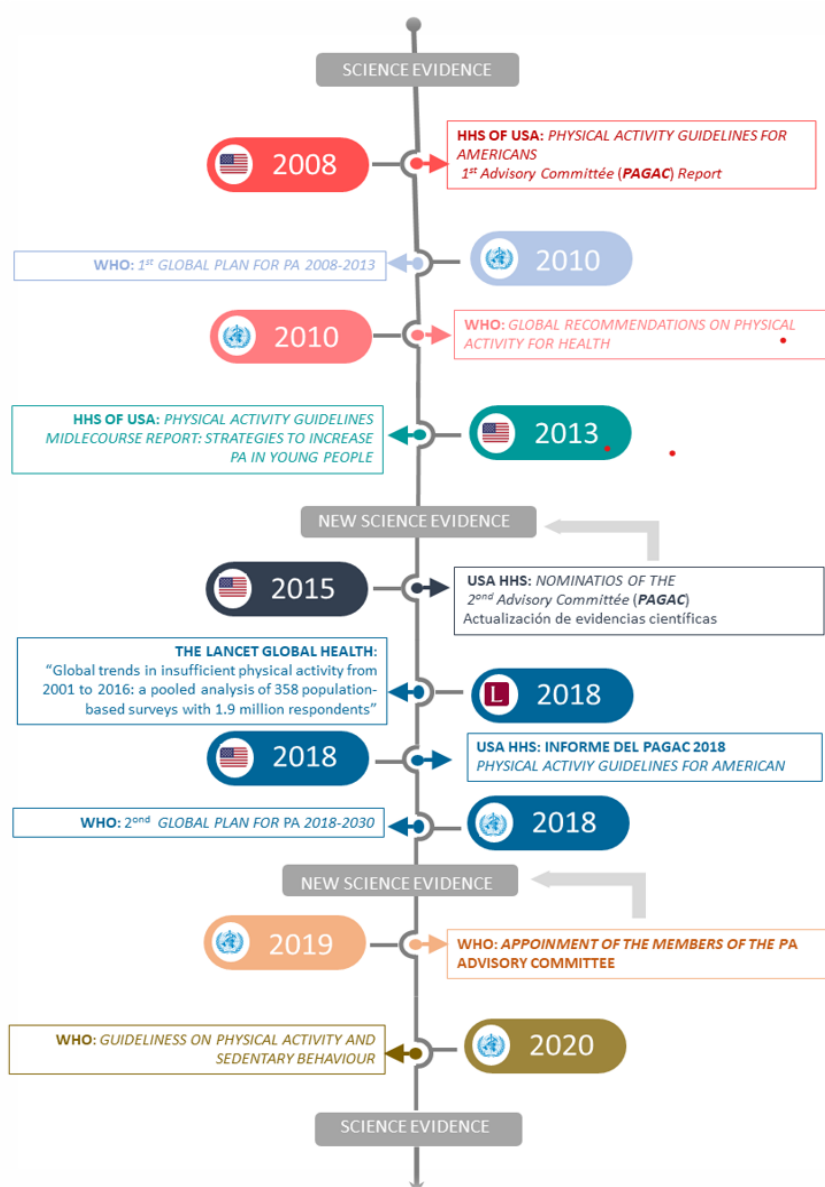


Figura 1.1: Cronograma de actuaciones y programas mundiales para la promoción de la práctica de AF. Fuente: elaboración propia

Los informes científicos elaborados por los comités asesores sintetizan el compendio exhaustivo de evidencia científica actualizada en las diferentes temáticas relativas al tipo y dosis de diferentes intervenciones de AF y sus efectos, proporcionando una base sólida y reputada que contribuya a una mayor adopción de la AF orientada a la salud por parte de la población. Basándonos en el informe científico elaborado por el PAGAC-2018, se ha profundizado en la comprensión de la rigurosa metodología de revisión sistemática, extracción y calificación de evidencias de los resultados de los estudios de los trabajos científicos. En el **Anexo B** se detallan las líneas maestras del proceso que

conlleva la búsqueda, el cribado y selección de trabajos, el resumen y el análisis de los estudios recogidos, la evaluación tanto de la calidad, como de los potenciales sesgos en éstos, y que culmina con la elaboración de conclusiones y la asignación del grado de evidencia de las respuestas de cada pregunta de investigación.

En la promoción de la AF en la población por parte de los organismos e instituciones a niveles mundial y gubernamental subyace además un motivo de índole económica que apunta a los severos costes económicos que la inactividad física genera a los sistemas sanitarios. La OMS estima que entre 2020 y 2030 se producirán cerca de 500 millones de nuevos casos de Enfermedades No Transmisibles (ENT) que supondrá unos costes de tratamiento de algo más de USD 300,000 millones anuales si no tiene lugar ningún cambio en la prevalencia actual del sedentarismo y la inactividad física. Casi la mitad de estos nuevos casos (47 %) serán consecuencia de la Hipertensión Arterial (HTA), y el 43 % de la depresión. El mayor costo económico se producirá en los países de ingresos altos, que representarán el 70 % del gasto en atención de salud para el tratamiento de las enfermedades derivadas de la inactividad física.

Entre las numerosas recomendaciones¹ dirigidas a los gobiernos de los países en *Global Action Plan for Physical Activity (GAPPA)* de la OMS 2018-2023: *More active people for a healthier world* para la adopción de medidas encaminadas a alcanzar el ODS 3 salud y bienestar, la OMS destaca la relevancia de la vigilancia nacional de AF en personas adultas; Nuestro trabajo se alinea con esta directriz en el contexto de la población femenina en España.

Una segunda medida destacada es la de adoptar un enfoque integral, consistente en fomentar la colaboración entre profesionales de la salud, sistemas de atención médica y comunidades para lograr un impacto positivo en la salud a largo plazo y reducir la carga global de enfermedades relacionadas con la inactividad física. El informe ejecutivo técnico de la OMS de 2022 relativo al grado de adopción de las medidas en los países señala que menos del 50 % ha acometido ese enfoque global adoptando, en particular, la controvertida y demandada decisión de incorporar la prescripción médica de ejercicio desde los sistemas de atención primaria para la prevención y control de ENT.

1.1.1. Tipos y medición de la AF

La medición de la AF y del sedentarismo supone un desafío debido a la diversa tipología de actividad de tipo físico que las personas realizan diariamente en diferentes ámbitos: laboral, doméstico, desplazamiento, o recreativa, a diferentes intensidades y duraciones. Esta diversa tipología de AF suele ser clasificada atendiendo al efecto fisiológico dominante, distinguiendo entre los tipos siguientes:

AERÓBICA: técnicamente, la AF aeróbica incluye cualquier actividad que pueda mantenerse utilizando solo vías de energía metabólica apoyadas por oxígeno y que pueda continuarse durante no más de unos pocos minutos. Tanto en el uso común como en el científico, la actividad aeróbica ha llegado a significar la AF que se esperaba que mantuviera, o mejorase la **Función, Aptitud o capacidad Cardiorespiratoria (CRF)**

ANAERÓBICA: la AF anaeróbica se refiere a la actividad de alta intensidad que excede la capacidad del sistema cardiovascular para proporcionar oxígeno a las células musculares para las vías metabólicas habituales que consumen O₂. La actividad anaeróbica se puede mantener solo durante unos 2 a 3 minutos. Las carreras de velocidad y el levantamiento de pesas son ejemplos de actividad física anaeróbica

FORTALECIMIENTO MUSCULAR O FUERZA: actividades destinadas a mantener o mejorar la fuerza muscular, la resistencia, o la potencia. Las actividades de fortalecimiento muscular pueden incluir

¹Las actuaciones a promover por los países se organizan en torno a cuatro bloques de intervenciones: 1^a Crear sociedades activas, 2^a Crear entornos activos, 3^a Fomentar poblaciones activas, 4^a Crear sistemas activos

comportamientos cotidianos repetitivos como cargar compras pesadas, levantar niños, o subir escaleras, así como el uso de equipos de ejercitación apropiados

FORTALECIMIENTO ÓSEO: actividades dirigidas a generar movimientos que crean fuerzas de impacto y carga muscular en el hueso. Estas fuerzas estresan el hueso, que se adapta modificando su estructura o masa, aumentando así su resistencia a la fractura

ENTRENAMIENTO DEL EQUILIBRIO: actividades dirigidas a mejorar el equilibrio en las que ponen en práctica diversos movimientos que desafían, de forma segura, el control postural. Si se practica regularmente, mejoran la capacidad de resistir las fuerzas intrínsecas o ambientales que causan caídas

ENTRENAMIENTO DE FLEXIBILIDAD O ESTIRAMIENTOS: actividades dirigidas a la mejora del rango y la facilidad de movimiento alrededor de una articulación. Los estiramientos dinámicos, como los movimientos de *tai chi*, *qigong* y yoga, y los estiramientos estáticos son ejemplos de entrenamiento de flexibilidad. Estas actividades suelen combinar el fortalecimiento muscular, el entrenamiento del equilibrio, la actividad aeróbica de intensidad ligera y la flexibilidad

HIGH INTENSITY INTERVAL TRAINING (HIIT) método de entrenamiento novedoso con intervalos de alta intensidad para activar el metabolismo y aumentar el consumo de oxígeno que alternan con un tiempo breve de recuperación

En términos dosis-tratamiento, la dosis de AF consta de 3 componentes: la intensidad, la duración y la frecuencia, con la intensidad como componente crítica a medir. La intensidad² es la tasa de gasto energético necesaria para realizar una actividad y como tal se puede medir en kilocalorías, julios, o consumo de oxígeno. Sin embargo, se ha desarrollado una unidad metabólica propia para medir el gasto energético durante la AF denominada *Metabolic Equivalent of Task (MET)* que refiere la intensidad de una actividad en relación al gasto energético basal. Así, un MET representa el consumo mínimo de oxígeno que el organismo necesita para mantener sus constantes vitales que se estima en, aproximadamente en 3.5 mililitros de oxígeno por cada kilogramo y minuto. La estandarización por consumo MET se establece en el *Compendium of Physical Activities (CPA)*³ un esquema de codificación que califica registros de AF por tasa de gasto energético, desarrollado para mejorar la comparabilidad de los resultados en los estudios basados en encuestas. En la tabla 1.1 pueden consultarse los METs de diferentes actividades extraídas del trabajo [1].

En términos de calorías, un MET equivale a 0.0175 kilocalorías / kilogramo x minuto o, de forma práctica:

$$1 \text{ MET} = 1 \text{ kilocaloría por cada kilogramo de peso, en cada hora de actividad}^4.$$

Utilizando como medida el MET, y atendiendo al consumo energético requerido, la intensidad de la AF usualmente se clasifica en niveles distinguiendo, esencialmente tres:

1. **AF Leve o de baja intensidad (AFL):** su realización requiere o equivale a entre 1.6 y 3.0 METs
2. **AF de intensidad Moderada (AFM):** su realización requiere entre 3.0 y 6.0 METs
3. **AF Vigorosa (AFV):** su realización requiere por encima de 6.0 METs

Estos tres niveles de intensidad son universales y son conforme a ellos la mayoría de estudios epidemiológicos de AF y salud; no obstante, se ha constatado que en diversos estudios en trabajos

²se denomina absoluta en relación a la percepción individual del esfuerzo que se asimila a una intensidad relativa

³<https://sites.google.com/site/compendiumofphysicalactivities/>

⁴Por ejemplo, para una persona de 70 kilogramos, realizar una actividad (como correr a 8km/h) de 9.0 METs supone un consumo de kilocalorías/minuto = 9.0 x 0.0175 x 70=11 kilocalorías/minuto. Si la realiza durante una hora consumirá un total de 660 kilocalorías

científicos, así como en las recomendaciones *Physical Activity Guidelines for Americans (PAGA)*-2018, se utiliza un nivel adicional situado entre los dos niveles superiores, referido como **AF de intensidad de Moderada a Vigorosa (AFMV)** que está bien caracterizado con un consumo energético de entre 4.5 y 6.0 METs.

En función de la dosis, el volumen acumulado durante un periodo de tiempo se puede obtener multiplicando la duración de cada episodio de **AF** por la frecuencia y los METs correspondientes al tipo de actividad realizada. Se expresa en METs-minutos, días o semanales y al agregarlo para el conjunto de actividades físico-deportivas realizadas durante determinado intervalo de tiempo, se obtiene el volumen de **AF** acumulado⁵

Actividades deportivas y AF recreativa		
Paseo en bicicleta (2,9 MET)	Baile ligero (3,7 MET), danza moderna (4,7 MET), baile de salón (rumba) (5,8 MET), Ping-pong, golf, gimnasia (4 MET)	Montar a caballo, ballet, tenis, ciclismo, natación estilo braza o espalda (6-6,9 MET), Baloncesto, fútbol, natación crol (7-8 MET)
	Paseo rápido o trotar (4,1 MET)	Balonmano, patinaje, remo, alpinismo, squash (8-9 MET)
Petanca (3 MET)	Excursión con mochila (5,6 MET)	Judo, karate (10,6 MET)
	Ir en bicicleta (14km/h) (5,7 MET)	Natación mariposa (11,4 MET)
	Gimnasia, voleibol, nadar suave (5,8 MET)	Jogging, esquí intenso, carrera (12-16 MET)
Ligera (AFL) Moderada (AFM) Vigorosa (AFV)		
INTENSIDAD		

Cuadro 1.1: Consumo metabólico equivalente (METs) en diferentes actividades. *Fuente: Elaboración propia*

1.1.2. Dosis de **AF** recomendadas

Se recogen en este apartado las recomendaciones sobre dosis de **AF** en cada etapa de la vida formuladas en la *PAGA*-2018 estadounidense y la *OMS-2020: Guidelines on Physical Activity and Sedentary Lifestyle* ambas con vigencia actual, que servirán de referencia en el dictamen de valoración del estado actual de la práctica de **AF** en la mujer en España que se realiza en la última sección de este capítulo.

El rango objetivo de **AF** recomendado a los adultos estadounidenses en la *PAGA*-2018 es de 500 a 1000 METs-minutos-semana de **AFV**, lo que equivale a de 150 a 300 minutos semanales de **AFM**, y a su vez a de 30 a 50 minutos diarios durante de 3 a 5 días a la semana de **AFM**. El comité de 2008 avanzó cómo las personas actualmente inactivas pueden conseguir mejoras sustanciales en su salud y prevención de riesgos, aumentando su nivel de actividad aunque no lleguen a alcanzar el rango objetivo. Las conclusiones del informe 2018 confirman y amplían los beneficios de diversas dosis de **AF**, aseverando:

- en mujeres que, o no realizan **AFMV**, o la realizan en menor medida de la recomendada, la sustitución del comportamiento sedentario por **AFL** contribuye a reducir el riesgo de mortalidad por cualquier causa, la incidencia y mortalidad por **Enfermedades Cardiovasculares (ECV)** y la incidencia de diabetes tipo 2

⁵Por ejemplo, correr a un ritmo equivalente a 8 METs durante 30 minutos al día y repetirlo 4 veces a la semana da lugar a un total de **AF** de: 8 METs x 30 mins x 4 veces = 960 METs-semanales

- las mujeres que, o no realizan [AFV](#), o la realizan en menor medida, pueden reducir sus riesgos añadiendo dosis de [AFM](#).
- para las personas cuya [AFV](#) está por debajo del objetivo, incluso pequeños incrementos de la [AFM](#) aportan beneficios para la salud sin que exista un umbral para que empiecen a producirse
- para cualquier incremento en la [AFM](#), la ganancia relativa en beneficios es mayor para las personas que están por debajo del rango objetivo
- las personas que ya se encuentran dentro del intervalo prescrito de [AF](#) pueden aumentar sus beneficios realizando más [AFV](#)
- los episodios de [AFM](#) de cualquier duración pueden incluirse en el volumen acumulado de [AF](#)

En cuanto a las dosis recomendadas por la [OMS](#), ésta divide a la población en tres grupos de edad: niños y adolescentes (de 6 a 17 años), adultos (de 18 a 64 años) y personas mayores (de más de 65 años) y en la formulación de las pautas de [AF](#), diferencia tres grupos específicos: mujeres embarazadas y en puerperio, adultos con enfermedades crónicas y personas con algún grado de discapacidad. Las recomendaciones sobre dosis y tipo de [AF](#) para cada segmento de la población, junto con los beneficios evidenciados para la salud que se detallan en esta sección, pueden verse resumidos en la tabla 1.2

Las recientes directrices de la [OMS](#) recomiendan al menos 150 minutos de [AFMV](#) aeróbica, o alternativamente más de 75 minutos de [AFV](#) aeróbica a la semana para todas las mujeres adultas, incluidas las mujeres embarazadas y en puerperio y las mujeres con discapacidad. Para las mujeres adultas mayores recomienda igualmente al menos 150 minutos de [AFV](#) aeróbica, o bien más de 75 minutos de [AFV](#) semanalmente, mientras que se recomienda un promedio de 60 minutos al día, al menos 3 veces a la semana de [AFMV](#) aeróbica para niñas y adolescentes.

Las niñas y adolescentes deben realizar al menos una media de 60 minutos de [AFMV](#) diaria principalmente aeróbica a lo largo de la semana, al menos tres días por semana [*Recomendación fuerte, evidencia de certeza moderada*]. Las niñas o adolescentes que no practican [AF](#) deben comenzar con pequeñas dosis, para ir aumentando gradualmente la duración, la frecuencia e intensidad. La [OMS](#) remarca la importancia de ofrecer a todas las niñas y adolescentes oportunidades seguras y equitativas para participar en [AF](#) que les resulten placenteras, variadas y aptas para su edad y capacidad, a fin de alentarlas en la adquisición y consolidación el hábito en la etapa adulta. Las niñas y adolescentes deben limitar el tiempo que dedican a actividades sedentarias, especialmente el tiempo de ocio que pasan delante de una pantalla.

Todas las mujeres mayores, incluidas las que padecen afecciones crónicas, deben realizar [AF](#) con regularidad debiendo acumular a lo largo de la semana un mínimo de 150 minutos de [AFM](#) aeróbica o, alternativamente, entre 75 y 150 minutos de [AFV](#) aeróbica, o bien una combinación equivalente de [AFMV](#), con el fin de obtener beneficios notables para la salud [*Recomendación fuerte, evidencia de certeza moderada*]. Las mujeres mayores también deben realizar actividades de fortalecimiento muscular de intensidad moderada o elevada, para trabajar todos los grandes grupos musculares al menos dos días a la semana. Se les recomienda dentro de su [AF](#) semanal realizar actividades variadas (de tipo multicomponente) que den cabida al equilibrio funcional a fin de mejorar su capacidad funcional y evitar las caídas. Las mujeres mayores deben limitar el tiempo que dedican a actividades sedentarias, sustituyendo el tiempo de TV por una [AF](#) de cualquier intensidad (incluso [AFL](#)).

En las directrices se alienta a las mujeres a mantener una [AF](#) regular durante el embarazo y después del parto y se recomienda a mujeres embarazadas y en puerperio sin contraindicaciones realizar dosis de [AF](#) idénticas a la población. Para las mujeres que practicasen [AFV](#) aeróbica de forma habitual, o fueran físicamente activas, deben continuar dichas actividades durante el embarazo y el puerperio realizando como mínimo 150 minutos de [AFM](#) aeróbica a la semana e incorporar actividades de fortalecimiento muscular, pudiendo resultar beneficioso añadir estiramientos moderados.

Población	Dosis de AF	Tipo de AF	Beneficios
<i>Niñas y adolescentes (de 5 a 17)</i> <i>Incluye con discapacidad</i>	>60 mins. diarios - 3 días a la semana	AFMV aeróbica+ actividades de fortalecimiento muscular	ADQUISICIÓN Y DESARROLLO
	De 300 a 500 METs semanales	AFMV aeróbica	✓ Forma física, ✓ función muscular ✓ salud cardio-metabólica (tensión, dilipidemia, glucosa y resistencia a la insulina) ✓ salud osea, ✓ CRF, ✓ salud mental (menor presencia de síntomas de depresión), ✓ resultados cognitivos (función ejecutiva y resultados académicos), ✓ habilidades psico-sociales, ✓ postura y equilibrio, ✓ crecimiento saludable, ✓ menor adiposidad
<i>Adultas (de 18 a 64 años)</i> <i>Incluye con discapacidad, enfermedad crónica o embarazada</i>	>150 mins. semana ó >75 mins. semana	AFMV aeróbica AFV aeróbica + actividades de fortalecimiento muscular	CONSOLIDACIÓN
	De 500 a 1000 METs semanales	AFV aeróbica	Idem + reducción del riesgo de depresión y enfermedades crónicas
<i>Mayores (más de 65 años)</i> <i>Incluye con discapacidad o enfermedad crónica</i>	>150 mins. semana ó >75 mins. semana	AFMV aeróbica AFV aeróbica + actividades equilibrio funcional	MANTENIMIENTO
	De 500 a 1000 METs semanales	AFV aeróbica	Idem + reducción del deterioro cognitivo

Cuadro 1.2: Pautas sobre AF de la OMS-2020 y de la PAGA-2018. Fuente: Elaboración propia a partir de las respectivas guías

1.1.3. Beneficios de la dosis de AF recomendada en la salud de la mujer

Hace ya más de quince años que las conclusiones del primer informe PAGAC-2008 permitieron afirmar rotundamente que los individuos físicamente activos duermen mejor, se sienten mejor y presentan un mejor funcionamiento vital. El informe científico del PAGAC-2018 corroboró y añadió además otros beneficios evidenciados en la literatura científica relativos a la reducción del riesgo de padecer determinadas enfermedades.

BENEFICIOS EN EL BIENESTAR FÍSICO, LA SALUD MENTAL Y LA CALIDAD DE VIDA

- pruebas sólidas demuestran que la AF de cualquier intensidad mejora la calidad del sueño. Lo hace reduciendo el tiempo que se tarda en conciliar el sueño y también puede aumentar el tiempo de sueño profundo, reduciendo la somnolencia diurna,
- los episodios de AFL promueven mejoras sustanciales en la función cognitiva-ejecutiva durante un periodo de tiempo. Se ha demostrado que la AF, aún en episodios aislados y de baja intensidad, contribuye a mejorar otros componentes de la cognición como la memoria, la velocidad de procesamiento, la atención y el rendimiento profesional,

- la AF regular de cualquier intensidad, no sólo reduce el riesgo de depresión clínica, sino que también reduce los síntomas depresivos entre las personas con y sin depresión clínica,
- la AF regular de cualquier intensidad reduce los síntomas de ansiedad, incluidos tanto los niveles crónicos de ansiedad, como los sentimientos agudos de ansiedad ocasionales,
- también existen pruebas sólidas que demuestran que la AF regular mejora la calidad de vida percibida

RELEVANCIA DE LA AF EN EL DESARROLLO FÍSICO-EMOCIONAL DE NIÑAS Y ADOLESCENTES

El PAGAC-2018 informa de la existencia de pruebas sólidas que demuestran que una mayor cantidad de AF se asocia a un menor riesgo de un aumento excesivo del peso corporal y la adiposidad en niñas y adolescentes hasta 17 años, auspiciando además innumerables beneficios de salud en su desarrollo físico-emocional. En la misma línea, las conclusiones del informe de expertos de la OMS afirman que se ha demostrado que, en las niñas y adolescentes, la AF es beneficiosa por cuanto respecta a los siguientes resultados de salud: mejora de la forma física (la CRF y muscular), reducción de la adiposidad, mejora de la salud cardiometabólica (tensión, dislipidemia, glucosa y resistencia a la insulina) y la salud ósea, junto con una mejora en los resultados cognitivos (desempeño académico y función ejecutiva) y la salud mental (menor presencia de síntomas de depresión).

Junto con la importancia de la adquisición de unos hábitos de vida saludables desde las primeras etapas de la vida, numerosos expertos coinciden en resaltar los beneficios psicosociales y personales de la AF en la adolescencia por cuanto ha demostrado un impacto positivo en la adquisición de habilidades sociales, el fortalecimiento de la autoestima y el desarrollo de una sólida fortaleza mental.

LA AF COMO PROTECTOR DE SALUD Y BIENESTAR PSICOLÓGICO EN MUJERES ADULTAS

Las conclusiones del informe del PAGAC-2008 afirmaban con rotundidad que la AF reduce el riesgo de un gran número de enfermedades y en los diez años posteriores se ampliaron considerablemente tanto la evidencia alcanzada sobre éstas, como el conjunto de enfermedades y afecciones cuyo riesgo podría reducirse considerablemente con la adopción de dosis adecuadas de AF. Los principales resultados reportados en el informe del PAGAC-2018, aplicables en particular a mujeres adultas, fueron los siguientes:

- pruebas sólidas demuestran que un mayor volumen de AFM está asociado con la prevención del aumento excesivo de peso en mujeres adultas, el mantenimiento de la condición física y la CRF en un rango saludable, así como la prevención de los problemas de salud derivados de la obesidad,
- pruebas sólidas demuestran que un mayor volumen de AF reduce el riesgo de demencia, mejorando aspectos clave de la función cognitiva,
- el comité de 2018 corroborará con evidencia fuerte la incidencia de dosis de AF en la reducción en el riesgo de cáncer, ampliando la lista de cánceres localizados. Se evidenció que los niveles recomendados de AFV se asocian con un menor riesgo de cáncer de mama, vejiga, endometrio, esófago, riñón, pulmón y estómago.

En línea con estas conclusiones, el informe de la OMS-2020 afirmó que existe evidencia de que la AF en mujeres adultas es beneficiosa en relación, al menos a los siguientes resultados de salud: reducción en el riesgo de mortalidad por todas las causas (más acentuada en el caso de las ECV), menor incidencia de problemas derivados de HTA, reducción en la incidencia de cánceres localizados y de diabetes de tipo 2. Así mismo, quedaron demostrados sus beneficios para la salud mental de la mujer adulta con menor presencia de síntomas de ansiedad y depresión, así como para el mantenimiento de la salud cognitiva y una mejora notable en la calidad del sueño.

El informe asevera que las dosis recomendadas y los beneficios de la AF para la salud en mujeres adultas resultan aplicables a mujeres adultas con diversidad funcional o discapacidad, para las que

se suman nuevos beneficios. Como beneficios adicionales, la OMS señala los siguientes: en adultas con esclerosis múltiple una mejora de la función física y las dimensiones física, mental y social en su calidad de vida; en mujeres con una lesión medular, mejoras de la función motora de las piernas, la fuerza muscular y la función de las extremidades superiores, unidas a una mayor calidad de vida; en mujeres con trastornos que afectan a la función cognitiva, una mejora de las funciones física y cognitiva; en mujeres con enfermedad de Parkinson, beneficios cognitivos; posible mejora de la calidad de vida en adultas con esquizofrenia; una mejora de la función física en mujeres adultas con discapacidad intelectual, y una mejora de la calidad de vida para mujeres adultas con depresión mayor.

BENEFICIOS DE LA AF EN MUJERES EMBARAZADAS Y POST-PARTO

Se ha demostrado que las mujeres físicamente más activas tienen menos probabilidades de aumentar excesivamente de peso durante el embarazo. En las directrices de la OMS para las mujeres embarazadas y en puerperio, la AF durante el embarazo y el postparto aporta los siguientes beneficios para la salud materna y fetal: menor riesgo de preeclampsia⁶, HTA gestacional, diabetes gestacional, complicaciones en el parto y depresión postparto, con menos complicaciones neonatales, ausencia de efectos nocivos en el peso al nacer y se confirma la ausencia de un incremento del riesgo de muerte fetal.

CONTRIBUCIÓN AL BIENESTAR FÍSICO EN MUJERES MAYORES O CON PATOLOGÍAS CRÓNICAS

Una parte de las mujeres adultas ya padece una enfermedad o afección crónica. El comité PAGAC-2018 llegó a la conclusión de que, para estas personas, la AF regular puede reducir el riesgo de desarrollar una nueva afección crónica, reducir el riesgo de progresión de la afección que ya padecen, y contribuir a mejorar diversos parámetros de su calidad de vida y su función física. Las afecciones examinadas por el comité incluyeron algunas de las más prevalentes en mujeres mayores o con patologías previas, como la osteoartritis, la HTA y la diabetes tipo 2. En adultas mayores, existen pruebas sólidas que demuestran una reducción en el riesgo de caídas y lesiones relacionadas con éstas, contribuyendo además a mantener una capacidad cognitiva adecuada para una vida independiente. Concretando, la OMS informa que la AF es beneficiosa para la salud de mujeres mayores con o sin afecciones crónicas: se ha reportado que en *Women Survivors of Breast Cancer* (WSBC), la práctica de AF mejora la mortalidad por todas las causas, la mortalidad por cáncer y el riesgo de recidiva o segundos tumores primarios; en mujeres con HTA, la AFM mejora la mortalidad por ECV, la progresión de la enfermedad, la función física y la Calidad de Vida Relativa a la Salud (CVRS); en las personas con diabetes de tipo 2, se ha evidenciado que la AF reduce las tasas de mortalidad por ECV, mejorando los indicadores de progresión de la enfermedad, y en las personas que viven con el *Human Immunodeficiency virus* (VIH), la AF puede mejorar la forma física y la salud mental (menor presencia de síntomas de ansiedad y depresión), sin incidir negativamente en la progresión de la enfermedad, ni en la composición corporal.

RIESGOS PARA LA SALUD DE LA MUJER DE UN ESTILO DE VIDA SEDENTARIO

Técnicamente el comportamiento sedentario alude a cualquier comportamiento de vigilia caracterizado por un gasto energético de 1.5 METs o menos, mientras se está sentado, reclinado o tumbado. La mayoría de las investigaciones sobre AF en salud se venían centrado en la asociación entre la AF de mayor intensidad, AFV y los resultados de salud, sin embargo, el comportamiento sedentario ha venido recibiendo cada vez más atención en la investigación como parte de un problema de salud pública.

El Subcomité PAGAC de comportamiento sedentario (designado a tal efecto) examinó la relación entre el comportamiento sedentario y las principales causas de mortalidad, y también se evaluó su

⁶Complicación originada por una presión arterial alta en la madre que ocasiona una reducción en el suministro de sangre al feto

relación con el estado global de salud del individuo, además del exceso de incidencia de enfermedades crónicas comunes. Las conclusiones del informe científico 2018 que alcanzaron calificación de evidencia fuerte y moderada (se han descartado las de evidencia débil ó no asignable) fueron:

- existen pruebas sólidas que demuestran una asociación positiva entre el mayor tiempo dedicado a conductas sedentarias y las tasas de mortalidad por todas las causas. *Grado de evidencia PAGAC: Fuerte*,
- la inactividad física y una baja CRF son factores de riesgo muy incidentes en el desarrollo de ECV. *Grado de evidencia PAGAC: Fuerte*,
- diversos estudios constatan una asociación entre el mayor tiempo dedicado a conductas sedentarias y un riesgo más elevado de cáncer de mama. *Grado de evidencia PAGAC: Moderado*.

Por su parte, el informe de la OMS-2020 apunta a la inactividad física como el cuarto factor de riesgo, por detrás de la hipertensión (HTA), el tabaco y la hipoglucemia y alerta que, cerca del 47.5 % de los adultos y del 81 % de los adolescentes no cumplen con las dosis de AF recomendadas. Asimismo desvela la existencia de desigualdades significativas entre grupos o colectivos en la población: los datos evidencian cómo las niñas y las mujeres continúan siendo menos activas; se detecta además una diferencia amplia en los niveles medios de AF entre colectivos con mayor y menor nivel económico.

1.2. La AF y el cáncer de mama

La mayoría de los cánceres surgen de una etiología compleja que involucra factores genéticos, ambientales y de estilo de vida, y sus interacciones. La investigación sobre el origen del BC es un campo en evolución continua y aunque no se ha llegado a la comprensión de las causas que originan su desarrollo, los estudios continúan investigando los factores de riesgo y la interacción compleja entre éstos, junto con la medida de sus efectos, con el objetivo de mejorar la prevención.

En base a estudios amplios y correctamente diseñados (cohortes aleatorizadas y con grupo de control) en el transcurso de los años, se ha alcanzado un consenso sobre los factores que inciden en la causalidad de la enfermedad, algunos de los cuales no son modificables, mientras que otros si lo son.

FACTORES DE RIESGO MODIFICABLES:

mantenerse físicamente inactiva: se ha evidenciado que entre las mujeres que se mantienen físicamente activas se da una menor prevalencia de BC respecto de las que no realizan AF, siendo significativamente menor entre las deportistas de élite,

tener sobrepeso u obesidad después de la menopausia: la obesidad está asociada a procesos inflamatorios y a anomalías metabólicas y endocrinas, que favorecen el crecimiento celular y ejercen efectos anti-apoptóticos, es decir, que las células cancerosas no se auto-destruyan,

tomar hormonas: algunas formas de terapia de remplazo hormonal (que incluyen tanto estrógeno, como progesterona) que se prescriben durante la menopausia y que pueden incrementar el riesgo de BC si se toman por más de cinco años. Ciertos anticonceptivos orales también aumentan el riesgo de BC,

consumo de alcohol y tabaco: Diversos estudios concluyen que el riesgo está directa y altamente correlado con la cantidad de alcohol y/o tabaco acumuladas.

La tesis de que la AF contribuye a reducir el riesgo de BC tendría, junto con un soporte de evidencia empírica (moderada), cuenta con una base explicativa desde el punto de vista biológico, en tanto que se ha reportado que la AF contribuye a modular los niveles de las hormonas reproductivas, el estado menstrual, ayuda a mantener un peso corporal bajo, pudiendo incluso aumentar las defensas del sistema inmune frente al crecimiento tumoral.

1.2.1. Dosis y beneficios en la prevención

Un estilo de vida físicamente activo es un protector contra la aparición de algunos cánceres comunes como el BC con una evidencia constatada a partir del informe del PAGAC-2018. En la revisión sistemática se investigó la forma de las dosis-respuesta, planteando si la asociación variaba en función de variables como la edad, el nivel socio-económico, o si la asociación entre AF y riesgo de BC también se da en mujeres con alto riesgo, como aquellas con predisposición familiar al cáncer.

El comité PAGAC evaluó 45 revisiones sistemáticas y meta-análisis y análisis que comprenden cientos de estudios epidemiológicos con varios millones de participantes, determinando que, al comparar la incidencia entre individuos en la categoría más alta de AF con individuos en la más baja, se alcanza una evidencia sólida de reducción de riesgo en diversos cánceres, incluyendo el BC. Se concluye también que, independientemente de la edad, el nivel socio-económico, e incluso en mujeres con antecedentes familiares, se logran reducciones significativas de riesgo relativo que oscilan entre aproximadamente el 10 % y el 20 %.

La revisión sistemática actualizada en 2019 [6] ratifica cómo las dosis de AF recomendadas se asocian a una reducción en el riesgo de BC evidenciando que la AFMV genera reducciones del riesgo del 12 % (Riesgo Relativo (RR) 0.88; Intervalo de Confianza (IC) 95 % = [0.85, 0.92] con datos similares en mujeres pre-menopausia y post-menopausia. Un hallazgo relevante del trabajo es que evidencia que las dosis de AFV, en concreto la participación en actividades de más de 10 METs, se asociaron con una reducción del riesgo del 25 % en la mortalidad específica por BC (RR 0.75; IC 95 % = [0.65, 0.85]).

Se ha tenido acceso a un trabajo reciente de revisión sistemática (en pre-publicación) [9] destinado a actualizar el tamaño de los efectos en el riesgo de BC de los factores del estilo de vida en el que se concluye que la reducción del riesgo promovida por la AF fluctúa entre el 12 % y el 32 % dependiendo de los factores cardio-metabólicos de las participantes y el subtipo de cáncer.

No obstante, conviene precisar que las dosis de AF recomendadas por la OMS representan un mínimo para ayudar a mantener la salud cardio-metabólica y disminuir el riesgo de enfermedades; sin embargo, en la prevención del BC, para alcanzar un impacto significativo de los beneficios, se requieren dosis de AF más elevadas, tanto en intensidad, como en frecuencia y éstas deben ser mantenidas en el tiempo. Así, desde la *World Cancer Research Fund* y el *American Institute for Cancer Research* se enfatiza en cómo en la prevención del BC, se requieren niveles elevados de práctica de AF. Ambas organizaciones recomiendan entre 45-60 minutos diarios de AFM, o bien un equivalente de 30 minutos de AFV con el fin de beneficiarse de su efecto protector en la prevención del BC.

1.2.2. Dosis y beneficios durante el tratamiento

El tratamiento del BC combina diferentes intervenciones en razón al tipo, estadio y progresión de la enfermedad e incluyen: las terapias dirigidas, el tratamiento hormonal, la radiación y quimioterapia, y la cirugía. La quimioterapia y la radiación suelen causar alteraciones físicas y efectos secundarios fisiológicos que incluyen fatiga, insomnio, náuseas, vómitos, neuropatía, cardio-toxicidad y deterioro de la fuerza muscular, y que pueden ser altamente persistentes en el tiempo.

Para conocer los estudios más recientes sobre los beneficios de la AF en el tratamiento del BC, se ha realizado una actualización de trabajos de revisión sistemática en la literatura que reporta tres trabajos posteriores al informe científico PAGAC-2018 [16], [10] y [13] en cuyas conclusiones se ha sustentado el enunciado de evidencias en esta sección.

MEJORAS EN LA COMPOSICIÓN CORPORAL Y LA CALIDAD DEL SUEÑO:

- la intervención durante el tratamiento de 12 semanas de AFMV aeróbica, combinada con RT de intensidad moderada ha evidenciado contribuir a un incremento en la masa y la fuerza muscular, disminuyendo el porcentaje de grasa corporal [14] y [17],
- los estudios demuestran que incluso la AFL es la medida más eficiente para mejorar la calidad del sueño y el número de horas, contribuyendo a reducir sus desórdenes derivados entre las WSBC.

MEJORAS EN LA CRF Y LA FATIGA:

- en pacientes tratadas con antracilinas las intervenciones durante 6 semanas de AFV aeróbica han demostrado una mejora en los niveles de cardio-toxicidad, mostrando mejoras en la CRF [22],
- las intervenciones de AF de tipo HIIT contribuyen a una mejora en la CRF medida respecto del pico máximo de VO₂: [11] y [20],
- en WSBC con obesidad, el entrenamiento durante 16 semanas de AFMV ha demostrado una mejora significativa en la CRF de las participantes [3],
- las intervenciones de AF de pilates y yoga, con frecuencia de 3 sesiones semanales y un año de duración, evidencian [19] mejoras significativas en la CRF (+ 44.5 % y + 38.1 %),
- una intervención de 12 semanas de AFMV aeróbica combinada con ejercicio RT evidencia una disminución de la fatiga [17]; resultados similares se alcanzan en pacientes en tratamiento de edad avanzada con 16 semanas [2],
- en pacientes post-tratamiento, la comparación de intervenciones de 12 semanas de AFMV aeróbica con ejercicios de yoga y yoga aeróbico, concluye cómo los últimos logran mejores resultados en la reducción de la fatiga [21].

MEJORAS EN LA DENSIDAD ÓSEA:

- estudios realizados en mujeres tratadas con Inhibidores de aromatasa tras un año de AFM aeróbica combinada con ejercicios RT de intensidad moderada, evidenciaron cambios favorables en la renovación de la densidad ósea en la espina lumbar, así como un incremento en la masa muscular,
- la intervención con AFMV aeróbica durante el tratamiento contribuye a preservar la densidad ósea [18]; en el estudio se pudo constatar cómo el grupo de intervención con AFV combinada con RT preservó la densidad ósea de la medula espinal.

MEJORAS EN LA FUERZA MUSCULAR Y EL LINFEDEMA:⁷

- se reporta un incremento en la fuerza muscular después de 12 semanas de ejercicio RT de alta intensidad [14], [17],
- en WSBC de edad avanzada, después de 16 semanas de AFM combinada con ejercicios HIIT, se han alcanzado similares resultados.

El linfedema perjudica la CVRS ya que disminuye la función de la extremidad afectada; se ha encontrado que una ejercitación de RT intensa de 8 semanas durante el tratamiento disminuye los síntomas relacionados con el linfedema y una reducción en el volumen del brazo [5]. MEJORAS EN LA CVRS

- un régimen de entrenamiento de RT moderado de 6 semanas, sin combinar con otro tipo de AF contribuye a mejorar la CVRS en WSBC de edad avanzada [2],

⁷La extirpación de los ganglios linfáticos en la cirugía de cáncer de mama, ocasiona la dificultad del sistema linfático para drenar adecuadamente ocasionando una acumulación de líquido linfático denominada linfedema

- la combinación de ejercicio aeróbico a intensidad superior al 60 % y frecuencia cardíaca máxima (AFV) con ejercicios RT moderado ha demostrado mejorar la CVRS en mujeres con sobrepeso u obesidad [3],
- en pacientes en tratamiento, el entrenamiento de 9 semanas de ejercicio RT moderado, o bien de AFM aeróbica, inciden positivamente en la CVRS y en la composición corporal con una disminución (ver el trabajo [7] y el de [15] del Índice de Masa Corporal (IMC)).

Por último, en WSBC la asociación entre la AF, los factores de riesgo cardio-metabólico derivados de los tratamientos y la mortalidad ha sido analizada en el trabajo [4] en un estudio con más de 150.000 participantes de la *Women's Health Initiative* en Estados Unidos. Los resultados respaldan una asociación inversa entre las dosis de AFV y la tasa de mortalidad por todas las causas, que alcanza una reducción del 38 % (RR, 0.62; IC 95 % = [0.48,0.80]).

1.3. Estado de la AF en la mujer española

Para valorar el estado de la práctica de AF en la mujer española y estimar si los resultados son conformes a las dosis recomendadas por la OMS pero, sobre todo, si se alcanzan las dosis preventivas en el riesgo de BC, se requiere disponer de mediciones sobre AF en un número elevado de mujeres. Si bien uno de los objetivos del proyecto es el de proporcionar una tecnología que permita integrar diferentes fuentes de información y, en su caso, la recopilación de datos sobre práctica de AF adaptada a los aspectos preventivos del BC, inicialmente se ha podido contar con los datos facilitados y extraídos de la EHD-2022 del Ministerio de Cultura y Deporte con la colaboración del Instituto Nacional de Estadística (INE). Si bien ésta adolece de datos relativos a aspectos físicos, ginecológicos y de salud, nos ha proporcionado una valoración inicial del estado de la práctica de AF en la mujer española. Resulta de interés además para el proyecto conocer las posibles barreras en el acceso a la AF identificadas en las mujeres en relación a aspectos propios como la maternidad, la conciliación, o el cuidado de los mayores, entre otros. El cuestionario de la EHD plantea algún tipo de limitaciones generales para la práctica de AF, pero sin consideraciones de género.

1.3.1. Medición de la AF en la población

La obtención de mediciones precisas de AF resulta fundamental tanto en los estudios epidemiológicos sobre su influencia en la salud, como para la prescripción de pautas sobre AF. Para estimar el volumen de AF regular que realiza una población se utilizan dos tipos de aproximaciones: las mediciones auto-informadas, a través de cuestionarios o entrevistas y las mediciones objetivas, obtenidas mediante dispositivos que miden parámetros fisiológicos asociados a la AF.

MEDICIONES OBJETIVAS

Existen instrumentos con elevada precisión tales como la *técnica del agua doblemente marcada*, o la *calorimetría directa* y la *calorimetría indirecta* que, por su elevado coste económico y de tiempo, no resultan prácticos. Sin embargo hoy en día se cuenta con tecnologías y dispositivos *wereables* que permiten medir la AF tales como:

podómetros: dispositivos económicos y sencillos de usar, con sensores que registran el movimiento en una sola dirección midiendo el número de pasos del individuo mediante el cálculo indirecto de un algoritmo. Presentan como limitación la incapacidad para medir la actividad superior del cuerpo, así como la intensidad o la carga de la actividad, además suelen ser precisos para velocidades de marcha típicas, pero se ha comprobado que tienden a perder precisión en niños, o en personas con andares lentos. Los podómetros se pueden implementar como una intervención de AF, en tanto que el usuario puede motivarse para incrementarla,

acelerómetros: son dispositivos que miden el movimiento hasta en tres dimensiones, utilizando un dispositivo piezoeléctrico y microprocesadores que miden la aceleración del cuerpo y lo convierten en una señal digital. Estas señales permiten medir la intensidad, la duración y la frecuencia del movimiento registrado. El fundamento de la acelerometría es la existencia de una relación lineal entre la integral de la aceleración corporal y el consumo de oxígeno, hecho que permite el cálculo del gasto energético asociado al movimiento [12]. Diversos estudios señalan a la acelerometría como una de las técnicas más fiables en la medición de la AF con la ventaja de que se puede evaluar el tiempo de AF a diferentes intensidades y el propio sedentarismo, además de que pueden utilizarse (y se han demostrado efectivos) en todas las edades,

monitores de ritmo cardíaco (pulsómetros): la frecuencia cardíaca proporciona información indirecta acerca de la AF, el gasto de energía y la carga en el sistema cardio-respiratorio por la demanda física. Existe una relación lineal directa entre la frecuencia cardíaca y la energía consumida, pero hay una gran variación entre los individuos en función de su frecuencia cardíaca en reposo. Su precisión se limita a la AFV, debido a que el pulso es relativamente estable en bajos o moderados estímulos cardíacos; la frecuencia cardíaca además puede verse afectada por factores tales como el consumo de medicamentos, la presencia de alguna patología, o por los niveles de estrés. Por todo ello, los medidores de frecuencia cardíaca pueden no resultar precisos para medir la AF por sí solos,

medidores de pico máximo de VO_2 : se considera uno de los mejores indicadores de aptitud cardiovascular y resistencia aeróbica ya que está determinado por la capacidad del corazón, los pulmones y los músculos para captar y utilizar el oxígeno durante la AF. Actualmente los relojes inteligentes de alta gama de los fabricantes más populares ya estiman el consumo de pico máximo de VO_2 .

La principal ventaja de las mediciones objetivas es que no presentan problemas de interpretación en los procesos cognitivos del sujeto, ni de la denominada *deseabilidad social*⁸, o incluso la memoria. Por contra, estas últimas resultan bastante más sencillas de administrar y recoger.

MEDICIONES SUBJETIVAS

Se han desarrollado en la literatura diversos cuestionarios dirigidos a obtener medidas de la práctica de AF en la población. Los cuestionarios más utilizados y referidos en los estudios, de origen estadounidense son:

el IPAQ (*International Questionnaire of Physical Activity*): desarrollado en Ginebra en 1998, es uno de los más utilizados en los trabajos científicos para medir el nivel de AF. Existen estudios sobre su precisión y validez en 12 países, por lo cual está ampliamente verificado en investigaciones sobre prevalencia de participación en AF. El cuestionario divide la AF en 4 dominios (trabajo, desplazamiento, actividades en el hogar y tiempo libre), donde se interroga la frecuencia y duración de la práctica de AF en los tipos: AFM, AFV y caminata. En la última parte, se interroga sobre el tiempo de inactividad en un día a la semana y durante el fin de semana,

el GPAQ versión 2.0 *Global Physical Activity Questionnaire* es el actualmente recomendado y utilizado por la OMS en el método de vigilancia de enfermedades crónicas y promoción de la salud, siendo un instrumento confiable y con validación internacional para la medición de la AF. Al igual que el anterior, tiene en cuenta la intensidad, la frecuencia, la duración y el tipo de AF en diferentes dominios o escenarios, así como el comportamiento sedentario. Consta de 16 preguntas y el periodo de tiempo que evalúa es de una semana. Cabe mencionar que en él se concreta cómo la AFM equivale a actividades de 4 METs y es dos veces más alto en AFV (recogiendo actividades de 8 METs).

⁸Se trata de un concepto asociado al contexto experimental, acuñado desde la psicología consistente en el hecho de que la persona entrevistada tiende a elegir respuestas que favorecen su reputación en lugar de las más realistas, lo cual introduce un sesgo en la experimentación

1.3.2. Resultados arrojados por la EHD-2022

El porcentaje de mujeres en la muestra que no ha practicado AF durante el último año es de 51.25 %; en base a éste, el intervalo de confianza para el porcentaje de mujeres en la población que no realiza AF regularmente es: $IC_{95\%}(p) = [49.9\%, 52.6\%]$ que permite concluir que más de la mitad de las mujeres no han practicado AF en el último año. Un estudio gráfico más detallado (véase **Figura 1.2**), muestra cómo los porcentajes de práctica de AF y su frecuencia difieren según la edad, el nivel de estudios y la situación laboral, no así la situación personal (salvo en las mujeres solteras).

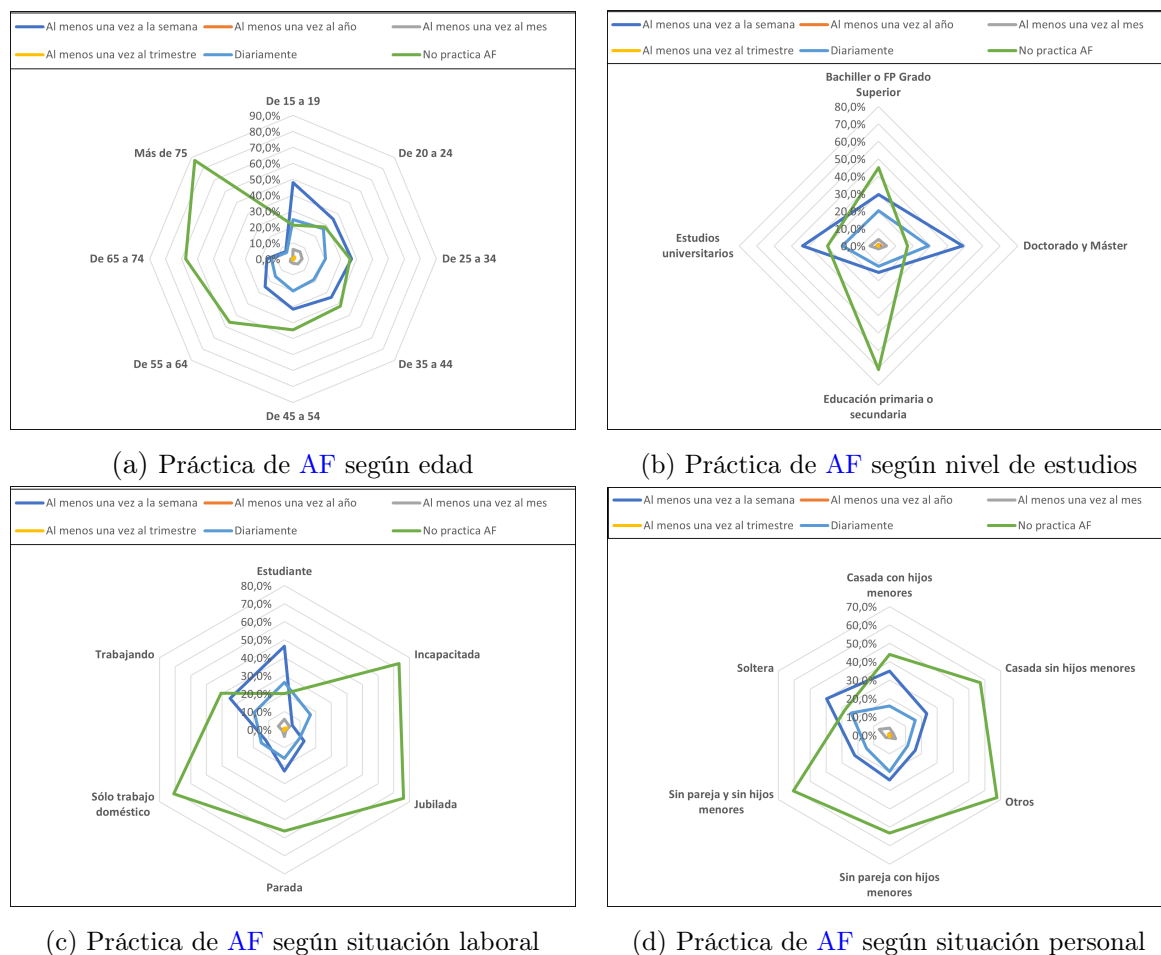


Figura 1.2: Distribución del porcentaje y frecuencia de AF en la población femenina atendiendo a su características. *Fuente: Elaboración propia*

El análisis inferencial ha permitido concluir que, a un nivel de significación del 5 %, existen diferencias significativas en los porcentajes de práctica de AF regular entre los diferentes grupos de edad, a excepción de los dos grupos de menor edad, en los que es indistinguible. Igualmente, se detecta la existencia de diferencias significativas en la práctica de AF regular según el nivel de estudios alcanzado, siendo el porcentaje creciente con éste, y fluctuando entre un 26.9 % para mujeres con estudios primarios, hasta un 77.3 % en mujeres con estudios de máster y/o doctorado. En lo referente a la situación laboral de la mujer, se evidencian diferencias significativas en todas las categorías, siendo las estudiantes con un porcentaje del 72.7 % las que más participan en la AF; opuestamente, en mujeres jubiladas, tan sólo un 22.2 % práctica AF de forma regular.

Centrándonos en la variable de interés en el proyecto, el tiempo destinado a la práctica de AF en la población femenina, se ha obtenido su distribución que, junto con otros gráficos usuales se muestran en la **Figura 1.3**. Como se observa, la distribución presenta una forma marcadamente asimétrica a derecha, exhibiendo un elevado número de valores atípicos. Un estudio convencional de

atípicos identifica cien registros de mujeres cuya práctica de AF aparece vinculada a sus estudios o actividad laboral, de los que se decide prescindir a fin de evitar una sobre-estimación. Con los datos depurados se obtiene que el tiempo medio destinado a la práctica de AF semanalmente en la muestra asciende a 92.9 minutos con el que se obtiene que el tiempo medio estimado en la población se encuentre en el $IC_{95\%}(\mu) = [95, 102]$. Éste permite concluir (con una garantía del 95 %) que en media, la población femenina no alcanza las dosis de AF recomendadas.

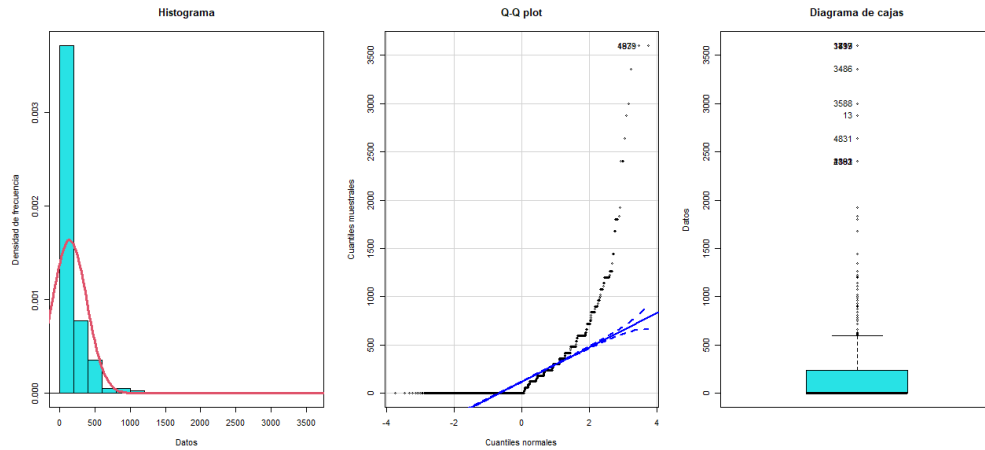


Figura 1.3: Histograma, gráfico Q-Q plot y gráfico Box-Plot del tiempo semanal de AF en la mujer
Fuente: Elaboración propia

No obstante, es probable que del mismo modo que se ha observado con los porcentajes, el tiempo dedicado a la AF difiera en función de las características de la mujer. Para analizar su dependencia respecto de las variables analizadas, se ha representado la distribución del tiempo de AF y el gráfico box-plot para el tiempo semanal de AF en función de la edad (**Figuras 1.4 y 1.5**, respectivamente), del nivel de estudios (**Figura 1.6**) y de la situación personal (**Figura 1.7**)

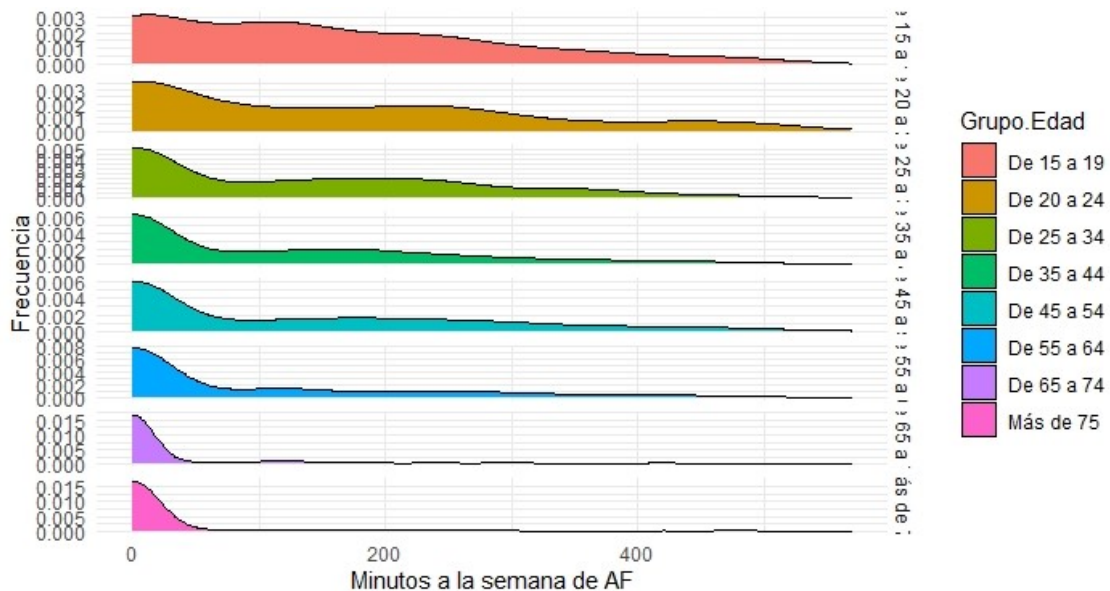


Figura 1.4: Distribución del tiempo semanal de AF según edad. Fuente: Elaboración propia

Se observa cómo la distribución del tiempo de práctica de AF varía en los diferentes grupos determinados por las covariables analizadas, y el análisis inferencial ha permitido concluir que los tiempos medios son significativamente diferentes en todos los grupos de edad, a excepción de los dos

grupos de menor edad en los que es indistinguible, superando los 150 minutos entre las mujeres más jóvenes; ninguno de los otros rangos de edad los alcanzan.

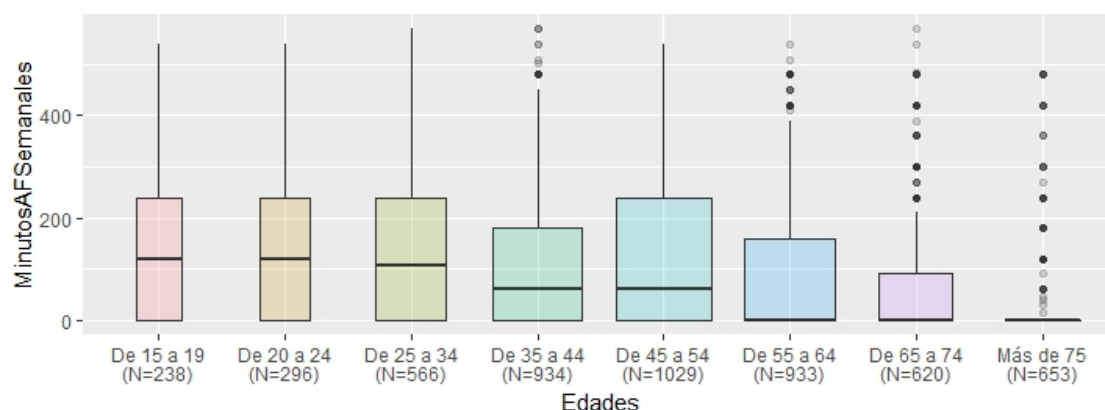


Figura 1.5: Box-Plot para el tiempo semanal de AF según edad. *Fuente: Elaboración propia*

Existen diferencias significativas según el nivel de estudios, alcanzando una media de 178 minutos entre mujeres doctoras ó máster, de 144 en mujeres con estudios universitarios, y significativamente menores en los niveles inferiores de formación. Respecto a la situación personal, el análisis inferencial concluye que no existen diferencias significativas entre el tiempo medio de AF entre las modalidades con o sin hijos menores, y con o sin pareja, difiriendo únicamente en la modalidad soltera, en la que es significativamente más elevado. En cuanto a la situación laboral, cabe señalar como dato curioso el que no se detectan diferencias significativas en el tiempo medio de práctica entre las mujeres en situación laboral trabajando y las que están en paro.

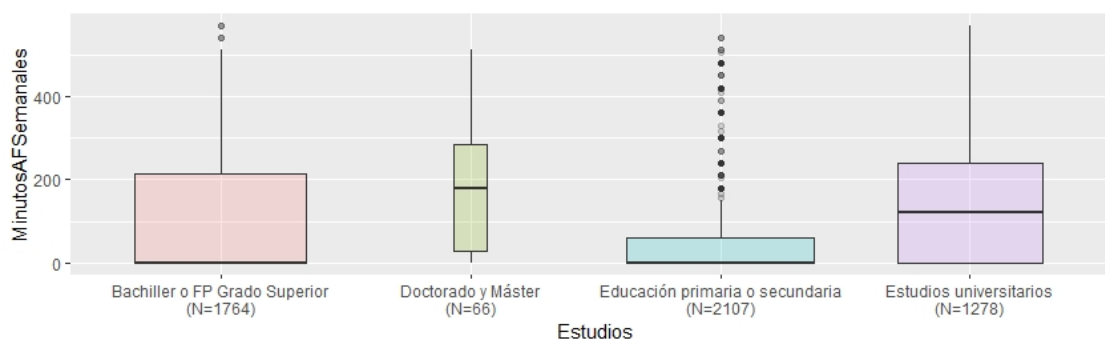


Figura 1.6: Diagrama de cajas para el tiempo semanal de AF por estudios. *Fuente: elaboración propia*

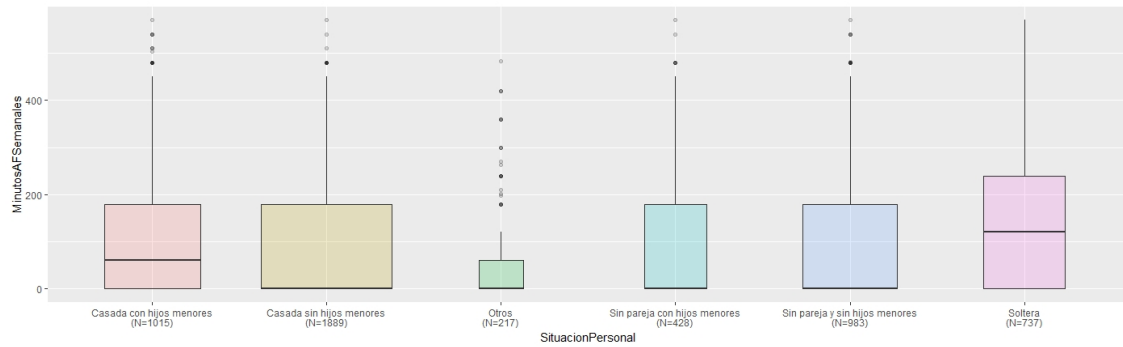


Figura 1.7: Diagrama de cajas para el tiempo semanal de AF por situación personal.

Por último, se han calculado, y se recogen en la **Tabla 1.3** las medidas descriptivas del tiempo de AF en función de las variables analizadas, así como sus respectivos intervalos de confianza para el tiempo medio de práctica de AF estimado en la población femenina.

Tiempo de AF según edad				Tiempo de AF según estudios			
Edad	Media	Desv.	$IC_{95\%}(\mu)$	Estudios	Media	Desv.	$IC_{95\%}(\mu)$
De 15 a 24	151	154	[131,168]	Primarios	56	138	[51,61]
De 25 a 34	126	138	[114,136]	Bachillerato	113	138	[106,119]
De 35 a 44	113	137	[103,121]	Universitarios	144	137	[136,152]
De 45 a 54	119	144	[110,128]	Doctorado	178	144	[141,215]
De 55 a 64	89	133	[81,99]				
De 65 a 74	66	124	[56,76]				
Más de 75	27	87	[20,33]				

Tiempo de AF según situación personal				Tiempo de AF según situación laboral			
S.Personal	Media	Desv.	$IC_{95\%}(\mu)$	S.Laboral	Media	Desv.	$IC_{95\%}(\mu)$
Soltera	142	138	[132,153]	Trabajando	120	138	[114,125]
Sin pareja ni menores	84	153.8	[76,92]	Parada	85	154	[74,96]
Sin pareja con menores	94	138	[82,107]	Jubilada	49	138	[42,55]
En pareja sin menores	89	137	[83,96]	Incapacidad	41	137	[16,65]
En pareja con comenores	110	144	[102,119]	Doméstico	64	144	[53,75]
				Estudiante	163	134	[150,178]

Cuadro 1.3: Medidas e IC para el tiempo semanal de AF (en minutos) según características

Para obtener una perspectiva de género para el nivel de práctica de AF en la población en España, se ha realizado y se presenta en el **Anexo C** el estudio de los principales resultados de la EHD en la población masculina. El estudio, similar al realizado en esta sección, nos ha permitido además realizar un análisis comparativo.

IDENTIFICACIÓN DE MOTIVACIONES Y DE BARRERAS EN LA ADOPCIÓN

La principal motivación para la práctica de AF identificada es la de "estar en forma" con un 27.5 % ($IC_{95\%} = [26.2\%, 28.7\%]$), seguida con un 26.7 % ($IC_{95\%} = [24.9\%, 27.4\%]$) por la "orientada a la salud". Estas dos principales motivaciones se mantienen en todos los grupos de edad, si bien se observa que, en mujeres de entre 35 a 55 años el orden se invierte, y lo mismo ocurre en los niveles

de estudios universitarios y máster o doctorado. Las modalidades de las situaciones personal y la situación laboral señalan a los motivos principales en el orden global de la encuesta. Por otra parte, la edad media de comienzo de la práctica de AF es de 17.5 años.

La principal barrera para no realizar AF (o más de la realizada) identificada es la *falta de tiempo*, que se elige como primer o segundo motivo por un 89.1 % ($IC_{95\%} = [88.2\%, 89.9\%]$) de las encuestadas, seguido de *la falta de interés* que se apunta como primer o segundo motivo para un 52.3 % ($IC_{95\%} = [50.9\%, 53.7\%]$) de las mujeres. Por edades no existen diferencias en cuanto a la barrera principal, a excepción de las mayores de 65 años en las que predomina *la falta de interés*. Todos los niveles de estudios coinciden en las dos barreras principales, al igual que ocurre entre las modalidades de la situación personal y laboral, con excepción de la situación de incapacidad en la que se apunta a "los problemas de salud".

1.4. Conclusiones y formulación de objetivos

Tras una revisión exhaustiva del estado del arte sobre el análisis de los hábitos deportivos y salud en la mujer, los artículos que se han revisado sistemáticamente hacen referencia a estudio clínicos con muestras relativamente pequeñas (menos de 100). Así mismo, se ha consultado a Nuria Garatachea y Lucía Sagarra que son investigadoras reconocidas en el campo de la medicina deportiva por la existencia de trabajos más amplios fuera del entorno clínico y no les consta ninguno.

Además, los estudios analizados sobre el estado de la práctica de AF se centran en el análisis de datos con técnicas de estadística descriptiva sin establecer modelos de inducción que permitan predecir la predisposición de la población femenina, en función de sus hábitos, a padecer enfermedades, en especial el BC y, mucho menos, poder determinar cuáles serían las medidas preventivas para poder reducir su incidencia.

El objetivo final de la línea de investigación abierta en este proyecto es el de superar las limitaciones en relación al tamaño de los estudios, obteniendo y analizando muestras suficientemente grandes; y en relación al enfoque estadístico, aplicando a esas muestras técnicas estadísticas de tipo inferencial y de modelización que nos permitan conocer la asociación con los hábitos saludables para estimar la magnitud de su efecto en la reducción del riesgo de desarrollar un BC.

Alineado con nuestro objetivo final, el objetivo parcial para este proyecto de Trabajo Fin de Máster es el de proporcionar las herramientas tecnológicas que den soporte a la automatización de los procesos, facilitando y contribuyendo al logro del objetivo final perseguido.

Adquiriendo datos relevantes

Conocido el dominio y con los objetivos formulados ha de procederse a la adquisición de los datos relevantes para su consecución, para lo cual es preciso evaluar el alcance, el volumen y la calidad de los datos requeridos, con el fin último de valorar si algunos de los recursos existentes pueden resultar adecuados, caso de contar además con la completitud y calidad requeridas.

En este capítulo se describen los dos tipos de actuaciones habituales en la adquisición de datos que se han acometido en el proyecto. En primer lugar, se comenzó realizando una búsqueda exhaustiva de fuentes de información fiables que divulgasen, o en su defecto, proporcionasen con propósitos de investigación, datos recientes sobre la medición de la práctica de AF en la mujer en España, junto con medidas relativas a sus hábitos saludables, y la de aquellas variables ginecológicas que se ha evidenciado ser relevantes en el riesgo de un diagnóstico de cáncer de mama. A nivel nacional e institucional, tanto el Sistema Nacional de Salud (SNS) que engloba los servicios de salud de la Administración del Estado y las comunidades autónomas, como el INE elaboran informes periódicos e indicadores de salud en los que incorporan variables de interés para el proyecto, tales como los estilos de vida y, en particular, la práctica de AF, junto con otros parámetros dedicados a evaluar el estado de salud de la población. No obstante, el carácter agregado de este tipo de indicadores los invalida para el propósito perseguido de análisis y evaluación de la asociación existente entre la práctica de AF y el riesgo de BC.

Para alcanzar los objetivos del proyecto se va a necesitar contar con datos y fuentes de información primarios⁹. Los requisitos de fiabilidad, rigor, representatividad y volumen de los conjuntos de datos, nos dirigieron al INE, como organismo encargado de realizar las operaciones estadísticas de gran envergadura y las macro-encuestas recogidas en su plan estadístico, en los diferentes ámbitos de la actividad humana en nuestro país; todo ello, con el atractivo de sus mecanismos de divulgación al proporcionar el acceso a datos primarios. En el dominio sociedad y el ámbito la salud, el INE realiza dos encuestas, algunas de cuyas variables están alineadas con los objetivos del proyecto: la Encuesta Europea de Salud en España (ESEE), que tiene como objetivo obtener datos sobre el estado de salud, la utilización de los servicios sanitarios y los factores determinantes de salud, todo ello de forma armonizada a nivel europeo. Su periodicidad es quinquenal, alternándose cada dos años y medio con la Encuesta Nacional de Salud (ENS), una operación estadística nacional colaborada por el Ministerio de salud, bienestar y consumo, con la que comparte grupo de variables armonizadas. Cabe mencionar, que en las dos últimas, que datan de los años 2020 y 2017, respectivamente, se muestrearon aproximadamente 37.500 viviendas distribuidas en 2.500 secciones censales en todo el territorio nacional, generando un total de más de 60.000 registros.

A pesar del importante volumen y la completitud en el número de variables de salud medidas, los datos de la ESEE presentan inconvenientes y limitaciones fundamentales para los objetivos del proyecto. Para el desarrollo de un Sistema de Vigilancia de la práctica de AF que permita evaluar el grado de cumplimiento de las pautas de la OMS, una desventaja proviene de que interroga directamente sobre el tipo de AF, en lugar de asignarlo en función del consumo MET de la actividad deportiva como se hace en los trabajos de investigación a fin de evitar el efecto de la *deseabilidad social*. Y cabe reseñar cómo los resultados ofrecidos por el INE son de carácter meramente descriptivo, sin proporcionar estimaciones estadísticas que, vinculadas a las recomendaciones de la OMS, permitan

⁹Se denominan datos primarios a aquellos que provienen directamente del informante, que los ha facilitado a través de un determinado servicio o en una encuesta, sin haber recibido tratamiento o elaboración alguna

formular conclusiones para la población.

Los datos de la [EESE](#) presentan una limitación insuperable para estimar el tamaño de la reducción en el riesgo de cáncer de mama asociado a la práctica de [AF](#) y los hábitos saludables en tanto no interroga sobre las variables ginecológicas relevantes, ni por los antecedentes familiares de cáncer¹⁰. Adelece también de la orientación con perspectiva de género que el proyecto persigue; no interroga sobre la motivación para la práctica de [AF](#), ni sobre los factores que podrían limitarla en la mujer por su condición¹¹ y que podrían explicar la menor participación en la [AF](#) de la mujer señalada en los informes de la [OMS](#).

2.1. Recolección de datos: El Cuestionario

Para avanzar en el proyecto, ante la necesidad de disponer de datos alineados y adaptados a los objetivos, se adoptó una primera decisión relevante para su desarrollo consistente en recabar los datos directamente a partir de las informantes, mediante un cuestionario diseñado a medida para obtener las mediciones de las variables requeridas.

Debido al carácter sensible de algunas de las variables de salud a recopilar, con el fin de dotar de las mayores garantías de seguridad y protección a los datos consignados por las encuestadas, se adoptó una segunda decisión significativa para el proyecto: la de prescindir del uso de los extendidos sistemas de cuestionarios externos, como *Google Questionnaire*, en favor de implementar una solución de almacenamiento y servidor propios. Al optar por una infraestructura interna, logramos ejercer un mayor control sobre la gestión de la información sensible, reduciendo los riesgos de la exposición a terceros. Se ha pretendido con ello contribuir al cumplimiento de los estándares de privacidad y las regulaciones, minimizando los riesgos asociados a la exposición de datos confidenciales y sensibles. En más detalle, se ha solicitado asesoramiento a la delegada de protección de datos de la Universidad de Zaragoza, la cual nos ha guiado a través del procedimiento para dar cumplimiento al contexto legislativo vigente en Europa (ver **Anexo D**).

2.1.1. Definición del contenido

La elaboración de un cuestionario efectivo para los objetivos del proyecto debe contemplar la inclusión de preguntas capaces de proporcionar las mediciones de la práctica de [AF](#) en lo relativo a su frecuencia, duración y tipología, así como del conjunto de variables incidentes en el riesgo del desarrollo del [BC](#) identificadas en el conocimiento del dominio. Concretamente junto a las características socio-demográficas de las participantes y aquellas que describen sus estilos de vida y hábitos saludables, se deben formular preguntas de características físicas y algunas variables ginecológicas de la mujer, así como otras, como la presencia de antecedentes familiares de cáncer. Se ha de preguntar finalmente si se ha recibido un diagnóstico de [BC](#), indagando en su caso¹², acerca de la prescripción oncológica de práctica de [AF](#) y la participación en [AF](#) durante y con posterioridad al diagnóstico.

Para la obtención de las medidas que cuantifican la práctica de [AF](#) y en la medición del tiempo de sedentarismo, se ha diseñado el mínimo número de cuestiones que proporcionen su volumen semanal, permitiendo categorizar el tipo de [AF](#) de acuerdo con la clasificación utilizada en los trabajos y por la [OMS](#). En la determinación de la intensidad nos hemos servido de las tablas de equivalencia [MET](#)

¹⁰La medición de esas variables es imprescindible desde el punto de vista de la formulación de un modelo estadístico a fin de evitar la confusión de efectos de los distintos factores incidentes

¹¹tales como la conciliación familiar y laboral, las responsabilidades domésticas, o el cuidado y atención de los hijos y/o personas mayores; aspectos éstos que tradicionalmente venían recayendo sobre ellas

¹²En conexión con los beneficios reportados de la práctica de [AF](#) durante y con posterioridad al tratamiento.

de las diferentes actividades físicas y deportes¹³ por cuya práctica se indaga.

La determinación y formulación de preguntas pertinentes proporciono un total de 45 preguntas con las que se diseñó una estructura del cuestionario definida en una secuencia lógica de agrupación de éstas entorno a las cinco secciones siguientes:

- *Sección I: ASPECTOS SOCIO-DEMOGRÁFICOS (SD)*
- *Sección II: ACTIVIDAD FÍSICO DEPORTIVA RECREATIVA (AFR)*
- *Sección III: FACTORES LIMITANTES (FL)*
- *Sección IV: HÁBITOS SALUDABLES (HS)*
- *Sección V: ASPECTOS FÍSICOS, GINECOLÓGICOS Y CÁNCER DE MAMA (AFGC)*

La redacción de cada pregunta se realizó cuidadosamente, evitando sesgos y ambigüedades, e incorporando opciones de respuesta acordes a su interés en el estudio. Un buen número de las preguntas ofrecen categorías alternativas para la respuesta, algunas de elección única y otras de elección múltiple. El resto de preguntas son de respuesta directa no elegible, o de puntuación en escala de *Likert*), en las que se indica con claridad las unidades o su rango. Se realizaron pruebas con un grupo reducido de mujeres con el propósito de asegurar la comprensión de las preguntas y certificar la exhaustividad de las respuestas ofrecidas¹⁴. El detalle completo de las preguntas formuladas y las respuestas ofrecidas se puede consultar en el **Anexo E**.

2.1.2. Diseño técnico

En el diseño y desarrollo del cuestionario *online* se ha seguido un proceso técnico y sistemático que garantiza contar con un instrumento de encuesta sólido, a la vez que sencillo de utilizar. Partiendo de la estructura en código *HTML* que define el diseño del cuestionario, se han ido incorporando los diferentes elementos característicos como *checkbox* y *radio buttons* para consignar la información de acuerdo con la tipología de la pregunta y sus opciones de respuesta. Se hace un uso meticuloso de hojas de estilo *CSS* definidas para lograr la estandarización de la apariencia visual y un diseño coherente. Se definieron clases específicas para los elementos conformando una base de código modular y organizada. La integración del marco *Bootstrap v5.0* con sus componentes prediseñadas y el sistema de retícula, contribuyó a facilitar el desarrollo y optimizar el diseño para la mejora de la experiencia de usuario y la capacidad de respuesta. Se implementaron diversas funcionalidades *JavaScript*, usuales en un formulario para lograr y mejorar la interactividad, tales como medidas de validación, generación dinámica de mensajes de error basados en las entradas del usuario, así como la gestión fluida de los envíos. El desarrollo ha sido implementado para adaptarse a diferentes tamaños de dispositivos, es decir se trata de un desarrollo Web *responsive*.

En la etapa final del desarrollo del cuestionario se efectuaron diversas evaluaciones con usuarios y expertos dirigidas a identificar y subsanar potenciales errores, o problemas de usabilidad. Incorporadas una serie de mejoras sugeridas por Sandra Baldassari, experta en usabilidad y a la que agradecemos su colaboración, se considera superada con éxito la fase de pruebas. Se ha prestado atención a la diversidad ofreciendo una herramienta accesible por cuanto en su diseño y desarrollo se han observado las directrices de las *Web Content Accessibility Guidelines (WCAG) 2.0* referidas a formularios accesibles.

Se puede acceder al cuestionario *online* en <https://cuantos01.unizar.es/tfm/cuestionario.php>, mientras que el código de su implementación está disponible en <https://github.com/LAURAMUNOZGARATACHEA/ProyectoAFySalud/>.

¹³En cualquier caso, con el objetivo de poder integrar en el sistema sus datos, las preguntas relativas a la práctica de **AF** han sido armonizadas con las del cuestionario de la **EHD**.

¹⁴Se contó con la colaboración de 12 mujeres en cada uno de los tres ciclos de revisión.

2.2. Almacenamiento

Con independencia de que finalmente se opte por integrar otras fuentes de información ¹⁵ que contribuyan al Sistema de Vigilancia, para el almacenamiento de los datos que provienen del cuestionario, en base a las características de la información recopilada, se ha optado por una base de datos relacional. El motivo principal es que permite dotar a los datos de una estructura normalizada evitando la redundancia y asegurando su consistencia. Además, se facilita la gestión de relaciones entre diferentes tablas de datos, en particular simplificando la consulta de datos relacionados, posibilitando un análisis más eficiente. Por otra parte, la integridad relacional garantiza la coherencia de las tablas evitando, en particular, la inserción o actualización de datos que quiebren estas relaciones; todo ello resulta esencial para mantener la precisión de la información, y por ende su calidad, máxime en un contexto tan sensible como el de salud. Finalmente, a nivel de seguridad, los [Sistemas Gestores de Bases de Datos \(SGBD\)](#) ofrecen actualmente capacidades avanzadas y control de acceso, permitiendo definir permisos específicos para usuarios o roles, garantizando la protección y confidencialidad en la custodia de datos sensibles.

Para alcanzar un diseño óptimo del esquema de la base de datos, se ha decidido evaluar el rendimiento de determinadas consultas¹⁶, de esquemas alternativos para éste. Por las características del cuestionario se han evaluado los siguientes tres esquemas ¹⁷ alternativos :

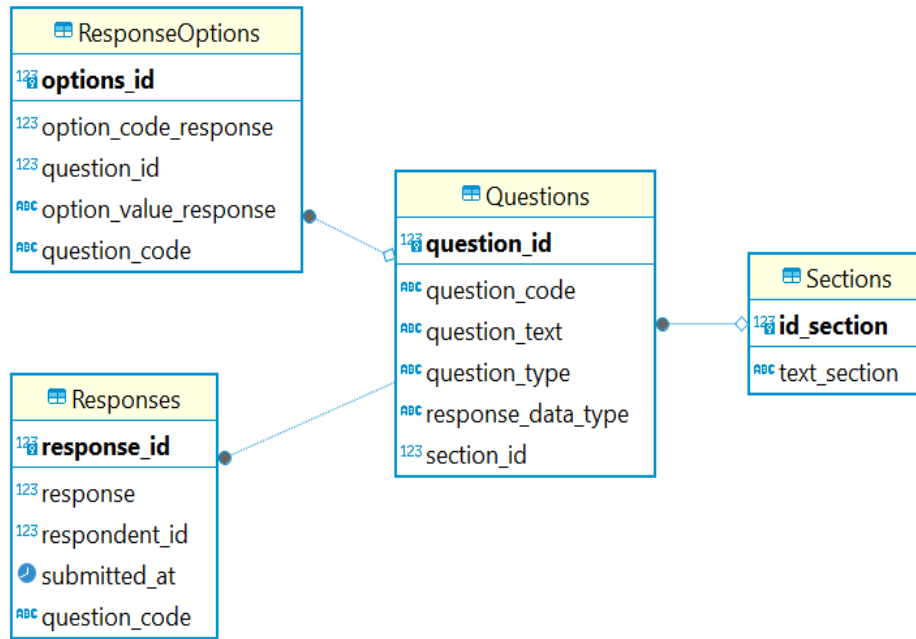
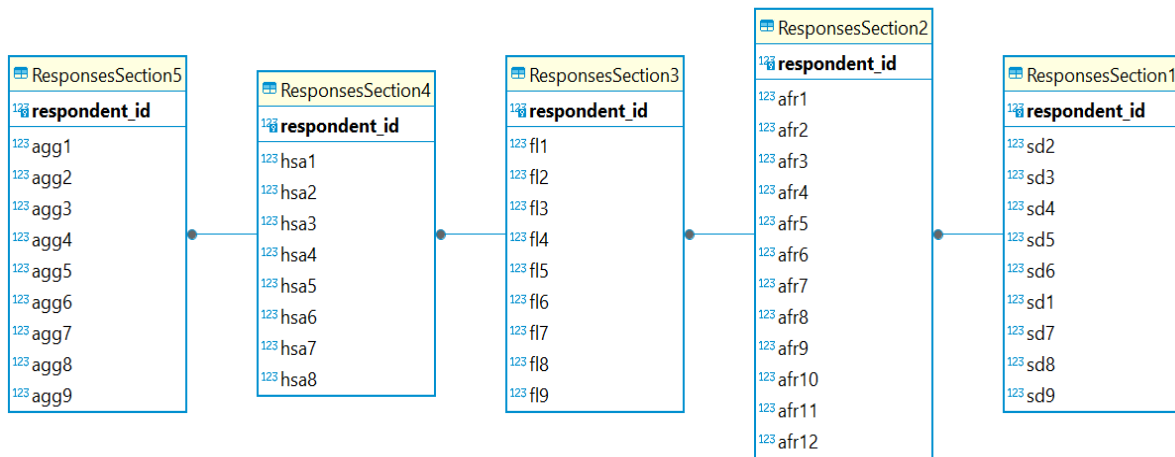
- *esquema 1- Respuesta Por Pregunta*: Consistente en un esquema normalizado en el que, para un identificador de usuario, cada registro almacena la respuesta a una pregunta identificada por su código. (véase la **Figura 2.1 a**))
- *esquema 2-Respuestas Por Sección*: Consistente en un diseño que, para un identificador de usuario, almacena las respuestas a las cuestiones de cada sección en sendas tablas (véase la **Figura 2.1 b**))
- *esquema 3-Respuestas Completas*: para un identificador de usuario, almacena las respuestas en una tabla única con tantas columnas como preguntas. Este esquema es similar al *esquema 2*, con una mayor desnormalización. ¹⁸

¹⁵La [EHD](#) con cuyas preguntas y modalidades para la respuesta se han alineado las del Cuestionario

¹⁶en lo que a eficiencia en el acceso a las tablas se refiere

¹⁷Cabe señalar que debido al elevado número de preguntas, no se ha considerado un posible esquema consistente en almacenar cada pregunta en una tabla dedicada

¹⁸Por su elevado tamaño y obviedad del diseño éste no se recoge se la **Figura 2.1**.

(a) Esquema *Respuestas Por Sección*(b) Esquema *Respuesta por Pregunta***Figura 2.1:** Esquemas relacionales de los **Esquemas 1 y 2**

En las pruebas de rendimiento se han utilizado 50.000 registros simulados de respuestas (acordes a su tipología y rango) de forma programática mediante *scripts* en **Phyton** con los que se crean y pueblan las tablas requeridas en cada uno de los esquemas (ver **Anexo F**). Nos hemos servido de la herramienta *Knime* que (además de facilitar la integración de datos de diferentes orígenes) ha permitido computar el tiempo de ejecución del flujo del proceso de recuperación de datos de las consultas formuladas.

La elección del esquema final se sustentó en el rendimiento de las consultas más frecuentes a realizar por el sistema para nutrir a la herramienta de visualización y analítica. Con el análisis de la **EHD** se evidenció que la práctica de **AF** difiere según características socio-demográficas. En la evaluación del grado de asociación entre la tasa de prevalencia del **BC** y los hábitos saludables y, en particular, la práctica de **AF**, se van a realizar de forma frecuente consultas que, en términos generales, consisten en recuperar datos de una variable para determinadas respuestas a otra, u otras preguntas. Ha sido éste el tipo de consultas evaluadas, de forma parametrizada, para la comparación

de los tiempo de respuesta los tres esquemas testados.

Es conveniente señalar, no obstante, cómo en el uso de técnicas multivariantes de modelización en la analítica de datos, se deben recuperar los datos de todas las variables presentes en el modelo. Por este motivo, se ha evaluado el tiempo de respuesta para esta consulta.

En la **Figura 2.2** se muestran los flujos en los dos tipos de consultas formuladas en cada uno de los tres esquemas, destacando el rendimiento¹⁹ del esquema elegido. Se observa que, para las consultas de las las herramienta de visualización y analítica bivalente, el *esquema 1* es el más eficiente. Como era de esperar, para el propósito específico de un analítica de datos multivariante, el *esquema 3* resultaría ser el más eficiente. Por su propósito más específico y mayor complejidad la analítica multivariante, resulta de menor interés a nivel de usuario y se determina que el diseño de la base de datos a utilizar sea el *esquema 1*.

2.3. Tecnologías utilizadas y alojamiento Web

Inicialmente, la preferencia para el desarrollo era *Java JDK* con el entorno de ejecución *JEE* debido a sus múltiples ventajas como la portabilidad, la seguridad, la integración de servicios, el manejo de la concurrencia, de las transacciones y las excepciones, sus herramientas de desarrollo avanzado y, fundamentalmente por ser el más familiar para la autora. En lo referente a la elección del **SGBD** la elección apuntaba a *PostgreSQL* con un posición actualmente destacada por diferentes aspectos como su rica variedad de tipos, ser considerado más estricto en cuanto a la integridad referencial, soporte más avanzado para procedimientos almacenados y funciones, la capacidad de usar múltiples lenguajes, la gestión avanzada de concurrencia, la escalabilidad y una licencia de código abierto.

No obstante, la elección final de la tecnología utilizada en el desarrollo del sistema de información, ha venido condicionada por la elección del proveedor de *hosting*: el **Servicio de Informática y Comunicaciones de la Universidad de Zaragoza (SICUZ)**. Por variadas razones tales como el respaldo y soporte de la Universidad de Zaragoza al sistema, o la condición de **Personal Docente e Investigador (PDI)** de la autora, se valoró con prioridad la oferta del **SICUZ** dirigida al **PDI** de la universidad de un servidor virtual privado (una máquina virtual, con gestión y mantenimiento), junto con el soporte para el despliegue Web. Se nos comunicó cómo, a fin de evitar los riesgos de seguridad, tanto el servidor, cómo el despliegue del sistema, limitan el desarrollo al *stack* de tecnologías con las que se dotan, a saber: *PHP*²⁰ como lenguaje de programación *scripting* y *MaríaDB* como **SGBD**.

Con el propósito de poder contar con el alojamiento del **SICUZ** sin por ello renunciar a la elección de tecnologías, se abrió un periodo de conversaciones, en el que se ofreció proporcionarles la tecnología en un contenedor *docker*. Finalmente, se constató cómo el alojamiento del **SICUZ** estaba muy restringido a determinados contextos de la actividad académica y de investigación. A pesar de todo ello, por la disponibilidad de un servidor²¹ en el departamento de Economía Aplicada al que pertenece la autora, con mantenimiento propio, pero con respaldo y despliegue por el **SICUZ**, se consideró la alternativa finalmente adoptada como solución para el despliegue del sistema. Éste igualmente limita las tecnologías a la utilizada en las máquinas del **SICUZ**.

En el desarrollo y las pruebas del sistema se ha utilizado el *framework XAMPP*, un conjunto de software libre y de código abierto que incorpora **Apache**, **MySQL**, **MariaDB**, **PHP** y **Perl** y proporciona la herramienta *PhPAdmin* para la administración de bases de datos. Este entorno nos ha ofrecido una solución todo en uno que ha permitido desarrollar y testar de manera eficiente el sistema previamente

¹⁹Las pruebas se realizaron en un equipo informático con un procesador Intel Core i7 de 8ª generación con 2.6 GHz y turbo hasta 4.5 GHz.

²⁰Software de código libre, compilado, ampliamente presente en el desarrollo Web, se integra con servidores como **textitApache** y **Nginx**; es especialmente conocido por ser integrado con *HTML* en la creación de páginas web dinámicas

²¹el servidor da soporte a las actividades del Grupo de Decisión Multicriterio (GDMZ)

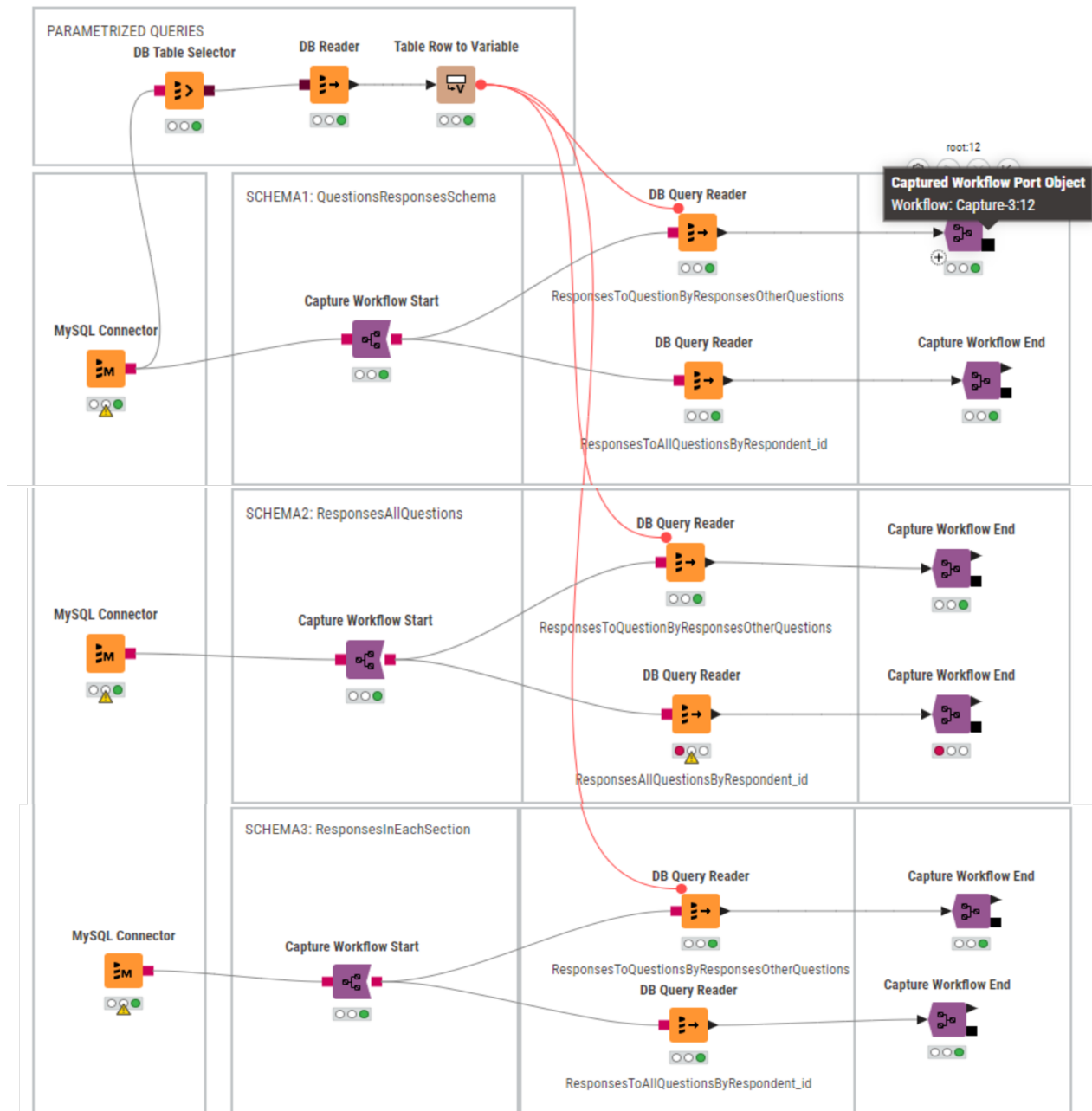


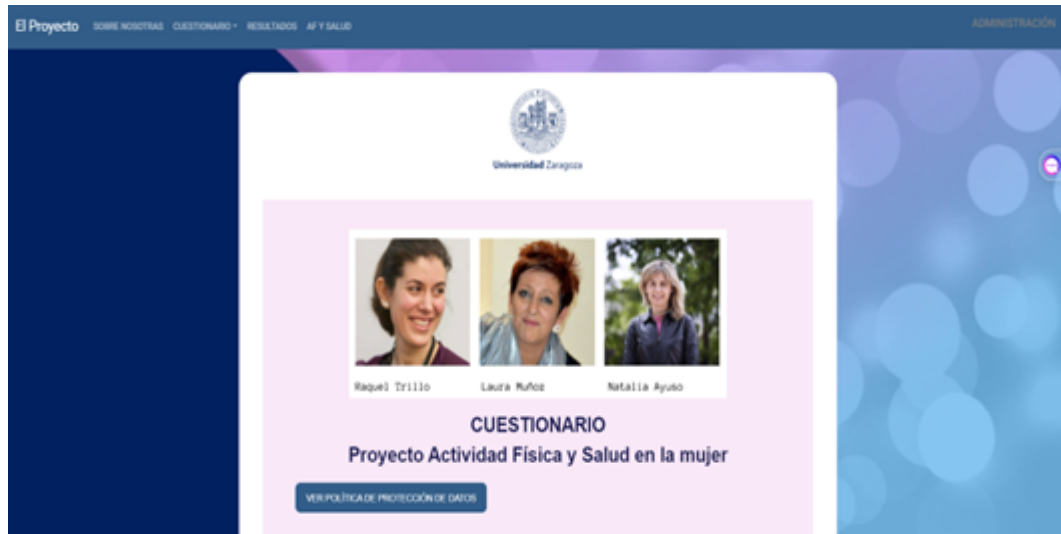
Figura 2.2: Evaluación con Knime de tiempos consumidos en los flujos de trabajo de las consultas en los tres esquemas alternativos

a su despliegue. Como *IDE's* de desarrollo se han utilizado indistintamente *Visual Studio Code* y *PHPStorm* de *JetBrain*.

Complementando al sistema de información que da soporte a la gestión de los envíos del Cuestionario, se ha configurado una administración del sistema que facilita el acceso a los datos garantizando su seguridad. El sistema ofrece otras funcionalidades de tipo informativo, como la presentación del proyecto, la información sobre el tratamiento y la protección de los datos, o el enlace a recursos de interés; y ofrecerá la funcionalidad de integración de la herramienta de visualización y analítica de resultados del Sistema de Vigilancia.

La apariencia y el comportamiento adaptativo al dispositivo de acceso al cuestionario y del sistema **Proyecto AF y Salud** se muestra en la **Figura 2.3**. Éste se encuentra alojado y es accesible en el

enlace <https://cuantos01.unizar.es/tfm>.



(a) Ordenador personal



(b) Tablet



(c) Móvil

Figura 2.3: Apariencia *responsive* del Sistema de Información Web

El código completo de desarrollo del sistema de información Web **Proyecto AF y Salud** está disponible en <https://github.com/LAURAMUNOZGARATACHEA/ProyectoAFySalud/>.

El fragmento de código siguiente, por ejemplo, muestra el proceso programático que realiza el guardado (inserción) de registros en la base de datos.

```
<?php
session_start();
$respondentId = session_id() . '_' . $_SESSION['form_counter'];
echo "Session is set. Respondent ID: " . $respondentId . "<br>";
include("../administrador/config/bd.php");
if ($_SERVER["REQUEST_METHOD"] == "POST") {
    $sql = "SELECT question_code FROM questions_metadata";
    $stmt = $pdo->prepare($sql);
    $stmt->execute();
    $questionCodes = $stmt->fetchAll(PDO::FETCH_COLUMN);
```

```
foreach ($questionCodes as $code) {
    $fieldName = $code;
    if (isset($_POST[$fieldName]) || isset($_POST[strtolower($fieldName)])) {
        if (is_array($_POST[$fieldName] ?? $_POST[strtolower($fieldName)]) {
            $selectedOptions = implode(',', $_POST[$fieldName] ?? $_POST[strtolower($fieldName)]);
            $respuesta = NULL; // Set $respuesta to NULL for multiple-choice questions
        } else {
            $respuesta = $_POST[$fieldName] ?? $_POST[strtolower($fieldName)];
            $respuesta = ($respuesta === "") ? NULL : $respuesta;
            $selectedOptions = NULL; } }
    else {
        $respuesta = NULL;
        $selectedOptions = NULL;
    }
    $insertQuery = "INSERT INTO responses (response, respondent_id,
submitted_at, question_code, selected_options)
VALUES (?, ?, NOW(), ?, ?)";
    $insertStatement = $pdo->prepare($insertQuery);

    $insertStatement->bindParam(1, $respuesta);
    $insertStatement->bindParam(2, $respondentId);
    $insertStatement->bindParam(3, $code);
    $insertStatement->bindParam(4, $selectedOptions);
    $insertStatement->execute();
    //echo "Response $respuesta successfully inserted into the database.";
    } }
header('Location: mensajeEnvio.php');
session_write_close();
?>
```

Capítulo 3

Desarrollando una aplicación Web dinámica e interactiva para la Visualización y Analítica

3.1. Contexto, motivación y objetivos

Con el avance continuo tanto de la disponibilidad de datos, como de las herramientas de manipulación de grandes volúmenes de datos, se ha vuelto cada vez más relevante el desarrollo de herramientas de visualización igualmente poderosas que sirvan de soporte a la aplicación de herramientas estadísticas de analítica y modelización. Estas herramientas facilitan a los analistas y científicos de datos, de un lado comprender, y de otro, comunicar patrones o descubrimientos complejos de una manera más efectiva.

El propósito principal del desarrollo de nuestra aplicación del Sistema de Vigilancia es el de proporcionar una herramienta compacta que, combinada con las herramientas descritas en el **capítulo 2** que proporcionan los datos, plasme visualmente y numéricamente el comportamiento de éstos, facilitando examinar, analizar, descubrir y, en su caso, concluir, en base a la información que encierran. Todo ello debe ofrecerse en un contexto dinámico, nutrido de las llegadas de nuevos datos, e interactivo facilitando al usuario la comprensión de su comportamiento en los ámbitos global, o particular (para diferentes características de la mujer).

Los objetivos de este capítulo son exponer el proceso de elección de tecnologías, la descripción de sus características y funcionalidades principales, así como el proceso de diseño y desarrollo de la aplicación. Se presentará la opción de despliegue adoptada y, finalmente, mediante varios casos de uso, se proporcionará de forma breve algunos ejemplos de permitirán mostrar el diseño alcanzado para nuestro Sistema de Vigilancia y la implementación de algunos elementos.

3.1.1. Elección de tecnologías

En un contexto de diseño y desarrollo, caracterizado por un elevado nivel de presentación visual, capaz de proporcionar al usuario cierto grado de interactividad, la tecnología que le va a dar soporte deberá brindar capacidades robustas de visualización y análisis de datos, junto con una integración fluida con el sistema de información desarrollado.

A continuación presentamos una clasificación amplia, atendiendo a diferentes criterios, de las herramientas disponibles y que abarca desde bibliotecas de visualización de datos de código abierto, hasta *suites* completas de análisis empresarial, proporcionando una visión bastante exhaustiva de las características, capacidades y usos de cada tipo de herramienta. El objetivo es el de facilitar una selección informada de la alternativa que mejor se adecúa a los objetivos formulados en nuestro proyecto.

1. BASADA EN EL PROPÓSITO PRINCIPAL:

- **analítica de datos:** herramientas centradas principalmente en el análisis estadístico y matemático de los datos, como MATLAB, R, y Python con bibliotecas como NumPy y SciPy,

- **visualización de datos:** aplicaciones que se enfocan en la representación visual de los datos, como Tableau, Power BI, D3.js y Plotly,
- **minería de datos:** desarrollos que permiten implementar herramientas de minería de datos para descubrir patrones, relaciones en grandes conjuntos de datos o su clasificación tales como Orange, RapidMiner o Weka.

2. BASADA EN EL TIPO DE USUARIO:

- **profesionales de datos:** herramientas dirigidas a analistas de datos y científicos de datos con experiencia en programación, como R, Python, MATLAB o herramientas de línea de comandos,
- **cuadros de mando:** herramientas que se centran en la creación de informes y paneles de control para usuarios comerciales, como Tableau, Power BI, o Metabase,
- **usuarios no técnicos:** herramientas con interfaces gráficas de usuario intuitivas y sencillas de usar, como Orange y algunas funcionalidades de Excel y Google Sheets.

3. BASADA EN LA NATURALEZA DE LA INTERFAZ:

- **línea de comandos:** herramientas que se utilizan principalmente a través de una interfaz de línea de comandos, como R, MATLAB y algunas funciones de Python,
- **interfaz gráfica de usuario (GUI):** herramientas que ofrecen una interfaz gráfica de usuario para la interacción y manipulación de datos, como Tableau, Orange y RapidMiner,
- **desarrollo web:** herramientas que permiten crear aplicaciones web interactivas para análisis y visualización de datos, como Shiny, Dash o Bokeh.

4. BASADA EN EL LENGUAJE DE PROGRAMACIÓN:

- **lenguaje R:** herramientas que se desarrollan entorno al lenguaje de programación R, como RStudio, Shiny o ggplot2
- **lenguaje Python:** herramientas que se basan en el lenguaje de programación Python, como Pandas, Seaborn, Dash o Flask
- **otros lenguajes:** herramientas que utilizan otros lenguajes de programación, como MatLab con su propio lenguaje, y JavaScript, utilizado en bibliotecas como D3.js

Por la mayor adecuación para los objetivos del desarrollo de la aplicación del Sistema de Vigilancia como una aplicación Web interactiva, desarrollada con *software* libre, se va a concluir que la elección de tecnología lleva a decidir, esencialmente, entre las basadas en R y las basadas en Python. En la **Tabla 3.1** se muestra una comparativa de las características de las alternativas disponibles más reconocidas en cada lenguaje. Las diferencias en cuanto a ventajas e inconvenientes son escasas, apuntando a una mayor sencillez de Shiny. Con ello, es posible afirmar que la elección se ha basado en las preferencias de la programadora. Debido a un natural compromiso de la autora con el lenguaje estadístico R, aunque durante la realización del máster se ha trabajado y en la actualidad se cuenta con mayor experiencia en Python, se ha decidido optar por Shiny con el fin de abordar el reto de retomar la programación en R y de explorar el alcance y profundizar en el desarrollo de las aplicaciones Shiny de cuadro de mandos con vistas a futuros proyectos académicos y/o de investigación.

Determinada la tecnología de desarrollo base, es preciso apuntar cómo en el ecosistema Shiny pueden adoptarse dos enfoques técnicos de desarrollo diferenciados que, sin embargo resultan indistinguibles a efectos de la aplicación final. Un primer enfoque consiste en el desarrollo de una aplicación Web Shiny propiamente dicha, mediante el uso del paquete ShinyDashboard diseñado para facilitar el desarrollo de un panel. El enfoque alternativo consiste en servirse de R-Markdown combinado con el paquete Flexdashboard, estableciendo la ejecución como Shiny. Para decidir entre ambos enfoques, en la **Tabla 3.2** se muestra una comparativa de algunas de sus características relevantes para la decisión. En tanto que el segundo enfoque se adecúa a la característica principal de

Característica	Shiny	Dash	Bokeh
Lenguaje de Programación	R	Python	Python
Desarrollador Principal	RStudio	Plotly	Continuum
Licencia	GPL-3.0	MIT	BSD-3-Clause
Integración con otros lenguajes	HTML, JS, Python	HTML, JS, R	HTML, JS, R
Desarrollo de gráficos	R, ggplot2, plotly, etc.	Plotly, Dash core	Bokeh, Matplotlib
Comunidad	Grande	Grande	Grande
Facilidad de uso	Alta	Media	Media
Documentación y soporte	Amplia, con ejemplos	Completa	Completa
Desarrollo Activo	Sí	Sí	Sí
Gratuito para uso no comercial	Sí	Sí	Sí

Cuadro 3.1: Comparación de alternativas para el desarrollo de aplicaciones web interactivas

nuestra aplicación de elevado nivel de un elevado nivel de presentación visual, junto con la ventaja que supone la ejecución parcial y progresiva de los denominados *chunks*²² de código, además de su carácter gratuito, finalmente se ha decidido la tecnología **Flexdashboard** con **R-Markdown** y **Shiny**.

3.2. Descripción de tecnologías

Se describen en esta sección las características de las dos herramientas técnicas claves en el diseño y el desarrollo **Flexdashboard** y **Shiny**, respectivamente presentando sus funcionalidades principales.

3.2.1. Shiny App

Shiny es un paquete para **R** de código abierto que permite a los usuarios construir aplicaciones Web interactivas desde **R** enfocadas principalmente a la creación de **IUs** para programas de procesamiento de datos y tableros de visualización (ver la **Figura 3.1**). Usando funciones y objetos de **R** consigue generar el código de páginas **HTML**, estilos **CSS** y el código **JavaScript** necesario para construir la **IU** de la aplicación y el servidor que le da soporte. Para alcanzar el dinamismo, junto a la generación de código en base a componentes de alto nivel, **Shiny** se sirve de una implementación relativamente intuitiva de la programación reactiva, para procesar las entradas de usuario y generar las salidas requeridas.

El potencial mostrado por la herramienta la ha popularizado entre analistas de datos, estadísticos y matemáticos, que suelen disponer de amplios conocimientos de modelización con **R**, pero con una experiencia limitada en programación Web. Con el uso extensivo de **R** como lenguaje para proyectos científicos y estadísticos, **Shiny** ha crecido para convertirse, no solo en una herramienta de prototipado, sino en un *framework* de desarrollo de pleno derecho, enfocado en la simplicidad de uso. La implicación de organizaciones de referencia y la contribución desinteresada de desarrolladores alrededor del mundo²³, ha propiciado que surjan alrededor de **Shiny** una enorme variedad de paquetes con nuevas funcionalidades. .

La interactividad en **Shiny** se obtiene a través del paradigma de la programación reactiva, un concepto central de la metodología. Una variable reactiva (o, en términos de **Shiny** un *Reactive Producer*) en **Shiny** es un objeto similar al sujeto en el patrón observador [8], que notifica a sus

²²Sección de código que se ejecutada independiente, cuyo resultado puede ser renderizado.

²³Entre los socios destacados, por citar alguno, se encuentra *AgileSoDA* en Corea del Sur que proporciona un soporte integrado a las empresas para internalizar tecnologías de inteligencia artificial con modelos de *Machine Learning*, o la europea *Appsilon* que ha desarrollado algunos de los más avanzados paneles de mandos proporcionando soporte a más de 500 compañías en el desarrollo de sus aplicaciones **Shiny** comerciales

Característica	Flexdashboard	ShinyDashboard
Lenguaje de Programación	R	R
Desarrollador Principal	RStudio	RStudio
Licencia	GPL-3.0	GPL-3.0
Estructura y Diseño	– Organización en páginas, filas, columnas y tarjetas. – Diseño <i>responsive</i> con Bootstrap – Soporta archivos CSS personalizados	– Basado en un <i>layout</i> con barras laterales y cuerpos de página. – Diseño <i>responsive</i> con Bootstrap – Temas personalizados a través de CSS
Funcionalidad	– Combina texto narrativo con visualizaciones interactivas – Ideal para informes y documentos interactivos	– Fuerte enfoque en la creación de aplicaciones web interactivas – Integración con <i>Widgets</i> y elementos interactivos de Shiny
Reactividad	Implementación básica de reactividad. Ideal para cuadros de mando estáticos o con interactividad limitada.	Completa implementación de reactividad a través de Shiny . Ideal para aplicaciones interactivas complejas
Facilidad de Uso	Alta. Orientado a usuarios que desean crear cuadros de mando con una curva de aprendizaje baja	Media. Requiere mayor conocimiento de Shiny y su ecosistema
Documentación y Soporte	– Documentación completa. – Ejemplos y tutoriales disponibles	– Documentación completa. – Amplios recursos y comunidad de usuarios de Shiny
Integración con Otros Lenguajes	HTML, JavaScript, Python	HTML, JavaScript, Python
Desarrollo activo	Sí	Sí
Gratuito para uso no comercial	Sí	No

Cuadro 3.2: Comparación entre Flexdashboard y ShinyDashboard

observadores (o *Reactive Consumers*) cuando se produce un cambio sobre él. Un ejemplo gráfico del flujo reactivo puede verse en la **Figura 3.2**. **Shiny** proporciona dos tipos de Consumidores para manejar los datos reactivos:

- **Reactive:** La función *reactive* encapsula una expresión (reactiva o no) y devuelve un objeto reactivo. Al usarlas como consumidor sirven para realizar cálculos intermedios sobre un objeto reactivo proporcionando su resultado final. Lo interesante de esa evaluación es que es perezosa, es decir, solo se evalúa si un consumidor de reactividad lo requiere.
- **Observe:** La función *observe* encapsula una expresión en un objeto observador. Este objeto ejecuta la expresión pasada como parámetro cada vez que alguna de sus dependencias reactivas notifica un cambio. Estos objetos observadores tienen dos diferencias fundamentales sobre

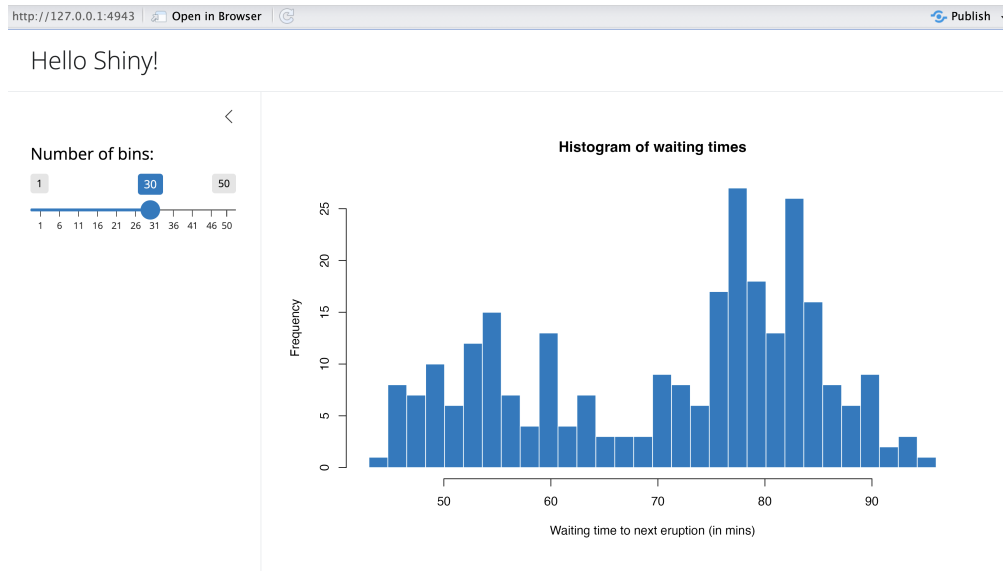


Figura 3.1: Ejemplo de una visualización interactiva básica con Shiny. *Fuente:* <https://shiny.posit.co/>.

los reactivos: no pueden devolver ningún valor y su evaluación no es perezosa, por lo que resultan útiles para modificar la IU. En Shiny todos los elementos de entrada están vinculados a *Reactive Produces* y todos los elementos de salida se asocian a *Reactive Consumers*, por lo que las operaciones de entrada/salida se limitan a grandes rasgos a una operación de lectura o asignación de una variable.

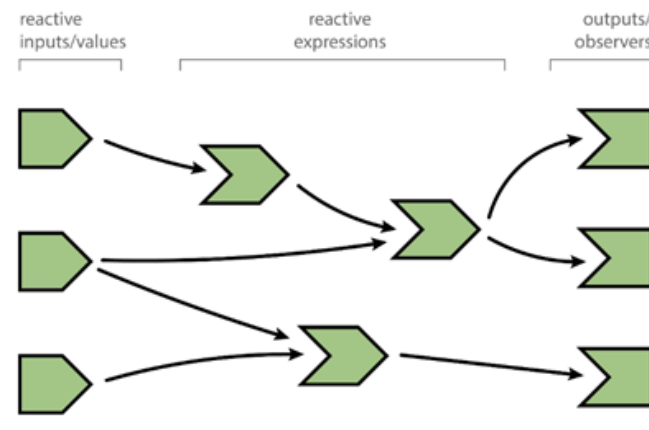


Figura 3.2: Características y funcionalidades de una aplicación Shiny. *Fuente:* elaboración propia

3.2.2. Flexdashboard con R-Markdown

Flexdashboard es un *framework* de desarrollo de tableros de control interactivos basado en R Markdown que ha sido diseñado para facilitar la creación de cuadros de mando en R. La funcionalidad principal de Flexdashboard es la capacidad de combinar visualizaciones de datos dinámicas y flexibles, con texto descriptivo o explicativo. Ello se logra con la integración mediante bloques de código R que generan gráficos y otros elementos interactivos con el diseño de lienzo determinado que con la compilación convierte en el código HTML de un panel de mandos interactivo que puede ser visualizado en un navegador.

El mayor atractivo de **Flexdashboard** es su capacidad para organizar el contenido en un diseño flexible y adaptable en el que los desarrolladores pueden estructurar el diseño de un panel en el cuadro de mandos dividiendo el lienzo en retículas de filas y/o columnas, a la vez que proporciona unas clases predeterminadas para las tarjetas que acogen al elemento visual, ajustando éste, de forma estilizada, a las dimensiones del contenido. La estructura modular proporcionada por **Flexdashboard** facilita alcanzar una disposición organizada y clara de los elementos visuales en la composición de la aplicación o cuadro de mandos a desarrollar.

La combinación de **Flexdashboard** con **Bootstrap**, el popular *framework* de **CSS**, facilita gestionar y desarrollar un diseño estético, estructurado y adaptable para un cuadro de mandos. **Bootstrap** proporciona una amplia gama de componentes predefinidos y temas, junto con un sistema de grillas que facilita la creación de *layouts* adaptativos, garantizando que los paneles desarrollados se adaptarán correctamente a diferentes tamaños de pantalla y dispositivos, mejorando con ello la experiencia del usuario. Junto con esta amplia gama de opciones de diseño predefinidas, es posible la personalización avanzada mediante la definición e incorporación de archivos **CSS** de estilos, con los que desarrolladores pueden modificar las clases predefinidas en **Flexdashboard** diseñando y ajustando la presentación de los elementos visuales a sus necesidades o su gustos específicos.

3.2.3. Otros paquetes y tecnologías complementarias

Una tecnología esencial para la creación de documentos dinámicos del ecosistema **R** es **knitr**²⁴ que ejecuta los bloques de código, extrae sus resultados y los muestra en el informe. Soporta diversos lenguajes como **Python**, **SQL** o **Bash**, proporcionando salidas en diferentes formatos como **HTML**, **PDF** o **Word** mediante el uso de diferentes motores de renderizado como **Latex** y **Pandoc**.

Nuestra aplicación va a extraer los datos de un repositorio **MySQL** para cuya conexión será necesario contar con los paquetes de **R** específicamente desarrollados para generar una lista de conexiones de acceso a la base de datos. Las operaciones de preparación, depuración y limpieza de los datos, en particular la recuperación de las respuestas originales que fueron codificadas en la inserción, van a requerir determinadas operaciones de inyección directa de código **SQL** o, de su ejecución contra la base de datos mediante *scripts* **R** o **Phyton**.

Para la manipulación de datos los paquetes **dplyr** y **tidyverse** ofrecen herramientas diseñadas para facilitar las operaciones avanzadas de transformación de datos, como la selección, filtrado, agrupación o resumen de datos, con una sintaxis intuitiva y legible.

Toda **app** de visualización de datos debe contar con el soporte de uno (o varios) paquetes gráficos avanzados y versátiles con funciones que faciliten la elaboración de un amplio abanico de representaciones gráficas. La librería **ggplot2** es una implementación de la gramática de gráficos en **R**, que permite crear una gama de visualizaciones avanzadas y personalizadas de forma intuitiva. Una de las principales ventajas de **ggplot2** es su capacidad para manejar datos complejos y grandes conjuntos de datos, proporcionando herramientas para facetas (subconjuntos de gráficos), transformaciones estadísticas, así como gráficos combinados y especializados como **ggpairs**. La librería es altamente extensible, permitiendo a los usuarios añadir sus propias funciones y temas, adaptando a sus gustos y necesidades, múltiples aspectos de la visualización.

La creación y renderizado con **Shiny** de diversos tipos de tablas estadísticas de tipo descriptivo o inferencial, tablas de contingencia o de asociación, de resolución de contrastes u obtención de tasas de prevalencia y [odds-ratio probability/razón de probabilidades](#), entre otros recursos, que nuestra aplicación va a proporcionar, se apoyará en funciones proporcionadas por varias de éstas librerías especializadas.

²⁴Se trata de un paquete de software para **R** que facilita la integración de código y texto en documentos reproducibles en un único documento.

3.3. Estructura de la Aplicación

3.3.1. Componentes del cuadro de mandos

La estructura de la aplicación se organiza entorno a varios componentes esenciales, cada uno cumpliendo una función específica en el funcionamiento y la presentación del cuadro de mandos, los cuales se detallan a continuación.

IU: constituye el componente visual y funcional de la aplicación. Se genera utilizando las bibliotecas **Shiny** y **Flexdashboard** de R. **Shiny** utiliza la función `FluidPage()` para crear una pantalla que se ajusta automáticamente a las dimensiones de la ventana del navegador y el diseño de la **IU** consiste en colocar los elementos en la `FluidPage`.

SERVER: responsable de ejecutar la lógica de la aplicación, lo que en el contexto de un cuadro de mandos, va a conllevar el procesamiento de los datos, las operaciones de cálculos estadísticos, y la elaboración de las representaciones gráficas y tablas apropiadas para ello. En un contexto de interactividad, el servidor deberá gestionar las peticiones del usuario, garantizando que las salidas se actualicen en conformidad, siendo el artífice de la experiencia interactiva en los elementos visuales del cuadro de mandos proporcionada por la aplicación.

FUNCIONES REACTIVAS: la reactividad es un aspecto central en la estructura de la aplicación, permitiendo que todos los elementos en los diferentes paneles del cuadro de mandos respondan automáticamente con la interacción del usuario. Las funcionalidades reactivas se implementan utilizando el paradigma de programación reactiva de **Shiny** que utiliza observadores, expresiones reactivas y funciones de filtrado dirigidas a capturar y comunicar las entradas de usuario a las que deben responder los elementos visuales del cuadro de mandos.

BASE DE DATOS: nuestra aplicación interactúa con un **SGBD MySQL** que constituye el repositorio centralizado del que se nutre, e interactúa mediante operaciones **SQL**.

VISUALIZACIÓN Y ANALÍTICA DE DATOS: el cuadro de mandos ofrece diversos tipos de elementos visuales de datos acordes a su tipología y al objetivo de la visualización contribuyendo a reflejar sus aspectos relevantes, de forma comprensible y que proporcione descubrimientos claros. Los elementos visuales de la analítica de datos dan soporte a la formulación de conclusiones precisas.

3.3.2. Secciones y elementos interactivos

Debido al elevado número de elementos visuales en el Sistema de Vigilancia, se va a dotar al cuadro de mandos de una estructura de paneles accesibles desde la barra de navegación. Los elementos del menú se han hecho corresponder con cada una de las secciones en el cuestionario, salvo la última que, por extensión y relevancia, se ha dividido en dos paneles. Dependiendo del número de elementos visuales a presentar, cada panel va a constar de una o varias subsecciones de contenido, accesibles desde el desplegable en el menú de la barra de navegación.

El elemento encargado de proporcionar la interactividad a la aplicación del Sistema de Vigilancia se va a conformar mediante varios selectores (*Select-Inputs* en terminología **shiny**). Por razones evidentes en un sistema de vigilancia, un primer selector va a ser el periodo temporal sobre el que se aplica la vigilancia. Se van a ofrecer otros selectores cuyo interés deriva de los hallazgos del estudio estadístico de la **EHD** y del interés de particularizar los elementos visuales a diversas características socio-demográficas de la mujer; los selectores ofrecidos van a proporcionar al usuario informes por edad, formación académica, situación laboral y situación personal.

En tanto que el Sistema de Vigilancia cuenta con un par de paneles de carácter estático, sirviéndonos de *Javascript*, se ha dotado de dinamismo a los propios selectores, permitiendo que el usuario pueda

mostrarlos u ocultarlos, en función de la tipología del panel.

La utilización del paquete `Plotly`²⁵ (o en su defecto la función `ggplotly()` de `ggplot2`) ha posibilitado proporcionar mayor interactividad a las representaciones gráficas mostrando información en el propio gráfico, con funcionalidades extras, como el *paneo*, *zoom*, captura de imagen, entre otras.

Igualmente cuentan con un comportamiento dinámico las diferentes tabulaciones que el Sistema de Vigilancia ofrece en la analítica de datos, que son generadas conformes a la elección de usuario. Éstas pueden presentar funcionalidades adicionales como la ordenación o la búsqueda. Finalmente, diversas conclusiones estadísticas formuladas en el Sistema de Vigilancia responden, igualmente de forma dinámica, a las peticiones de los usuarios.

3.3.3. Organización del código

La naturaleza reactiva `Shiny` ha abocado al desarrollo de nuestra aplicación a una estructura de código monolítica; no obstante se puede considerar una aproximación de dos capas: el `SGBD` y el servidor de la aplicación. La reactividad dificulta una modularización del código porque implica que las variables y funciones dependen de cambios en las entradas y sus elementos reactivos, lo que complica realizar una separación clara de componentes sin perder la interactividad y cohesión del sistema. Aunque la modularización es una práctica recomendada para mejorar la mantenibilidad y escalabilidad del código, en el contexto de aplicaciones `Shiny` la implementación modular suele resultar compleja²⁶.

Se ha codificado un archivo `Shiny R Markdown` que comienza con la especificación `YAML`:

```
title: "Proyecto AF y Salud"
author: "TFM Laura Muñoz"
output:
  flexdashboard::flex_dashboard:
    orientation: columns
    vertical_layout: scroll
    self_contained: true
    css: styles.css
    theme:
      version: 5
      bootswatch: slate
    source_code: embed
runtime: shiny
resource_files:
- config.yml
```

²⁵`Plotly` es una reputada librería de gráficos interactivos de código abierto cubriendo una amplia gama de casos de uso estadísticos, financieros, geográficos, científicos y tridimensionales que proporciona funciones para combinar representaciones para datos numéricos y datos categóricos

²⁶Se está trabajando, no obstante y, recientemente han aparecido paquetes en el ecosistema `Shiny` como `shinyModules` y `golem` dirigidos a facilitar la creación de aplicaciones modulares, con el peaje de perjudicar su sencillez

Y continua incorporando las dependencias de librerías utilizadas en el desarrollo:

```
library(flexdashboard)
library(readr)
library(ggplot2)
library(plotly)
library(treemap)
library(treemapify)
library(scico)
library(shiny)
library(shinyjs)
library(shinydashboard)
library(tidyverse)
library(knitr)
```

```
library(viridis)
library(DT)
library(ggmosaic)
library(GGally)
library(kableExtra)
library(ggthemes)
library(RColorBrewer)
library(rgdax)
library(magrittr)
library(sf)
library(dplyr)
library(rlang)
```

```
library(sjPlot)
library(summarytools)
library(kableExtra)
library(corrplot)
library(sjlabelled)
library(epiR)
library(DBI)
library(RMySQL)
library(odbc)
library(shinythemes)
library(sjmisc)
```

Se han definido funciones auxiliares para el cálculo de las distribuciones de frecuencias y de frecuencias condicionadas, que son requeridas de forma recurrente en diversas representaciones y tabulaciones de datos; y funciones para el cálculo de IC para la media y para la proporción. Éstas se utilizan ampliamente en la estimación de parámetros de la analítica de datos en la aplicación del Sistema de Vigilancia.

```
get_frequencies <- function(dataset, variable) {
  if (nrow(dataset()) == 0 || is.null(dataset()[[variable]])) {
    return(data.frame(Modalidad = character(), Porcentaje = numeric()))}
  data <- dataset() %>% filter(!is.na(.data[[variable]]))
  freq_table <- table(data[[variable]])
  prop_table <- prop.table(freq_table) * 100
  result <- data.frame(
    Modalidad = names(freq_table), Porcentaje = as.numeric(prop_table),
    stringsAsFactors = FALSE )
  result$Porcentaje <- as.numeric(sprintf("%.2f", result$Porcentaje))
  return(result)
}

get_conditional_frequencies <- function(dataset, variable1, variable2) {
  if (is.reactive(dataset)) {
    dataset <- dataset()}
  dataset <- dataset %>%
    filter(!is.na(.data[[variable1]]) & !is.na(.data[[variable2]]))
  if (nrow(dataset) == 0) {
    return(data.frame(!variable1 := character(),
      !!variable2 := character(), count = integer(),
      Porcentaje = numeric(), stringsAsFactors = FALSE))}
  freq_table <- table(dataset[[variable1]], dataset[[variable2]])
  prop_table <- prop.table(freq_table, margin = 2) * 100
  freq_df <- as.data.frame(freq_table)
  prop_df <- as.data.frame(prop_table)
  colnames(freq_df) <- c(variable1, variable2, "count")
  colnames(prop_df) <- c(variable1, variable2, "Porcentaje")
  result <- merge(freq_df, prop_df, by = c(variable1, variable2))
  result$Porcentaje <- round(as.numeric(result$Porcentaje), 2)
  result <- result[result$count > 0, ]
  return(result)

  calc_ci <- function(x, conf = 0.95) {
    se <- sd(x) / sqrt(length(x))
    alpha <- 1 - conf
    error_margin <- qt(1 - alpha / 2, df = length(x) - 1) * se
    mean_x <- mean(x)
    lower_bound <- mean_x - error_margin
    upper_bound <- mean_x + error_margin
    return(c(lower_bound, upper_bound))
  }
```

```

}
calc_ci_pi <- function(x, conf = 0.95) {
  p_hat <- mean(x)
  n <- length(x)
  alpha <- 1 - conf
  se <- sqrt(p_hat * (1 - p_hat) / n)
  z <- qnorm(1 - alpha / 2)
  error_margin <- z * se
  lower_bound <- p_hat - error_margin
  upper_bound <- p_hat + error_margin
  return(c(lower_bound, upper_bound))
}

```

Tras la creación de los elementos necesarios para la conexión y acceso al repositorio, la definición de los elementos interactivos en Shiny, y sus funciones de filtrado, que se describen en la siguiente sección, en el código comienza el proceso de creación de secciones y sus paneles, continua con la elaboración de elementos visuales y finaliza con su acomodación en la tarjeta adecuada del panel. El proceso es laborioso en el aspecto de generación de los elementos visuales, representaciones gráficas y tablas aunque sencillo en cuanto al diseño como se ilustra en el fragmento de código siguiente:

```

Socio-Demografia { data-navmenu="El Sistema"}
=====
## column1{data-width=280}
-----
### <span style="font-size: 80%;"> SD5- </span> Vive CON PAREJA
gaugeOutput("gauge_plot_pareja")

### <span style="font-size: 80%;"> SD1- </span> Distribución por EDAD
plotlyOutput("bar_chart_age")

## column2
-----
### <span style="font-size: 80%;"> SD2- </span> Distribución por SITUACIÓN LABORAL
plotlyOutput("treemap_laboral", height = "100%")

```

3.4. Desarrollo de la aplicación

En esta sección se detallan los dos pasos iniciales en el desarrollo de la aplicación, la preparación inicial del entorno de trabajo y la gestión y manipulación de los datos que va a utilizar.

3.4.1. Preparación del entorno de trabajo

El desarrollo de la aplicación comienza con la preparación del entorno de trabajo, la instalación de R, un lenguaje de programación y entorno de software ampliamente utilizado para el análisis de datos y la visualización gráfica. La versión más reciente de R puede descargarse desde el sitio oficial del Proyecto R en <https://cran.r-project.org/>. El entorno de desarrollo integrado (EDI) para R utilizado es RStudio que proporciona una IU amigable, facilitando la escritura y revisión del código, la visualización de variables y la gestión de proyectos.

RStudio ofrece varias funcionalidades avanzadas que lo hacen una herramienta esencial para el desarrollo con R como:

- **conectividad:** RStudio permite la conexión a una amplia variedad de bases de datos y fuentes de datos, como SQL, NoSQL, CSV, Excel, Google Sheets, entre otros. Esto facilita la importación y manipulación de datos de diversas fuentes,
- **consola interactiva:** permite ejecutar código línea por línea en una consola interactiva, lo cual es ideal para pruebas rápidas, exploración de datos y depuración de errores al mostrar los resultados de forma inmediata,
- **ejecución de *scripts*:** junto con la consola interactiva, RStudio permite la ejecución de *scripts* completos que permiten automatización de tareas y modularización de código. Éstos pueden ejecutarse de manera parcial (mediante puntos de parada de ejecución de hitos), permitiendo almacenar y reutilizar sus resultados,
- **renderización de documentos:** con el soporte de knitr y Rmarkdown, RStudio permite la creación y renderización de documentos dinámicos que combinan texto, código y resultados en un solo documento. Ello resulta especialmente útil para la generación de informes reproducibles y documentos interactivos como los cuadros de mando creados con Flexdashboard y Shiny,
- **diversos orígenes de datos:** RStudio facilita la importación de datos desde diversas fuentes y formatos, permitiendo a los usuarios trabajar tanto con datos locales, como en remoto.
- **trabajo en segundo plano:** permite la ejecución de tareas en segundo plano, lo que posibilita continuar trabajando mientras se ejecutan análisis complejos o durante el procesamiento de grandes volúmenes de datos.
- **integración con Git y GitHub:** RStudio ofrece integración nativa con sistemas de control de versiones como Git y plataformas de colaboración como GitHub que facilitan el seguimiento de cambios en el código, la colaboración y la gestión de versiones del proyecto. RStudio proporciona interfaces gráficas para realizar operaciones comunes de Git, como el *stage* y *unstaging* de cambios, la visualización del historial de *commits*, la resolución de conflictos de *merge*, y la comparación de diferencias entre versiones (*diffs*).

Se realizan las configuraciones iniciales en RStudio para optimizar el flujo de trabajo (configuración del directorio de trabajo predeterminado, ajuste de preferencias y comprobación sobre la actualización de componentes) y se procede a la instalación de los paquetes de soporte al desarrollo. Como en otros lenguajes, en R, los paquetes son colecciones de funciones, datos y código que extienden las capacidades del lenguaje base. Se ha hecho uso de numerosos paquetes con Flexdashboard, Shiny y Knitr como paquetes básicos mientras que se instalaran los paquetes requeridos para dar soporte a los procesos de desarrollo presentados.

3.4.2. Preprocesamiento, acceso y carga de datos

Los datos almacenados en nuestro repositorio, que han viajado codificados y cifrados por seguridad, son preprocesados desde la app en un proceso de decodificación que recupera las respuestas originales mediante una operación de *join* con las tablas de metadatos, como se muestra en el fragmento de código siguiente:

```

1 UPDATE responses r
2 LEFT JOIN responseoptions_metadata rom
3 ON r.question_code = rom.question_code
4 AND r.selected_options = rom.option_code_response
5 SET r.response_text = CASE
6     WHEN rom.option_value_response IS NOT NULL THEN rom.option_value_response
7     ELSE CAST(r.response AS VARCHAR(255))
8 END;
```

En la preparación de los datos se ha realizado una limpieza consistente en eliminar los registros en los que todos los campos son *NULL*. Para aprovechar al máximo los datos disponibles, se ha decidido mantener inicialmente los registros con alguna respuesta, realizando el filtrado de datos ausentes en la generación de elementos visuales. La depuración de datos se ha resuelto con la exploración de las modalidades en las variables categóricas, y el estudio de atípicos en las variables numéricas; este último nos ha permitido identificar (para verificar o descartar) los valores extremos en cada variable.

A partir de las variables originales se han definido nuevas variables que operan sobre aquéllas para codificar, evaluar y asignar los valores que las definen. Por ejemplo, la variable relevante del *Tipo de AF*, se ha obtenido mediante la aplicación de la tabla de equivalencia de *METs* de la actividad por la duración y frecuencia; la variable *IMC* se obtiene como $IMC = \frac{\text{Peso (kg)}}{\text{Altura (m)}^2}$; para el *Estilo de vida* se han definido categorías en base a parámetros en variables de los hábitos saludables y la combinación del tipo y tiempo de práctica de *AF*.

Tanto el preprocesamiento, la definición de variables y la recuperación de datos en el repositorio, se ha realizado definiendo y manteniendo un *pool* de conexiones al repositorio. Nos hemos servido de las funcionalidades de los paquetes *RMySQL*, *DBI* y *yaml* para definir una función de conexión con los parámetros de las configuraciones local y de producción requeridos en el despliegue *ShinyApps.io* recogidos en el *CONFIG.YML* como se ilustra en los fragmentos de código siguientes:

```

1 config <- yaml.load_file("config.yml")
2 env <- "default"
3 mysql_config <- config[[env]]$mysql
4 databaseName <- "laura_db"
5 table <- "responses"
6 loadData <- function() {
7   db <- dbConnect(MySQL(),
8     dbname = databaseName, host=mysql_config$host, port =mysql_config$port,
9     user = mysql_config$user, password = mysql_config$password)
10   query <- sprintf("SELECT EdadSD, FormacionSD, LaboraSD, PersonalSD, TiempoAF, TipoAF FROM %s",
11     ↪ table)
12   data <- dbGetQuery(db, query)
13   dbDisconnect(db)
14   data
15 }

```

```

1 CONFIG.YML
2 default:
3   dataconnection:
4     driver: 'xxxxxxxxxxxxxxxxxxxxxx'
5     server: 'yyyyyyyyyyyyyyyy'
6     uid: 'root'
7     port: 3306
8     database: 'laura_db'
9
10 shinyapps:
11   dataconnection:
12     driver: 'MySQL'
13     server: 'zzzzzzzzzzzzzzzz'
14     uid: 'root'
15     port: 3306
16     database: 'laura_db'

```

3.5. Diseño de la IU

Como se señala en la descripción, el diseño de la IU en Shiny se maneja a través de funciones de *layout* que organizan los elementos visuales facilitando la creación de *layouts* fluidos y *responsive*. El uso de las clases predeterminadas que ofrece **Flexdashboard** combinado con **Bootstrap** ha permitido diseñar en la IU una barra de navegación que acoge los menús de la cinco secciones proporcionando (mediante desplegaables) el acceso a sus paneles temáticos de una forma intuitiva para el usuario.

3.5.1. Lienzo del cuadro de mandos

Junto con la estructuración de diseño facilitada **Flexdashboard**, en el diseño del *layout* de nuestra aplicación se han utilizado las funciones `fluidPage()`²⁷ y `sidebarLayout()` para organizar de forma personalizada y estructurada los *inputs*, *outputs* y el código HTML de los elementos visuales generados en paneles.

De una parte, aprovechando la facilidad de acomodación de elementos en el panel ofrecida por **Flexdashboard** en el Sistema de Vigilancia se han diseñado un total de ocho paneles temáticos agrupados entorno a las cinco secciones, con los que se ha pretendido proporcionar una presentación pormenorizada de la información contenida en los datos del cuestionario. De otra, se ha utilizado `fluidPage()` combinado con diversas clases **Bootstrap** y estilos **CSS** propios para diseñar dos paneles adicionales cuyo propósito difiere, son paneles informativos: **Presentación** e **Informe Final** de resultados. A continuación se describe el proceso e implementación del diseño de los dos tipos de paneles.

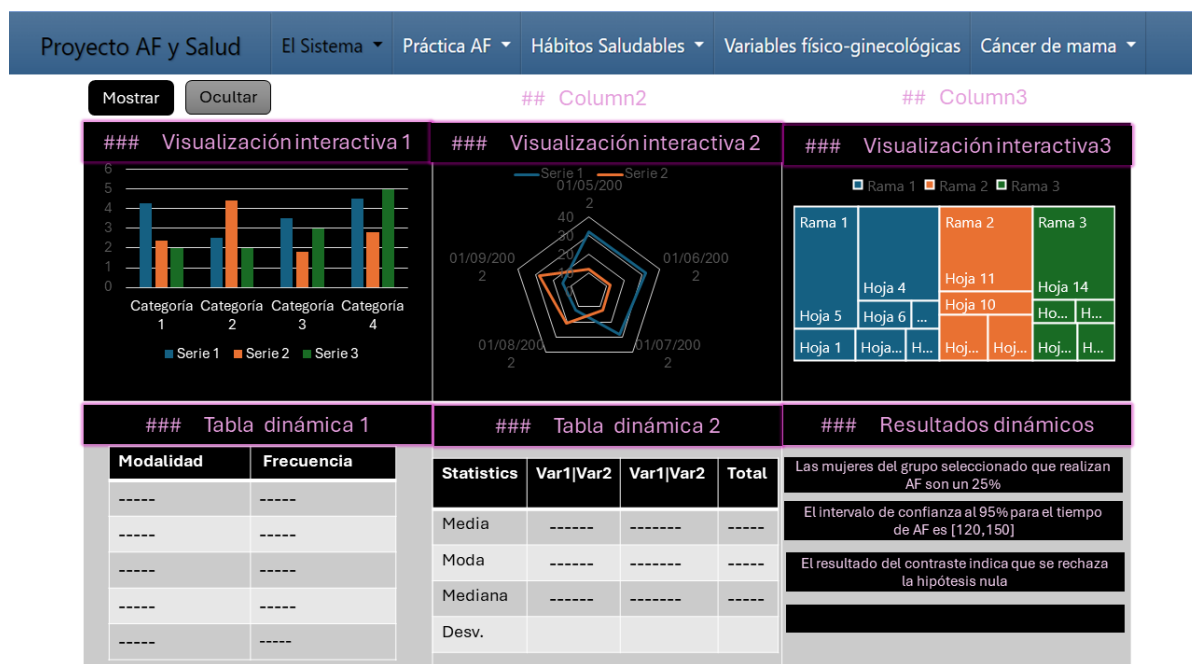


Figura 3.3: Esbozo del diseño de la barra de navegación y la la estructura de un panel de visualización en el *layout*

²⁷Permite crear una página con una disposición flexible que se ajusta al tamaño de la ventana del navegador

PANELES VISUALES INTERACTIVOS PREDETERMINADOS

En la **Figura 3.3** puede verse un boceto de la disposición de elementos visuales en un panel²⁸. La disposición diseñada, como cualquier otra, se alcanza con **Flexdashboard** que habilita la acomodación y estilizado de elementos en el tablero mediante un código programado para crear divisiones flexibles del *layout* con orientación por filas o columnas y tarjetas visuales que alojan a los diferentes elementos de una forma flexible y adaptada a su contenido.

Por ejemplo, la creación del panel titulado **Socio-Demografía** que colgará de la sección **El Sistema** en la barra de navegación, se realiza con el código inferior (de muestra). Éste código origina que los elementos visuales alojados en el panel se distribuyan en dos columnas, la primera con dos elementos visuales, la segunda con un único *plot*.

```
#Socio-Demografía { data-navmenu="El Sistema"}
=====
##column1{data-width=280}
-----
### Tamaño de la muestra
valueBoxOutput("women_box", width = "100%")

### SD5- Vive CON PAREJA
gaugeOutput("gauge_plot_pareja")

## column2{data-width=450}
-----
### SD2- Distribución por SITUACIÓN LABORAL
plotlyOutput("treemap_laboral", height = "100%")
```

PANELES VISUALES INTERACTIVOS PERSONALIZADOS

Fuera del entorno de diseño definido por **Flexdashboard**, se han diseñado los dos paneles de carácter informativo. La apariencia del panel de **Presentación** se muestra en la **Figura 3.4**, y para la implementación del diseño personalizado de elementos visuales en el **Informe final** de resultados mediante **FluidPage()** con clases **Bootstrap**, un fragmento de su código es el siguiente:

```
INFORME FINAL {data-navmenu="El Sistema" data-orientation=rows}
=====
fluidPage(
  div(class = "container",
    div(class = "row",
      tags$div(class = "degradado", style = " color: white; padding: 10px; text-align:center",
        "INFORME DEL ESTADO de la PRÁCTICA DE AF en la mujer en España en el periodo seleccionado"),
      tags$div(class = "panel-heading", style = "margin-top: 10px;", "PARTICIPACIÓN EN AF de la
        mujer"),
      tags$div(class = "resume-txt", style = "text-align: justify; font-size: 12pt;",
        "La participación en AF recreativa....."),
      div(class="row spaced-row",
        div(class = "col-md-6",
          style = "text-align: center;",
          HTML("<p style='text-align:center; font-size:normal; color: #337AB7;'><em>Figura 1:</em>
          Porcentaje de participación en AF por Edad</p>"), plotlyOutput("radialPlot_edad")
        )))
```

²⁸diseñado en *Power-Point*



Figura 3.4: Diseño y contenido del panel *Presentación* en el menú **El Sistema** en el cuadro de mandos del Sistema de Vigilancia

3.5.2. Diseño de elementos interactivos

Como se indica en la descripción, la interactividad de la aplicación consiste en la oferta de varios selectores, *inputs* en terminología *Shiny* situados en la IU. A nivel de diseño, los selectores ideados para la interactividad del cuadro de mandos se han alojado en un panel contenedor a medida, que *Flexdashboard* acomoda a la izquierda del *layout*. El panel de selectores se ha personalizado definiendo estilos propios para las clases definidas que han permitido ajustar el panel lateral al contenido, a la visibilidad de los elementos, al uso del *scroll* y su integración en el *layout* como se muestra en la **Figura 3.5**.

La definición de cada uno de los cinco selectores, junto con su incorporación en el panel de cada sección se realiza mediante el fragmento de código:

```
Inputs{.sidebar data-width=190}
=====
# Creación de selectores
### <span style="font-size:14pt;font-family:system-ui;color:white"> SELECCIONAR AÑO </span>
selectInput("anyo_selector", label = " ",
  choices = c("Todos años", levels(unique(responses_factors$AnyoEncuesta))))
### <span style="font-size:14pt;font-family:system-ui;color:white"> FILTRAR SEGÚN </span>
selectInput("edad_selector", "Edad",
  choices = c("Todas edades", levels(unique(AF_dataset$EdadSD))))
selectInput("formacion_selector", "Formación",
  choices = c("Todos niveles", levels(unique(AF_dataset$FormacionSD))))
selectInput("laboral_selector", "Situación Laboral",
  choices = c("Todas situaciones", levels(unique(AF_dataset$SituacionLaboralSD))))
selectInput("personal_selector", "Situación Personal",
  choices = c("Todas situaciones", levels(unique(AF_dataset$PersonalSD))))
```

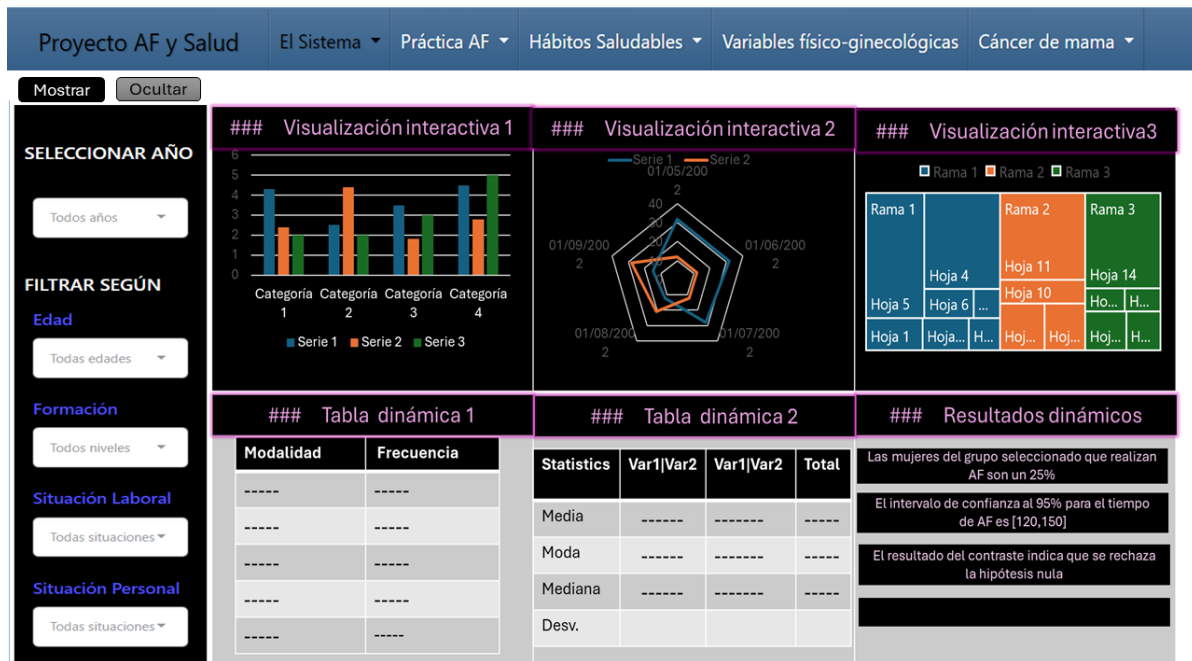


Figura 3.5: Esbozo del lienzo mostrando el panel de selectores en una sección del cuadro de mandos

3.6. Implementación de la lógica de la aplicación

En esta sección se expone la programación de la lógica del servidor, la elaboración de los cálculos y la generación de los elementos visuales. Y se ofrece la descripción de las funciones y variables reactivas definidas en la aplicación para dar soporte al comportamiento dinámico actuando (reaccionando) ante cambios en las entradas de usuario, garantizando que el servidor genere las salidas conformes.

3.6.1. Manejo de entradas, salidas y funciones reactivas

Es en el manejo de entradas y salidas dónde reside el fundamento del desarrollo de aplicaciones web interactivas con **Shiny**. Cada *input* en **Shiny** está asociado a un identificador único que es utilizado para referir a cada entrada particular de usuario; los cambios en ese *input* deben trasladarse al servidor que reaccionará de forma automática y eficiente a ellos. Las salidas en **Shiny** son los elementos visuales en la IU que muestran al usuario el contenido propio este tipo de aplicación. Al igual que los *inputs*, éstos también cuentan con identificadores únicos, y van a ser generados dinámicamente por el servidor a la reactividad.

La reactividad en **Shiny** se maneja mediante funciones y variables reactivas y observadores tales como `reactive()`, `observe()` y `observeEvent()` que permiten definir cómo deben actualizarse los *outputs* en respuesta a los cambios en los *inputs*. Si, por ejemplo, una representación gráfica debe actualizarse por la selección del usuario de un año, el uso de la función `reactive()` permite capturar la entrada y proporcionarla al servidor que elaborará los cálculos y elementos visuales de forma consecuente. Por su parte, los observadores se utilizan para ejecutar código en respuesta a los cambios en las entradas que generan una acción y no un valor directamente, lo cual no se va a dar en nuestro desarrollo. La implementación de la reactividad del cuadro de mandos hace uso de la función definida sobre un conjunto de datos que se obtiene en el fragmento de código siguiente:

```
filtered_data <- function(dataset) {
  data <- dataset
  if (input$edad_selector != "Todas edades") {
    data <- data[data$EdadSD == input$edad_selector, ]
  }
}
```

```

}
if (input$formacion_selector != "Todos niveles") {
  data <- data[data$FormacionSD == input$formacion_selector, ]
}
if (input$laboral_selector != "Todas situaciones") {
  data <- data[data$SituacionLaboralSD == input$laboral_selector, ]
}
if (input$anyo_selector != "Todos años") {
  data <- data[data$AnyoEncuesta == input$anyo_selector, ]
}
if (input$personal_selector != "Todas situaciones") {
  data <- data[data$PersonalSD == input$personal_selector, ]
}
data
}

```

La función actuará de forma reactiva generando una variable reactiva, mediante el fragmento de código siguiente (aplicado al `dataset=AF-dataset`). En ejecución, el uso de la variable reactiva asegura que el servidor actuará sobre ésta de forma conforme a la elección realizada por el usuario.

```

# Reactive AF_dataset
AF_filtered_dataset<-reactive({filtered_data(AF_dataset)})

```

3.6.2. Generación y renderizado de elementos visuales

La capacidad de generar elementos visuales para elaborar un cuadro de mandos es uno de los principales atractivos de **Shiny**. Para lograrlo, junto con los *inputs*, *outputs*, las funciones reactivas y observadoras, **Shiny** proporciona una serie de funciones que realizan el renderizado de los objetos, específicas para su tipología .

Se dispone de las funciones `renderPlot()` y `renderPlotly()` que intervienen en la definición de cada *output* gráfico en el servidor y que se renderiza en la **IU** esencialmente, mediante las **Shiny** funciones `plotOutput()` y `plotlyOutput()` que actúan mostrando (en su correspondiente ubicación de diseño) el gráfico dinámico generado.

Como complemento a la visualización gráfica, en el ámbito de la analítica de datos, el Sistema de Vigilancia prevé la incorporación en paneles temáticos y en el informe final de diversas tabulaciones de tipología y contenido variado. **Shiny** proporciona funciones para el renderizado de tablas interactivas mediante los paquetes **DT** o **Reactable**. Una tabla dinámica se define como *output* en el servidor mediante funciones específicas de **Shiny** como `renderDataTable()`, o `renderReactable()` y se renderiza en su ubicación en la **IU** mediante las funciones `dataTableOutput()`, o `reactableOutput()`.

La presentación de elementos de texto suele resultar un aspecto complementario en el desarrollo de aplicaciones interactivas con **Shiny** en lo que aporta a la contextualización, la interpretación dinámica de resultados, la formulación de conclusiones, e incluso la presentación de modelos (sus ecuaciones) estimados de forma dinámica para la elección de usuario. Nuestro Sistema de Vigilancia ofrece la formulación de diversas conclusiones de carácter dinámico.

Shiny posibilita la generación de texto dinámico mediante el uso de las funciones: `renderText()`, `renderUI()` y `renderPrint()`, cada una adecuada para diferentes tipos de contenido textual. La función `renderText()` se utiliza para generar textos alternativos, o incluyendo valores numéricos

dinámicos, que son actualizados en respuesta a cambios en los *inputs* y se muestran en la IU mediante el uso de la función `textOutput()`; `renderPrint()` resulta un opción especialmente útil para resultados de modelos, resúmenes de datos u otras salidas textuales más complejas; `renderUI()` es extremadamente versátil y devuelve un objeto de tipo *output* que puede ser prácticamente cualquier elemento generado mediante funciones Shiny. Su uso posibilita la generación de elementos dinámicos genéricos que incluyen combinaciones de texto con otros elementos, como formularios o informes y que son renderizados en la IU mediante la Shiny función `uiOutput()`.

3.6.3. Robustez de los elementos visuales

En el Sistema de Vigilancia se ha abordado un aspecto que es crítico en la generación dinámica de elementos visuales en contextos interactivos basados en selectores, en tanto que conllevan un filtrado selectivo cuya aplicación sucesiva es, potencialmente degenerativa. Por ejemplo, si en nuestra aplicación se desea obtener una visualización del *Tipo de práctica de AF durante el tratamiento del cáncer de mama* en mujeres empresarias de menos de 30 años, es bastante probable que el *dataset* resultante sea vacío, provocando un error en ejecución para la petición.

Para garantizar la robustez global del cuadro de mandos en relación a todos sus elementos visuales, es imprescindible acometer la comprobación en la definición de cada salida. La solución no ha resultado evidente, dado que, como se ha señalado, el renderizado de objetos se realiza mediante funciones específicas para su tipología; siendo ésta el resultado de un proceso de filtrado que es dinámico, por ello es a priori desconocida. La solución se ha alcanzado modificando la definición de cada objeto mediante el uso de la función de renderizado más genérica: `renderUI()`. Esta función, posibilita dotar de una definición inicial del objeto, dentro de la cual, y en función del resultado de la comprobación, se establecerá su definición definitiva mediante la función de renderizado apropiada. Dado que la Shiny función `uiOutput()` permite mostrar el objeto generado con independencia de su tipología, ha sido posible alcanzar la robustez para nuestra aplicación.

3.7. Despliegue y ejemplos de implementación

Una vez completamente desarrollada y testada su funcionalidad en nuestro entorno local, para el despliegue de la aplicación Web se va utilizar la tecnología proporcionada por ShinyApps.io. ShinyApps.io es un servicio en la nube ofrecido por RStudio que permite desplegar y alojar las aplicaciones Shiny directamente desde el entorno de RStudio en un proceso relativamente sencillo. El despliegue de una aplicación Shiny con ShinyApps.io supone una solución efectiva para compartir aplicaciones interactivas en la web, sin necesidad de gestionar servidores propios y con el atractivo de ofrecer numerosas ventajas. El servicio maneja automáticamente la escalabilidad, asegurando que la aplicación admite múltiples usuarios, sin perjudicar su rendimiento. ShinyApps.io ofrece capacidades de monitorización y administración, proporciona estadísticas de uso, registro de histórico de errores, manteniendo y ofreciendo la gestión de versiones. La aplicación Web del Sistema de Vigilancia está desplegada y es accesible en el enlace <https://afysaludtfmlaura.shinyapps.io/AFySaludOptimizedv2/>.

Se han planteado una serie de casos de uso con el objeto de mostrar la apariencia final alcanzada para el Sistema de Vigilancia, e ilustrar brevemente su implementación. En cada caso se presentará la imagen del panel que proporciona la respuesta y se mostrará y explicará el fragmento de código utilizado en su implementación.

- **Caso de uso 1:** se desea conocer qué porcentaje de mujeres ha practicado AF recreativa en 2024 entre mujeres con estudios de máster o doctorado que viven en pareja con hijos menores, y cuyo trabajo o estudios no están vinculados a la AF

- **Caso de uso 2:** se quiere conocer entre qué valores del tiempo semanal de *AFV* se encuentra el 50 % central de las mujeres de idénticas características

Navegando en el Sistema de Vigilancia en el menú **Práctica AF** se accede al panel *Visualización I* que se presenta en la **Figura 3.6**. Mostrando el panel de selectores para elegir las opciones de éstos en cuanto a las características deseadas, automáticamente el panel se actualiza. Un movimiento del ratón en el gráfico interactivo rotulado de forma concisa, ofrece la respuesta.

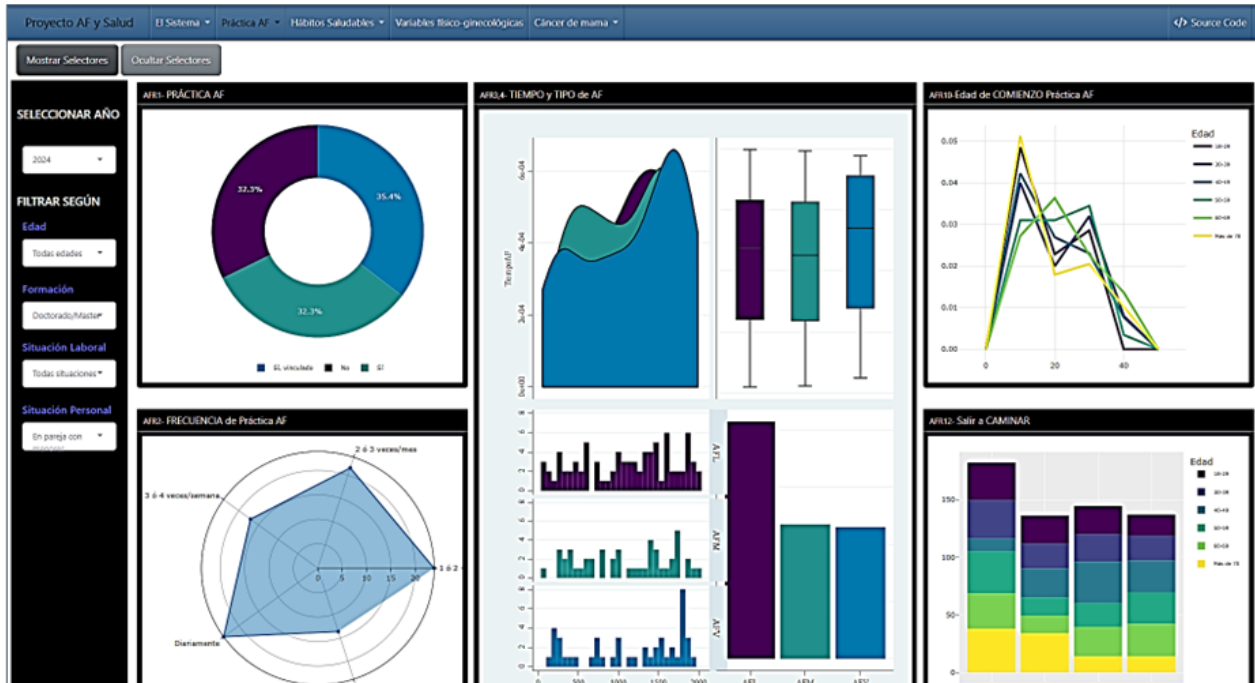


Figura 3.6: Sección **Práctica AF**; *Visualización I* en el cuadro de mandos del Sistema de Vigilancia.

La creación del panel y la generación de las representaciones gráficas se realiza mediante los fragmentos de código de implementación que se muestran y explican a continuación. El primer fragmento genera de forma dinámica las cuatro representaciones gráficas en el panel de tipología variada elaboradas con funciones de los paquetes *Plotly*, *ggplot2* y *GGgally*. Sirva éste para ilustrar el incremento notable en líneas de código que conlleva el garantizar la robustez de todos los elementos visuales.

```
# Output plots in page Visualization AF
```

```
output$pie_chart_ui<-renderUI({
  data<-get_frequencies(AF_filtered_dataset,"PracticaAF")
  if (nrow(data) == 0) {
    return(HTML("<p>No hay datos disponibles para las categorías seleccionadas.</p>"))
  } else{plotlyOutput("pie_chart_render")}} )
output$pie_chart_render <- renderPlotly({
  data <- get_frequencies(AF_filtered_dataset, "PracticaAF")
  plot_ly(data, labels = ~Modalidad, values = ~Porcentaje, type = "pie", hole = 0.5,
    marker = list(colors = paleta_viridis_laura)) %>%
  layout(title = "",
    xaxis = list(showgrid = FALSE, zeroline = FALSE, showticklabels = FALSE),
    yaxis = list(showgrid = FALSE, zeroline = FALSE, showticklabels = FALSE),
    legend = list(orientation = "h", x = 0.5, y = -0.1, xanchor = 'center', traceorder = 'normal',
      font = list(size = 10))))
output$radial_chart_ui<-renderUI({
  data<-get_frequencies(AF_filtered_dataset,"FrecuenciaAF")
  if (nrow(data) == 0) {
```

```

    return(HTML("<p>No hay datos disponibles para las categorías seleccionadas.</p>"))}
    else{ plotlyOutput("radial_chart_render")}}

output$radial_chart_render <- renderPlotly({
  data <- get_frecuencias(AF_filtered_dataset,"FrecuenciaAF")
  max_y <- max(as.numeric(gsub("%", "", data$Porcentaje)))
  plot_ly(data, type = "scatterpolar", fill = "toself",
    r = ~as.numeric(gsub("%", "", Porcentaje)), theta = ~Modalidad,
    line = list(color = "rgb(31, 119, 180)", marker = list(color = "rgb(31, 119, 180)")) %>%
    layout(polar = list(radialaxis = list(visible = TRUE, range = c(0, max_y))),
      showlegend = FALSE)})

output$histoplot_ui<-renderUI({
  data <- AF_filtered_dataset()
  data <- data[!is.na(data$EdadComienzoAF), ]
  if (nrow(data) == 0) {
    return(HTML("<p>No hay datos disponibles para las categorías seleccionadas.</p>"))
  } else{ plotlyOutput("histoplot_render")}})

output$histoplot_render <- renderPlotly({
  data <- AF_filtered_dataset()
  data <- data[!is.na(data$EdadComienzoAF), ]
  data$EdadSD <- as.factor(data$EdadSD)
  histoplot <- ggplot(data, aes(x = EdadComienzoAF, y =density, group = EdadSD, color = EdadSD)) +
    geom_freqpoly(binwidth = 10, size = 1) +
    scale_color_viridis_d(name = "Edad") + # Cambiar el nombre de la leyenda aquí
    labs(x = "", y = "") + theme_minimal() +
    theme(legend.position = "right", legend.text = element_text(size = 6),
      legend.key.width = unit(0.3, "cm"),
      legend.spacing.x = unit(0.1, "cm")) +
    guides(fill = guide_legend(nrow = 1, title.position = "right", title.hjust = 0.5))
  ggplotly(histoplot)})

output$caminarplot_ui<-renderUI({
  data<-get_conditional_frequencies(AF_filtered_dataset,"CaminarAF", "EdadSD")
  if (nrow(data) == 0) {
    return(HTML("<p>No hay datos disponibles para las categorías seleccionadas.</p>"))}
  else{plotlyOutput("caminarplot_render")}} )

output$caminarplot_render <- renderPlotly({
  data <- get_conditional_frequencies(AF_filtered_dataset, "CaminarAF", "EdadSD")
  caminarplot <- ggplot(data, aes(x = CaminarAF, y = Porcentaje, fill = EdadSD)) +
    geom_bar(stat = "identity", position = "stack") +
    labs(x = "", y = "", fill = "Edad") +
    scale_fill_viridis_d() +
    theme(legend.position = "right", legend.text = element_text(size = 6),
      legend.key.width = unit(0.3, "cm"),
      legend.spacing.x = unit(0.1, "cm")) +
    guides(fill = guide_legend(nrow = 1, title.position = "right", title.hjust = 0.5))
  return(ggplotly(caminarplot)})

output$ggpairs_plot_ui<-renderUI({
  data<- AF_filtered_dataset()
  data<- subset(data, !is.na(data$TiempoAF) & !is.na(data$TipoAF))
  if(nrow(data)==0){
    return(HTML("<p>No hay datos disponibles para las categorías seleccionadas.</p>"))}
  else{ plotlyOutput("ggpairs_plot_render")}})

output$ggpairs_plot_render<-renderPlotly({
  data <- AF_filtered_dataset()
  data <- subset(data, !is.na(data$TiempoAF) & !is.na(data$TipoAF))
  data_to_plot <- subset(data, select = c(TiempoAF, TipoAF))
  ggpairs_plot <- ggpairs(data_to_plot,

```

```

columns = c("TiempoAF", "TipoAF"), aes(color = TipoAF),
lower = list(continuous = wrap("density", bins = 10)), upper = list(continuous = "boxplot"),
palette = "viridis" +
  scale_fill_manual(values = paleta_viridis_laura) +
  theme( legend.position = "top") + theme_stata() +
  scale_color_manual(values = c("grey"))
return(ggplotly(ggpairs_plot)))

```

La disposición de los elementos generados en el panel se realiza con escaso esfuerzo de código aprovechando las clases predeterminadas por Flexdashboard como se muestra en el fragmento de código siguiente:

```

Visualización I {data-navmenu="Práctica AF"}
=====
column1{data-width=300}
-----
TARJETA 1 ### <span style="font-size: 80%;"> AFR1- </span> PRÁCTICA AF
uiOutput("pie_chart_ui")

TARJETA 2 ### <span style="font-size: 80%;"> AFR2- </span> FRECUENCIA de Práctica AF
uiOutput("radial_chart_ui")

column2{data-width=400}
-----
TARJETA 3 ### <span style="font-size: 80%;"> AFR3,4- </span> TIEMPO y TIPO de AF
uiOutput("ggpairs_plot_ui")

column3{data-width=300}
-----
TARJETA 4 ### <span style="font-size: 80%;"> AFR10- </span> Edad de COMIENZO Práctica AF
uiOutput("histoplot_ui")

TARJETA 5 ### <span style="font-size: 80%;"> AFR12- </span> Salir a CAMINAR
uiOutput("caminarplot_ui")

```

- **Caso de uso 3:** determinaremos si se puede concluir con un nivel de confianza del 95 % que las mujeres de menos de 30 años en España en 2024 alcanzan la recomendaciones de la [OMS](#) de práctica semanal de [AFV](#), o en su defecto de [AFM](#). Sabiendo que los trabajos de investigación concluyen que se necesitan al menos 300 minutos semanales de [AFV](#) para conseguir su beneficio protector para en el riesgo de [BC](#), concluiremos si el grupo de menor edad se beneficiaría de éste (supuesto mantuviera el nivel actual)
- **Caso de uso 4:** averiguaremos entre qué valores del tiempo semanal de [AFV](#) se encuentra el 50 % central de las mujeres de idénticas características

La formulación en términos de obtención de conclusiones del caso de uso debería llevar al usuario de forma intuitiva al panel *Informe Final* del Sistema de Vigilancia. Navegando en el panel que se muestra en la **Figura 3.7**, se accede a las tablas en las que se han elaborado los intervalos de confianza al 95 % para el tiempo de práctica de cada tipo de [AF](#) por edades. Éstas van a proporcionar la respuesta en la formulación de las conclusiones del caso de uso.

El diseño del panel del **Informe Final** de resultados del Sistema de Vigilancia se ha ilustrado en la **sección 5.3** de este capítulo. La generación dinámica y robusta de las tablas de [IC](#) y la representación de los gráficos *box-plot* para el tiempo de práctica de [AF](#) por tipo y edad se implementa mediante el fragmento de código siguiente:

```

filtered_dataset <- AF_dataset %>%
  filter(!is.na(PracticaAF_cat)) %>%
  select(EdadSD, PracticaAF_cat)
output$radialPlot_edad <- renderPlotly({
  grouped_data<-get_conditional_frequencies(filtered_dataset,"EdadSD","PracticaAF_cat")
  grouped_data$Porcentaje <- as.numeric(sub("%", "", grouped_data$Porcentaje))
  plot_ly(data = grouped_data, type = 'scatterpolar', mode = 'lines',
    r = ~Porcentaje, theta = ~EdadSD, fill = 'toself',
    marker = list(color = 'viridis')) %>%
  layout(polar = list(radialaxis = list(visible = TRUE,
    range = c(0, max(grouped_data$Porcentaje))), tickformat = "%"),
    showlegend = FALSE)})
output$tablep_edad <- renderTable({
  filtered_dataset <- AF_dataset %>%
  filter(!is.na(PracticaAF_cat))%>%
  mutate(PracticaAF_bin = ifelse(PracticaAF_cat == "Si", 1, 0))%>%
  dplyr::group_by(EDAD = sjlabelled::as_label(EdadSD)) %>%
  dplyr::summarise(
    Ext.Inf = calc_ci_pi(PracticaAF_bin)[1]*100,
    Ext.Sup = calc_ci_pi(PracticaAF_bin)[2]*100
  ) %>%
  mutate(
    Ext.Inf = paste0(round(Ext.Inf, 2), "%"),
    Ext.Sup = paste0(round(Ext.Sup, 2), "%"))},digits=2)

```

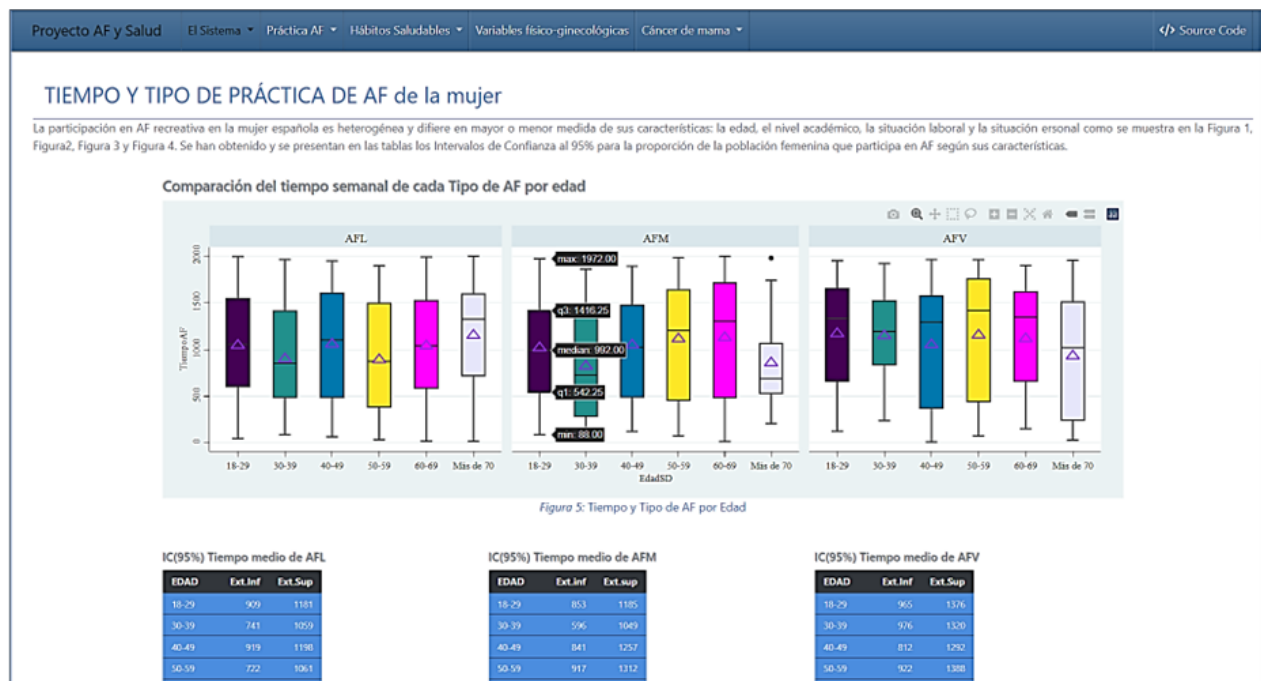


Figura 3.7: Sección panel INFORME FINAL en el cuadro de mandos del Sistema de Vigilancia

El código completo de implementación de la aplicación Web del Sistema de Vigilancia esta disponible en el enlace <https://github.com/LAURAMUNOZGARATACHEA/SistemaVigilancia/>

Capítulo 4

Conclusiones y futuras líneas de trabajo

En este proyecto se ha desarrollado una plataforma compacta, versátil y de utilidad para la obtención de datos, su visualización y analítica en contextos en los que se requiera disponer de un almacenamiento propio. Aunque en el proyecto se ha dirigido al estudio de la práctica de AF, los hábitos saludables y el riesgo de BC en la mujer, su versatilidad reside en que podría fácilmente adaptarse a otros estudios de investigación por encuesta, ya sea en salud u otros ámbitos que precisen la recopilación de datos de carácter sensible.

La herramienta desarrollada del Sistema de Vigilancia del estado de la práctica de AF presenta una utilidad natural como instrumento de soporte tecnológico a la recogida de datos en las unidades muestrales y la elaboración automatizada de informes de resultados, en la operación periódica de la EHD²⁹.

Unido al interés de carácter preventivo de los resultados que ofrecen rangos de estimación de diversos parámetros del estado de la práctica de AF con validez para la población general y particular (de diferentes características), la prescripción oncológica de AF y su estado entre mujeres supervivientes de BC puede resultar de utilidad para profesionales e investigadores en medicina deportiva.

Finalmente, un uso mantenido en el tiempo de nuestra aplicación de vigilancia de la AF brindaría la oportunidad de realizar estudios sobre su evolución permitiendo, en particular, evaluar la eficacia de medidas de promoción de la AF adoptadas por el sistema nacional de salud, o los sistemas de salud autonómicos. En este sentido, por ejemplo, es pionera y prometedora la iniciativa de la Comunidad de Madrid en el ámbito de la salud pública permitiendo a los médicos prescribir la práctica de AF como parte del tratamiento integral del paciente y ofreciendo programas estructurados de AF.

El desarrollo de la herramienta informática de *software* libre y multi-plataforma implementado en este proyecto, que se sitúa en el ámbito de la Ciencia de Datos, se ha sustentado en los conocimientos y habilidades tecnológicas adquiridos en las asignaturas del grado y máster. Junto con diversas herramientas conocidas de gestión de Bases de Datos, de diseño, arquitectura y desarrollo de Sistemas de Información Web, o las adquiridas en el Manejo de grandes Volúmenes de Datos, ha resultado necesaria la adquisición de nuevas herramientas y tecnologías, ampliando así el acervo de conocimientos proporcionados por el máster. Para mayor detalle véase en **Anexo G**.

El planteamiento, la orientación y los objetivos auspiciados y diseñados por las directoras de este trabajo, supusieron desde el primer momento un reto profesional muy estimulante. La consecución de los objetivos planteados en la defensa de este trabajo, con la que culmina este apasionante reto, ha supuesto una experiencia enormemente enriquecedora a nivel profesional y muy satisfactoria a nivel personal.

En la formulación de los objetivos en la memoria, se expresa claramente la intención de dar continuidad a la línea de investigación iniciada con este proyecto, con el objetivo de proporcionar soporte técnico a diversas direcciones de trabajo:

²⁹cabe señalar, cómo hasta la fecha se realiza mediante cuestionarios tradicionales en papel

1. Se dará continuidad a la línea principal de obtención de estimaciones precisas del tamaño del efecto de la práctica de AF y los hábitos saludables en la disminución del riesgo de BC mediante la aplicación de herramientas de modelización estadística³⁰. Cuando se disponga de una muestra suficientemente amplia³¹, la modelización permitirá contar además con estimaciones de mayor precisión, respecto de los estudios clínicos actuales, del efecto de las variables ginecológicas y de la presencia de antecedentes familiares de cáncer. Todo ello nos permitirá obtener predicciones rigurosas y fiables para el riesgo de BC en la mujer en función de sus covariables físicas, ginecológicas, familiares y de su estilo de vida.

Para ello, desde que se cuente con el dictamen favorable del CEICA se pondrá en marcha un plan de difusión del cuestionario con el que se pueda alcanzar una muestra amplia y representativa. Con el fin de lograr una elevada participación se prevé buscar la colaboración por parte de la organización de eventos multitudinarios dirigidos a mujeres, como la carrera de la mujer, ya que sería posible enlazar al cuestionario en la inscripción de participantes en las carreras que se organizan a lo largo y ancho de la geografía española.

Se pretende con todo ello contribuir a una mayor concienciación de la mujer que propicie la adopción de la práctica de AF y los hábitos saludables como estilo de vida.

2. La obtención e incorporación de medidas objetivas de la práctica de AF basadas en sensores con el fin de mejorar la precisión en la medición, tanto de la intensidad como de su duración. La incorporación creciente de diferentes tipos de sensores en dispositivos portátiles como móviles, relojes inteligentes o bandas de actividad proporciona datos de forma continua que podrían ser utilizados para realizar un seguimiento de la AF, detectar patrones de actividad o inactividad, contribuyendo al objetivo de la evaluación del impacto de la AF en la salud de la mujer, y en particular en el riesgo de BC.
3. Se quiere explorar, por la versatilidad de la herramienta desarrollada, otros usos del Sistema de Vigilancia de la AF para contextos de patologías físicas, como Enfermedades Cardiovasculares (ECV), hipertensión (HTA) o la diabetes, y enfermedades y problemas de salud mental como la depresión y el riesgo de suicidio, entre otras. En todos estos contextos se adaptará la herramienta manteniendo nuestro objetivo de alcanzar estimaciones precisas de la disminución en el riesgo de la patología asociado a la práctica de AF.

³⁰El modelo de *hazard proportional* de Cox, un modelo de regresión logística, caracterizado porque los efectos relativos de las variables explicativas (covariables) en el riesgo de mantienen constantes en el tiempo, ha sido ampliamente utilizado y testado en los trabajos

³¹Para estimar, por ejemplo, el tiempo medio de práctica de AF en la población femenina en España con un margen de error de 10 minutos y un nivel de confianza del 95 % el tamaño muestral necesario es de 1083 mujeres

Bibliografía

- [1] Ainsworth B.E., Haskell W.L., Whitt M.C., Irwin M.L., Swartz A.M., Strath S.J., O'Brien W.L., Bassett D.R., Schmitz K.H., Emplaincourt P.O., Jacobs D.R., and Jr Leon A.S. Compendium of physical activities: a second update of activity codes and met intensities. *Med Sci Sports Exerc*, 43(8):1575–81, 2011.
- [2] Serra M. C., Ryan A. S., Ortmeyer H. K., Addison O., and Goldberg A. P. Resistance training reduces inflammation and fatigue and improves physical function in older breast cancer survivor. *Menopause*, 25:211–216, 2018.
- [3] Dieli-Conwright C.M., Courneya K.S., Demark-Wahnefried W., Sami N., Lee K., and Sweeney F. C. Aerobic and resistance exercise improves physical fitness, bone health, and quality of life in overweight and obese breast cancer survivors: a randomized controlled trial. *Breast Cancer Res*, 20:124–133, 2018.
- [4] Dieli-Conwright C.M., Nelson R.A., and Simon M.S. Cardiometabolic risk factors, physical activity, and postmenopausal breast cancer mortality: results from the women's health initiative. *BMC Women's Health*, 22, 2022.
- [5] Panchik D., Masco S., Zinnikas P., Hillriegel B., Lauder T., and Suttman E. Effect of exercise on breast cancer-related lymphedema: what the lymphatic surgeon needs to know. *J. Reconstr. Microsurg*, 35:37–45, 2019.
- [6] McTiernan F., Christine M., Katzmarzyk P.T., Powell K.E., and Macko R. Physical activity in cancer prevention and survival: A systematic review. *Medicine and Science in Sports and Exercise*, 51(6):1252–1261, 2019.
- [7] Thomas G.A., Cartmel B., Harrigan M., Fiellin M., Capozza S., and Zhou Y. The effect of exercise on body composition and bone mineral density in breast cancer survivors taking aromatase inhibitors. *Obesity*, 25:346–351, 2017.
- [8] E. Gamma. *Patrones de diseño: elementos de software orientado a objetos reutilizable*. Addison-Wesley professional computing series. Pearson Educación, 2002.
- [9] Ilir H., Fitim S., Lot H., Midhet N., Buteau M.A., Krenare G., Mary D., and Chamberlin J. Breast cancer and lifestyle factors: Umbrella review. *Hematology/Oncology Clinics of North America*, 38(1):137–170, 2024.
- [10] Lee J. and Lee M.G. Effects of exercise interventions on breast cancer patients during adjuvant therapy: A systematic review and meta-analysis of randomized controlled trials. *Cancer Nurs*, 43(2):115–125, 2022.
- [11] Campbell K., Winters-Stone K., Wiskemman J., Mayo A., Schwartz A., Courneya K., Zucker D., Matthews C., Ligibel J., Gerber L., Morris G., Patel A., Hue T., Perna F., and Schmitz K. Exercise guidelines for survivors of cancer: consensus statement from an international multidisciplinary round table. *Medicine and Science in Sports and Exercise*, 51(11), 2019.

- [12] Garatachea N., Torres L., and González J. Physical activity and energy expenditure measurements using accelerometers in older adults. *Nutr. Hosp.*, 25(2):224–230, 2010.
- [13] Gebruers N., Camberlin M., Theunissen F., Tjalma W., Verbelen H., Van Soom T., and Van Breda E. The effect of training interventions on physical performance, quality of life, and fatigue in patients receiving breast cancer treatment: a systematic review. *Support Care Cancer*, 27(1):109–122, 2019.
- [14] Cešeiko R., Eglitis J., Srebnis A., Timofejevs M., Purmalis E., and Erts R. The impact of maximal strength training on quality of life among women with breast cancer undergoing treatment. *Exp. Oncol.*, 41:166–122, 2019.
- [15] Paulo T. R., Rossi F. E., Viesel J., Tosello G.T., Seidinger S. C., and Simões R. R. The impact of an exercise program on quality of life in older breast cancer survivors undergoing aromatase inhibitor therapy: a randomized controlled trial. *Health Qual. Life Outcomes*, 17, 2017.
- [16] Ficarra S., Thomas E., Bianco A., Gentile A., Thaller P., Grassadonio F., Papakonstantinou S., Schulz T., Olson N., Martin A., Wagner C., Nordström A., and Hofmann H. Impact of exercise interventions on physical fitness in breast cancer patients and survivors: a systematic review. *Breast Cancer*, 29(3):402–418, 2022.
- [17] Santagnello S., Martins F., Oliveira G., Sousa J., Nomelini R., and Murta E. Improvements in muscle strength, power, and size and self-reported fatigue as mediators of the effect of resistance exercise on physical performance breast cancer survivor women: a randomized controlled trial. *Support. Care Cancer*, 28:6075–6084, 2020.
- [18] Zaidi S., Hussain S., Verma S., Vegar Z., Khan A., and Nazir S. Efficacy of complementary therapies in the quality of life of breast cancer survivors. *Front. Oncol.*, 7:326, 2018.
- [19] Odyne T., Briskin Y., and Todorova V. Effects of different exercise interventions on quality of life in breast cancer patients: a randomized controlled trial. *Integr Cancer Ther.*, 18, 2019.
- [20] Möller U.O., Beck I., Rydén L., and Malmström M. A comprehensive approach to rehabilitation interventions following breast cancer treatment – a systematic review of systematic reviews. *BMC Cancer*, 19:472–491, 2019.
- [21] Liu Y.C., Hung T.T., Konara S.P., Wang C.J., and Lin M.F. Beneficial exercises for cancer-related fatigue among women with breast cancer: A systematic review and network meta-analysis. *Cancers*, 15, 2023.
- [22] Ma Z., Yao S., and Shi Y. Effect of aerobic exercise on cardiotoxic outcomes in women with breast cancer undergoing anthracycline or trastuzumab treatment: a systematic review and meta-analysis. *Support Care Cancer*, 30:10323–10334, 2022.

Acrónimos

- AF** Actividad Física. [1–22](#), [24](#), [40](#), [46](#), [47](#), [49](#), [51](#), [52](#), [58–62](#), [66–69](#), [73](#), [c](#), [I](#)
- AFL** AF Leve o de baja intensidad. [4–7](#), [12](#)
- AFM** AF de intensidad Moderada. [4–6](#), [8](#), [9](#), [11–14](#), [49](#)
- AFMV** AF de intensidad de Moderada a Vigorosa. [5–7](#), [11](#), [12](#)
- AFV** AF Vigorosa. [4–9](#), [11–14](#), [47](#), [49](#)
- AMSTAR** *A Measurement Tool to Assess Systematic Reviews*. [64](#)
- BC** *Breast Cancer*. [10](#), [11](#), [13](#), [19–21](#), [24](#), [49](#), [51](#), [52](#), [c](#)
- CEICA** Comité de Ética de la Investigación en la Comunidad autónoma de Aragón. [52](#), [70](#), [71](#), [II](#)
- CPA** *Compendium of Physical Activities*. [4](#)
- CRF** Función, Aptitud o capacidad Cardiorespiratoria. [3](#), [8](#), [10](#), [12](#)
- CVRS** Calidad de Vida Relativa a la Salud. [9](#), [12](#), [13](#)
- ECV** Enfermedades Cardiovasculares. [5](#), [8–10](#), [60](#)
- ESE** Encuesta Europea de Salud en España. [20](#), [21](#)
- EHD** Encuesta de Hábitos Deportivos. [13](#), [15](#), [18](#), [22–24](#), [35](#), [51](#), [66](#), [I](#), [II](#)
- ENS** Encuesta Nacional de Salud. [20](#)
- ENT** Enfermedades No Transmisibles. [3](#)
- GAPPA** *Global Action Plan for Physical Activity*. [3](#)
- GRADE** *Grading of Recommendations Assessment, Development, and Assessment*. [65](#)
- HHS** *Health and Human Services Department*. [1](#), [60](#)
- HIIT** *High Intensity Interval Training*. [4](#), [12](#)
- HTA** Hipertensión Arterial. [3](#), [8–10](#)
- IC** Intervalo de Confianza. [11](#), [13](#), [15](#), [16](#), [18](#), [19](#), [37](#), [49](#), [59](#), [66–69](#)
- IMC** Índice de Masa Corporal. [13](#)
- INE** Instituto Nacional de Estadística. [13](#), [20](#)
- IP** Internet Protocol. [71](#)
- IU** Interfaz de Usuario. [31](#), [33](#), [35](#), [38](#), [41](#), [43–46](#), [I](#)
- MET** *Metabolic Equivalent of Task*. [4](#), [5](#), [7](#), [9](#), [11](#), [14](#), [20](#), [21](#), [40](#), [59](#)
- ODS** Objetivos de Desarrollo Sostenible. [3](#)

OMS Organización Mundial de Salud. [1–3](#), [5–11](#), [13](#), [14](#), [20](#), [21](#), [49](#), [59–62](#), [c](#), [II](#)

PA *Physical Activity*. [c](#)

PAGA *Physical Activiyy Guideliness for Americans*. [5](#), [7](#), [59–61](#)

PAGAC *Physical Activiyy Guideliness Advisor Comitte*. [1](#), [2](#), [7–11](#), [60](#), [61](#), [63–65](#), [II](#)

PDI Personal Docente e Investigador. [25](#)

PICO *Patient, Intervention, Comparation and Outcome* . [63](#)

RR Riesgo Relativo. [11](#), [13](#)

RS Revisión Sistemática. [63–65](#)

SGBD Sistemas Gestores de Bases de Datos. [23](#), [25](#), [35](#), [36](#)

SICUZ Servicio de Informática y Comunicaciones de la Universidad de Zaragoza. [25](#)

SNS Sistema Nacional de Salud. [20](#)

TCL Teorema Central del Límite. [68](#)

VIH *Human Immunodeficiency virus*. [9](#)

WCAG *Web Content Accessibility Guidelines*. [22](#)

WHO *World Health Organization*. [c](#)

WSBC *Women Survivors of Breast Cancer*. [9](#), [12](#), [13](#)

Glosario

Antracilinas : antibióticos citotóxicos utilizados en el tratamiento de diferentes cánceres. Los fármacos más utilizados de este grupo son la daunorrubicina y la doxorrubicina. [12](#)

Atípico : dato caracterizado por diferir del grueso de los datos en más de 1.5 el recorrido intercuartílico (atípico débil), o más de 3 veces (atípico extremo). [15](#)

Calorimetría directa : técnica de medición de la actividad física consistente en una cámara herméticamente sellada por cuyas paredes pasa una tubería con agua tal que el calor generado por la persona al ejercitarse en su interior, se transfiere al aire y a las paredes de la cámara incrementando la temperatura del agua, y ese cambio de calor es una medida del ritmo metabólico del ejercicio. Debido a su validez, confianza y objetividad esta tecnología se considera como patrón de referencia y se utiliza en la comparación para validar otras técnicas más prácticas pero con menor precisión. [13](#)

Calorimetría indirecta : técnica que permite medir el calor en forma indirecta a partir de la medición del intercambio gaseoso entre el oxígeno y el CO₂. Las evaluaciones directas e indirectas del consumo máximo de oxígeno están directamente relacionadas con la capacidad funcional cardio-respiratoria y metabólica. El consumo máximo de oxígeno (VO_{2max}) puede ser medido de forma bastante exacta en un laboratorio de fisiología del ejercicio . [13](#)

Inhibidores de aromatasa : bloquean la acción de la aromatasa, una enzima responsable de un paso fundamental en la biosíntesis de los estrógenos. Debido a que los estrógenos son importantes factores mediadores en ciertos cánceres, los inhibidores de aromatasa son frecuentemente usados para tratar dichas enfermedades . [12](#)

Odds-ratio probability/razón de probabilidades : es el cociente de probabilidades de una categoría entre la exposición y no exposición a un tratamiento. [34](#)

Pico máximo de VO₂ : volumen máximo de oxígeno que puede procesar el organismo durante un ejercicio. Se expresa en mililitros de oxígeno por kilogramo corporal y minuto (ml/kg/min) y la población general suele tener unos valores de alrededor de 40-50 ml/kg/min, mientras que los deportistas profesionales suelen rondar los 70-80 ml/kg/min. [12](#), [14](#)

RT : ejercicios de resistencia para el fortalecimiento muscular que pueden realizarse a una intensidad baja o alta. [12](#), [13](#)

Técnica del agua doblemente marcada técnica utilizada para la medición precisa del gasto energético total y la tasa metabólica durante la AF basada en el principio de utilizar dos isótopos estables del agua, generalmente el deuterio y el oxígeno-18, para rastrear los cambios en la concentración de estos marcadores isotópicos en el cuerpo humano a lo largo del tiempo, en particular durante la AF. [13](#)

Lista de Figuras

1.1	Cronograma de actuaciones y programas mundiales para la promoción de la práctica de AF. <i>Fuente: elaboración propia</i>	2
1.2	Distribución del porcentaje y frecuencia de AF en la población femenina atendiendo a su características. <i>Fuente: Elaboración propia</i>	15
1.3	Histograma, gráfico Q-Q plot y gráfico Box-Plot del tiempo semanal de AF en la mujer <i>Fuente: Elaboración propia</i>	16
1.4	Distribución del tiempo semanal de AF según edad. <i>Fuente: Elaboración propia</i>	16
1.5	Box-Plot para el tiempo semanal de AF según edad. <i>Fuente: Elaboración propia</i>	17
1.6	Diagrama de cajas para el tiempo semanal de AF por estudios. <i>Fuente: elaboración propia</i>	17
1.7	Diagrama de cajas para el tiempo semanal de AF por situación personal.	18
2.1	Esquemas relacionales de los Esquemas 1 y 2	24
2.2	Evaluación con Knime de tiempos consumidos en los flujos de trabajo de las consultas en los tres esquemas alternativos	26
2.3	Apariencia <i>responsive</i> del Sistema de Información Web	27
3.1	Ejemplo de una visualización interactiva básica con Shiny . <i>Fuente: https://shiny.posit.co/</i>	33
3.2	Características y funcionalidades de una aplicación Shiny . <i>Fuente: elaboración propia</i>	33
3.3	Esbozo del diseño de la barra de navegación y la la estructura de un panel de visualización en el <i>layout</i>	41
3.4	Diseño y contenido del panel <i>Presentación</i> en el menú El Sistema en el cuadro de mandos del Sistema de Vigilancia	43
3.5	Esbozo del lienzo mostrando el panel de selectores en una sección del cuadro de mandos	44
3.6	Sección Práctica AF ; <i>Visualización I</i> en el cuadro de mandos del Sistema de Vigilancia. . . .	47
3.7	Sección panel INFORME FINAL en el cuadro de mandos del Sistema de Vigilancia	50
		Pág.
III.1	Distribución del porcentaje y frecuencia de AF en la población masculina según características <i>Fuente: Elaboración propia</i>	66
IV.2	Estado favorable para la solicitud a CUSTOS del proyecto del Sistema de Vigilancia	71
IV.3	Cabecera del Cuestionario para la política de protección de datos	71
VII.4	Cronograma del trabajo realizado. <i>Fuente: Elaboración propia</i>	78

Lista de Tablas

1.1	Consumo metabólico equivalente (METs) en diferentes actividades. <i>Fuente: Elaboración propia</i>	5
1.2	Pautas sobre AF de la OMS-2020 y de la PAGA-2018. <i>Fuente: Elaboración propia a partir de las respectivas guías</i>	7
1.3	Medidas e IC para el tiempo semanal de AF (en minutos) según características	18
3.1	Comparación de alternativas para el desarrollo de aplicaciones web interactivas	31
3.2	Comparación entre Flexdashboard y ShinyDashboard	32
	Pág.	
III.1	Medidas e IC para el tiempo (en minutos) semanal de AF en los grupos	68
III.2	IC para la diferencia en los tiempos medios de AF semanal en hombres y mujeres en cada grupo	69

ANEXO A: Cronología de actuaciones de los PAGAC y los comités de expertos de la OMS

En la primera década del siglo XX el estilo de vida en cuanto a salud de los estadounidenses presentaba desafíos significativos ya que, si bien había una creciente creencia sobre la importancia de la AF y una alimentación equilibrada, se enfrentaban a problemas de salud relacionados con un estilo de vida poco saludable adoptado mayoritariamente por la población.

Con el objetivo de abordar la creciente preocupación por los niveles insuficientes de AF y los problemas de obesidad asociados a un riesgo más elevado de enfermedades crónicas como la diabetes tipo 2, enfermedades cardiovasculares y algunos tipos de cáncer, en 2008, el HHS estadounidense publicó la primera edición de las PAGAs. Estas directrices se sustentaron en una fuerte evidencia científica constatada de que la práctica de AF regular promueve la salud y reduce el riesgo de muchas enfermedades crónicas, proporcionando por 1ª vez recomendaciones y pautas con sólida base científica. Las PAGAs sirven como punto de referencia y voz principal y autorizada que formula recomendaciones concretas y específicas sobre AF, la forma física y la salud para los estadounidenses, proporcionando una base sobre la que proceder al desarrollo de programas de AF para niños, adolescentes y adultos estadounidenses.

El aumento significativo de enfermedades no transmisibles y los cambios de hábitos de la vida moderna con una reducción en la AF ocupacional, fue igualmente identificado en muchos otros países del primer mundo. A nivel de salud, se comienza a constatar cómo cuatro categorías de enfermedades –las ECV, las respiratorias crónicas, el cáncer y la diabetes– son las principales responsables de la morbilidad y la mortalidad por enfermedades no transmisibles en el mundo. Las cuatro categorías pueden prevenirse, o controlarse en gran parte, mediante intervenciones eficaces que aborden factores de riesgo comunes y conocidos, a saber: el consumo de tabaco, la dieta, la inactividad física, junto con la intervención temprana.

Unos meses después de la publicación de la PAGA, la OMS elabora su **1er Plan de Acción Mundial sobre AF 2008-2013** con el objetivo de proporcionar a los gobiernos el marco de políticas, programas, actuaciones y medidas de intervención concretas. Este primer plan se centró en áreas clave como la promoción de AF en entornos escolares, lugares de trabajo y comunidades a través de campañas de concienciación y políticas públicas. En 2010, tomando como referencia el informe del PAGAC-2008, el comité asesor de la OMS elaboró sendas directrices - similares a las PAGA- para la población mundial, las primeras pautas sustentadas en evidencia científica: *The Global Recommendations for Physical Activity-2010*

Unos años después de las directrices PAGA-2008, en 2013 el departamento HHS de US elaboró el informe denominado *Mid-term report of the Guidelines on Physical Activity in young people*, que en base a las directrices de 2008, diseñó estrategias eficaces para ayudar a que los jóvenes alcanzasen los 60 minutos de AF diaria -recomendados en las PAGA - en una variedad de entornos desde la guardería o el colegio, la comunidad, la familia, llegando incluso al ámbito de la atención sanitaria primaria. Las cifras que en 2013 manejaba la Asamblea Mundial de Salud (AMS) indicaban que unos 36 millones de defunciones, el 63 % de los 57 millones de muertes registradas en el mundo ese año, se debieron a enfermedades no transmisibles, principalmente enfermedades cardiovasculares

(48 %), cánceres (21 %), enfermedades respiratorias crónicas (12 %) y diabetes (3,5 %). Tomando como antecedente «*un severo problema de salud pública que supone la carga mundial de enfermedades no transmisibles y que socava el desarrollo social y económico en todo el mundo*», en marzo de 2013, la OMS pone en marcha su **2º Plan de Acción Mundial 2013-2025** para el control y prevención de enfermedades no transmisibles atendiendo a lo recogido en la resolución WHA64.11 de la AMS.

Los planes globales de promoción de AF de la OMS, así como las pautas sobre AF, suponen actuaciones de carácter orientativo o recomendaciones a nivel gubernamental y para la ciudadanía, cuya eficacia (medida en términos de incorporación de hábitos de AF en la población), es imprescindible supervisar o monitorizar. Resulta conveniente evaluarla, no sólo en términos del avance hacia los objetivos, sino también con otros propósitos, como el de identificar poblaciones de alto riesgo, o comprobar o, en su caso refutar, la efectividad de las políticas implementadas guiando la planificación de futuras políticas y diseño de programas de promoción eficaces.

En esta línea de evaluación, en el año 2018 vio la luz el estudio *Global Lancet on Physical Activity Practice 2001-2016* conformando una investigación a gran escala, con un elevado grado de exhaustividad, en la que se midieron y analizaron los niveles de inactividad física en la población mundial entre los años 2001 y 2016. Los resultados y conclusiones del estudio se publicaron en la prestigiosa revista médica *The Global Lancet* con el título *Worldwide trends insufficient physical activity from 2001-2016: a pooled análisis of 358 population-based surveys with 1.9 million participant*. El estudio destaca varios hallazgos significativos: de una parte, evidencia una elevada prevalencia de inactividad en la población, estimando que hasta un 65 % de la población no alcanza los niveles recomendables de AF. En la investigación se detectan diferencias significativas en cuanto a los niveles de AF entre regiones y países. El estudio examinó además la evolución temporal, estimando las tendencias a lo largo del tiempo con la conclusión de que, en términos globales, había habido escasas mejoras significativas en los niveles de AF desde inicios del S.XX hasta 2016.

En ese mismo año 2018, los miembros del 2º comité PAGAC-2018 emitieron su informe en base a la actualización de evidencias científicas, tras revisar el nutrido acervo de trabajos científicos en los diez años transcurridos desde el anterior informe. En base a las conclusiones del informe del PAGAC-2018, el departamento de salud estadounidense redactó la que es actualmente la guía más reciente sobre AF en US: *The US Physical Activity Guidelines-2018* (PAGA-2018).

Nuevamente, la OMS con la colaboración del departamento de HSS de US, y en respuesta a las peticiones de los países de recibir tanto una orientación actualizada, como un marco de medidas normativas efectivas y viables destinadas a reducir la inactividad física en la población, en la 71ª Asamblea Mundial de la Salud, aprobó la decisión de poner en marcha un **3º Plan Mundial sobre Actividad Física 2018-2030: más personas activas para un mundo más sano**. El plan acordado por los líderes mundiales de la salud marco como objetivo el de reducir la inactividad física en un 15 % para 2030. La última y más reciente actuación de la OMS, conforme a las revisiones sistemáticas del PAGAC-2018 US actualizadas por el comité de expertos de Institutos de Cáncer y prestigiosas universidades en 2019, en 2020 elaboró las directrices globales sobre AF-2020 con vigencia hasta la actualidad.

El plan mundial 2018-2030 de promoción de AF de la OMS

A pesar de las sólidas evidencias, la OMS continúa alertando de que una gran parte del mundo mantiene niveles de AF insuficientes, además de que en las sociedades más avanzadas, la población de todas las edades está adoptando de forma progresiva estilos de vida más sedentarios; se constata que conforme las sociedades más se desarrollan (económica y tecnológicamente), los niveles de inactividad física en la población se incrementan de forma alarmante y también comprometedor para los sistemas sanitarios, llegando a alcanzar hasta un 70 % de población.

El bautizado como **Plan de Acción Mundial sobre AF 2018-2030: Personas más activas para un mundo más sano** persigue un objetivo claro y cuantitativo: reducir la inactividad física en un 10 % para 2025 y un 15 % para 2030 y la OMS ha puesto a disposición de los países un paquete de intervenciones técnicas denominado «ACTIVE», que brinda orientaciones sobre cómo promover la AF a lo largo de todo el ciclo de la vida en entornos de todo tipo. El plan se articula en torno a cuatro dimensiones claves en cada una de las cuales se proporcionan medidas concretas y soluciones eficaces al conjunto de instancias gubernamentales:

1. **CREAR UNA SOCIEDAD ACTIVA:** Se debe trabajar por lograr un cambio de paradigma en toda la sociedad mediante la mejora de la información, el conocimiento, la comprensión y la valoración de los múltiples beneficios que reporta la actividad física regular, en función de la capacidad y a todas las edades.
 - Medidas como la organización de eventos deportivos con gran participación popular, promocionar los beneficios indirectos y potenciar la formación y aumentar las capacidades de los recursos humanos necesarios.
2. **CREAR ENTORNOS ACTIVOS:** Crear y mantener espacios y lugares que promuevan y salvaguarden el derecho de todas las personas, de todas las edades, a gozar de un acceso equitativo a lugares y espacios seguros en sus ciudades y comunidades en los que puedan realizar una actividad física regular acorde a sus posibilidades.
 - Medidas como reforzar la seguridad vial, los espacios públicos al aire libre y sus accesos. Aumentar las redes de vías peatonales y ciclistas, integrar políticas de transporte y planificación urbanística. Aplicar políticas de construcción inmobiliaria en espacios abiertos y centros deportivos.
3. **FOMENTAR POBLACIONES ACTIVAS:** Implementar programas y promover el acceso a oportunidades, en múltiples entornos, para ayudar a las personas de todas las edades y capacidades a participar regularmente en actividades físicas, solas o junto a sus familias y comunidades.
 - Medidas como instituir programas transversales en múltiples entornos, mejorar la oferta para las personas mayores, potenciar la educación física y los programas escolares o integrar la actividad física en los servicios de salud y sociales.
4. **CREAR SISTEMAS ACTIVOS:** Elementos propiciadores en la gobernanza y sus políticas. Crear y reforzar el liderazgo, las alianzas multisectoriales, las capacidades del personal, la promoción y los sistemas de información en todos los sectores, con el fin de alcanzar la excelencia en la movilización de recursos y la aplicación de medidas internacionales y nacionales coordinadas dirigidas a aumentar la AF y reducir el sedentarismo.
 - Medidas como instituir campañas de mercadotecnia social y ampliar las actividades de promoción o fomentar actividades de investigación y desarrollo en el ámbito deportivo, así como mejorar e integrar los sistemas de datos.

Por último, según señala la OMS, el conjunto de medidas normativas que propician la AF y su adopción por los países, se encuentran alineadas en mayor o menor medida con hasta 13 de los objetivos de desarrollo sostenible (ODS) para 2030.

ANEXO B: Metodología en la extracción de evidencias científicas de los **PAGACs**

El informe (de carácter público) elaborado por el comité **PAGAC-2018**, en cuya metodología se profundiza en este anexo, utilizó la metodología más avanzada, las denominadas revisiones sistemáticas, caracterizadas por una revisión pormenorizada de la bibliografía pertinente para dar respuesta a una serie de preguntas de investigación de **Revisión Sistemática (RS)** previamente diseñadas a tal efecto; en este caso se formularon un total de 38 preguntas con 104 sub-preguntas. El proceso metodológico para la evaluación de evidencias se realiza siguiendo una serie de etapas, que se exponen a continuación y cuyos procedimientos se describen de forma sintética.

*Etapas 1.- Desarrollo de preguntas **RS***

Basándose en los informes previos disponibles, junto con un catálogo de áreas científicas de interés con un rápido desarrollo, se identifican una serie de temáticas clave con investigación disponible; en el caso que nos ocupa: niños menores de 6 años, adultos mayores, salud cerebral a lo largo de la vida, dosis-respuesta y comportamiento sedentario, que no se habían incluido (o no se habían alcanzado evidencias suficientes) en las anteriores revisiones. Una vez decididas las temáticas y conformados los respectivos subcomités, fueron éstos los encargados de desarrollar (en su caso, perfeccionar) las preguntas de **RS** enfocadas a cada temática que resultarán clave para conducir la búsqueda, selección y revisión de trabajos.

*Etapas 2.- Desarrollo de una estrategia de **RS***

La sistematización de un proceso de búsqueda de trabajos científicos, suele comenzar con el desarrollo de los denominados “marcos analíticos”. Los marcos analíticos son una representación formal de la búsqueda que permite focalizar el contexto de cada estrategia y se utilizan para definir con precisión las variables y los términos clave, para determinar los criterios de inclusión y exclusión, para el desarrollo de la estrategia de búsqueda bibliográfica y para acotar, o limitar el alcance. Se utilizó uno de los marcos analíticos más utilizado y reconocido *Patient, Intervention, Comparison and Outcome* (**PICO**).

Toda estrategia de revisión exige formular con la mayor precisión posible los criterios de inclusión de trabajos, siendo su objetivo fundamental garantizar que los artículos que se van a tener en consideración resultan pertinentes o relevantes en la respuesta de la pregunta **RS**. Estos criterios deciden aspectos como el tipo de investigación (considerando artículos originales, revisiones existentes -incluyendo las revisiones sistemáticas-, las meta-análisis -MA- o, los trabajos combinados); el idioma de publicación (habitualmente a texto completo en inglés), o la calidad de las revistas. Según los propósitos se pueden considerar criterios adicionales que permitan rescindir los trabajos susceptibles de proporcionar respuesta a una pregunta de **RS** concreta (la edad, el estado de salud previo de los

sujetos, la presencia o no de grupos de control en los estudios, o el tipo de intervención/exposición abordada).

Se procede a elaborar unas estrategias de búsqueda que permitan identificar toda aquella bibliografía que conformará el acervo de trabajos científicos a revisar, analizar y evaluar para cada pregunta RS. Ello incluye formular unos términos de búsqueda, una lógica booleana que combine los términos, establecer las bases de datos a consultar, así como los límites y filtros establecidos. Las bases de publicaciones utilizadas por el PAGAC-2018 fueron PubMed, CINAHL y Cochrane que representan depósitos exhaustivos de citas, resúmenes y artículos completos en los campos relevantes para las preguntas RS formuladas³².

Etapas 3.- Revisión y cribado de trabajos

La selección de trabajos en una revisión sistemática es un proceso crítico y complejo, cuyo objetivo es el de identificar de forma objetiva, garantizando completitud y exhaustividad, el conjunto de trabajos más completo, pertinente y a su vez con capacidad de dar respuesta a una pregunta de RS.

Para una selección óptima, se recomienda el denominado “método de pares” consistente en que son dos o más miembros en el equipo de selección los encargados de revisar de forma independiente los trabajos. En este método, cuando se identifican discrepancias, un revisor adicional examina los títulos o resúmenes con el fin de alcanzar un consenso. Se comienza revisando los títulos y los resúmenes para comprobar si cumplen con los criterios objetivos de inclusión y se recupera luego el texto completo.

Etapas 4.- Resumir y evaluar la calidad y el riesgo de sesgo de los trabajos

Se requiere un enfoque objetivo de abstracción para resumir las características de los estudios realizados y analizados en los trabajos, así como su concordancia con cada pregunta de RS. La abstracción de datos requiere identificar con precisión y describir los elementos clave de cada estudio, al tiempo que se debe capturar información coherente del trabajo en el conjunto de artículos. Una vez recuperado el texto completo de los artículos seleccionados, para una tarea tan relevante y compleja, en el caso del comité PAGAC-2018 se contrató, se formó y se certificó a los “resumidores” de pruebas, utilizando estrictos procedimientos de control de calidad, mediante un equipo de expertos que orientó y supervisó los resúmenes realizados.

Además de resumir la información clave, en esta etapa se debe evaluar la calidad de los trabajos. Para este objetivo se suele utilizar una herramienta desarrollada para apoyar (mediante la valoración de diferentes ítems), la apreciación sobre la calidad de un artículo denominada *A Measurement Tool to Assess Systematic Reviews (AMSTAR)*. El PAGAC-2018 utilizó una adaptación de ésta, la denominada AMSTARExBP en base a la cual, la pareja de resumidores evaluaron de forma independiente la calidad de cada trabajo.

El riesgo de sesgos es una consideración crucial en la evaluación de la evidencia a partir de trabajos científicos en tanto puede afectar a la validez y fiabilidad de los resultados/conclusiones. Los sesgos en los estudios empíricos, dirigidos o no, pueden provenir de la presencia de errores sistemáticos en el diseño de las pruebas, en su ejecución, en la interpretación de los resultados, o en la formulación o extrapolación de conclusiones y debe ser valorado (preferiblemente con el apoyo de herramientas técnicas). En el contexto que nos ocupa, se utilizó una versión adaptada de la Herramienta de Evaluación de Sesgos (BAT³³, por sus siglas en inglés) una evaluación basada en dominios: sesgos de

³²El marco analítico, los criterios de inclusión y exclusión y la estrategia de búsqueda para cada pregunta de la RS se encuentran en los portafolios de evidencia específicos y pueden consultarse en www.health.gov/paguidelines

³³La NEL BAT se adapta según el diseño del estudio, con diferentes conjuntos de preguntas que se aplican a los ensayos controlados aleatorizados (ECA) (14 preguntas), a los ensayos controlados no aleatorizados (14 preguntas) y a los estudios observacionales (12 preguntas)

selección, de rendimiento, de detección y de deserción que ayudan a los revisores a determinar la existencia de algún error sistemático en el estudio que pudiera sobreestimar o subestimar los efectos evidenciados.

Etapas 5.- Calificar el grado de evidencia de las pruebas

La calificación de la solidez de las pruebas se basa en su aplicabilidad a poblaciones, las exposiciones o tratamientos aplicados y los resultados investigados; su generalización a la población de interés, los riesgos de sesgos valorados, las limitaciones del estudio, la cantidad y coherencia de los hallazgos entre los trabajos, así como la magnitud y precisión de los efectos. El sistema *Grading of Recommendations Assessment, Development, and Assessment* (GRADE) es la metodología más utilizada para evaluar el grado de evidencia y la fuerza de las recomendaciones en la investigación médica y de salud al proporcionar un marco estructurado diseñado para certificar el grado de evidencia en los estudios clínicos y epidemiológicos. En base a los parámetros que integra, GRADE clasifica la calidad de la evidencia en cuatro niveles: alta, moderada, baja o muy baja. Cabe mencionar que para el análisis de la calidad de las evidencias, los evaluadores PAGAC-2018 contaron con una rúbrica adaptada por el comité del sistema GRADE: *PAGAC's qualification criteria* para guiar la calificación de la fuerza de la evidencia que apoyaba las conclusiones de cada estudio³⁴

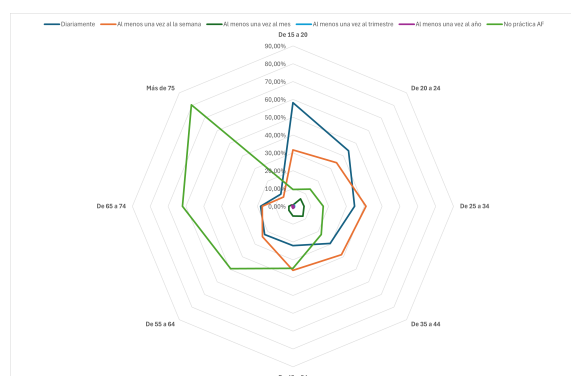
Etapas 6.- Elaborar un resumen narrativo y formular las conclusiones

Tras la calificación de evidencias de todos los trabajos científicos seleccionados y resumidos para una pregunta RS, y la puesta en común, se procede a elaborar un resumen narrativo que precede a la declaración de conclusiones fundamentadas en la revisión sistemática, enfatizando en los hallazgos identificados en las respuestas a las preguntas RS. En esta etapa de conclusiones, se formulan además diversas recomendaciones de investigación, enumerando las áreas claves en las que la investigación adicional podría ayudar a mejorar la base de evidencia permitiendo abordar determinadas lagunas identificadas en la investigación disponible.

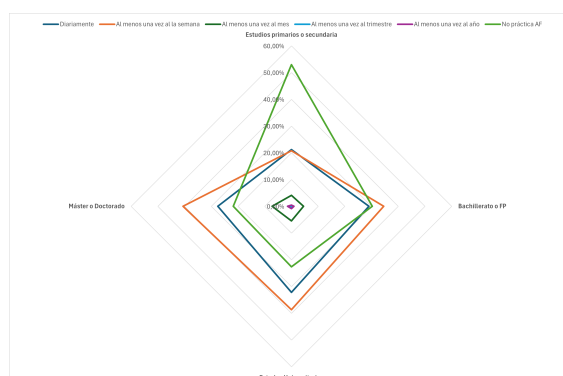
³⁴La información del catálogo de evidencias de los trabajos en cada pregunta RS se puede consultar en el enlace www.health.gov/paguidelines

ANEXO C: Perspectiva de género: Análisis de resultados de la EHD en la población masculina

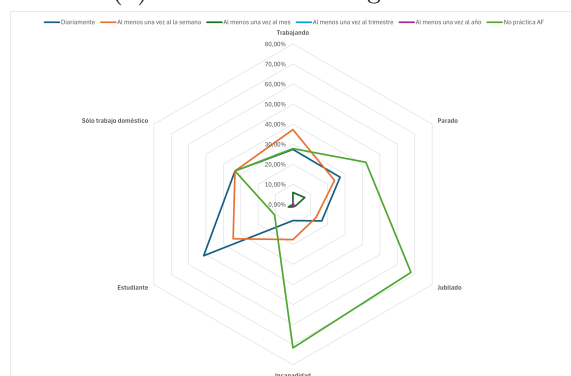
El porcentaje de hombres en la muestra en la EHD de 2022 que practicó AF durante ese año es de 62.32 % , lo que supone un 13 % más que el obtenido en la población femenina; en base a éste, el intervalo de confianza para el porcentaje de hombres en la población que realizó AF regularmente durante el 2022 es: $IC_{95\%}(p) = [61.2\%, 63.6\%]$ y permite concluir que, al menos el 61 % de la población masculina practica AF regular. Un estudio gráfico más detallado (véase figura III.1), muestra cómo los porcentajes de práctica de AF y su frecuencia difieren según la edad, el nivel de estudios y la situación laboral, no así la situación personal (salvo en los hombres solteros).



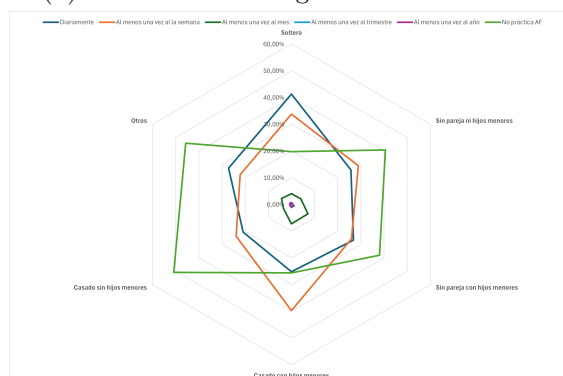
(a) Práctica de AF según edad



(b) Práctica de AF según nivel de estudios



(c) Práctica de AF según situación laboral



(d) Práctica de AF según situación personal

Figura III.1: Distribución del porcentaje y frecuencia de AF en la población masculina según características *Fuente: Elaboración propia*

El análisis inferencial ha permitido concluir que, a un nivel de significación del 5 %, existen diferencias significativas en los porcentajes de práctica de AF regular entre los diferentes grupos de edad. A diferencia del caso de las mujeres, sí se detectan diferencias significativas entre los porcentajes de práctica de AF entre los dos grupos de menor edad. Igualmente, se detecta la existencia de diferencias significativas en la práctica de AF regular según el nivel de estudios alcanzado, con excepción del grupo con estudios universitarios y el de con bachillerato o formación profesional, entre los que son indistinguibles. Al igual que en el caso de las mujeres, el porcentaje es creciente con el nivel académico, fluctuando entre un 21.4 % para hombres con estudios primarios, hasta un 32.5 % entre hombres con estudios universitarios (una diferencia mucho menos acentuada que en el caso de las mujeres). En lo referente a la práctica en función de la situación laboral, en hombres se evidencian diferencias significativas en todas las categorías, a excepción de las de trabajando y parado en las que los porcentajes son indistinguibles; conclusión ésta también obtenida en el caso de las mujeres. Los hombres estudiantes con un porcentaje del 51.3 % son los que más participan en la AF destacando, no obstante, cómo este porcentaje resulta 20 puntos inferior al de las estudiantes; opuestamente, en hombres jubilados, tan sólo un 16.6 % practica AF de forma regular, resultando de nuevo un porcentaje algo inferior al de las mujeres jubiladas.

Analizando el tiempo destinado a la práctica de AF en la población masculina, con los datos depurados, se obtiene que el tiempo medio de práctica en la muestra asciende a 198 minutos semanales, con un $IC_{95\%}(\mu) = [187, 202]$ como estimación del tiempo medio de práctica en la población; éste, permite concluir que, en media, la población masculina alcanza, sobradamente la dosis de AF recomendada.

El estudio inferencial ha permitido concluir que existen diferencias significativas en el tiempo de práctica de AF en todos los grupos de edad, incluyendo los dos de menores edades (a diferencia de lo que ocurre en las mujeres); ésta fluctúa entre 390 minutos en el grupo de 15 a 19 años, hasta 63 minutos en el grupo de más de 75 años. Igualmente, se detectan diferencias significativas en el tiempo de práctica en función del nivel de estudios, alcanzando la media más elevada (279 minutos) entre los hombres con estudios universitarios, seguidos de los que cuentan con estudios de Bachiller o FP Superior, con 218 minutos, resultando menor en hombres con estudios de máster o doctorado (194 min), y con la más baja entre hombres con estudios primarios o secundaria (146 min). Resulta un patrón que difiere del caso de las mujeres con estudios de máster o doctorado.

Respecto a la situación personal, el análisis inferencial permite concluir (al igual que en el caso de las mujeres) que no existen diferencias significativas entre el tiempo medio de AF entre las modalidades con, o sin, hijos menores y con, o sin pareja, difiriendo únicamente, en la modalidad de solteros en la que resulta significativamente superior a las demás. Se han calculado, y se recogen en la **Tabla III.1** las medidas descriptivas del tiempo de AF en función de las variables analizadas, junto con sus intervalos de confianza en cada grupo.

Tiempo de AF según edad				Tiempo de AF según estudios			
Edad	Media	Desv.	$IC_{95\%}(\mu)$	Estudios	Media	Desv.	$IC_{95\%}(\mu)$
De 15 a 24	390	333	[344,431]	Primarios	146	273	[138,253]
De 25 a 34	310	326	[270, 348]	Bachillerato	218	269	[210,225]
De 35 a 44	246	245	[228,266]	Universitarios	249	323	[240,257]
De 45 a 54	180	247	[200, 232]	Doctorado	194	218	[188,200]
De 55 a 64	167	290	[150,187]				
De 65 a 74	158	275	[134,180]				
Más de 75	73	63	[41,84]				

Tiempo de AF según situación personal				Tiempo de AF según situación laboral			
S.Personal	Media	Desv.	$IC_{95\%}(\mu)$	S.Laboral	Media	Desv.	$IC_{95\%}(\mu)$
Soltero	284	309	[262,303]	Trabajando	208	264	[199,218]
Sin pareja ni hijos	205	319	[185,228]	Parado	189	333	[156,222]
Sin pareja con menores	94	138	[82,107]	Jubilado	124	285	[109,140]
Con pareja sin menores	89	137	[138,165]	Doméstico	340	345	[74,605]
Con pareja con menores	110	144	[175,204]	Estudiante	346	317	[313,378]

Tabla III.1: Medidas e IC para el tiempo (en minutos) semanal de AF en los grupos

ALGUNAS CONCLUSIONES NUMÉRICAS EN EL ESTUDIO DE GÉNERO

Se han estimado las diferencias medias en el tiempo de práctica de AF entre hombres y mujeres, en términos globales y, según las características. El elevado tamaño muestral, merced al [Teorema Central del Límite \(TCL\)](#), permite servirse de la normalidad de la media muestral en la obtención de los intervalos de confianza aproximados obtenidos.³⁵

En términos globales, se estima que la diferencia entre el tiempo medio de práctica de AF en hombres y mujeres se encuentra en el $IC_{95\%} = [88, 105]$ que permite concluir que, en media, la población masculina realiza al menos 88 minutos semanales por encima del tiempo de práctica de AF de la población femenina. Se han obtenido y se muestran en la **Tabla III.2** sendos intervalos de confianza para la diferencia en el tiempo medio de práctica en cada grupo generados por las variables analizadas: edad, nivel de estudios, situación personal y situación laboral.

Particularizando las diferencias de tiempo semanal de AF entre hombres y mujeres en los diferentes grupos de edad, las diversas situaciones laborales y personales consideradas, así como el nivel de formación académica, la lectura de los resultados obtenidos permite formular algunas conclusiones del estudio de género tales como:

1. Independientemente de su edad, su situación personal y su formación académica, se concluye que los hombres que trabajan realizan una media de al menos 1 hora semanal más de AF que las mujeres que trabajan. En los estudiantes, se concluye que ellos practican por encima de dos horas más de lo que lo hacen las estudiantes. Así mismo, la media de práctica de los jubilados es casi una hora superior al tiempo medio dedicado por las mujeres jubiladas. También los parados dedican más tiempo y lo hace en media en más de 100 minutos del tiempo que dedican las mujeres paradas; los hombres que sólo realizan trabajo doméstico, practican una media

³⁵Es posible utilizar técnicas no paramétricas para refinar la estimación de esta distribución y tratar de mejorar la precisión de los intervalos estimados. No obstante, se ha comprobado en la implementación que la ganancia en precisión es en decimales y resulta irrelevante en este contexto.

Diferencia de AF según edad		Diferencia de AF según formación	
Edad	$IC_{95\%}(dif.)$	Estudios	$IC_{95\%}(dif.)$
De 15 a 19	[193,286]	Primarios	[25,53]
De 20 a 24	[116,202]	Bachillerato	[74,108]
De 25 a 34	[99, 143]	Universitarios	[27,110]
De 35 a 44	[99,143]	Doctorado	[83,191]
De 45 a 54	[42,80]		
De 55 a 64	[59,100]		
De 65 a 74	[65,116]		
Más de 75	[13,58]		
Diferencia de AF según situación personal		Diferencia de AF según situación laboral	
Situación Personal	$IC_{95\%}(dif.)$	Situación Laboral	$IC_{95\%}(dif.)$
Soltero	[150,194]	Trabajando	[62,142]
Sin pareja ni menores	[40,87]	Parado	[106,174]
Sin pareja con menores	[72,150]	Jubilado/Pensionista	[54,113]
Con pareja sin menores	[84,116]	Doméstico	[74,605]
Con pareja con menores	[61,92]	Estudiante	[146,217]

Tabla III.2: IC para la diferencia en los tiempos medios de AF semanal en hombres y mujeres en cada grupo

de al menos 75 minutos más que la media de tiempo de las mujeres que solo realizan trabajo doméstico.

- En todas las edades, los hombres dedican un tiempo considerablemente superior al que dedican las mujeres siendo reseñables los casos de los hombres de 25 a 34 años y de los de 15 a 19 por cuanto realizan más de 1,5 horas y 3 horas, respectivamente por encima del tiempo que destinan las mujeres de sus edades.
- Por nivel académico, las menores diferencias, con respecto del tiempo dedicado por las mujeres con igual formación académica, se dan en hombres con estudios primarios y con estudios universitarios, con una diferencia de al menos 25 minutos semanales. Los hombres tanto con estudios de bachillerato, como los que cuentan con estudios de máster o doctorado, realizan en media como mínimo una hora más de AF del tiempo que practican las mujeres en los niveles de formación respectivos.
- Para todas la situaciones personales, se concluye que los hombres realizan una media de al menos 40 minutos semanales más de AF que las mujeres en su misma situación personal. Entre los solteros, la media excede al menos en 150 minutos a la de las mujeres solteras. Los hombres que viven en pareja y con hijos menores realizan en media más de 1 hora por encima del tiempo que dedican las mujeres que viven en pareja con hijos menores; incluso los hombres que no viven en pareja, pero tienen hijos menores, realizan una media de al menos 70 minutos semanales más que las mujeres que no viven en pareja y tienen hijos menores. Los hombres con pareja y sin hijos menores superan en más de 80 minutos semanales de práctica a las mujeres con pareja y sin hijos menores.

ANEXO D: Solicitudes de Conformidad de Protección de Datos y de Evaluación de proyecto en el CEICA

PROTECCIÓN DE DATOS UNIVERSIDAD DE ZARAGOZA: CUSTOS

La información sobre el Tratamiento de la información que se presenta a continuación se proporciona en la *app* web del sistema de información cuantos01.unizar.es/tfm/terminos.php

LEA AQUÍ LOS TÉRMINOS DE PROTECCIÓN DE LOS DATOS

Tratamiento de la información

1. Se trata de un **cuestionario totalmente anónimo**, no incluye datos que permitan identificarle y sus **respuestas serán confidenciales**.
2. A diferencia de otros formularios, las respuestas no se envían ni alojan en ningún proveedor de servicios, sino en un **servidor propio custodiado, gestionado y con uso exclusivo por parte de las investigadoras**.

En relación con esta encuesta, Ud. puede ejercer sus derechos en materia de privacidad directamente si lo desea, dirigirse al investigador principal (nayuso@unizar.es), o al Delegado/a de Protección de Datos de la Universidad de Zaragoza (dpd@unizar.es) o, finalmente, en reclamación, a la Agencia Española de Protección de Datos (www.aepd.es).

En proteccion.de.datos.unizar podrá encontrar amplia información respecto de este tratamiento y de su política de protección de datos, así como formularios para el ejercicio de sus derechos.

Si usted es una **mujer mayor de edad**, puede colaborar rellenando el cuestionario y así contribuir al conocimiento que esperamos adquirir en este proyecto.

MUCHAS GRACIAS POR COLABORAR CON ESTA INVESTIGACIÓN

Se tramitó y se ha obtenido el visto bueno de la Universidad de Zaragoza sobre la política de protección de datos CUSTOS (véase la **Figura IV.2**)

CUSTOS 1.3.9

MIS AUTODECLARACIONES

Desde esta pantalla podrás visualizar las distintas autodeclaraciones y generar un PDF con el informe completo de las mismas.

Datos del solicitante

NIF: 110643	Nombre y apellidos: María Guzmán Lora	DNI/Pasaporte: 9401 114430351
Correo electrónico principal: marigu@unizar.es	Correo electrónico secundario: marigu@unizar.es	Teléfono: 913129411 / 900130205
Centro/Departamento: Depto. Economía Aplicada / Facultad de Economía y Empresa	Servicio/Unidad: Área Métodos Cuantitativos Económicos / Área Métodos Cuantitativos Económicos F y E	Extensión telefónica:

ID	Código UPD	Función UNIZAR	Tipo	Presentada	Estado	Fecha estado
100601	2024-190	LZ Responsable	TQM	Un sistema para la vigilancia del estado de la actividad física: la mujer española y su influencia en el riesgo de cáncer de mama	13/05/2024 16:42:17	Aprobado

En total hay 1 autodeclaraciones que cumplen los criterios buscados.

© 2023 Unidad de Protección de Datos | E-mail de soporte: dpt@unizar.es
© 2023 Universidad de Zaragoza (Pedro Carbono 1,2, Edificio Interfacultades, 50009 ZARAGOZA - ESPAÑA)

Figura IV.2: Estado favorable para la solicitud a CUSTOS del proyecto del Sistema de Vigilancia

El visto bueno, por ejemplo, ha requerido incorporar en la cabecera del cuestionario de forma explícita, tanto la información del tratamiento de datos, como con la solicitud de acuerdo explícita de la aceptación y conformidad con la política de protección de datos de la Universidad de Zaragoza tal como se muestra en la **Figura IV.3**

CUESTIONARIO
Proyecto Actividad Física y Salud en la mujer

Tratamiento de la información

1º.- Se trata de un cuestionario totalmente anónimo que no incluye datos que permitan identificarle y sus respuestas serán confidenciales.

2º.- A diferencia de otros formularios, las respuestas no se envían ni alojan en ningún proveedor de servicios, sino en un servidor propio custodiado, gestionado y con uso exclusivo por parte de las investigadoras.

En relación con esta encuesta, Ud. puede ejercer sus derechos en materia de privacidad directamente si lo desea, dirigiéndose al investigador principal (ayuso@unizar.es), o al Delegado/a de Protección de Datos de la Universidad de Zaragoza (dpt@unizar.es) o, finalmente, en reclamación, a la Agencia Española de Protección de Datos (www.aepd.es)

En [Política de Privacidad](#) podrá encontrar amplia información respecto de este tratamiento y su política de protección de datos, así como formularios para el ejercicio de sus derechos.

Si usted es una mujer mayor de edad puede colaborar rellenando el cuestionario y así contribuir al conocimiento que esperamos adquirir en este proyecto.

AGRADECEMOS POR ANTICIPADO SU COLABORACIÓN!!

¿Ha leído y acepta la Política de Privacidad de la Universidad de Zaragoza? [Ver Política de Privacidad](#)

☐ Sí ☐ No

Debe marcar "Sí" para acceder a la encuesta si está de acuerdo con la política de privacidad.

EMPEZAR LA ENCUESTA

Figura IV.3: Cabecera del Cuestionario para la política de protección de datos

SOLICITUD DE EVALUACIÓN AL CEICA

El Cuestionario *online* no recoge ningún tipo de dato identificativo de las encuestadas, ni siquiera la **Internet Protocol (IP)** del equipo en el que se realiza y se envía el cuestionario. A pesar de ello, la puesta en marcha del proyecto requiere de su evaluación por parte del **CEICA**, en tanto se enmarca en uno de los casos previstos de *Investigaciones con encuestas que recojan datos personales (aunque no sean identificativos) o de salud realizadas en persona, online, en papel por medio de terceros, por correo electrónico, por teléfono o por cualquier medio.*

Debido a que el proyecto se realiza en la Universidad es preceptivo contar con el informe favorable de Protección de Datos de la Universidad de Zaragoza. Una vez que se cuenta con éste, está previsto tramitar en julio la solicitud de evaluación al **CEICA**.

ANEXO E: El Cuestionario

SECCIÓN I: ASPECTOS SOCIODEMOGRÁFICOS (SD)

SD1: Seleccione por favor el rango de su edad

1: 18-29 2: 30-39 3: 40-49 4: 50-59 5: 60-69 6: +70

SD2: ¿Cuál es su situación laboral actual?. Elija la opción que mejor la refleje

1: Empleada a tiempo completo 2: Empleada a tiempo parcial 3: Estudiante
4: Pensionada 5: Autónoma 6: Empresaria
7: Desempleada (En búsqueda de empleo) 8: Desempleada (Sin búsqueda de empleo)

SD3 Señale la formación académica (alcanzada o en curso)

1: Estudios Primarios, Secundaria o FP 2: Bachillerato o FP Grado Superior 3: Licenciatura, Diplomatura o Grado
4: Máster o Doctorado

SD4: Tecle los dígitos de su código postal

SD5: Díganos, ¿vive con una pareja?

1: Sí 2: No

SD6: Indique, por favor, si residen menores de 18 años en su domicilio habitual

1: Ninguno 2: Un menor 3: Dos o más menores

SD7: ¿Tiene a su cargo alguna persona mayor, o alguna persona con cierto grado de dependencia?

1: Sí 2: No

SECCIÓN II: PRÁCTICA DE ACTIVIDAD FÍSICO DEPORTIVA RECREATIVA (AFR)

AFR1: ¿Ha practicado algún deporte o realizado actividad física recreativa durante el último año?

1: No 2: Sí 3: Sí vinculada a mi trabajo o estudios

AFR2: Díganos, por favor: ¿con qué frecuencia practica actividad físico deportiva?

1: Diariamente (casi todos los días) 2: Más de dos veces por semana 3: Una o dos veces por semana
4: Dos o tres veces al mes 5: Estacionalmente

AFR3: ¿Cuánto tiempo (en minutos) de actividad físico deportiva suele realizar a lo largo de una semana?

AFR4: Señale, por favor, los tipos de actividad físico deportiva que practica de forma regular

1: Gimnasia, actividad físico-deportiva suave (pilates, mantenimiento, yoga, taichi...)
2: Gimnasia, actividad físico-deportiva intensa (elíptica, cinta, aeróbic, spinning, body power, crossfit,...)
3: Actividad física con acompañamiento musical (zumba, gym-jazz, hip-hop, kapoira,...)
4: Ejercicios de resistencia, músculo, halterofilia...

AFR5: Si ha practicado algún deporte en el último año de forma regular (semanalmente), indique cuál o cuáles

1: Fútbol 2: Baloncesto 3: Balonmano 4: Rugby
5: Natación 6: Ciclismo 7: Motociclismo 8: Boxeo
9: Duatlón, triatlón 10: Marcha, carrera 11: Escalada 12: Patinaje artístico
13: Senderismo, alpinismo 14: Bolos, petanca 15: Artes Marciales 16: Piragüismo, remo..
17: Gimnasia deportiva/rítmica 18: Otros

AFR6: ¿Qué tipo de espacios o instalaciones utiliza habitualmente para hacer deporte y/o actividad física?

1: Espacios abiertos de uso libre (urbanos: parque, calles; montaña; acuáticos: ríos, pantanos, mar)
2: Instalaciones deportivas públicas (gimnasios, centros deportivos o espacios deportivos público)
3: Instalaciones deportivas privadas (gimnasios, centros deportivos o espacios deportivos privados)

AFR7: ¿Suele practicar actividad físico deportiva de forma individual, o acompañada?

1: Individual 2: Acompañada 3: Ambos

AFR8: ¿Participa en competiciones deportivas?

1: No 2: Sí

AFR9: ¿Suele practicar de forma libre o dirigida?

1: Libre 2: Dirigida 3: Ambas

AFR10: ¿A qué edad aproximada comenzó a realizar deporte o actividad física?

AFR11: Señale, por favor, los dos motivos principales por los que realiza actividad física o practica algún deporte

- | | |
|---|---|
| 1: Por diversión o entretenimiento | 2: Por mantener o mejorar su forma física |
| 3: Por relajarse (aliviar estrés, tensiones...) | 4: Por fortalecer relaciones sociales |
| 5: Por salud (prevenir enfermedades) | 6: Por su profesión |

AFR12: ¿Suele ir a caminar (más o menos rápido) de forma intencional, durante más de 30 minutos al día, con el fin de mejorar su salud? Señale el rango que mejor refleje el tiempo semanal que dedica a esa actividad

- | | |
|----------------------------|-----------------------|
| 1: No suelo ir a caminar | 2: Menos de dos horas |
| 3: Entre dos y cinco horas | 4: Más de cinco horas |

SECCIÓN III: FACTORES QUE LIMITAN LA PRÁCTICA DE AF (FL)

FL1: ¿Le gustaría realizar más actividad físico deportiva de la que actualmente practica?

- 1: Sí 2: No. Considero que mi práctica es suficiente 3: No. La actividad física no me interesa

FL2: Desde la perspectiva del tiempo disponible, seleccione por favor, el motivo principal por el que no realiza, o no realiza más actividad físico deportiva

- 1: Mi trabajo/estudios me limitan y no soy capaz de generar tiempo para mí
- 2: Mis horarios son incompatibles
- 3: En la gestión de mi tiempo libre, considero la práctica de actividad física como secundaria
- 4: El tiempo no condiciona mi práctica de actividad físico deportiva

FL3: Desde la perspectiva de la conciliación familiar, seleccione por favor, el motivo principal por el que no realiza, o no realiza más actividad físico deportiva

- 1: Porque las tareas domésticas ocupan una buena parte de mi tiempo libre
- 2: Porque debo ocuparme de la atención y cuidado a menores (reuniones escolares, citas médicas,...)
- 3: Porque me ocupo de la atención física y/o emocional de alguna persona mayor (o con discapacidad)
- 4: La conciliación familiar no condiciona mi práctica de actividad física

FL4: Desde la perspectiva de la oferta de actividades, seleccione por favor, el motivo principal por el que no realiza, o no realiza más actividad físico deportiva

- 1: Porque no cuento con un espacio próximo en el que se oferten actividades de mi interés
- 2: Porque las actividades me resultan pesadas, o no me enganchan y pierdo el interés
- 3: Por una cuestión económica
- 4: La oferta de actividades no limita mi práctica de actividad física

FL5: Indíquenos, por favor, cuántas horas semanales suele dedicar a actividades de ocio personal (cine, teatro, lectura, cenas, amistades) excluyendo las horas destinadas a la práctica de actividad física

- 1: Menos de cuatro horas 2: Entre cuatro y ocho horas 3: Más de ocho horas

SECCIÓN IV: HÁBITOS SALUDABLES Y ALIMENTACIÓN (HSA)

HSA1: ¿Ha fumado durante el último año?

- | | |
|--|--------------------------------------|
| 1: No | 2: He fumado, pero lo dejé |
| 3: Sí, menos de 10 cigarrillos diarios | 4: Sí, más de 10 cigarrillos diarios |

HSA2: ¿Con qué frecuencia consume alcohol?

- 1: Diariamente 2: Semanalmente 3: Mensualmente 4: Esporádicamente

HSA3: Díganos, ¿es vegetariana o vegana?

1: No 2: Sí

HSA4: Puntúe cómo es de saludable su alimentación en base a las recomendaciones nutricionales

HSA5: Díganos, por favor, si está motivada, y en su caso, señale la principal motivación en el cuidado de su alimentación

1: No estoy motivada 2: Verse bien físicamente
3: Cuidar la salud 4: Mantener o perder peso

HSA6: ¿Cree que su alimentación sería más saludable si dispusiese de más tiempo para cocinar?

1: Sí 2: No

HSA7: En un día normal, ¿cuánto tiempo suele pasar sentada delante de una pantalla?

1: Menos de dos horas 2: Entre dos y cuatro horas 3: Más de ocho horas

HSA8: ¿Suele sentir estrés?

1: Nunca 2: Puntualmente, en alguna ocasión
3: Sí, a diario (excepto el fin de semana) 4: Sí, todos los días

SECCIÓN V: ASPECTOS FÍSICOS, GINECOLÓGICOS Y CÁNCER DE MAMA (AFGC)

AFGC1: Indique, por favor, su estatura en centímetros

AFGC2: Indique, por favor, su peso aproximado en kilogramos

AFGC3: ¿A qué edad tuvo su primera menstruación?

AFGC4: Si procede, díganos, por favor, a qué edad tuvo su primer embarazo

AFGC5: Indique, en su caso, el número de embarazos que haya completado

AFGC6: Por favor, seleccione el rango de su edad cuando entró en la menopausia, si procede

1: Todavía no 2: Entre los 45 y 50 años 3: Menopausia inducida 4: Con más de 50 años

AFGC7: ¿Hay algún antecedente directo de cáncer en su familia (abuelos, padres, hermanos, tíos, sobrinos e hijos)?

1: No 2: Sí 3: No está segura

AFGC8: Deseando que la respuesta sea negativa, necesitamos saber si ha sido diagnosticada de cáncer de mama (mediante prueba histológica)

1: No 2: Sí

AFGC9: ¿Puede por favor, indicarnos el tipo de cáncer, si lo conoce?

1: Luminal A 2: Luminal B 3: Triple negativo
4: Subtipo Her2 6: Otro/No lo sabe con certeza

AFGC10: Indíquenos, por favor, cuál o cuáles de estos tratamientos se le administraron (o le están administrando, o le van a administrar)

1: Radioterapia 2: Inmunoterapia 3: Quimioterapia
4: Cirugía 5: Tratamiento Hormonal 6: Otro/ No lo sabe con certeza

AFGC11: Tras el diagnóstico, ¿su oncólogo le prescribió o recomendó realizar actividad físico deportiva?

1: No 2: Sí

AFGC12: Realizó actividad físico deportiva durante la administración de los tratamientos?

1: No realicé actividad físico deportiva 2: Sí, realicé actividad físico deportiva suave
3: Sí, realicé actividad físico deportiva de intensidad moderada 3: Sí, realicé actividad físico deportiva intensa

AFGC13: ¿Y tras finalizar el tratamiento (o los tratamientos), y en la actualidad?

1: No practiqué, ni practico actualmente 2: Practiqué, pero actualmente no
3: Actualmente realizo actividad físico deportiva suave o moderada 4: Actualmente realizo actividad físico deportiva intensa

ANEXO F: Scripts Python

```
1  \textit{Simulate.py}
2  # This code simulate and poblate the database for 50.000 responses to the questionnaire
3  import mysql.connector
4  import random
5  try:
6      conn = mysql.connector.connect(
7          host="localhost",
8          user="root",
9          password="*****"
10         database= "QUESTIONNAIRE")
11     cursor = conn.cursor()
12     query = """ SELECT question_id, question_code, response_data_type FROM Questions"""
13     cursor.execute(query)
14     rows = cursor.fetchall()
15     quest_id, quest_code, quest_datatype = zip(*rows) # Unpack the data into separate lists
16     response_id = 1
17     respondent_id = 1 # Generating random responses according to the question
18     while respondent_id < 50000:
19         for qid, qcode, qdatatype in zip(quest_id, quest_code, quest_datatype):
20             if qcode == 'sd1':
21                 response = random.randint(18, 89)
22             elif qcode == "sd2":
23                 response = random.randint(130, 198)
24             elif qcode == "sd3":
25                 response = random.randint(45, 100)
26             elif qdatatype == "boolean":
27                 response = random.randint(0, 1)
28             else:
29                 response = random.randint(1, 4)
30         newquery = """INSERT INTO Responses (response_id, response, respondent_id, submitted_at,+
31         question_code) VALUES (%s, %s, %s, CURRENT_TIMESTAMP, %s)"""
32         cursor.execute(newquery, (response_id, response, respondent_id, qcode))
33         response_id += 1
34         respondent_id += 1
35     conn.commit()
36 except mysql.connector.Error as e:
37     print("Error:", e)
38 finally:
39     if conn is not None:
40         cursor.close()
41         conn.close()
```

```

1
2 import mysql.connector
3 # Function to populate response tables for each section
4 def populate_section_responses(section_number, question_codes):
5     try:
6         conn = mysql.connector.connect(
7             host="xxxxxx",
8             user="root",
9             password="*****",
10            database="QUESTIONNAIRE" )
11        cursor = conn.cursor()
12        insert_query = f"""
13            INSERT INTO ResponsesInEachSection.ResponsesSection{section_number} (respondent_id, +
14            {'', '.join(f"{code}" for code in question_codes)}
15            VALUES ({', ' .join(['%s'] * (len(question_codes) + 1))})"""
16        fetch_responses_query = f"""
17            SELECT respondent_id, {'', '.join(question_codes)}
18            FROM Responses
19            WHERE question_code IN ({', ' .join(['%s'] * len(question_codes))})"""
20        cursor.execute(fetch_responses_query, question_codes)
21        responses = cursor.fetchall()
22        # Insert responses into the section-specific response table
23        cursor.executemany(insert_query, responses)
24        conn.commit()
25        cursor.close()
26        conn.close()
27    except (Exception, mysql.connector.Error) as error:
28
29 def fetch_section_question_codes():
30     try:
31         conn = mysql.connector.connect(
32             host="xxxxxx",
33             user="root",
34             password="*****",
35             database="QUESTIONNAIRE")
36         cursor = conn.cursor()
37         query = """SELECT section_code, GROUP_CONCAT(question_code) AS question_codes
38             FROM Questions
39             GROUP BY section_code
40             ORDER BY section_code;"""
41         cursor.execute(query)
42         question_codes_for_sections = cursor.fetchall() # Fetch as list of tuples
43         cursor.close()
44         conn.close()
45         return question_codes_for_sections
46     except (Exception, mysql.connector.Error) as error:
47         return None
48 section_question_codes = fetch_section_question_codes()
49 if section_question_codes:
50     for section_code, question_codes_array in section_question_codes:
51         populate_section_responses(section_code, question_codes_array)

```

```

1  import mysql.connector
2
3  # Establish connection to the original database
4  conn_orig = mysql.connector.connect(
5      host="localhost",
6      user="root",
7      password="*****",
8      database="QUESTIONNAIRE")
9
10 # Establish connection to the new database schema
11 conn_new = mysql.connector.connect(
12     host="localhost",
13     user="root",
14     password="*****",
15     database="QUESTIONNAIREALL")
16 cursor_orig = conn_orig.cursor()
17 cursor_new = conn_new.cursor()
18
19 # Fetch questions from the original table
20 cursor_orig.execute("""SELECT * FROM Questions""")
21 questions = cursor_orig.fetchall()
22
23 # Fetch responses from the original table
24 cursor_orig.execute("""SELECT * FROM Responses""")
25 responses = cursor_orig.fetchall()
26
27 # Create a new schema in the new database
28 cursor_new.execute("""CREATE SCHEMA IF NOT EXISTS
29 ResponsesCompleSchema""")
30
31 # Create a new table in the new schema for consolidated responses
32 cursor_new.execute("""
33     CREATE TABLE ResponsesAll(
34         respondent_id INT,
35         question_id INT,
36         question_code VARCHAR(255),
37         response INT)""")
38 # Iterate through responses and insert into the new table
39
40 for response in responses:
41     response_id, question_id, *individual_responses, respondent_id, submitted_at = response
42     for i, r in enumerate(individual_responses):
43         # question_id is in the first position in the questions tuple
44         cursor_new.execute("""
45             INSERT INTO ResponsesAll (respondent_id, question_id, question_code, response)
46             VALUES (%s, %s, %s, %s) """ , (respondent_id, questions[i][0], questions[i][1], r))
47 conn_new.commit()
48 cursor_orig.close()
49 cursor_new.close()
50 conn_orig.close()
51 conn_new.close()
52

```

ANEXO G: Cronograma del proyecto

La distribución de las principales tareas llevadas a cabo en este proyecto, cuyo desarrollo temporal y esfuerzo necesario se representa en el cronograma de la Figura VII.4, se presentan a continuación. Se pueden diferenciar 5 tareas principales:

- 1. Revisión sistemática de la bibliografía y estado del arte. Para obtener el conocimiento del dominio necesario para interpretar de la forma adecuada los datos y resultados con los que se ha trabajado, ha sido necesario llevar a cabo una revisión pormenorizada del estado del arte en actividad física y cáncer de mama.
- 2. Familiarización con las tecnologías de desarrollo e implementación. Para poder abordar el diseño, desarrollo e implementación conforme al contexto actual y recursos disponibles en la Universidad de Zaragoza, se consideraron diferentes alternativas tecnológicas. Finalmente, se optó por las más adecuadas y accesibles teniendo en cuenta su coste.
- 3. Desarrollo del cuestionario y el Sistema de Vigilancia. En esta tarea en primer lugar se realizaron diversos diseños de prototipos en papel para definir los requisitos del sistema. Posteriormente, se abordó la implementación del código y la realización de pruebas unitarias. Finalmente, se procedió al despliegue del sistema para poder abordar una fase de pruebas y validación con usuarios finales.
- 4. Pruebas y validación. Actualmente el sistema está disponible online y se continúa validando con usuarias con las que se ha contactado.
- 5. Documentación del trabajo. A lo largo del proyecto se han ido documentando las diferentes etapas. Por últimos, se ha elaborado esta memoria.

	Noviembre de 2023	Diciembre de 2023	1-15 Enero de 2024	15-30 Enero de 2024	Febrero de 2024	Marzo de 2024	Abril de 2024	Mayo de 2024	Junio de 2024
Revisión sistemática de bibliografía									
Diseño de preguntas									
Familiarización con las tecnologías de desarrollo e implementación									
Desarrollo del cuestionario y el Sistema de Vigilancia									
Pruebas y Validación									
Documentación de Trabajo									

Figura VII.4: Cronograma del trabajo realizado. Fuente: Elaboración propia.

El desarrollo completo de este proyecto se ha extendido desde octubre de 2023 hasta julio de 2024 y ha supuesto una carga total de trabajo de aproximadamente 600 horas/persona.