

CRITERIOS DE BEERS Y STOPP/START. REVISIÓN SISTEMÁTICA DE LA LITERATURA.

REALIZADO POR:

CAMILO ERNESTO NEMEGUÉN ARIAS

LAURA MANUELA JARAMILLO BUILES

TUTORES

CARLOS ALBERTO CALDERÓN OSPINA

CAMILO ALBERTO DOMÍNGUEZ DOMINGUEZ

TRABAJO REALIZADO PARA OPTAR POR EL TÍTULO DE TOXICÓLOGO CLÍNICO

2020

Identificación del proyecto

Institución académica: Universidad del Rosario

Dependencia: Escuela de Medicina y Ciencias de la Salud

Título de la investigación: Criterios de BEERS y STOPP/START. Revisión sistemática de la literatura.

Tipo de investigación: Revisión sistemática.

Investigador principal: Camilo Ernesto Nemeguén Arias

Laura Manuela Jaramillo Builes

Investigadores asociados: Carlos Alberto Calderón Ospina

Asesor clínico o temático: Carlos Alberto Calderón Ospina

Asesor metodológico: Camilo Alberto Domínguez Domínguez

“La Universidad del Rosario no se hace responsable de los conceptos emitidos por los investigadores en su trabajo, solo velará por el rigor científico, metodológico y ético del mismo en aras de la búsqueda de la verdad y la justicia”.

Agradecimientos:

A nuestras familias, amigos, compañeros de estudio, por la paciencia, el acompañamiento y orientación en cada uno de nuestros actos, lo cual podemos reflejar en este trabajo.

Contenido

1.	Introducción.....	10
1.1.	<i>Planteamiento del problema.....</i>	<i>10</i>
1.2.	<i>Justificación.....</i>	<i>11</i>
2.	Marco Teórico	12
3.	Pregunta de investigación	13
4.	Objetivos.....	13
4.1.	<i>Objetivo general.....</i>	<i>13</i>
4.2.	<i>Objetivos específicos.....</i>	<i>14</i>
5.	Metodología.....	14
5.1.	<i>Tipo y diseño de estudio.....</i>	<i>14</i>
5.2.	<i>Población.....</i>	<i>14</i>
5.3.	<i>Tamaño de muestra.....</i>	<i>14</i>
5.4.	<i>Criterios de selección.....</i>	<i>14</i>
5.4.1.	<i>Criterios de inclusión.....</i>	<i>14</i>
5.4.2.	<i>Criterios de exclusión.....</i>	<i>15</i>
5.5.	<i>Variables.....</i>	<i>15</i>
5.5.1.	<i>Variables dependientes.....</i>	<i>15</i>
5.5.2.	<i>Variables independientes.....</i>	<i>16</i>
5.6.	<i>Hipótesis.....</i>	<i>19</i>
5.6.1.	<i>Hipótesis nula.....</i>	<i>19</i>
5.6.2.	<i>Hipótesis alterna.....</i>	<i>19</i>
5.7.	<i>Plan de análisis.....</i>	<i>19</i>
5.8.	<i>Proceso de recolección de la información.....</i>	<i>19</i>
6.	Aspectos éticos.....	20
7.	Administración del proyecto.....	21
7.1.	<i>Cronograma.....</i>	<i>21</i>
7.2.	<i>Presupuesto.....</i>	<i>23</i>
8.	Métodos.....	23

8.1	<i>Evaluación de sesgos</i>	26
9.	Resultados	29
10.	Discusión	37
11.	Referencias	47
12.	Anexos	56

Resumen:

Datos de la OMS en 2014, afirman que entre el 2015 y 2050, la proporción de la población mundial con más de 60 años de edad pasará de 900 millones hasta 2000 millones. Los reportes han indicado que la prescripción inadecuada de medicamentos aumenta con la edad, llegando a ser tan alta como 70% (3), y entre estos pacientes, se ha encontrado que un 10 a 15% de los ingresos hospitalarios se relacionan con eventos adversos a medicamentos (4), con el consecuente incremento de la morbilidad (5–8). La aparición de herramientas, cuyo fin sea identificar formulaciones potencialmente inapropiadas, abre una luz a la práctica clínica en pro de la salud del anciano y la disminución de los riesgos a los cuales puede estar expuesto por sus tratamientos. A pesar de contar con muchas de estas herramientas su aplicación en ciertas poblaciones está limitada como consecuencia de la diferencia en los medicamentos a los que se pueden acceder de acuerdo a las condiciones en cada país. Es por esto que, al determinar la utilidad de estas herramientas, se espera abrir una puerta para la realización de estudios que permitan llegar a un consenso sobre la aplicabilidad de las mismas en la población geriátrica colombiana.

Metodología:

Se realizó una revisión sistemática de ensayos clínicos controlados.

La búsqueda inicial arrojó 4.291 resultados. Después de eliminar los duplicados, y evaluar los criterios de inclusión y exclusión, 24 estudios fueron seleccionados en texto completo. Luego fueron sometidos a evaluación de sesgos para encontrar la mejor evidencia disponible. Posteriormente se realizó una completa extracción de datos basados en variables como edad, sexo, número de medicamentos, estancia hospitalaria, consultas a urgencias, reingresos hospitalarios, reacciones adversas, mortalidad, costos de la atención. Cuando se encontraron datos que no estaban descritos en los estudios, se procedió a solicitarlos a los respectivos autores. Se encontró una alta heterogeneidad en la realización de los estudios y el reporte de resultados, lo cual limitó nuestro análisis.

Resultados

Luego del análisis de cada uno de los artículos y teniendo en cuenta las limitaciones encontradas en cada uno de ellos, encontramos que la aplicación de herramientas como los criterios de Beers y los criterios de STOPP/START, no mostró una clara repercusión en los resultados de parámetros como la mortalidad; no se halló una diferencia estadísticamente significativa la disminución del número de medicamentos posterior a la intervención, los días de estancia hospitalaria, los reingresos hospitalarios y el número de consultas a urgencias. Sin embargo se encontraron resultados favorecedores respecto a la disminución en el número de eventos adversos, número de reacciones adversas, porcentaje de caídas, disminución en el número de interacciones medicamentosas y disminución en los costos de atención en salud

Conclusión:

Nuestra revisión sistemática señaló que no existen estudios clínicos controlados de buena calidad metodológica que hayan abordado la ocurrencia de desenlaces como número de medicamentos, estancia hospitalaria, consultas a urgencias, reingresos hospitalarios, reacciones adversas, mortalidad, costos de la atención en pacientes ancianos como desenlace primario. Respecto a las dos herramientas analizadas (Beers y STOPP/START), no existe claridad si su implementación da como resultado desenlaces clínicamente significativos. Se requieren estudios más robustos, con informes detallados, que permitan sacar conclusiones más contundentes, con el objetivo de orientar hacia la o las intervenciones con resultados más favorables, y su incorporación a la práctica clínica, de acuerdo al grupo poblacional al cuál serán aplicadas.

Palabras claves: Beers Criteria, STOPP/START criteria, potentially inappropriate medication, elderly, geriatric.

Abstract

WHO data of 2014, state that between 2015 and 2050, the proportion of the world's population over 60 years of age will go from 900 million to 2000 million. Reports have indicated that inadequate prescription of medications increases with age, becoming as high as 70% (3), and among these patients, it has been found that 10 to 15% of hospital admissions are related to adverse events to medications (4), with the consequent increase in morbidity and mortality (5–8). The emergence of tools, whose purpose is to identify potentially inappropriate formulations, opens a light to clinical practice for the health of the elderly and the reduction of the risks to which they may be exposed by their treatments. Despite having these tools, their application in certain populations is limited as a result of the difference in the medications that can be accessed according to the conditions in each country. That is why, when determining the usefulness of these tools, it is expected to open a door for studies that allow reaching a consensus on their applicability in the Colombian geriatric population.

Methodology:

A systematic review of controlled clinical trials was performed.

The initial search yielded 4,291 results. After eliminating duplicates, and evaluating inclusion and exclusion criteria, 24 studies were selected in full text. The risk of bias evaluation was made in order to find the best available evidence. Subsequently, a complete data extraction was carried out based on variables such as age, sex, number of medications, hospital stay, emergency consultations, hospital readmissions, adverse reactions, mortality, care costs. When the data were not described in the studies, they were requested from the respective authors. A high heterogeneity was found in the conduct of the studies and the report of results, which limited our analysis.

Results

After analyzing each of the articles and taking into account the limitations found in each of them, we found that the application of tools such as the Beers criteria and the STOPP / START

criteria did not show a clear impact on the results of parameters such as mortality; No statistically significant difference was found in the decrease in the number of medications after the intervention, the days of hospital stay, hospital readmissions and the number of emergency consultations. However, favorable results were found regarding the decrease in the number of adverse events, number of adverse reactions, percentage of falls, decrease in the number of drug interactions and decrease in health care costs

Conclusion:

Our systematic review indicated that there are no controlled clinical studies of good methodological quality that have addressed the occurrence of outcomes such as number of medications, hospital stay, emergency consultations, hospital readmissions, adverse reactions, mortality, costs of care in elderly patients as an outcome primary. Regarding the two tools analyzed (Beers and STOPP/START), there is no clarity if their implementation results in clinically significant outcomes. More robust studies are required, with detailed reports, which allow more conclusive conclusions to be drawn, with the aim of orienting towards the best intervention (s), and their incorporation into clinical practice, according to the population group to which they will be applied.

Keywords: Beers Criteria, STOPP/START criteria, potentially inappropriate medication, elderly, geriatric.

1. Introducción

1.1. Planteamiento del problema

La Organización Mundial de la Salud estimó, en el año 2014, que una de cada nueve personas era anciana (tenía 60 años o más) y se espera que este valor se incremente a una de cada cinco personas en el año 2050 (1), tanto en países desarrollados como en vía de desarrollo. Teniendo en cuenta que la población mundial está envejeciendo rápidamente, entre el 2015 y 2050, la proporción de la población mundial con más de 60 años de edad pasará de 900 millones hasta 2000 millones, lo que representa un aumento del 12% al 22% (2). Este aumento progresivo en la proporción de población anciana, va acompañado de un aumento en el número de medicamentos prescritos, viéndose expuestos a una polifarmacia – definida como el uso de cinco o más medicamentos de manera diaria en un mismo paciente (1,4), y/o la administración de otros medicamentos fuera de los requeridos para el manejo de la condición clínica del paciente (1) - y por ende la posibilidad de la prescripción inadecuada de los medicamentos. (1, 2), la cual puede comprometer la seguridad del paciente, al incrementar el riesgo de presentar diferentes eventos adversos, entre los cuales podemos encontrar el compromiso de la capacidad cognitiva, el delirium, las caídas, incontinencia urinaria, entre otras (4); y la existencia de interacciones medicamentosas (5,11,12).

Estudios han reportado que la prescripción inadecuada de medicamentos aumenta con la edad, llegando a ser tan alta como 70% en algunos hospitales geriátricos (3). Entre estos pacientes, se ha encontrado que un 10 a 15% de los ingresos hospitalarios se relacionan con eventos adversos a medicamentos (4), con el consecuente incremento de la morbimortalidad (5–8).

Con mayor frecuencia observamos una tendencia a tratar farmacológicamente todas las patologías, esto puede desencadenar una "cascada" de efectos en el organismo, los cuales finalmente también resultan medicados (10); además todas aquellos cambios secundarios al proceso de envejecimiento, terminan dentro del listado de condiciones a tratar, lo cual sumado a la anteriormente descrito, termina convirtiéndose en un gran reto para en

profesional clínico (4). Así, la concepción de salud en la actualidad y más aún en el paciente anciano va condicionada a la necesidad del uso de un medicamento para mantener su estado de bienestar (9). Además, las grandes farmacéuticas, en su afán de innovar, crean medicamentos o modifican los ya existentes, aportando a la medicalización de lo ya reconocido como fisiológico (20).

Hay múltiples factores presentes en las personas ancianas que aumentan el riesgo de polifarmacia. Entre dichos factores asociados, se encuentran los cambios de las propiedades farmacodinámicas y farmacocinéticas de los medicamentos, con lo cual se puede ver alterado su comportamiento dentro del organismo, afectando a su vez los efectos terapéuticos y pudiendo incrementar las reacciones adversas (11). Adicionalmente las funciones fisiológicas en el anciano se encuentran deterioradas, y esto, acompañado de múltiples comorbilidades que generalmente se sobreponen, suele aumentar el riesgo de reacciones adversas a medicamentos para esta población en particular (1, 11).

1.2. Justificación

Es conocido que la pirámide poblacional actualmente apunta a un incremento cada vez mayor de la población anciana, con una proporción de mayores de 60 años de 12% en el año 2015 hasta un 22% al año 2050 (2), y que este grupo poblacional es más susceptible a una mayor polifarmacia, lo cual conlleva a más riesgo de eventos adversos e incremento de la morbimortalidad (4 - 8). Esto, sumado a los cambios fisiológicos que se dan en el anciano, influye negativamente en las propiedades farmacodinámicas y farmacocinéticas de los medicamentos, lo cual se añade a los riesgos anteriormente mencionados (1, 11).

Aunque se conocen muchas herramientas creadas para la detección de prescripciones inadecuadas de medicamentos (por ejemplo los criterios de Beers y los criterios STOPP/START), no son utilizadas por todo el personal médico que atiende a pacientes adultos mayores, esta falta de utilización de las herramientas pueden deberse al desconocimiento de su existencia, a la complejidad de algunas de ellas o incluso a la dificultad para su acceso.

Por otra parte, se desconoce la efectividad de estas herramientas para impactar positivamente (o no) desenlaces clínicamente relevantes en ancianos como los eventos adversos (p. ej. Caídas), la mortalidad y los reingresos hospitalarios entre otros. Es por esto que, al determinar la utilidad de estas herramientas, se espera conocer un poco más sobre la verdadera utilidad de las mismas, promoviendo la realización de estudios adicionales que permitan llegar a un consenso sobre la aplicabilidad de ellas (o la generación de nuevas herramientas) en la población geriátrica colombiana.

2. Marco Teórico

Las herramientas más comunes para evaluar la prescripción inapropiada de medicamentos, son los criterios de Beers y los criterios STOPP/START (16–19).

Los criterios de Beers fueron desarrollados en 1991, siendo acogidos y actualizados por la *American Geriatrics Society* para ser aplicada en pacientes geriátricos de Estados Unidos y corresponde a un listado de medicamentos considerados potencialmente inapropiados, y que deben ser evitados, prescritos a dosis reducidas o con monitoreo cuidadoso durante su uso. La aplicación de esta herramienta tiene una mayor prevalencia en la población norteamericana y hasta la fecha los estudios han demostrado beneficios en variables como la disminución reacciones adversas y eventos adversos, sin embargo la calidad metodológica y cantidad de estudios realizados haciendo uso de ella, no han permitido ser conclusivos respecto al potencial de su uso.

Los criterios STOPP/START se hicieron públicos en el año 2008 por Gallagher et al. en Irlanda y en el 2009 se publica la primera versión en español. Estos son avalados y actualizados por la *European Union Geriatric Medicine Society*. La intención es recoger los errores de medicación más comunes, pero además comprende errores de omisión en las personas mayores, detectando la falta de uso de medicamentos potencialmente indicados y haciendo relación con los diagnósticos de los pacientes, agrupando estos errores por sistemas

fisiológicos (21). Esto incluye 81 indicadores (STOPP) dirigidos a identificar interacciones fármaco – fármaco y fármaco – enfermedad (que potencialmente puedan conducir a efectos adversos) y duplicación terapéutica (16). A diferencia de los criterios de Beers, esta guía cuenta con el subtítulo START que incorpora además 34 oportunidades de prescripción basadas en la evidencia (16, 21). Esta herramienta se encuentra mayormente difundida en países del viejo mundo, y algunos estudios han logrado identificar resultados positivos en variables como la detección de medicación potencialmente inadecuada, el número de interacciones medicamentosas, el número de reacciones adversas, o los costos para el sistema de salud. Pero, al igual que con los criterios de Beers, la calidad y cantidad de los estudios realizados no es la mejor para dar datos conclusivos respecto a su potencial utilidad.

Los estudios clínicos realizados solo están circunscritos a población Europea (23, 24, 26, 28 - 31, 35 - 40, 43 - 46) y de América del Norte (25, 27, 32, 33, 34, 41, 42), faltando estudios en población Latinoamericana. En nuestro país no se contamos con ningún estudio que valide estos criterios, por lo que desconocemos su verdadero impacto en desenlaces clínicamente significativos en nuestros pacientes ancianos, además aún contamos con medicamentos que en otras latitudes ya han salido del mercado.

3. Pregunta de investigación

¿Cuál es la utilidad de los criterios de BEERS y los criterios STOPP/START como herramientas para mejorar la prescripción en pacientes mayores de 65 años?

4. Objetivos

4.1. Objetivo general

Describir la utilidad de los criterios de BEERS y los criterios STOPP/START como herramientas para mejorar la prescripción en pacientes mayores de 65 años.

4.2. Objetivos específicos

1. Describir las características socio-demográficas de los pacientes objeto del estudio.
2. Describir el impacto de la aplicación de los criterios Beers y/o STOPP/START en los siguientes parámetros: mortalidad, estancia hospitalaria, eventos adversos, reingresos hospitalarios, y polifarmacia.
3. Comparar la eficacia de los criterios Beers Vs. STOPP/START para mejorar la prescripción en pacientes ancianos de acuerdo a los siguientes parámetros: mortalidad, estancia hospitalaria, eventos adversos, reingresos hospitalarios, y polifarmacia.

5. Metodología

5.1. Tipo y diseño de estudio:

Revisión sistemática de la literatura.

5.2. Estrategia PICO

Pacientes adultos de ambos sexos de 65 años o más (P). La intervención a evaluar fueron los Criterios de Beers y/o STOPP/START (I) para optimizar la medicación en ancianos. Como comparador (C) se tomó el cuidado habitual y/o los Criterios de Beers y/o STOPP/START según el estudio empleado. Como desenlaces (O) evaluados se incluyeron los siguientes: reacciones adversas, mortalidad, reingresos hospitalarios...

5.3. Tamaño de muestra

Se incluyeron todos los artículos de estudios clínicos controlados que hayan utilizado los criterios de Beers y/o STOPP/START como intervención.

5.4. Criterios de selección

5.4.1. Criterios de inclusión

- Estudios clínicos aleatorizados

- Artículos que incluyan la aplicación de los criterios de Beers y/o los criterios STOPP/START para optimizar la medicación en pacientes ancianos
- Estudios realizados en pacientes de 65 años o más (población).

5.4.2. Criterios de exclusión

- Estudios observacionales.
- Estudios clínicos no aleatorizados.
- Artículos en los que no se hayan empleado los criterios de Beers y/o STOPP/START.
- Estudios sin resultados publicados.

5.5. Variables

5.5.1. Variables dependientes

- Número de medicamentos al ingreso
- Número de medicamentos posterior a la intervención/hospitalización
- Presencia de eventos adversos al ingreso
- Presencia de eventos adversos luego de la intervención.
- Días de estancia hospitalaria secundario a evento adverso
- Número de reacciones adversas
- Índice de mortalidad posterior a la aplicación de los criterios.
- Reingresos hospitalarios.
- Costos definidos por los autores de los estudios.
- Número de interacciones medicamentosas clínicamente significativas encontradas antes y después de la intervención
- Comparador más usado en los estudios
- Inclusión de pacientes con demencia
- Porcentaje de pacientes que consultaron al servicio de urgencias antes de la terminación del estudio

- Tiempo promedio de los estudios

5.5.2. Variables independientes

- Edad.
- Género.
- Uso de los criterios de Beers y STOPP/START el cual fue definido como la mención explícita de los autores del uso de dichos criterios en la metodología de los estudios.

Tabla 1. Definición de variables

VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERATIVA	NATURALEZA	NIVEL DE MEDICIÓN	POSIBLES VALORES
Edad	Tiempo transcurrido a partir del nacimiento de un individuo.	Para el estudio clínico la edad mínima para ser incluido dentro de la población a estudio es de 65 años de edad. Se tomarán los datos de los ensayos clínicos para agrupar por quinquenios a los pacientes y con estos datos evaluar variables como promedio de medicamentos utilizados, promedio de eventos adversos/interacciones encontrados	Cualitativa	Ordinal	65 - 69 años 70 – 74 años 75 – 79 años >80 años
Sexo	Identidad sexual de los seres vivos, la distinción que se hace entre Femenino y Masculino, hombre/mujer	Se tomaran datos para determinar por género el número de pacientes participantes en los estudios	Cualitativa	Nominal	Masculino/ femenino.
Número de medicamentos formulados	Cantidad de xenobióticos	Se tomarán los datos del promedio de medicamentos	Cualitativa	Ordinal	1-3 3-5 >5

	consumidos por el paciente	potencialmente inapropiados formulados antes de la intervención.			
Presencia de eventos adversos	Daño no intencionado al paciente, que ocurre como consecuencia de la utilización de un dispositivo médico.	Se cuantificará la cantidad de eventos adversos antes de la intervención	Cualitativa	Ordinal	0 1-3 >3
Presencia de eventos adversos luego de la intervención	Daño no intencionado al paciente, que ocurre como consecuencia de la utilización de un dispositivo médico.	Se cuantificará la cantidad de eventos adversos posterior a la aplicación de los criterios	Cualitativa	Ordinal	0 1-3 >3
Días de estancia hospitalaria secundario a evento adverso	Se refiere al espacio de tiempo que invierte un paciente en condición de hospitalizado en las instalaciones de un hospital.	Se identificará el promedio de días de hospitalización como consecuencia de la medicación potencialmente inapropiada, teniendo en cuenta la influencia de estos en calidad de vida del pacientes, costos para el paciente y para el sistema de salud)	Cualitativa	Ordinal	0 1-3 3-5 >5
Índice de mortalidad posterior a la aplicación de los criterios.	Valor numérico que expresa la relación entre el número de muertes que se producen en un período de tiempo y el número total de individuos de una población.	Se tomarán datos de numero de eventos que finalizaron en mortalidad luego de la intervención	Cuantitativa	Razón	Número de pacientes intervenidos/ Mortalidad luego de la intervención
Reingresos hospitalarios.	Considerado como aquel ingreso programado o urgente que se	Cantidad de reingresos hospitalarios secundarios a la toma de medicación potencialmente	Cualitativa	Ordinal	0 1-3 >3

	produce tras el alta del paciente en un plazo de tiempo determinado y por un motivo o no relacionado con el diagnóstico de su primer ingreso.	inapropiada antes de la intervención			
Costos definidos por los autores de los estudios.	Valor de un paciente hospitalizado/ manejo ambulatorio, un día (pte/día). Equivalente paciente día como el valor de la unidad relativa.	Se tomará el promedio de costos que implica el manejo de complicaciones en pacientes con medicación potencialmente inapropiada antes y después de la intervención	Cuantitativa	Razón	Promedio de costos en pesos colombianos invertidos en el tratamiento por paciente
Número de medicamentos posterior a la intervención/ hospitalización	Cantidad de xenobióticos consumidos por el paciente	Se identificará el promedio de número de medicamentos formulados posterior a la intervención.	Cualitativa	Ordinal	0 1-3 >3
Número de interacciones encontradas antes de la intervención	Cantidad de reacciones entre dos (o más) medicamentos o entre un medicamento y un alimento o bebida. Puede disminuir o aumentar la acción de un medicamento o causar efectos secundarios no deseados.	Se tomarán datos para cuantificar la cantidad de interacciones medicamentosas presentes en las formulaciones antes de la intervención	Cualitativa	Ordinal	0 1-3 >3

5.6. Hipótesis

5.6.1. Hipótesis nula

La aplicación de herramientas como los criterios de Beers y STOPP/START para identificación de medicación potencialmente inapropiada, en mayores de 65 años, tienen impacto en desenlaces clínicamente significativos (p. ej. mortalidad).

5.6.2. Hipótesis alterna

La aplicación de herramientas como los criterios de Beers y STOPP/START para identificación de medicación potencialmente inapropiada, en mayores de 65 años, no tienen impacto en desenlaces clínicamente significativos (p. ej. disminución en el número de eventos adversos o mortalidad).

5.7. Plan de análisis

Se realizó la extracción de datos de cada uno de los artículos seleccionados al finalizar la revisión general de las bases de datos utilizadas.

5.8. Proceso de recolección de la información

Se realizó la búsqueda en las siguientes bases de datos con fecha límite de búsqueda de Abril 30 de 2019: US National Library of Medicine (PubMed); MEDLINE (Ovid SP); Cochrane Central Register of Controlled Trial (CENTRAL); EBSCO e- journals; Clinicaltrials.gov; Clinicalkey; Springer Link; Web of Science; Biblioteca Virtual en Salud (BVS); EMBASE (Ovid SP); Scientific Electronic Library Online (SCIELO); EBSCO MedicLatina; Effective Practice and Organisation of Care (EPOC); The Cumulative Index to Nursing and Allied Health Literature (CINAHL). No hubo restricción de idioma.

6. Aspectos éticos

No se tuvieron personas sino artículos médicos como unidad de análisis, por lo que no se incurrió en aspectos éticos relacionados con intervención en pacientes. Se limitó el acceso de los instrumentos de investigación únicamente a los investigadores según Artículo 8 de la Resolución 008430 de 1993 del Ministerio de Salud.

Todos los integrantes del grupo de investigación estuvimos prestos a dar información sobre el estudio a entes organizados, aprobados e interesados en conocerlo siempre y cuando sean de índole académica y científica, preservando la exactitud de los resultados y haciendo referencia a datos globales y no a pacientes o instituciones en particular.

El estudio se realizó con un manejo estadístico imparcial y responsable.

Se incluye la tabla de las bases de datos en las cuales se realizó la búsqueda, registrando en ellas todos y cada uno de los títulos de artículos que fueron encontrados y que fueron sometidos a los criterios de inclusión y exclusión, buscando asegurarnos que se encontraron todos los estudios susceptibles de ser analizados.

No existe ningún conflicto de interés por parte de los autores del estudio que deba declararse.

7. Administración del proyecto

7.1. Cronograma

Semana \ Mes		Enero				Febrero				Marzo				Abril				Mayo				Junio				Julio				Agosto				Septiembre				Octubre				Noviembre				Diciembre			
Actividad		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4				
1	Obtener personal.																																																
2	Formación del personal																																																
3	Escritura del protocolo.																																																
4	Búsqueda de estudios y extracción de datos.																																																
5	Análisis																																																
6	Elaboración de informe final																																																
8	Envío del informe final a los pares evaluadores																																																
Tiempo total requerido:		88 semanas																																															

Semana \ Mes		Enero				Febrero				Marzo				Abril				Mayo				Junio				Julio				Agosto				Septiembre				Octubre				Noviembre				
Actividad		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4									
1	Obtener personal.																																													
2	Formación del personal																																													
3	Escritura del protocolo.																																													
4	Búsqueda de estudios y extracción de datos.																																													
5	Análisis																																													
6	Elaboración de informe final.																																													
8	Envío del informe final a los pares evaluadores.																																													
Tiempo total requerido:		88 semanas																																												

7.2. Presupuesto

RUBROS	FUENTES DE FINANCIACION	FUENTES DE FINANCIACION	TOTAL
	INVESTIGADOR 1	INVESTIGADOR 2	
Personales	\$8.000.000	\$8.000.000	\$16.000.000
Software	\$100.000	\$100.000	\$200.000
Papelería	\$100.000	\$100.000	\$200.000
Servicios Técnicos	\$200.000	\$200.000	\$400.000
Internet	\$100.000	\$100.000	\$200.000
Material Bibliográfico	\$2.500.000	\$2.500.000	\$5.000.000
Publicaciones	\$750.000	\$750.000	\$1.500.000
Asesorías	\$750.000	\$750.000	\$1.500.000
Traducción	\$750.000	\$750.000	\$1.500.000
TOTAL	\$13.250.000	\$13.250.000	\$26.500.000

8. Métodos.

Esta revisión sistemática fue realizada bajo la metodología Cochrane, siguiendo para ello los pasos descritos en el Cochrane Handbook for Systematic Reviews of Interventions Versión 5.1.0 (22). Dos revisores independientes (CN y MJ) realizaron la búsqueda en las bases de datos antes relacionadas, utilizando como términos de búsqueda MeSH los siguientes: Beers Criteria, STOPP/START criteria, potentially inappropriate medication, elderly, y geriatric, con fecha límite abril 30 de 2019. La búsqueda inicial arrojó un total de de 4291 artículos distribuidos de la siguiente manera: (Ver Anexo 1. Tabla base de datos).

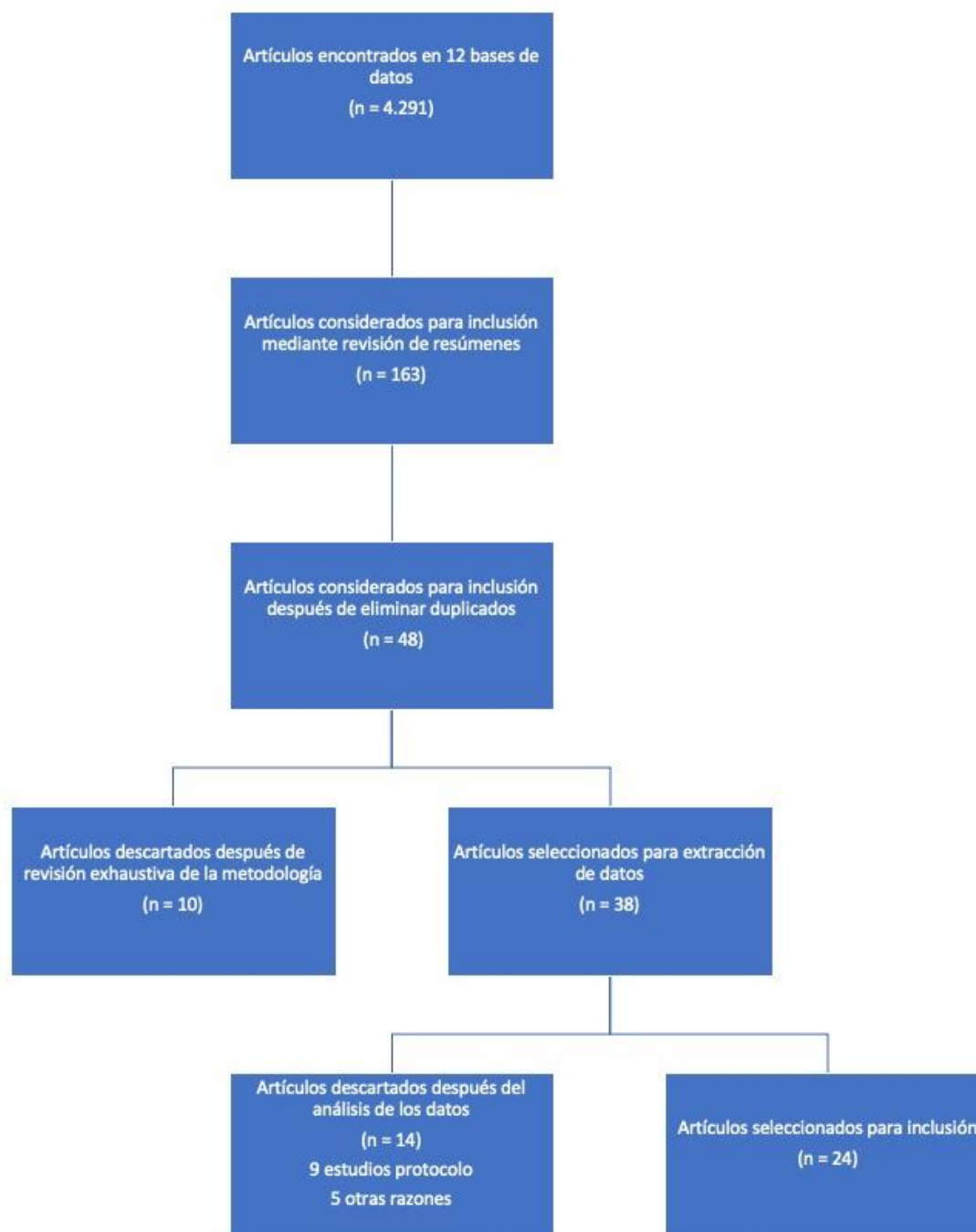
- US National Library of Medicine (PubMed): total 929 artículos.
- MEDLINE (Ovid SP): total 589 artículos.
- Cochrane Central Register of Controlled Trial (CENTRAL): total 116 artículos.
- EBSCO e- journals: total 416 artículos.
- Clinicaltrials.gov: total 69 artículos.
- Clinicalkey: total 261 artículos.
- Springer Link: total 67 artículos
- Web of Science: total 1221 artículos.
- Biblioteca Virtual en Salud (BVS): total 374 artículos.

- EMBASE (Ovid SP): total 211 artículos.
- Scientific Electronic Library Online (SCIELO): total 22 artículos.
- EBSCO MedicLatina: total 16 artículos.
- Effective Practice and Organisation of Care (EPOC): total 0 artículos.
- The Cumulative Index to Nursing and Allied Health Literature (CINAHL): no se encontraron artículos libres, ya que requerían previa suscripción a la revista.

Después de eliminar los duplicados, y posterior a la revisión individual de cada uno de los *abstract* en busca de artículos que cumplieran con los criterios de selección, se obtuvieron 48 artículos. 24 artículos más fueron eliminados por no corresponder a un ensayo clínico aleatorizado o, no haber empleado los criterios de Beers y/o STOPP/START, por lo que finalmente fueron seleccionados 24 artículos en texto completo para el análisis cualitativo. La selección de artículos en todas las fases fue realizada por dos evaluadores independientes (CN y MJ), y cualquier discrepancia fue resuelta por un tercer evaluador (CC). En la figura 1 se presenta el diagrama PRISMA para la selección de estudios en la presente revisión sistemática.

Para aquellos estudios que contaban con información incompleta se realizó la solicitud de los datos faltantes al autor de correspondencia por correo electrónico. (ver Anexo 2. Formato de carta solicitud de información adicional y Anexo 3. Formato carta solicitud artículo completo). Posteriormente se realizó la extracción de datos empleando para ello un formato diseñado para ese fin por los cinco autores del estudio. (Ver Anexo 4. Tabla de extracción de datos). Seguidamente se realizó la evaluación del riesgo de sesgos para cada uno de los estudios incluidos (ver Anexo 5. Tabla de evaluación de sesgos) basados en el Cochrane Handbook for Systematic Reviews of Interventions Versión 5.1.0 (22). Toda la información obtenida fue depositada en el programa REVMAN 5-3 de libre acceso, para su respectivo análisis.

Figura 1. Diagrama PRISMA para la selección de estudios en la presente revisión sistemática



8.1 Evaluación de sesgos

Se realizó el análisis de sesgos de selección, realización, detección, desgaste, notificación y otros sesgos encontrados (p. Ej. la coincidencia de ser autor de las herramientas evaluadas y del estudio en cuestión) en los 24 estudios seleccionados, basados en el Cochrane Handbook for Systematic Reviews of Interventions Versión 5.1.0 (22). Los sesgos de selección y detección fueron incluidos en la sección de resultados.

Sesgo de realización

De los 24 estudios incluidos, solo cuatro describieron las medidas de cegamiento a los participantes y el personal del estudio, los cuales fueron el de García-Gollarte et al. (24), el de Van der Linden et al. (26), el de Terrell et al. (27), y el de Pitkälä et al. (30). En el estudio de Santolaya-Perrín et al. (23), participaron 4 centros, de los cuales solo uno tuvo protección de la identidad del paciente participante, pero no especifican el enmascaramiento de los demás centros. De los 19 estudios restantes, en 10 de ellos no se describió si se realizaron medidas de cegamiento [O'Brien et al. (28), Schmader et al. (32), Ramsdale et al. (33), Griffey et al. (34), Saltvedt et al. (36), O'Connor et al. (37), Raebel et al. (41), Dalleur et al. (44), Pope et al. (45), Gallagher et al. (46)], y en los 9 restantes, estas no fueron realizadas [Price et al. (25), Wouters et al. (29), Spinewine et al. (31), Frankenthal et al. (35), O'Sullivan et al. (38), O'Connor et al. (39), Gallagher et al. (40), Cossette et al. (42), Campins et al. (43)].

Sesgo de desgaste

En 9 de los 24 estudios no se describen los abandonos ni exclusiones del análisis del estudio (23, 28, 32 - 34, 37, 41, 45, 46). De los 15 estudios restantes, en el estudio de García-Gollarte F et al. (24) que contó con un total de 1018 participantes, durante los 6 meses de intervención, 302 pacientes no pudieron seguirse (101 murieron, 130 fueron dados de alta y 71 no quisieron participar o necesitaban cuidado paliativo). En el estudio de Price M, et al. (25) una de las clínicas elegidas se retiró por cuestiones técnicas, pero las restantes 8 clínicas participaron activamente. En el estudio de Van der Linden L, et al. (26) de los 438 pacientes

seleccionados, sólo 172 (39.26%) participó en el análisis, de ellos 91 (52.9%) conformaron el grupo de intervención. En el estudio de Terrell K.M, et al. (27) sólo el 69,21% del total participaron en el estudio, de los cuales el 51,27% pertenecía al grupo de intervención. En el estudio de Wouters H, et al (29), de los 426 participantes, el 54.69% pertenecía al grupo de intervención (n= 233), la pérdida al momento del seguimiento fue de 14% en el grupo control y 14.6% en el grupo intervención. En el estudio de Pitkälä KH, et al. (30), 41 pacientes (18.1%) se perdieron al seguimiento a los 6 meses, y el desgaste total del estudio a los 12 meses fue de 63 residentes (27.8%) debido a 63 muertes (39 en el grupo intervención y 24 en el grupo control). Respecto al estudio de Spinewine A, et al. (31) el 92,7% de los integrantes del grupo de intervención y el 92,2% de los del grupo control, completaron el seguimiento a un año. En el estudio de Frankenthal D, et al (35), de los 359 pacientes aleatorizados, el grupo final de intervención comprendió 160 (44.56%) y 146 (40.66%) en el grupo final de control, el 14,76% restante está comprendido por pacientes que murieron o dejaron el estudio. En el estudio de Saltvedt I, et al (36) todos los análisis se realizaron dos veces: primero incluyendo todos los pacientes incluidos en el estudio, y segundo excluyendo aquellos que murieron durante la estancia hospitalaria (6% del grupo de unidad geriátrica y 13% de la unidad de medicina general), presentando únicamente el segundo análisis. En el estudio de O'Sullivan D, et al (38), de los 737 pacientes incluidos, 34 pacientes murieron, 17 (4,52%) en el grupo control, 17 (4,70%) en el grupo intervención; se incluyeron estos pacientes en el análisis final en base a la adherencia al principio intención-de-tratar. En el estudio de O'Connor MN, et al. (39), de los 732 pacientes incluidos, 20 fallecieron, 11 (3.05%) en el grupo intervención y 9 (2.41%) en el grupo control, los datos fueron analizados en base a intención-de-tratar, para que ningún dato de los pacientes inscritos fuera excluido. Para el estudio de Gallagher PF, et al (40) se incluyeron 400 pacientes, 200 en cada grupo, 18 pacientes murieron antes de que se realizara la primera medición de resultados, por lo que los datos relevantes fueron excluidos del análisis, finalizaron el estudio 89% del grupo de control y el 90% del grupo de intervención. En el estudio de Cossette B, et al (42) 15% de los pacientes aleatorizados solicitaron remoción de sus datos del análisis; el mayor número de exclusiones en el grupo intervención se debió a siete hospitalizaciones con altas en el mismo día, ya que las

intervenciones no pudieron realizarse en estos pacientes, esta exclusión no fue aplicada al grupo control. 126 hospitalizaciones fueron analizadas en el grupo intervención y 128 en el grupo control. En el estudio de Campins L, et al. (43), entre los pacientes reclutados, 13 (2.58%) murieron durante los 12 meses de seguimiento (6 en el grupo control y 7 en el grupo intervención) y fueron excluidos del análisis de costos. Para el estudio de Dalleur O, et al (44), de los 158 pacientes inicialmente reclutados, 12 pacientes debieron ser excluidos , 9 (11.1%) del grupo de control y 3 (3.89%) del grupo de intervención, resultando en 146 pacientes para análisis.

Sesgo de notificación.

Los resultados primarios y secundarios definidos previamente fueron reportados en su totalidad en 22 de los 24 estudios (24-36, 38-46); en el estudio de Santolaya-Perrin R, (23) en el proyecto se definió como variable principal la visita a urgencias e ingresos hospitalarios por paciente/año a los 12 meses de la intervención, pero en el artículo solo se presentaron los resultados de una variable secundaria que era los pacientes con prescripciones potencialmente inapropiadas (PPI). Para la estimación de la prevalencia de PPI solo se tuvo en cuenta el grupo intervención.

Otros sesgos.

Consideramos que se presenta un sesgo por autor en el momento de evaluar la efectividad de la herramienta utilizada, teniendo en cuenta que Gallagher describió los criterios STOPP/START y se encuentra participando como autor a Gallagher P, et al. en 3 artículos (37, 39, 40), Gallagher J, et al. en 2 artículos (28, 46) y a ambos autores en 1 artículo (38).

9. Resultados

Nuestro protocolo fue registrado en PROSPERO, ID 14319.

En este apartado se incluyen los resultados de los sesgos de selección y detección.

Sesgo de selección.

De los 24 estudios, 12 de ellos (23, 24, 27,30 - 32, 35, 40 – 42, 44, 45), cumplieron con una adecuada secuencia de la aleatorización, así como una adecuada ocultación de la asignación a los grupos del estudio. De los 12 estudios restantes, los estudios de Price et al. (25) y el de O'Connor et al. (39), cumplieron con la secuencia de aleatorización, pero no tuvieron cegamiento de los participantes del estudio. Los estudios de Van der Linden et al. (26), de Griffey et al. (34), y el de O'Sullivan et al. (38), no cumplieron con una adecuada secuencia de aleatorización y no describen la ocultación. Los estudios de O'Brien et al. (28), de Wouters et al. (29), el de Saltvedt et al. (36), y el de Gallagher et al. (46) cumplieron con la secuencia de aleatorización, pero no describen si hubo ocultación de la asignación. En el estudio de O'Connor et al. (37) no se logró obtener datos de las características de asignación ya que se cuneta únicamente con el *abstract* del estudio. Por último, los estudios de Ramsdale et al. (33) y el de Campins et al. (43) no describen ninguna de las características de la asignación.

Sesgo de detección.

De los 24 estudios incluidos, seis describieron el cegamiento realizado a los evaluadores del estudio sobre la intervención que recibieron los participantes, los cuales fueron el de Wouters et al. (29), Schmader et al. (32), Ramsdale et al. (33), Griffey et al. (34), Dalleur et al. (44) y el de Pope et al. (45). En el estudio de Spinewine et al. (31), el investigador principal evaluó la prescripción de acuerdo al puntaje MAI (Medication Appropriateness Index) de manera enmascarada solo al inicio del estudio, pero no lo hizo al finalizar el mismo, adicionalmente la medición de los criterios Beers y ACOVE se realizó de manera enmascarada por dos evaluadores no involucrados en el manejo del paciente. 13 estudios no describieron el cegamiento de los evaluadores [Santolaya-Perrín et al. (23), García-Gollarte et al. (24), Price et al. (25), Van der Linden et al. (26), O'Brien et al. (28), Pitkälä et al. (30), Frankenthal

et al. (35), Saltvedt et al. (36), O'Connor et al. (37), Raebel et al. (41), Cossette et al. (42), Campins et al. (43), Gallagher et al. (46)]. Finalmente, 4 estudios no realizaron el cegamiento de los evaluadores [Terrell et al. (27), O'Sullivan et al. (38), O'Connor et al. (39), Gallagher et al. (40)].

Figura 2. Resumen de "Riesgo de sesgo": revisión de los juicios de los autores sobre cada elemento de "Riesgo de sesgo" para cada estudio incluido.

ESTUDIOS						
Campins 2017	(-)	(-)	(?)	(-)	(+)	(+)
Cossette 2017	(+)	(-)	(?)	(-)	(+)	(+)
Dalleur 2014	(+)	(?)	(+)	(+)	(+)	(+)
Frankenthal 2014	(+)	(-)	(?)	(+)	(+)	(+)
Gallagher J 2016	(?)	(?)	(?)	(?)	(+)	(-)
Gallagher P 2011	(+)	(-)	(-)	(+)	(+)	(-)
García 2014	(+)	(+)	(?)	(-)	(+)	(+)
Griffey 2011	(-)	(?)	(+)	(?)	(+)	(+)
O'Brien 2018	(?)	(?)	(?)	(?)	(+)	(-)
O'Connor 2012	(?)	(?)	(?)	(?)	(?)	(-)
O'Connor 2016	(-)	(-)	(-)	(+)	(+)	(-)
O'Sullivan 2015	(-)	(-)	(-)	(+)	(+)	(-)
Pitkälä 2014	(+)	(+)	(?)	(+)	(+)	(+)
Pope 2011	(+)	(?)	(+)	(?)	(+)	(+)
Price 2017	(-)	(-)	(?)	(+)	(+)	(+)
Raebel 2007	(+)	(?)	(?)	(?)	(+)	(+)
Ramsdale 2018	(-)	(?)	(+)	(?)	(+)	(+)
Saltvedt 2005	(?)	(?)	(?)	(-)	(+)	(+)
Santolaya 2016	(+)	(-)	(?)	(?)	(-)	(+)
Schmader 2004	(+)	(?)	(+)	(?)	(+)	(+)
Spinewine 2007	(+)	(-)	(+)	(+)	(+)	(+)
Terrell 2009	(+)	(+)	(-)	(+)	(+)	(+)
Van der Linden 2016	(-)	(+)	(?)	(+)	(+)	(+)
Wouters 2017	(?)	(-)	(+)	(+)	(+)	(+)
SESGOS	Selección	Realización	Detección	Desgaste	Notificación	Otros sesgos

Características basales de los pacientes

Edad y género:

La edad de los pacientes en promedio fue de 79.6 años en el grupo control, con rangos entre 65 y 89 años; y en el grupo intervención fue de 80.1 años, con rangos entre 65 y 89,1 años de edad. El sexo no se estableció en dos estudios (25, 37). Este fue predominantemente femenino en la mayoría de los estudios, con un porcentaje promedio de participación de mujeres en el 56.8% de los grupos control y en el 58.1% de los grupos intervención. En seis estudios, el grupo control estuvo compuesto predominantemente por hombres (28, 32, 33, 38, 39, 44); en cuatro estudios, el grupo intervención estuvo compuesto predominantemente por hombres (26, 32, 44, 46%); y en un estudio (33), el grupo intervención estuvo compuesto por hombres y mujeres en igual proporción.

Desenlaces evaluados

Mortalidad:

9 estudios no reportaron la mortalidad de los participantes (23, 25, 27, 32 - 34, 37, 41, 44). 15 estudios reportaron la mortalidad: el estudio de Pitkälä et al. (30), en el cual se usaron ambos criterios (Beers y STOPP/START), describe en sus resultados una mortalidad de 33% de los residentes en el grupo intervención y del 22% en el grupo control, pero de acuerdo al modelo de riesgo proporcional de Cox, al ajustarlo por edad, sexo y comorbilidades, el aplicar la intervención no se relacionó con el incremento en la mortalidad. En el estudio de Saltvedt et al. (36), en el cual se usaron los criterios Beers como intervención, se reportaron muertes en el 6% de los pacientes del grupo control y 13% de los pacientes en el grupo intervención, lo cual fue estadísticamente significativo, con una $p=0.002$, sin embargo, no aparece descrita la causa de la muerte. En los demás estudios la diferencia en el porcentaje de muerte no fue estadísticamente significativa (24, 26, 28, 29, 31, 35, 38 - 40, 42, 43, 45, 46).

Presencia de eventos adversos antes y después de la intervención.

El número de eventos adversos por medicamentos antes de la intervención no fue descrito en 21 de los 24 estudios (23, 25-33, 35-41, 43-46). En 2 estudios (24, 42), se identificó el

delirium como el principal evento adverso por medicamento, presentándose en el grupo control con un promedio de 9.7% en un rango de 3.8% a 15.6%, en el grupo de intervención el promedio fue de 12.6%, con un rango de 6.1% a 19.1%. En el estudio restante (34) se describe un programa de cómputo diseñado para la aplicación de los criterios, el cual al estar encendido presento una tasa de detección de eventos adversos por medicación en 8 de 237 casos analizados (3.4%) y cuando el programa se encontraba apagado, la tasa de detección de eventos adversos por medicación en 31 de 436 casos analizados (7,1%), sin establecer una diferencia entre tasa de detección antes y después de la intervención, sino con y sin intervención. En lo concerniente al número de eventos adversos por medicamentos posterior a la intervención no fue descrito en 19 de los 24 estudios (23, 25 - 33, 35 - 37, 39 - 41, 43, 44, 46). En el estudio de García-Gollarte F, et al. (24) se detectó nuevamente el delirium como uno de los principales eventos adversos presentándose en el 9.1% de los pacientes del grupo control y en un 3.8% de los del grupo de intervención; en el estudios de O'Sullivan D, et al. (38) refieren la presencia de 38 reacciones adversas por problemas relacionados por medicamentos en el grupo control y 15 reacciones adversas en el grupo de intervención; en el estudio de Cossette B, et al. (42) se realizan control a un número de cohortes diferentes encontrando la presencia de delirium en 5 de 10 pacientes (50%) en el grupo de control, 3 de 12 pacientes (25%), en el grupo de intervención; hipoglicemia en 3 de 6 pacientes (50%) en el grupo de control, 5 de 8 pacientes (62,5%) en el grupo de intervención; trastornos neurocognitivos en 14 de 27 pacientes (51.9%) del grupo control, 6 de 24 pacientes (25%) del grupo de intervención; y enfermedad de parkinson en 5 de 11 pacientes (45,5%) del grupo control, 4 de 8 pacientes (50%) del grupo de intervención. Finalmente en el estudio de Pope G, et al (45) evidenciaron 1 episodio de sangrado gastrointestinal -(0.86%) en el grupo de control, y 1 episodio de gastritis (0.2%), 1 episodio de toxicidad por digoxina (0.92%) en el grupo de intervención.

Porcentaje de caídas.

Se realizó el análisis de todos los artículos incluidos en el estudio, en búsqueda del porcentaje de caídas antes y después de la intervención, encontrándose ausencia de este dato en 16

estudios (23, 25, 27, 28, 30, 32, 34, 36 - 41, 43, 45, 46). El promedio del porcentaje de pacientes con caídas al ingreso fue de 30,55 % con un rango entre 7,5% (26) y 65,1% (31) en el grupo control; en el grupo de intervención el porcentaje fue de 31,48% con un rango entre 4,5% (26) y 58,6 (31). El promedio del porcentaje de pacientes con caídas posterior a la intervención fue de 37,42% con un rango entre 27,7% (29) y 60,3% (31) para el grupo de control; para el grupo de intervención, el porcentaje fue de 27,64% con un rango de 14,3% (42) y 44,8% (31).

Número de medicamentos formulados al inicio y posterior a la intervención.

El número de medicamentos formulados tanto al inicio como luego de la intervención, no fue descrito en 8 de los 24 estudios seleccionados (25, 27, 29, 34, 41 - 43, 46). En 7 estudios no se describió el número de medicamentos posterior a la intervención (30, 31, 33, 37, 39, 40, 44). En estudios que lo incluyeron, el promedio de medicamentos formulados al inicio del estudio en los pacientes del grupo control fue de 8,11, con un rango entre 4 y 11; y en los pacientes del grupo intervención fue de 8,38, igualmente con rango entre 4 y 11, sin encontrar diferencias significativas. El promedio de medicamentos formulados luego de la intervención en los estudios que incluyeron este dato fue de 6,76 con un rango entre 1 y 11 en el grupo de control y de 6,78 con un rango de 1 a 11 en el grupo de intervención, no encontrando diferencias significativas.

Interacciones medicamentosas.

En 22 de los 24 artículos seleccionados (23-30, 32-35, 37-46) no se describió el número de interacciones encontradas o el porcentaje de pacientes que presentaban interacciones; en el estudio de Spinewine A, et al. (31) se encontró al momento de la admisión, un porcentaje de interacciones medicamentosas en el grupo control de 7,4% de los pacientes y en el grupo intervención de 7,3%. Al momento del alta el porcentaje de interacciones era de 6,7% de los pacientes en el grupo control y de 1,3% en el grupo intervención; en el estudio de Saltvedt I, et al. (36) se encontraron DDI (Drug-Drug interactions) potenciales al ingreso en 56 (44%) pacientes de la sala de cuidado geriátrico (GEMU) y 66 (52%) pacientes en sala de medicina

general (MW) y en 35 (28%) pacientes de la sala de cuidado geriátrico y 55 (43%) pacientes de la sala de medicina general al momento del alta. Ningún paciente en ninguno de los grupos tenía DDI de clase 1 (combinaciones que deben evitarse).

Días de estancia hospitalaria.

15 estudios no reportaron los días de estancia hospitalaria como parte de sus resultados (23, 25 - 29, 31 - 33, 35 - 37, 41, 43, 44). 9 estudios si los reportaron con resultados variables: en el estudio de García-Gollarte et al. (24), en el cual se usaron como intervención los criterios STOPP, el número de días de estancia hospitalaria aumentaron significativamente en el grupo control, pero permanecieron estables en el grupo intervención, al final del estudio. En el estudio de Pitkälä et al. (30), en el cual se usaron ambos criterios (Beers y STOPP/START) como intervención, el tiempo de estancia hospitalaria fue significativamente mayor en el grupo control vs el grupo intervención (2.3 vs 1.4 días/persona/año, respectivamente), con un IRR de 0.60, $p < 0.001$. El estudio de Pope et al. (45), por el contrario, mostró un mayor número de días de estancia hospitalaria en el grupo intervención (101 días) que en el grupo control (63 días). En los demás estudios que reportaron esta variable, la diferencia no es significativa entre ambos grupos (34, 38, 39, 40, 42, 46).

Criterio más usado.

Dentro de los 24 estudios incluidos para análisis, el comparador más frecuentemente usado fueron los criterios STOPP/START, los cuales se usaron en 10 estudios como único comparador (23 - 25, 28, 35, 37, 39, 40, 43, 44); en 6 artículos fueron utilizados los criterios STOPP/START junto con los criterios de Beers (26, 29, 30, 38, 42, 46); en los 8 artículos restantes (27, 31 - 34, 36, 41, 45), los criterios Beers fueron usados como único comparador.

Reingresos hospitalarios.

De los 24 estudios analizados, en 17 de ellos (23, 25, 27, 28, 29, 32, 33, 35, 36, 37, 38, 39, 41, 43, 44, 45, 46) no se encontraron datos de readmisiones hospitalarias, ya sea porque no hacía parte de los objetivos del estudio o porque no se tuvo en cuenta esa información. En 6

de los 7 estudios restantes se encontró en los grupos de control, un promedio de reingresos en el 28.89% (9.5%-39.2%), encontrándose el porcentaje más bajo de reingresos en el estudio de García-Gollarte F, et al. (24) en el cual se utilizaron los criterios STOPP/START, y el porcentaje más alto de reingresos en el estudio de Van der Linden L, et al. (26) donde se utilizaron los criterios STOPP y Beers. En los grupos de intervención el promedio de reingresos fue de 27.57% (8.7%-37,22%), siendo el porcentaje más bajo en el estudio de García-Gollarte F, et al. (24) en el cual se utilizaron los criterios STOPP/START y el más alto en el estudio de Gallagher PF, et al. (40) en el cual también se utilizaron los criterios STOPP/START. En el estudio de Griffey RT, et al. (34) reportan un 3.62% de readmisiones del total de la población estudiada sin especificar si eran del grupo en la cual el programa de computador que realiza la aplicación de los criterios, estaba encendido o apagado, en este estudio se utilizó los criterios de Beers.

Costos.

Luego de la revisión de los 24 estudios, se encontró que en 18 de ellos (23 – 27, 29 - 34, 36 - 38, 40 - 42, 44) no se reportaron los costos. En el estudio de O'Brien GL, et al. (28) haciendo uso de los criterios STOPP/START, se encontró un aumento no estadísticamente significativo de 983 dólares en el costo promedio de la atención de salud comparado con el grupo de control. En el estudio de Frankenthal D, et al. (35) que utilizó los criterios de STOPP/START, se encontró un promedio de disminución de costos en 29 dólares por participante por mes. Datos similares se encuentran en el estudio de O'Connor MN, et al. (39), en el cuál también se utilizó los criterios de STOPP/START, con una disminución promedio de 19,57 dólares en el grupo de intervención por mes. Un ahorro menor en costos se encontró en el estudio de Campins L, et al. (43), en el que nuevamente son utilizados los criterios de STOPP/START, allí el ahorro fue de 71,95 dólares paciente/año atribuible a la intervención. Pope G, et al. (45), haciendo uso de los criterios de Beers, encontró una reducción neta en el total de costos de medicación de 23.466,62 dólares, pero se consideró que fue secundario a una reducción en los participantes del grupo de intervención. Finalmente el estudio de Gallagher J, et al. (46), en el cuál se utilizaron tanto los criterios Beers como los de STOPP/START, el costo medio del

cuidado de un paciente de intervención durante una sola admisión fue 14.826,09 dólares, mientras que el grupo control mostró un costo medio de 17.304,56 dólares, sin embargo no especifica con cuál de las dos herramientas se obtuvo un mejor resultado.

Reacciones Adversas a Medicamentos (RAM)

En 16 de los 24 estudios analizados no se especifican el número, tipo o porcentaje de pacientes que presentaron RAM (23 - 25, 27, 29, 30, 31, 33 - 36, 40 - 44); en el estudio de Van der Linden L, et al. (26) durante la estancia hospitalaria se presentó delirium en el 13,3 % de los pacientes en el grupo control y en el 13,2 % de los pacientes en el grupo de intervención, posterior al egreso no se tienen datos de pacientes con delirium, pero se encontró 1 o más caídas en el 28,2 % de los pacientes del grupo de control y en el 29,3% de los pacientes del grupo de intervención; en el estudio de O'Brien GL, et al. (28) el 18,01% de los pacientes del grupo de control y el 10,83% de los pacientes del grupo de intervención experimentaron una RAM, 2,96% del grupo de control y 0,83% de los pacientes del grupo de intervención experimentaron 2 RAM, con una media de RAM por paciente (media (DE)) de 0.239 (0.492) del grupo control, y de 0.125 (0.356) del grupo de intervención; en el estudio de Schmader KE, et al. (32) reportan 128 RAM durante el periodo hospitalario y 171 RAM en el periodo ambulatorio, sin especificar cuantas pertenecen al grupo control y cuantas al grupo de intervención; en el estudio de O'Connor M, et al. (37) reportan 60 RAM en el grupo control y 38 RAM en el grupo intervención, sin que se pudieran obtener datos adicionales; en el estudio de O'Sullivan D, et al. (38) reportan 91 RAM en 78 pacientes del grupo de control (20,7%) y 61 RAM en 50 pacientes del grupo de intervención (13,9%); en el estudio de O'Connor MN, et al. (39) se reportó la presencia de una o más RAM en 78 de los 372 pacientes del grupo de control (21.0%) y en 42 de los 360 pacientes del grupo de intervención (11.7%); el estudio de Pope G, et al. (45) reportó RAM en 1 paciente de los 115 del grupo de control (0.86%) y 2 pacientes de los 110 del grupo de intervención (1.81%); finalmente el estudio de Gallagher J, et al. (46) reportó RAM en 78 pacientes de los 376 del grupo de control (20.74%) y en 50 pacientes de los 361 del grupo de intervención (13.85%).

Consulta al servicio de urgencias.

Se analizó en los 24 estudios el porcentaje de pacientes que tuvo que consultar al servicio de urgencias durante la realización del estudio y se encontró que no se especificó esta información en 17 estudios (23, 25, 27, 28, 32 - 34, 36 - 44, 46). Del grupo de control, un promedio del 15,22% del total de pacientes requirió consultar por urgencias, con un rango entre 0.34% (35) y 48.1% (26), en el grupo de intervención un promedio del 13.57% del total de pacientes requirió consultar por urgencias, con un rango de 0.31% (35) a 29.8% (26).

Tiempo promedio de los estudios.

Luego del respectivo análisis, se encontró que del estudio de O'Connor M, et al. (37) se desconoce el tiempo de duración del estudio ya que sólo se cuenta con el Abstract del mismo; el estudio de Ramsdale e, et al. (33), el tiempo de duración se limitó al dedicado por el oncólogo para la valoración del paciente y transcripción de la fórmula; 8 estudios tuvieron una duración menor a 12 meses (24 - 26, 31, 34, 40, 42, 45); 12 tuvieron una duración entre 12 y 19 meses (23, 28, 30, 32, 35, 36, 38, 39, 41, 43, 44, 46); y 2 estudios (27, 29) tuvieron una duración mayor a 20 meses.

10. Discusión.

Es clara la asociación entre polifarmacia con la presencia de mayor riesgo de eventos adversos y prescripciones potencialmente inapropiadas, los cuales van aumentando con la edad, llegando a ser tan alta como 70% en algunos hospitales geriátricos. Las intervenciones sobre este aspecto se espera impacten de forma favorable sobre las condiciones de salud y calidad de vida en los pacientes adultos mayores, motivo por el cual se han utilizado múltiples herramientas cuyo fin es ayudar al clínico a disminuir estos riesgos, especialmente en pacientes mayores de 65 años. Nuestro objetivo con este estudio fue realizar una revisión sistemática de los estudios clínicos aleatorizados existentes para abordar de una forma crítica y objetiva la efectividad de dos de estas herramientas (Beers y STOPP/START) sobre diferentes desenlaces de importancia médica, como la detección de interacciones de

relevancia clínica, la repercusión de éstas en la presentación de eventos adversos, disminución en los días de estancia hospitalaria, la cantidad de consultas al servicio de urgencias, reingresos hospitalarios, su impacto sobre la mortalidad y los costos de la atención en salud. Encontramos más de 4000 artículos que las usan como herramienta única o como comparador con otras herramientas, sin embargo, solo 24 de ellos cumplieron con los criterios de inclusión que consideramos necesarios para nuestra revisión. En término general, los estudios incluidos en esta revisión estaban limitados por su alto grado de heterogeneidad con respecto a la forma de aplicación de los criterios, la falta de resultados reportados y su baja calidad.

Las intervenciones fueron hasta cierto punto complejas para su análisis comparativo y en muchos casos se ajustaron de acuerdo a la conveniencia del examinador para la adecuada progresión de la investigación, encontrando estudios que aunque, se basaban en una misma herramienta, tenían un elevado grado de heterogeneidad, haciendo que las estimaciones de los resultados fueran muy variables, pareciendo así que la(s) herramienta(s) utilizada, no presentara(n) la misma eficacia en los diferentes estudios.

Al revisar la literatura, podemos encontrar que se han realizado múltiples revisiones sistemáticas, en las cuales se utilizaron diferentes herramientas de detección de medicación potencialmente inapropiada. Sin embargo, ante la gran variedad de metodologías de estudio y la inclusión de varias de estas herramientas en una misma investigación, optamos por seleccionar sólo aquellas en las cuales se utilizaran los criterios de Beers y/o STOPP/START, obteniéndose para su análisis un total de veintinueve revisiones sistemáticas, cuatro de ellas publicadas en 2019 (Thomas (47), Xing (48), Hukins (49) y Aguiar (50)), tres publicadas en 2018 (Motter (51), Dörks (52) y Thompson (53)), cinco publicados en 2017 (Monteiro (54), Nothelle (55), Lucchetti (56), Storms (57) y Redston (58)), seis publicadas en 2016 (Hytinen (59), Hill-Taylor (60), Walsh (61), Todd (62), Morin (63) y Muhlack (64)), cuatro publicados en 2015 (Gebara (65), Johnell (66), Cooper (67) y Tommelein (68)), una publicada en 2014 (Patterson 2014 (69)), una publicada en 2013 (Hill-Taylor (70)), una publicada en 2012

(Opondo (13)), dos publicadas en 2011 (Dimitrow (71), Guaraldo (72)), una publicada en 2009 (Kaur (73)) y una publicada en 2007 (Jano (74)). Encontramos que en su gran mayoría estos estudios proporcionaron muy pocos detalles respecto a cómo fueron desarrolladas las intervenciones, lo cual limita la discusión a las revisiones sistemáticas similares a la nuestra.

En nuestra revisión encontramos que en su gran mayoría los estudios incluidos no tuvieron como fin la optimización de las habilidades de los prescriptores, aunque si permitían un adiestramiento previo sobre el conocimiento de la herramienta antes de la intervención. Varios de los estudios se centraron en buscar reducción de medicamentos, en lugar de mejorar la idoneidad general de la prescripción, incluida la prescripción insuficiente. Se realizó el análisis de las variables propuestas en el protocolo, teniendo en cuenta que ninguno de los estudios tuvo los datos completos y que la información sobre una misma variable estaba reportada y analizada de forma heterogénea, lo cual influye en la interpretación final de las mismas.

En el análisis de variables encontramos que sólo en dos estudios (23, 24 ambos utilizaron los criterios STOPP/START) reportaron el número de medicamentos que los pacientes tomaban a su ingreso al estudio, y, la diferencia entre los grupos control e intervención, fue estadísticamente significativa siendo este valor mayor en el grupo intervención. Solo en uno de ellos (24) se realizó el análisis comparativo al egreso y se encontró que la diferencia en el promedio del número de medicamentos por paciente posterior a la intervención fue significativa. Este mismo resultado final se encontró en otros dos estudios (26 y 32, en el primero utilizaron los criterios de Beers y STOPP/START, y en el segundo utilizaron los criterios de Beers únicamente). Similar resultado se encontró en las revisiones sistemáticas de Cooper (67) y Patterson (69). En otros 4 estudios (uno de ellos utilizando criterios STOPP/START (35), dos utilizando criterios de Beers (36, 45) y uno de ellos utilizando criterios de Beers, STOPP/START Y Priscus), no mostraron diferencia significativa posterior a la intervención, resultado similar a la revisión sistemática de Walsh (61). Esto nos permite

deducir que posiblemente existan otras variables que influyan en la disminución en el número de medicamentos, y no depende solamente de la herramienta utilizada.

Respecto al número de interacciones medicamentosas clínicamente significativas encontradas antes y después de la intervención, solo dos estudios (31, 36) en los cuales se utilizaron criterios de Beers reportaron esta variable en sus resultados, encontrando con el seguimiento una disminución en el número de interacciones en el grupo de intervención estadísticamente significativa ($P = 0.009$ en el estudio de Saltvedt y col), sin embargo se debe tener en cuenta que el estudio de Spinewine (31) tiene un alto riesgo de sesgo de realización y el estudio de Saltvedt (36) un alto riesgo de sesgo de desgaste. Los demás estudios no reportaron estos datos. No se logró hacer un análisis comparativo con las revisiones sistemáticas seleccionadas, ya que ninguna de ellas incluyó esta información –que pudo ser muy relevante-, en sus resultados, considerándose entonces como información insuficiente para definir si la aplicación de las herramientas Beers y/o STOPP/START tengan un impacto en el número y la relevancia de las interacciones medicamentosas.

Un total de once estudios reportaron la inclusión de pacientes con demencia (23, 24, 26, 29, 30, 31, 35, 36, 40, 42, 44), pero las diferencias entre los grupos de control e intervención no fueron estadísticamente significativas. Datos similares se obtuvieron en las revisiones sistemáticas de Walsh (61), Cooper (67) y Patterson (68). Con esta información no podemos determinar si la presencia o no de demencia en un paciente-, influye en la presentación de complicaciones y si éstas pueden ser modificadas por la aplicación de los criterios de Beers y/o STOPP/START.

Al analizar la repercusión en los días de estancia hospitalaria, ocho estudios (28, 30, 38 – 40, 42, 45, 46) describen en sus resultados esta variable. No se encontraron diferencias estadísticamente significativas después de aplicar como intervención los criterios STOPP/START (28, 40), o aplicando las dos herramientas (30, 42), lo cual coincide con lo reportado por Hyttinen (59) y Hill-Taylor (60). En el estudio de García - Gollarte (24), el cual

tiene un alto sesgo de desgaste, se utilizaron los criterios STOPP/START y se encontró una diferencia significativa con disminución del tiempo de estancia hospitalaria, pero en el grupo de control Vs el de intervención ($P = 0.011$, $P = 0.822$ respectivamente), este dato no se encontró en ninguna de las revisiones sistemáticas con las que se realizó la comparación de resultados. En dos estudios (38, 39), utilizando criterios de Beers, STOPP/START y PRISCUS en el primero y sólo criterios STOPP/START en el segundo, se encontró una disminución estadísticamente significativa en los días de estancia hospitalaria en aquellos pacientes que no presentaron reacciones adversas a medicamentos, comparado con aquellos que si las presentaron, pero no es claro si la diferencia es a favor del grupo de intervención ($P < 0.001$ en ambos grupos). Sin embargo, se debe tener en cuenta el alto riesgo de sesgo que se presentan en ambos estudios (selección, realización, detección, otros). Se encontró en el estudio de Pitkälä (30), considerado de bajo riesgo, que esta diferencia es estadísticamente significativa después de usar los criterios STOPP/START y Beers como intervención, siendo menores los días de estancia hospitalaria en el grupo intervención que en el grupo control (tasa de incidencia 0.60, IC 95% 0.49 - 0.75; $P < .001$, ajustado por edad, sexo y comorbilidades). Ninguna de las revisiones sistemáticas mostró datos estadísticamente significativos respecto a este último resultado.

Del total de los estudios, solamente cuatro de ellos evaluaron la reducción en el número de RAM y los cuatro estudios fueron considerados como de alto riesgo de sesgos (realización, reporte selectivo, y otros sesgos) (32, 38, 39, 45). Dos de los trabajos incurrieron en reporte selectivo (32, 45), ya que uno de ellos no informó el número de pacientes con RAM sino el número total de reacciones, y en el otro estudio, solamente se reportaron tres casos en 225 pacientes lo cual señaló un sub-registro muy importante.

En siete estudios en los cuales se usaron los criterios STOPP/START (24, 40), los criterios Beers y STOPP/START (26, 30, 42), y los criterios Beers (31, 34) no se encontraron diferencias significativas en los ingresos hospitalarios (cuando los pacientes iniciaron el estudio en grupo ambulatorio) o reingresos (cuando los pacientes incluidos estaban hospitalizados al inicio del

estudio). Al comparar con las revisiones sistemáticas seleccionadas, se encontró que la asociación entre las prescripciones potencialmente inapropiadas y las readmisiones fueron inconclusas y no concluyente (Hyttinen (59)), mientras que se encontraron resultados no significativos en las otras cuatro (60, 61, 67, 69). Lo que permite concluir que la aplicación de las herramientas, no ha demostrado por el momento, una repercusión importante en la disminución de los reingresos hospitalarios independiente de su causa (reacciones adversas, interacciones, etc).

Al comparar el efecto de las intervenciones con la necesidad de consulta en los servicios de urgencias, se encontró en un estudio (26) en el cual se usaron los criterios Beers y STOPP, una diferencia estadísticamente significativa en los pacientes con una o más visitas al departamento de emergencias sin requerir readmisión ($P = 0.021$); pero la diferencia en el número de pacientes con visitas a emergencias en total no fue estadísticamente significativa entre los grupos ($P = 0,189$). Si tenemos en cuenta el alto riesgo de sesgo (selección) presente en este estudio, no es posible concluir si la aplicación de las dos herramientas tiene una repercusión favorable en la disminución del número de consultas en el servicio de urgencias. En tres estudios en los cuales se usaron los criterios STOPP/START (24) y los criterios de Beers (31, 45), respectivamente, la diferencia entre la aplicación de la intervención con la necesidad de consulta en los servicios de urgencias, no fue estadísticamente significativa. Estos resultados contrastan con lo encontrado en las revisiones de Walsh (61) y Patterson (69), en los cuales, si se encontraron diferencias significativas, sin embargo, en ambas revisiones, solo uno de los artículos incluidos, -mostró dicho resultado, con lo cual no se encuentra peso suficiente para considerar que las intervenciones se asocien a una disminución en el número de consultas a urgencias.

En trece estudios en los cuales se reportó la relación entre la intervención y la disminución en la mortalidad utilizando los criterios STOPP/START (24, 28, 35, 39, 40, 43), Beers y STOPP/START (26, 30, 42, 46), Beers, STOPP/START y PRISCUS (38) y sólo criterios de Beers (31, 45), la diferencia en la mortalidad entre ambos grupos de estudio no fue significativa.

Resultado similar fue encontrado en las revisiones de Hill-Taylor (60) -y Walsh (61). Sólo un estudio (36), en el que se aplicaron los criterios de Beers, encontró una diferencia estadísticamente significativa en reducción de la mortalidad ($P = 0.002$), sin embargo, se debe tener en cuenta que este estudio presenta alto riesgo de sesgo por desgaste. Este resultado contrasta con los de los otros estudios que también realizaron esa misma intervención, considerándose entonces insuficiente la información respecto a esta variable de gran importancia clínica.

Solo O'Brien (28) describe los costos de aplicar como intervención STOPP/START, reportando que la intervención se asoció con un aumento no significativo de los costos asociados a la aplicación de la intervención, pero con una disminución estadísticamente significativa en los costos asociados a las reacciones adversas a medicamentos evitadas; sin embargo, no aportan los valores de P y no se cuenta con las cifras necesarias para determinar el mismo. En cinco estudios (35, 39, 43, 45, 46) la aplicación de las herramientas STOPP/START (35, 39, 43), Beers (45) y la combinación de ambos criterios (46) mostró una reducción estadísticamente significativa en los costos en el grupo de intervención, pero debe tenerse en cuenta que a excepción del estudio de Pope (45) los demás presentan alto riesgo de sesgos (Realización, selección, detección, desgaste, otros). Dichos resultados van en contra de lo encontrado en las revisiones sistemáticas seleccionadas como comparador, ya que éstas sugieren que, aunque las PIMs están asociadas con aumento en los costos en el cuidado de la salud, debido principalmente a la utilización de los servicios de salud entre adultos mayores, los resultados no son estadísticamente significativos.

En algunos estudios se reportó de manera diferencial frente a las reacciones adversas medicamentosas, la presencia de eventos adversos al ingreso y luego de la intervención. En el estudio de García-Gollarte (24) en el cual utilizan los criterios STOPP/START describen la diferencia entre los eventos adversos al inicio y final del estudio para cada grupo, la cual fue estadísticamente significativa en el grupo control ($P = 0,001$), así como en el grupo intervención ($P = 0,035$), pero se debe tener precaución al interpretar estos resultados ya

que el estudio presenta alto riesgo de sesgo de desgaste. En el estudio de Griffey (34), el cual tiene un alto riesgo de sesgo de selección, se utilizaron los criterios de Beers y se reportó una diferencia significativa en el grupo de intervención, en contraposición con el estudio de Pope (45). Por último, en el estudio de Cossette (42) en el cuál se utilizaron las dos herramientas, se encontraron 4 eventos adversos (delirium, hipoglicemia, enfermedad de parkinson / parkinsonismo y trastorno neurocognitivo) sin embargo en sólo el último de estos eventos adversos, se encontró una reducción en este evento ($P = 0,029$) al aplicar las intervenciones respecto al grupo control, con un resultado estadísticamente significativo, pero este estudio presenta alto riesgo de sesgo de realización y de desgaste. En las revisiones sistemáticas no se encontraron desenlaces significativos (Cooper (67), Patterson (69)). Los resultados respecto a esta variable son inconclusos ya que presentan un reporte significativo pero que en algunos casos le puede restar relevancia el hecho de que el estudio este acompañado de sesgos. Teniendo en cuenta lo anterior, no se logra concluir cual es el papel que tiene(n) la(s) intervenciones en la disminución de los eventos adversos.

Como vemos, los resultados de los estudios no permiten dar una apreciación objetiva sobre el beneficio o desventaja de la aplicación de estas herramientas para un adecuado abordaje de los pacientes adultos con edad $>$ a 65 años intervenciones. No existe una diferencia significativa que permita detectar si una de las herramientas es mejor que la otra para beneficio de los desenlaces o si, contrario a esto, se asocian a un efecto deletéreo en la salud de los pacientes. Se encontró un resultado alentador respecto a lo concerniente a costos de la atención y tratamiento, lo cual beneficia al sistema de salud y al paciente mismo.

Consideramos que nuestra revisión, respecto a otras revisiones realizadas previamente, se diferencia al intentar hacer un análisis incluyendo todas aquellas variables que nos permitan una apreciación objetiva de cada una de las herramientas analizadas y que en muchos estudios no fueron tomadas en cuenta. Sin embargo encontramos como limitaciones la falta de uniformidad de los estudios en la forma de implementar la(s) intervención(es), la periodicidad de seguimientos, los desenlaces buscados, etc. La ausencia de una gran

cantidad de datos, a nuestro parecer relevantes, hacen que la gran mayoría de ellos sean considerados de baja calidad, orientándonos así, en futuras oportunidades, a realizar estudios más rigurosos, enfocados hacia aquellos desenlaces clínicos que tienen relevancia en la calidad de vida del paciente y que a su vez puedan tener un impacto benéfico sobre los sistemas de salud. Adicionalmente, la existencia de un gran número de herramientas creadas para la detección de medicación potencialmente inapropiada (Método MAI - *Medication Appropriateness Index*; Criterios de Beers, Criterios STOPP/START, Listado PRISCUS, Criterios IPET, Criterios LESS-CHRON, Listado del BJGP -*British Journal of General Practice*, Proyecto ACOVE - *Assessing Care of Vulnerable Elders*, etc), las cuales incluyen una alta variabilidad de criterios, limita las posibilidades de extraer datos que puedan ser traslocados de una intervención a otra. La presencia de posibles sesgos en la gran mayoría de los estudios, puede influir en la apreciación de los resultados encontrados.

CONCLUSIONES

Nuestra revisión sistemática señaló que no existen estudios clínicos controlados de buena calidad metodológica que hayan abordado la ocurrencia de desenlaces como RAM, interacciones medicamentosas, mortalidad, tiempos de hospitalización, calidad de vida, etc en pacientes ancianos como desenlace primario, las cuales tienen una alta relevancia clínica. Respecto a las dos herramientas analizadas (Beers y STOPP/START), no existe claridad si su implementación da como resultado desenlaces clínicamente significativos, aunque en algunos estudios han mostrado un impacto positivo en una adecuada prescripción en pacientes ancianos, éstos resultados podrían cobrar mayor relevancia si este desenlace fuera incluido en todos los estudios existentes.

Existen múltiples factores que son inherentes al paciente o al entorno e independientes de la herramienta a utilizar, como la edad, el sexo, las comorbilidades, sitio de atención (intrahospitalario, servicio de urgencias, consulta externa, institucionalizado en hogar de

cuidado crónico/ancianato, oportunidad de la atención médica, recursos económicos, etc.) los cuales influirán de forma positiva o negativa en los desenlaces y cuya repercusión no es predecible y es difícil de manejar e interpretar, por lo que se debe tener especial precaución en el tipo de herramienta que se va a implementar.

El completo conocimiento y entendimiento sobre la función y la aplicabilidad de las diferentes intervenciones por parte del clínico, puede aumentar la efectividad de las mismas. Se requieren estudios más robustos, con informes detallados, que permitan sacar conclusiones más contundentes, con el objetivo de orientar hacia la o las intervenciones cuyos resultados sean más favorables, y su incorporación a la práctica clínica, de acuerdo al grupo poblacional al cuál serán aplicadas.

Si bien consideramos que no es posible concluir objetivamente si el resultado clínico obtenido se puede globalizar ya que se debe tener en cuenta la baja calidad metodológica de los estudios incluidos en el estudio, si es relevante afirmar que la implementación de estas herramientas de una forma adecuada sirve como guía importante para un acercamiento al paciente, orientado a disminuir los errores en medicación y es tal vez la única forma en la actualidad de objetivizar el actual médico.

11. Referencias

1. Dagli RJ, Sharma A. Polypharmacy: A Global Risk Factor for Elderly People. *J Int Oral Heal* [Internet]. 2014;66(66):6–7. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4295469&tool=pmcentrez&rendertype=abstract>
2. Organización Mundial de la Salud (OMS). Informe mundial sobre el envejecimiento y la salud. *J Chem Inf Model*. 2015;53(9):1689–99.
3. Frankenthal D, Lerman Y, Kalendaryev E, Lerman Y. Potentially inappropriate prescribing among older residents in a geriatric hospital in Israel. *Int J Clin Pharm*. 2013;35(5):677–82.
4. Löffler C, Drewelow E, Paschka SD, Frankenstein M, Eger J, Jatsch L, et al. Optimizing polypharmacy among elderly hospital patients with chronic diseases--study protocol of the cluster randomized controlled POLITE-RCT trial. *Implement Sci* [Internet]. 2014;9:151. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4192341&tool=pmcentrez&rendertype=abstract>
5. Maeda K. Systematic review of the effects of improvement of prescription to reduce the number of medications in the elderly with polypharmacy. *Yakugaku Zasshi*. 2009;129(5):631–45.
6. Garfinkel D, Mangin D. Feasibility study of a systematic approach for discontinuation of multiple medications in older adults: addressing polypharmacy. *Arch Intern Med* [Internet]. 2010;170(18):1648–54. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20937924>
7. Vass M, Hendriksen C. Polypharmacy and older people -: The GP perspective. *Z Gerontol Geriatr*. 2005;38(SUPPL. 1):14–7.
8. Ostini R, Hegney D, Jackson C, Williamson M, Mackson JM, Gurman K, et al. Systematic review of interventions to improve prescribing. *Ann Pharmacother*. 2009;43(3):502–13.

9. Gavilán-Moral E, Villafaina-Barroso A, Jiménez-de Gracia L, Gómez Santana M del C. Ancianos frágiles polimedicados: ¿Es la deprescripción de medicamentos la salida? *Rev Esp Geriatr Gerontol.* 2012;47(4):162–7.
10. Tomás F, Núñez P, Tejada F. Prescripción en Cascada y Desprescripción. *Rev Clín Med Fam.* 2012;5(2):111–9.
11. Maher RL, Hanlon J, Hajjar ER. Clinical consequences of polypharmacy in elderly. *Expert Opin Drug Saf.* 2014;13(1):57–65.
12. Walsh EK, Cussen K. Take Ten Minutes: A dedicated ten minute medication review reduces polypharmacy in the elderly. *Ir Med J.* 2010;103(8):1–4.
13. Opondo D, Eslami S, Visscher S, de Rooij SE, Verheij R, Korevaar JC, et al. Inappropriateness of medication prescriptions to elderly patients in the primary care setting: A systematic review. *PLoS One.* 2012;7(8).
14. Denneboom W, Dautzenberg MGH, Grol R, De Smet PAGM. Treatment reviews of older people on polypharmacy in primary care: Cluster controlled trial comparing two approaches. *Br J Gen Pract.* 2007;57(542):723–31.
15. Grindrod KA, Patel P, Martin JE. What interventions should pharmacists employ to impact health practitioners' prescribing practices? *Ann Pharmacother.* 2006;40(9):1546–57.
16. O'mahony D, O'sullivan D, Byrne S, O'connor MN, Ryan C, Gallagher P. STOPP/START criteria for potentially inappropriate prescribing in older people: Version 2. *Age Ageing.* 2015;44(2):213–8.
17. American Geriatrics Society. American Geriatrics Society 2015 updated beers criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc.* 2015;63(11):2227–46.
18. Osakidetza. Revisando la medicación en el anciano. ¿Qué necesito saber? *Infac.* 2015;23(2):6–15.
19. Delgado Silveira E, Muñoz García M, Montero Errasquin B, Sánchez Castellano C, Gallagher PF, Cruz-Jentoft AJ. Prescripción inapropiada de medicamentos en los pacientes mayores: los criterios STOPP/START. *Rev Esp Geriatr Gerontol.* 2009;44(5):273–9.
20. Wyatt WJ. Behavior analysis in the era of medicalization: the state of the science and

recommendations for practitioners. *Behavior Analysis in Practice*. 2009;2(2):49–57. doi: 10.1007/BF03391748

21. Delgado Silveira E, Montero Errasquín B, Muñoz García M, Vélez-Díaz-Pallarés M, Lozano Montoya I, Sánchez-Castellano C, Cruz-Jentoft AJ. Mejorando la prescripción de medicamentos en las personas mayores: una nueva edición de los criterios STOPP-START. *Rev Esp Geriatr Gerontol*. 2015 Mar-Apr;50(2):89-96. DOI: 10.1016/j.regg.2014.10.005

22. 46. J.P.T. Higgins, S. Green (Eds.) *Cochrane handbook for systematic reviews of interventions* version 5.1.0 (updated March 2011). Cochrane Collaboration. ; 2011 www.handbook.cochrane.org

23. Santolaya-Perrin R, Jimenez-Díaz G, Galan-Ramos N, Moreno Carvajal MT, Rodriguez-Camacho JM, Sierra-Sanchez JF, et al. A randomised controlled trial on the efficacy of a multidisciplinary health care team on morbidity and mortality of elderly patients attending the Emergency Department. Study design and preliminary results. *Farmacia Hospitalaria* 2016;40(5):371-84. DOI: 10.7399/fh.2016.40.5.10465

24. García-Gollarte Fermín, Baleriola-Júlvez José, Ferrero-López Isabel, Cuenllas-Díaz Álvaro, Cruz-Jentoft Alfonso J. An Educational Intervention on Drug Use in Nursing Homes Improves Health Outcomes Resource Utilization and Reduces Inappropriate Drug Prescription. *Journal of the American Medical Directors Association*. 2014;15(12):885–891. doi: 10.1016/j.jamda.2014.04.010.

25. Price M, Davies I, Rusk R, et al. Applying STOPP guidelines in primary care through electronic medical record decision support: randomized control trial highlighting the importance of data quality. *JMIR Med Inform* 2017; 5: e15. DOI: 10.2196/medinform.6226.

26. Van der Linden L, Decoutere L, Walgraeve K, Milisen K, Flamaing J, Spriet I, Tournoy J. Combined use of the rationalization of home medication by an adjusted STOPP in older patients (RASP) list and a pharmacist-led medication review in very old inpatients: impact on quality of prescribing and clinical outcome. *Drugs Aging*. 2017;34(2):123–133. doi: 10.1007/s40266-016-0424-8.

27. Terrell K.M., Perkins A.J., Dexter P.R., Hui S.L., Callahan C.M., Miller D.K. Computerized decision support to reduce potentially inappropriate prescribing to older emergency

department patients: A randomized, controlled trial. *J. Am. Geriatr. Soc.* 2009;57:1388–1394. doi: 10.1111/j.1532-5415.2009.02352.x.

28. O'Brien GL, O'Mahony D, Gillespie P, Mulcahy M, Walshe V, O'Connor MN, et al. Cost-effectiveness analysis of a physician-implemented medication screening tool in older hospitalised patients in Ireland. *Drugs Aging.* 2018;35(8):751–62. DOI: 10.1007/s40266-018-0564-0

29. Wouters H, Scheper J, Koning H, Brouwer C, Twisk JW, van der Meer H, et al. Discontinuing inappropriate medication use in nursing home residents. A cluster randomized controlled trial. *Ann Intern Med.* 2017;167(9):609–17. doi: 10.7326/M16-2729.

30. Pitkälä KH, Juola AL, Kautiainen H, et al. Education to reduce potentially harmful medication use among residents of assisted living facilities: a randomized controlled trial. *J Am Med Dir Assoc* 2014;15:892–8. DOI: 10.1016/j.jamda.2014.04.002

31. Spinewine A, Swine C, Dhillon S, et al. Effect of a collaborative approach on the quality of prescribing for geriatric inpatients: a randomized, controlled trial. *J Am Geriatr Soc.* 2007;55(5):658–665. doi: 10.1111/j.1532-5415.2007.01132.x.

32. Schmader KE, Hanlon JT, Pieper CF, et al. Effects of geriatric evaluation and management on adverse drug reactions and suboptimal prescribing in the frail elderly. *Am J Med.* 2004;116(6):394–401. doi: 10.1016/j.amjmed.2003.10.031.

33. Ramsdale, Erika; Lemelman, Tatyana; Loh, Kah Poh et al. (2018) Geriatric assessment-driven polypharmacy discussions between oncologists, older patients, and their caregivers. *J Geriatr Oncol* 9:534-539. DOI: 10.1016/j.jgo.2018.02.007

34. Griffey RT, Lo HG, Burdick E, Keohane C, Bates DW. Guided medication dosing for elderly emergency patients using real-time, computerized decision support. *J Am Med Inform Assoc.* 2012;19:86–93. doi: 10.1136/amiajnl-2011-000124.

35. Frankenthal D, Lerman Y, Kalendaryev E, Lerman Y. Intervention with the screening tool of older persons potentially inappropriate prescriptions/screening tool to alert doctors to right treatment criteria in elderly residents of a chronic geriatric facility: a randomized clinical trial. *J Am Geriatr Soc.* 2014;62(9):1658–1665. doi: 10.1111/jgs.12993.

36. Saltvedt I, Spigset O, Ruths S, Fayers P, Kaasa S, Sletvoll O. Patterns of drug

prescription in a geriatric evaluation and management unit as compared with the general medical wards: a randomised study. *Eur J Clin Pharmacol* 2005; 61: 921–928. DOI: 10.1007/s00228-005-0046-2

37. O'Connor M, O'Sullivan D, Gallagher P, Byrne S, Eustace J, O'Mahony D (2012) Prevention of adverse drug events in hospitalized older patients: a randomized controlled trial using STOPP/START criteria. *Eur Geriatr Med* 3(S1):132. DOI: 10.1016/j.eurger.2012.07.328

38. O'Sullivan D, O'Mahony D, O'Connor MN, et al. Prevention of adverse drug reactions in hospitalised older patients using a software-supported structured pharmacist intervention: a cluster randomised controlled trial. *Drugs Aging* 2016;33:63–73. DOI: 10.1007/s40266-015-0329-y

39. O'Connor MN, O'Sullivan D, Gallagher PF, et al. Prevention of hospital-acquired adverse drug reactions in older people using screening tool of older persons' prescriptions and screening tool to alert to right treatment criteria: a cluster randomized controlled trial. *J Am Geriatr Soc* 2016;64:1558–66. DOI: 10.1111/jgs.14312

40. Gallagher PF, O'Connor MN, O'Mahony D. Prevention of potentially inappropriate prescribing for elderly patients: a randomized controlled trial using STOPP/START criteria. *Clin Pharmacol Ther* 2011; 89: 845–854. DOI: 10.1038/clpt.2011.44

41. Raebel MA, Charles J, Dugan J, Carroll NM, Korner EJ, Brand DW, et al. Randomized trial to improve prescribing safety in ambulatory elderly patients. *Journal of the American Geriatrics Society*. 2007;55(7):977–85. DOI: 10.1111/j.1532-5415.2007.01202.x.

42. Cossette B, Ethier JF, Joly-Mischlich T, Bergeron J, Ricard G, Brazeau S, et al. Reduction in targeted potentially inappropriate medication use in elderly inpatients: a pragmatic randomized controlled trial. *European Journal of Clinical Pharmacology* 2017;73(10):1237-45. DOI: 10.1007/s00228-017-2293-4

43. Campins L, Serra-Prat M, Palomera E, Bolibar I, Martinez M A, Gallo P. Reduction of pharmaceutical expenditure by a drug appropriateness intervention in polymedicated elderly subjects in Catalonia (Spain). *Gaceta Sanitaria* 2017;Epub ahead of print:1-6. DOI: 10.1016/j.gaceta.2017.09.002

44. Dalleur O, Boland B, Losseau C, et al. Reduction of potentially inappropriate medications using the STOPP criteria in frail older inpatients: a randomised controlled study. *Drugs & Aging*. 2014;31(4):291–98. doi: 10.1007/s40266-014-0157-5.
45. Pope G, Wall N, Peters CM, O'Connor M, Saunders J, O'Sullivan C, Donnelly TM, Walsh T, Jackson S, Lyons D, Clinch D. Specialist medication review does not benefit short-term outcomes and net costs in continuing-care patients. *Age Ageing*. 2011;40:307–12. doi: 10.1093/ageing/afq095.
46. Gallagher J, O'Sullivan D, McCarthy S, et al. Structured pharmacist review of medication in older hospitalised patients: a Cost-Effectiveness analysis. *Drugs Aging* 2016;33:285–94. DOI: 10.1007/s40266-016-0348-3
47. Thomas RE, Thomas BC. A systematic review of studies of the STOPP/START 2015 and American Geriatric Society Beers 2015 criteria. *Curr Aging Sci*. 2019 doi: 10.2174/1874609812666190516093742.
48. Xing XX, Zhu C, Liang HY, Wang K, Chu YQ, et al. (2019) Associations Between Potentially Inappropriate Medications and Adverse Health Outcomes in the Elderly: A Systematic Review and Meta-analysis. *Ann Pharmacother* 25. DOI: 10.1177/1060028019853069
49. Hukins D, Macleod U, Boland JW. Identifying potentially inappropriate prescribing in older people with dementia: a systematic review. *Eur J Clin Pharmacol*. 2019;75(4):467–481. <https://doi.org/10.1007/s00228-018-02612-x>
50. Aguiar JP, Brito AM, Martins AP, Leufkens H, Costa FA. Potentially Inappropriate Medications with risk of Cardiovascular Adverse Events in the Elderly: a systematic review of tools addressing inappropriate prescribing". *Journal of Clinical Pharmacy and Therapeutics*. 2019. DOI: 10.1111/jcpt.12811
51. Motter FR, Fritzen JS, Hilmer SN, Paniz ÉV, Paniz VMV. Potentially inappropriate medication in the elderly: a systematic review of validated explicit criteria. *Eur J Clin Pharmacol*. 2018;74(6):679–700. <https://doi.org/10.1007/s00228-018-2446-0>
52. Dörks M, Allers K, Hoffmann F. Pro Re Nata Drug Use in Nursing Home Residents: A Systematic Review. *J Am Med Dir Assoc*. 2019 Mar;20(3):287–293.e7. doi:

10.1016/j.jamda.2018.10.024.

53. Thompson W, Lundby C, Graabaek T, et al. Tools for deprescribing in frail older persons and those with limited life expectancy: a systematic review. *J Am Geriatr Soc.* 2019;67(1):172-180. doi: 10.1111/jgs.15616
54. Monteiro L, Maricoto T, Solha IS, et al. Computerised decision to reduce inappropriate medication in the elderly: a systematic review with meta-analysis protocol. *BMJ Open* 2018; 8: e018988. DOI: 10.1136/bmjopen-2017-018988
55. Nothelle SK, Sharma R, Oakes AH, Jackson M, Segal JB. Determinants of potentially inappropriate medication use in long-Term and acute care settings: a systematic review. *J Am Med Dir Associ.* 2017;18(9):806.e1–806.e17. DOI: 10.1016/j.jamda.2017.06.005
56. Lucchetti, G., and Lucchetti, A. L. (2017). Inappropriate prescribing in older persons: a systematic review of medications available in different criteria. *Arch. Gerontol. Geriatr.* 68, 55–61. doi: 10.1016/j.archger.2016.09.003
57. Storms H, Marquet K, Aertgeerts B, et al. Prevalence of inappropriate medication use in residential long-term care facilities for the elderly: a systematic review. *Eur J Gen Pract* 2017;23:69–77. DOI: 10.1080/13814788.2017.1288211
58. Redston MR, Hilmer SN, McLachlan AJ, Clough AJ, Gnjidic D. Prevalence of potentially inappropriate medication use in older inpatients with and without cognitive impairment: a systematic review. *J Alzheimer's Dis.* 2018;61(4):1639–1652. doi:10.3233/JAD-170842
59. Hyttinen V, Jyrkkä J, Valtonen H. A systematic review of the impact of potentially inappropriate medication on health care utilization and costs among older adults. *Med Care.* 2016;54(10):950-964. <https://doi.org/10.1097/MLR.0000000000000587>
60. Hill-Taylor B, Walsh KA, Stewart S, et al. Effectiveness of the STOPP/START (Screening Tool of Older Persons' potentially inappropriate Prescriptions/Screening Tool to Alert doctors to the Right Treatment) criteria: systematic review and meta-analysis of randomized controlled studies. *J Clin Pharm Ther* 2016;41:158–69. DOI: 10.1111/jcpt.12372
61. Walsh, K. A., O'Riordan, D., Kearney, P. M., Timmons, S., and Byrne, S. (2016). Improving the appropriateness of prescribing in older patients: a systematic review and meta-analysis of pharmacists' interventions in secondary care. *Age Aging* 45, 201–209. doi:

10.1093/ageing/afv190

62. Todd A, Husband A, Andrew I, et al. Inappropriate prescribing of preventative medication in patients with life-limiting illness: a systematic review. *BMJ Support Palliat Care* 2017; 7(2): 113–121. DOI: 10.1136/bmjspcare-2015-000941
63. Morin L, Laroche ML, Texier G, Johnell K. Prevalence of potentially inappropriate medication use in older adults living in nursing homes: A systematic review. *J Am Med Dir Assoc*. 2016;17(9):862. <https://doi.org/10.1016/j.jamda.2016.06.011>
64. Muhlack DC, Hoppe LK, Weberpals J, et al. The association of potentially inappropriate medication at older age with cardiovascular events and overall mortality: a systematic review and meta-analysis of cohort studies. *J Am Med Dir Assoc* 2017; 18: 211–220. DOI: 10.1016/j.jamda.2016.11.025
65. Gebara MA, Lipsey KL, Karp JF, et al: Cause or effect? Selective serotonin reuptake inhibitors and falls in older adults: a systematic review. *Am J Geriatr Psychiatry* 2015. DOI:10.1016/j.jagp.2014.11.004
66. Johnell K. Inappropriate drug use in people with cognitive impairment and dementia: a systematic review. *Curr Clin Pharmacol* 2015;10:178–84. PMID: 26054404 PMCID: PMC4997919
67. Cooper, J. A., Cadogan, C. A., Patterson, S. M., Kerse N., Bradley, M. C., Ryan C., et al. (2015). Interventions to improve the appropriate use of polypharmacy in older people: a Cochrane systematic review. *BMJ Open* 5:e009235. doi: 10.1136/bmjopen-2015-009235
68. Tommelein E, Mehuys E, Petrovic M, Somers A, Colin P, Boussery K. Potentially inappropriate prescribing in communitydwelling older people across Europe: a systematic literature review. *Eur J Clin Pharmacol*. 2015;71(12):1415-1427. <https://doi.org/10.1007/s00228-015-1954-4>
69. Patterson SM, Cadogan CA, Kerse N, Cardwell CR, Bradley MC, Ryan C, Hughes C. Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database Syst Rev*. 2014;10:CD008165. DOI: 10.1002/14651858.CD008165.pub3
70. Hill-Taylor B, Sketris I, Hayden J, Byrne S, O’Sullivan D, Christie R. Application of the STOPP/START criteria: a systematic review of the prevalence of potentially inappropriate

prescribing in older adults, and evidence of clinical, humanistic and economic impact. *Journal of Clinical Pharmacy and Therapeutics*. 2013; 38(5):360–372. <https://doi.org/10.1111/jcpt.12059> PMID: 23550814

71. Dimitrow MS, Airaksinen MS, Kivelä SL, Lyles A, Leikola SN. Comparison of prescribing criteria to evaluate the appropriateness of drug treatment in individuals aged 65 and older: A systematic review. *J Am Geriatr Soc*. 2011;59:1521–30. DOI: 10.1111/j.1532-5415.2011.03497.x

72. Guaraldo L, Cano FG, Damasceno GS, Rozenfeld S. Inappropriate medication use among the elderly: a systematic review of administrative databases. *BMC Geriatr*. 2011;11:79. doi:10.1186/1471-2318-11-79

73. Kaur S, Mitchell G, Vitetta L, Roberts MS. Interventions that can reduce inappropriate prescribing in the elderly: a systematic review. *Drug Aging*. 2009;26(12):1013-28. DOI: 10.2165/11318890-000000000-00000

74. Jano E, Aparasu RR. Healthcare outcomes associated with beers' criteria: a systematic review. *Ann Pharmacother*. 2007;41(3):438–47. DOI: 10.1345/aph.1H473

12. Anexos

Anexo 1. Tabla base de datos

Anexo 2. Formato de carta solicitud de información adicional

Anexo 3. Formato carta solicitud artículo completo

Anexo 4. Tabla de extracción de datos

Anexo 5. Tabla de evaluación de sesgos