

**Caracterización de la literatura científica en la aplicación de ciencia de datos en  
farmacovigilancia**

Manuel Alejandro Jáuregui Acuña

Asesor

Andrés Felipe Solís Pino

Universidad Nacional Abierta y a Distancia UNAD  
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI  
Especialización en Ciencia de Datos y Analítica  
2025

**Nota de Aceptación**

Andrés Felipe Solís Pino  
Director de Trabajo de Grado

Rafael Roberto Ruiz  
Jurado

## Resumen

La farmacovigilancia es fundamental para garantizar la seguridad de los medicamentos. No obstante, la detección temprana y eficaz de reacciones adversas a medicamentos sigue siendo un reto considerable (Fornasier et al., 2018). Este trabajo revisará la literatura científica sobre la aplicación de la Ciencia de Datos en farmacovigilancia, enfocándose en caracterizar los avances, técnicas y desafíos actuales. Utilizando bases de datos clave como Vigibase, se describirá el estado actual de estas herramientas en el análisis de reacciones adversas, evaluando su efectividad y limitaciones (Gauffin et al., 2023; Kiryu, 2023). Se analizarán diversas técnicas de Ciencia de Datos documentadas en la literatura, destacando sus beneficios y las barreras que impiden su implementación efectiva, tales como problemas de calidad de datos, heterogeneidad y sesgos (Deimazar & Sheikhtaheri, 2023; Pilipiec et al., 2022). Además, se sintetizarán recomendaciones para superar estos obstáculos y optimizar el uso de la Ciencia de Datos en la farmacovigilancia (Kaas-Hansen et al., 2023; Kompa et al., 2022). Este estudio proporcionará un entendimiento más profundo de cómo las tecnologías de Ciencia de Datos pueden avanzar en la farmacovigilancia y sugerirá futuras direcciones para la investigación y aplicación práctica en este campo esencial.

**Palabras clave:** Farmacovigilancia, minería de datos, eventos adversos a medicamentos, informática médica, análisis de datos.

## Abstract

Pharmacovigilance is essential to ensure drug safety. However, the early and effective detection of adverse drug reactions remains a considerable challenge (Fornasier et al., 2018; Hamid et al., 2022). This paper will review the scientific literature on the application of Data Science in pharmacovigilance, focusing on characterizing current advances, techniques, and challenges. Using key databases such as VigiBase, the current status of these tools in the analysis of adverse reactions will be described, evaluating their effectiveness and limitations (Gauffin et al., 2023; Kiryu, 2023). Various Data Science techniques documented in the literature will be analyzed, highlighting their benefits and the barriers that impede their effective implementation, such as data quality issues, heterogeneity, and bias (Deimazar & Sheikhtaheri, 2023; Pilipiec et al., 2022). In addition, recommendations for overcoming these obstacles and optimizing the use of data science in pharmacovigilance will be synthesized (Kaas-Hansen et al., 2023; Kompa et al., 2022). This study will provide a deeper understanding of how data science technologies can advance pharmacovigilance and suggest future directions for research and practical application in this essential field.

**Keywords:** Pharmacovigilance, data mining, adverse drug events, medical informatics, data analysis.

## Contenido

Introducción .....	8
Justificación .....	10
Objetivos.....	12
Objetivo General.....	12
Objetivos Específicos.....	12
Marco Conceptual.....	14
Metodología .....	17
Resultados.....	19
Planificación del Mapeo Sistemático de Literatura .....	19
Estrategia de Búsqueda.....	19
Términos de Búsqueda.....	19
Criterios de Inclusión.....	20
Criterios de Exclusión.....	20
Proceso de Selección .....	20
Extracción y Análisis de Datos .....	21
Análisis Temporal de los Estudios Incluidos.....	21
Hallazgos Temáticos Identificados en la Revisión Sistemática.....	23
¿Cómo Han Evolucionado las Fuentes de Datos Utilizadas en Farmacovigilancia?.....	23
¿Qué Técnicas de Análisis se Han Aplicado en Farmacovigilancia y Cómo Ha Sido Su Transición del Enfoque Estadístico Tradicional Hacia el Aprendizaje Automático? .....	25
¿En Qué Ámbitos se Ha Aplicado la Ciencia de Datos Dentro de la Farmacovigilancia y Cómo Ha Influido en la Personalización del Seguimiento Clínico?.....	28

¿Qué Beneficios se Han Identificado en la Integración de la Ciencia De Datos a la Farmacovigilancia? .....	30
¿Qué Limitaciones Enfrenta Actualmente el Uso de Ciencia De Datos En Farmacovigilancia y Cómo Afectan su Implementación? .....	32
Conclusiones .....	35
Recomendaciones .....	37
Referencias Bibliográficas .....	38

## Lista de Figuras

<b>Figura 1</b> <i>Diagrama PRISMA en Farmacovigilancia y Ciencia de Datos</i> .....	21
<b>Figura 2</b> <i>Publicaciones por Año PubMed</i> .....	22
<b>Figura 3</b> <i>Publicaciones por Año Scopus</i> .....	22
<b>Figura 4</b> <i>Número de Publicaciones por Año Revisadas</i> .....	23
<b>Figura 5</b> <i>Cantidad &amp; Proporción de Técnicas Usadas para el Tratamiento de Datos en Farmacovigilancia</i> .....	28

## Introducción

En el ámbito de la salud, asegurar la seguridad y eficacia de los medicamentos es fundamental y requiere de métodos de monitoreo efectivos. La farmacovigilancia, definida como la ciencia y actividades relacionadas con la detección, evaluación, comprensión y prevención de efectos adversos o cualquier otro problema asociado con medicamentos (Fornasier et al., 2018), juega un papel crucial en la protección de la salud pública. Sin embargo, a pesar de los avances tecnológicos, la detección de reacciones adversas a medicamentos (RAM) enfrenta desafíos significativos que comprometen la capacidad de los sistemas de salud para proteger eficazmente a los pacientes. Los sistemas de farmacovigilancia tradicionales, que dependen de reportes voluntarios y estudios de cohorte, aunque útiles, están limitados por demoras en la acumulación de datos post-mercado y la subjetividad en los reportes de eventos. Estas limitaciones conllevan a retrasos en la identificación de RAM potencialmente graves, aumentando el riesgo de morbilidad y mortalidad, retiros de medicamentos del mercado y un impacto económico significativo en los sistemas de salud debido a litigios y compensaciones (Harpaz et al., 2013).

Adicionalmente, la calidad y heterogeneidad de los datos representan obstáculos notables. Los sistemas de farmacovigilancia deben manejar una variedad de formatos de datos y fuentes que a menudo son inconsistentes o incompletos. Bases de datos como VigiBase y JADER, aunque extensivas, varían en sus metodologías de recopilación y en la naturaleza de la información que almacenan, complicando los análisis y afectando la precisión de los resultados obtenidos (Gauffin et al., 2023). El volumen creciente de datos de salud también plantea desafíos, pues muchos sistemas existentes no están equipados para manejar eficientemente este flujo de datos, resultando en cuellos de botella que impiden análisis rápidos y confiables (Deimazar & Sheikhtaheri, 2023).

En respuesta a estos retos, la ciencia de datos ofrece soluciones prometedoras a través de técnicas avanzadas como el aprendizaje automático y la minería de datos, las cuales pueden mejorar la detección y gestión de RAM. Estas técnicas permiten un análisis sistemático y automatizado de grandes conjuntos de datos, facilitando la identificación de patrones y señales de posibles reacciones adversas (Kompa et al., 2022). Sin embargo, la implementación efectiva de estas técnicas en la farmacovigilancia enfrenta barreras significativas, incluyendo la necesidad de adaptar modelos a la especificidad y complejidad de los datos médicos, así como resistencia al cambio por parte de las instituciones debido a cuestiones de coste, formación y confianza en los nuevos sistemas (Pilipiec et al., 2022).

Abordar estos desafíos mediante la ciencia de datos no solo podría mejorar los resultados en salud pública, sino también mitigar los impactos económicos adversos asociados con las reacciones a medicamentos. A través de esta revisión, se espera contribuir a un mejor entendimiento de cómo las tecnologías de ciencia de datos pueden ser aplicadas para superar los obstáculos actuales y avanzar en la seguridad y eficacia de los medicamentos, ofreciendo una base sólida para futuras investigaciones y desarrollos en este campo vital.

## Justificación

La farmacovigilancia es crucial para garantizar la seguridad y eficacia de los medicamentos en el mercado, protegiendo así la salud pública. Sin embargo, el sistema actual de farmacovigilancia enfrenta retos significativos que pueden comprometer su capacidad para detectar y responder eficaz y oportunamente a las reacciones adversas (Hamid et al., 2022). La necesidad de mejorar los métodos existentes de detección y análisis de eventos adversos es urgente, dado que estos son esenciales para minimizar los riesgos para los pacientes y mejorar los resultados en salud (Harpaz et al., 2013).

**Avances en Ciencia de Datos:** La ciencia de datos, que incorpora técnicas avanzadas como el aprendizaje automático, la minería de datos y el análisis predictivo, ha demostrado su capacidad para transformar sectores completos ofreciendo comprensiones más profundas y predicciones precisas a partir de grandes volúmenes de datos. En farmacovigilancia, estas herramientas ofrecen potencial para detectar indicaciones de reacciones adversas de manera más rápida y precisa que los métodos tradicionales, superando limitaciones como la dependencia en informes voluntarios y los retrasos en la recopilación de datos (Kompa et al., 2022).

**Necesidad de eficiencia y precisión:** Con el incremento en la cantidad de medicamentos disponibles y la complejidad de los regímenes de tratamiento, es vital que los sistemas de farmacovigilancia procesen eficientemente grandes cantidades de datos para identificar rápidamente posibles riesgos. La aplicación de la ciencia de datos promete aumentar la eficiencia de estos sistemas, reduciendo el tiempo entre la identificación de una posible reacción adversa y las acciones necesarias para mitigar el riesgo (Deimazar & Sheikhtaheri, 2023).

**Mejora en la calidad y seguridad del paciente:** Integrar la ciencia de datos en la farmacovigilancia puede mejorar significativamente la seguridad del paciente. Al identificar

patrones y tendencias en los datos de reacciones adversas que podrían pasar inadvertidos mediante métodos convencionales, se puede obtener una comprensión más completa de los riesgos asociados con los medicamentos, permitiendo realizar ajustes más informados en las prescripciones y advertencias a los pacientes (Pilipiec et al., 2022)

Superación de barreras existentes: Los desafíos como la heterogeneidad de los datos, los sesgos en los informes y la subjetividad en la interpretación de los datos pueden abordarse mediante técnicas robustas de ciencia de datos, con métodos para normalizar, analizar y extraer significado de conjuntos de datos complejos. Esto puede aumentar la confianza en los hallazgos de farmacovigilancia y mejorar la toma de decisiones basada en evidencia (Kaas-Hansen et al., 2023).

Contribución al desarrollo de políticas de salud: Este estudio no solo tiene implicaciones directas para la seguridad del paciente, sino que también puede influir en las decisiones del sector salud. Al proporcionar métodos más efectivos y eficientes para el monitoreo de medicamentos, se apoya el desarrollo de regulaciones más robustas y respuestas más rápidas a los problemas de seguridad, resultando en un entorno regulatorio más ágil y adaptativo (Fornasier et al., 2018)

La caracterización de la literatura científica en este trabajo revelará tendencias emergentes, brechas de conocimiento y oportunidades para futuras investigaciones, contribuyendo sustancialmente al cuerpo de conocimiento en este campo vital.

## Objetivos

### Objetivo General

Analizar exhaustivamente la literatura científica sobre la aplicación de la ciencia de datos en la farmacovigilancia para identificar y describir los avances técnicos y metodológicos, así como los desafíos y oportunidades emergentes. Este análisis busca proponer mejoras en las prácticas de monitoreo, detección temprana, captación de datos, análisis profundo, reporte y toma de decisiones basadas en evidencia, con el fin de optimizar la seguridad y eficacia de los medicamentos.

### Objetivos Específicos

Analizar el estado actual de las principales bases de datos utilizadas en farmacovigilancia, como VigiBase y JADER, evaluando su contribución al análisis de reacciones adversas. Este análisis se centrará en cómo estas bases de datos mejoran la precisión y efectividad del monitoreo de medicamentos a través de datos estructurados, destacando la necesidad de sistemas de datos coherentes y completos para optimizar la detección temprana de RAM.

Evaluar las técnicas de ciencia de datos aplicadas actualmente en farmacovigilancia, documentadas en la literatura, determinando sus principales beneficios y limitaciones. Este objetivo busca proponer mejoras en su aplicación práctica y evaluar cómo estas técnicas pueden ser implementadas más eficazmente para superar los desafíos actuales y mejorar los procesos de toma de decisiones en farmacovigilancia.

Identificar y describir los desafíos específicos en la implementación de la ciencia de datos en farmacovigilancia, como la calidad de los datos, la heterogeneidad y los sesgos. Este objetivo se centrará en catalogar y examinar los problemas clave que impiden la efectividad de los

métodos de ciencia de datos en la farmacovigilancia. Se explorarán estudios y reportes existentes para comprender las limitaciones actuales y cómo estas afectan la precisión y confiabilidad de los análisis de seguridad de medicamentos.

Sintetizar las recomendaciones propuestas en la literatura científica para superar las barreras actuales y optimizar el uso de la Ciencia de Datos en farmacovigilancia. Este objetivo implica consolidar guías y prácticas recomendadas que faciliten la adopción de tecnologías avanzadas y mejoren las prácticas de seguridad de medicamentos. Se enfocará en cómo la integración efectiva de la ciencia de datos puede conducir a sistemas de vigilancia más ágiles y respuestas más rápidas a los problemas de seguridad

## Marco Conceptual

La farmacovigilancia, definida por la Organización Mundial de la Salud como "la ciencia y actividades relacionadas con la detección, evaluación, comprensión y prevención de efectos adversos o cualquier otro problema relacionado con medicamentos" (Fornasier et al., 2018), es fundamental para asegurar la seguridad en el uso de medicamentos. En un entorno de constante crecimiento en la variedad y volumen de medicamentos, la expansión del uso de estos y los desafíos asociados se hacen cada vez más complejos (Busfield, 2010). La necesidad de monitorear eficazmente una creciente cantidad de información sobre seguridad de medicamentos exige la adopción de tecnologías innovadoras que puedan manejar y analizar grandes volúmenes de datos con rapidez y precisión.

El big data en farmacovigilancia abarca grandes volúmenes de datos y una diversidad de datos provenientes de diferentes fuentes, como registros electrónicos de salud, reportes de pacientes y estudios clínicos. Este vasto conjunto de datos requiere de técnicas avanzadas de análisis para ser aprovechado efectivamente. El uso de técnicas avanzadas de aprendizaje automático, tales como bosques aleatorios, regresiones logísticas y máquinas de vectores de soporte (SVM), ha revolucionado la capacidad de procesar y analizar grandes conjuntos de datos en farmacovigilancia. Estas metodologías permiten detectar patrones y conexiones que serían inaccesibles mediante métodos tradicionales y ofrecen la posibilidad de realizar predicciones y detectar riesgos antes de que estos se conviertan en problemas de seguridad, como han demostrado (Kompa et al., 2022; Pilipiec et al., 2022)

Además, el procesamiento de lenguaje natural (NLP, por sus siglas en inglés), se ha convertido en una herramienta indispensable para mejorar la interpretación de los datos no estructurados. Técnicas como el análisis de sentimiento y la clasificación de textos permiten una

mejor interpretación y extracción de información valiosa de las narrativas clínicas y reportes de incidentes, facilitando la identificación y análisis de eventos adversos relacionados con medicamentos de una manera más eficiente (Allabun & Soufiene, 2023).

Estas herramientas de NLP son particularmente valiosas para monitorear discusiones en redes sociales y foros de pacientes, donde las experiencias individuales con medicamentos se comparten frecuentemente. Al analizar estas conversaciones, los profesionales de la salud pueden obtener alertas rápidas sobre posibles efectos secundarios no reportados en ensayos clínicos o en la literatura médica, ofreciendo una visión más amplia y temprana de los perfiles de seguridad de los medicamentos.

La eficacia de las técnicas de aprendizaje automático y NLP se ha demostrado en diversos estudios, como los de (Harpaz et al., 2013), que analizaron la capacidad de los algoritmos de aprendizaje automático para identificar problemas de seguridad de medicamentos utilizando el sistema de reporte de eventos adversos de la FDA. Estos estudios muestran que las técnicas de ciencia de datos pueden mejorar significativamente la detección de señales en comparación con los métodos tradicionales. Además, investigaciones como las de (Kiryu, 2021; 2022; 2023) han mostrado cómo el análisis de big data con inteligencia artificial puede prever interacciones medicamentosas peligrosas antes de que causen daño, resaltando la importancia de estas tecnologías no solo en la detección sino también en la prevención de eventos adversos.

En conclusión, la integración de la ciencia de datos en la farmacovigilancia es esencial para manejar efectivamente el big data inherente a este campo. Las técnicas de aprendizaje automático y procesamiento de lenguaje natural no solo optimizan la detección y análisis de eventos adversos, sino que también promueven la prevención y contribuyen significativamente a la mejora continua de la seguridad de los medicamentos. Estos avances tienen un impacto

transformador no solo en la industria farmacéutica, sino también en entornos hospitalarios y programas de salud pública, asegurando el uso seguro de medicamentos por parte de los pacientes y facilitando la dispensación informada.

## Metodología

La presente monografía se desarrolló siguiendo un enfoque cualitativo de tipo descriptivo, orientado a la revisión sistemática de literatura científica. El proceso metodológico se dividió en varias fases, basadas en las etapas propuestas por el plan de trabajo aprobado institucionalmente, las cuales permitieron establecer una ruta clara para el análisis del tema: la aplicación de la ciencia de datos en farmacovigilancia.

La primera fase del proceso consistió en la elección de un área de interés que integrara tanto los principios de la ciencia de datos como de la química farmacéutica, con el fin de garantizar una conexión directa con la formación profesional y el entorno laboral del autor. Como resultado de este análisis, se definió como línea principal la aplicación de ciencia de datos en farmacovigilancia, motivado por los desafíos actuales en la detección y gestión de reacciones adversas a medicamentos (RAM), así como el creciente volumen de datos disponibles en entornos clínicos y regulatorios. A partir de esta delimitación, se construyó un árbol de problemas que permitió precisar el enfoque, delimitar las variables clave y formular una pregunta orientadora de la revisión.

La segunda fase correspondió al levantamiento de información y análisis del contexto. En esta etapa se revisaron fuentes institucionales, bases de datos científicas y documentos regulatorios, con el fin de comprender el estado actual de la farmacovigilancia, las limitaciones metodológicas de los sistemas convencionales, y las oportunidades emergentes del uso de técnicas computacionales. Esto permitió identificar los elementos conceptuales fundamentales para estructurar el marco teórico y establecer los criterios de búsqueda documental.

Posteriormente, se elaboró la matriz de revisión bibliográfica como herramienta central de organización y análisis. Esta matriz permitió sistematizar los hallazgos relevantes de cada

estudio, categorizando las investigaciones según criterios como: año de publicación, tipo de fuente de datos utilizada, técnicas de ciencia de datos aplicadas, ámbito de aplicación, beneficios reportados, limitaciones encontradas y relevancia para el objetivo general del estudio. Aunque esta matriz fue construida conforme a los lineamientos del modelo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), el análisis detallado de dicho modelo se presenta en la sección de Resultados.

Durante la etapa final del proceso metodológico se diseñó la estructura de la monografía con base en los objetivos formulados. Se establecieron los capítulos correspondientes al marco conceptual, los hallazgos organizados por subtemas y una sección específica para responder a la pregunta de investigación. Adicionalmente, se definió un cronograma de actividades que permitió distribuir adecuadamente las fases de búsqueda, análisis, redacción y revisión.

Este proceso permitió garantizar la rigurosidad en la selección de fuentes, la coherencia entre objetivos y resultados, y la trazabilidad de las decisiones metodológicas adoptadas. El diseño metodológico buscó además propiciar un análisis crítico de los hallazgos, reconociendo tanto los avances como las brechas existentes en el uso de ciencia de datos para la farmacovigilancia.

## **Resultados**

### **Planificación del Mapeo Sistemático de Literatura**

Esta monografía se desarrolló bajo el enfoque de una revisión sistemática de la literatura, siguiendo los lineamientos de la declaración PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Pino et al., 2024). El objetivo fue identificar, seleccionar, analizar y sintetizar estudios científicos relevantes que abordaran la aplicación de técnicas de ciencia de datos en farmacovigilancia, particularmente en la detección y gestión de reacciones adversas a medicamentos (RAM).

### ***Estrategia de Búsqueda***

La búsqueda se llevó a cabo en las siguientes bases de datos científicas:

- ✓ PubMed
- ✓ Scopus
- ✓ VigiBase, JADER y FADER (Las cuales se usaron como fuente en los artículos revisados.)

### ***Términos de Búsqueda***

Se utilizaron los siguientes términos de búsqueda y combinaciones booleanas: ("pharmacovigilance") AND ("data science"). La búsqueda, realizada en diciembre del 2024 y enero 2025, se centró en artículos publicados en los últimos cinco años (2019–2024), priorizando estudios recientes que aplicaran técnicas de ciencia de datos en farmacovigilancia. No obstante, se incluyeron algunas publicaciones anteriores a este periodo con el fin de aclarar conceptos clave y proporcionar contexto teórico, especialmente en lo relacionado con los fundamentos de la farmacovigilancia y su evolución histórica. En todos los casos, se consideraron únicamente estudios disponibles en texto completo, en inglés o español.

### ***Criterios de Inclusión***

Se incluyeron estudios que cumplieran con las siguientes condiciones:

- ✓ Publicaciones científicas revisadas por pares.
- ✓ Artículos que abordaran la aplicación de ciencia de datos en farmacovigilancia.
- ✓ Estudios empíricos, revisiones sistemáticas, estudios de caso y reportes técnicos relevantes.
- ✓ Estudios que reportaran resultados sobre técnicas, beneficios, limitaciones o implementación de modelos basados en big data, machine learning o inteligencia artificial en contextos relacionados con la seguridad de medicamentos.

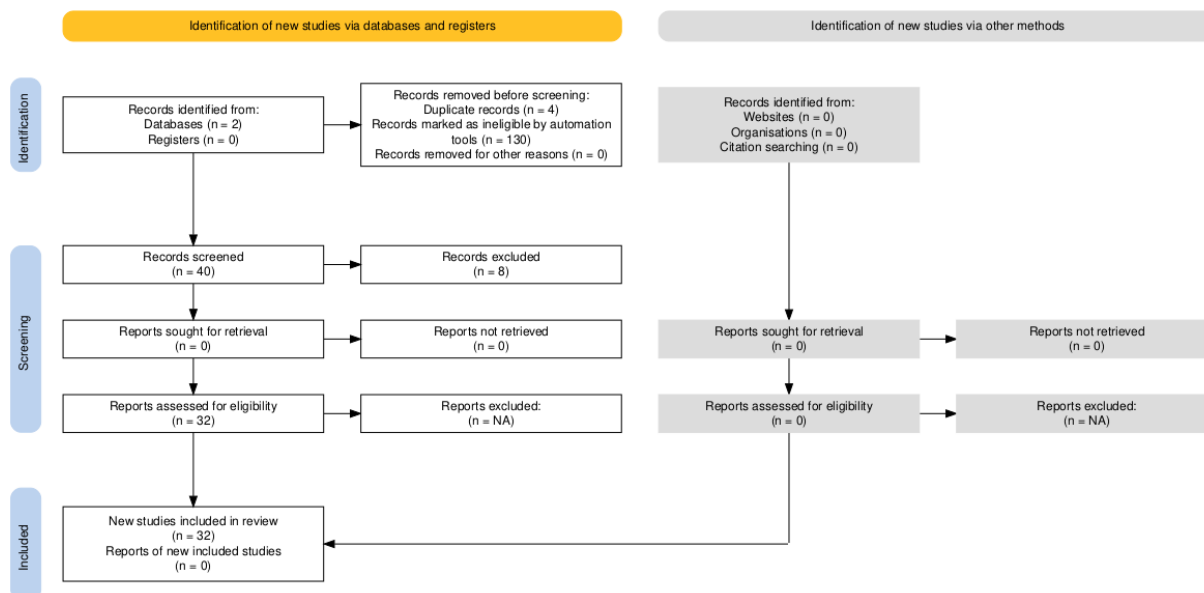
### ***Criterios de Exclusión***

Se excluyeron estudios que:

- ✓ No aplicaran técnicas de ciencia de datos.
- ✓ Se centrarán exclusivamente en áreas distintas a la farmacovigilancia.
- ✓ No presentaran resultados verificables o replicables.
- ✓ No estuvieran disponibles en texto completo.

### ***Proceso de Selección***

El proceso de selección de artículos se realizó de acuerdo con el diagrama PRISMA. Inicialmente se identificaron 150 artículos en las bases de datos mencionadas. Tras la eliminación de duplicados, quedaron 146 registros. Luego de la revisión por título y resumen, se seleccionaron 40 estudios para evaluación a texto completo. Finalmente, 32 artículos fueron incluidos en la síntesis final. De estos, 24 correspondieron a investigaciones empíricas que aplicaron técnicas de ciencia de datos a problemas de farmacovigilancia, y 8 artículos se utilizaron como soporte teórico y contextual, enriqueciendo el marco conceptual de la revisión.

**Figura 1***Diagrama PRISMA en Farmacovigilancia y Ciencia de Datos**Extracción y Análisis de Datos*

De cada artículo seleccionado se extrajeron las siguientes variables:

- ✓ Tipo de estudio
- ✓ Técnicas de ciencia de datos utilizadas
- ✓ Bases de datos analizadas (FAERS, JADER, VigiBase)
- ✓ Tipo de RAM estudiadas.
- ✓ Resultados, hallazgos, limitaciones y aportes

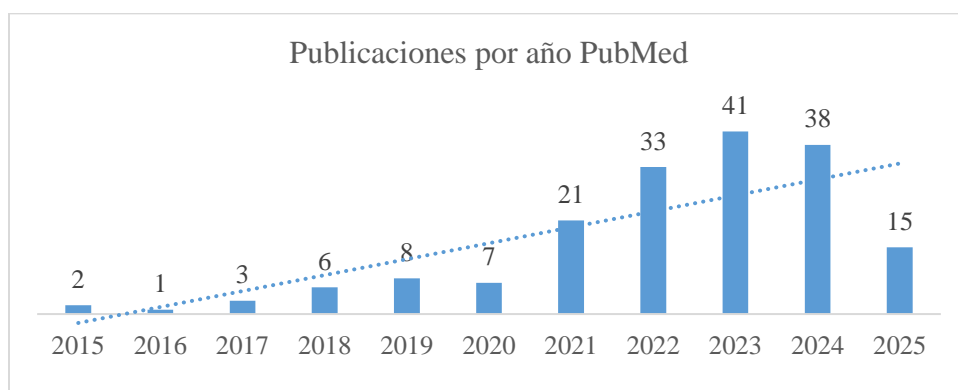
*Análisis Temporal de los Estudios Incluidos*

Es relevante destacar que, tras la búsqueda sistemática en bases de datos especializadas, se evidenció una tendencia creciente en el número de publicaciones relacionadas con ciencia de datos y farmacovigilancia a lo largo del tiempo. Aunque este comportamiento es esperable en múltiples disciplinas debido al crecimiento sostenido de la producción científica global, en este

caso particular, se observó que el año 2023 registró el mayor número de publicaciones tanto en la base de datos PubMed como en el conjunto de artículos seleccionados para esta revisión. En contraste, la base de datos Scopus mostró un volumen de publicaciones más constante y considerablemente menor en comparación. Esta distribución temporal, visualizada en los gráficos 2, 3 y 4, refuerza la idea de que la aplicación de la ciencia de datos en farmacovigilancia constituye un campo emergente, aún en fase de consolidación, pero con un claro aumento de interés y un notable potencial de desarrollo.

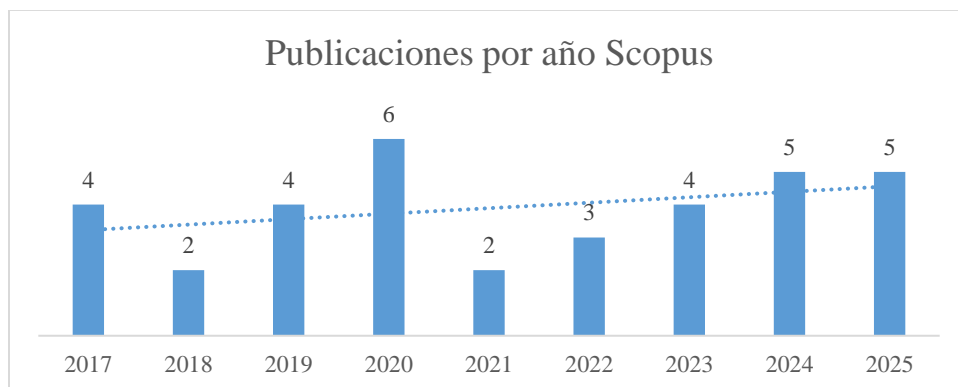
### Figura 2

*Publicaciones por Año PubMed*



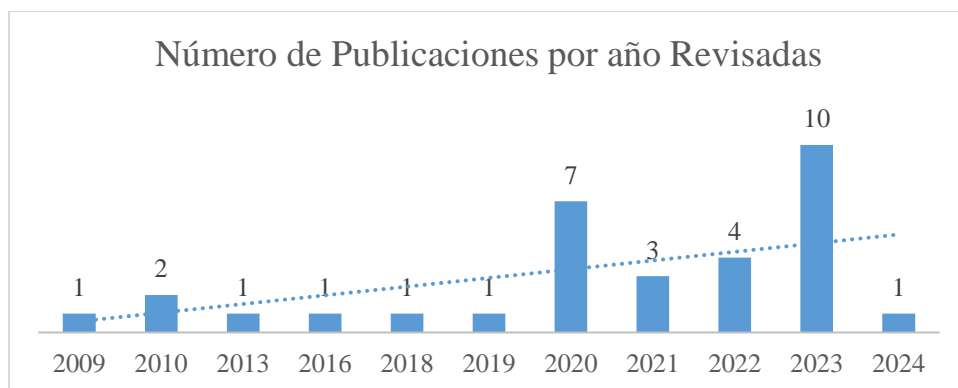
### Figura 3

*Publicaciones por Año Scopus*



## Figura 4

*Número de Publicaciones por Año Revisadas*



## Hallazgos Temáticos Identificados en la Revisión Sistemática

### *¿Cómo Han Evolucionado las Fuentes de Datos Utilizadas en Farmacovigilancia?*

La necesidad de contar con fuentes de datos adecuadas en farmacovigilancia ha surgido en respuesta directa al incremento sostenido en el consumo de medicamentos a lo largo del tiempo. Tal como lo plantea (Busfield, 2010), el fenómeno de la medicalización —es decir, la creciente tendencia a utilizar medicamentos para una amplia gama de situaciones humanas— ha generado un entorno donde se consume una mayor cantidad y variedad de fármacos. Este incremento, si bien representa avances terapéuticos, también ha traído consigo un aumento proporcional en la frecuencia de reacciones adversas a medicamentos (RAM), lo cual ha convertido a la farmacovigilancia en un componente esencial de los sistemas de salud.

A medida que crece la complejidad del uso de medicamentos, también lo hace la necesidad de contar con datos clínicos estructurados y oportunos que permitan detectar, caracterizar y prevenir efectos adversos. Sin embargo, esta necesidad no siempre ha podido satisfacerse, especialmente en las primeras etapas, donde predominaban registros manuales, fuentes fragmentadas y datos no estandarizados.

La evolución de las fuentes de datos en farmacovigilancia puede conceptualizarse en tres grandes etapas, cada una con características, oportunidades y desafíos particulares que han marcado el desarrollo de esta disciplina. Estas etapas reflejan el tránsito desde sistemas de notificación espontánea, hasta ecosistemas complejos de big data que integran información en tiempo real desde múltiples fuentes clínicas, regulatorias y tecnológicas.

Primera etapa: datos de reportes espontáneos:

En los inicios de la farmacovigilancia moderna, la principal fuente de información fueron los reportes espontáneos de reacciones adversas a medicamentos (RAM), como los contenidos en la base de datos FAERS (FDA Adverse Event Reporting System) o JADER en Japón. Este tipo de datos se caracteriza por ser voluntario, no estructurado y con alta latencia, lo que implica que no siempre se reportan todos los casos y que pueden existir retrasos significativos entre el evento y su registro. El desafío principal en esta etapa fue la subnotificación, la baja calidad de la información y la imposibilidad de realizar análisis sistemáticos. Como plantea (Kiryu, 2023), estas limitaciones motivaron el desarrollo de nuevos sistemas que integraran inteligencia artificial para mejorar la capacidad de detección de señales, compensando las debilidades de las bases tradicionales.

Segunda etapa: registros electrónicos de salud (EHR) y datos de ensayos clínicos:

A medida que los sistemas sanitarios adoptaron tecnologías digitales, emergieron nuevas fuentes como los registros clínicos electrónicos y los datos derivados de ensayos clínicos. Estos datos son más estructurados, longitudinales y disponibles en tiempo casi real, lo que mejoró la capacidad de seguimiento de pacientes y la identificación de patrones. No obstante, como mencionan (Tillman et al., 2021); (Litvinova et al., 2024) los principales desafíos en esta etapa han sido la interoperabilidad entre sistemas, la estandarización de variables clínicas y la

codificación de eventos, especialmente en contextos donde los sistemas de salud están fragmentados.

Tercera etapa: redes federadas y fuentes no tradicionales:

En los últimos años, el avance de las tecnologías de la información ha permitido integrar redes federadas de datos clínicos como EHDEN (European Health Data & Evidence Network) y aprovechar fuentes no tradicionales como redes sociales, dispositivos portátiles (wearables) y plataformas de salud digital. Esta etapa se caracteriza por el uso de datos con alto volumen, variedad y velocidad, propios del entorno de Big Data. Si bien este enfoque amplía el panorama de análisis y permite una farmacovigilancia más proactiva, también plantea desafíos complejos en términos de validación clínica, calidad del dato y riesgos éticos asociados a la privacidad y el uso secundario de la información.

En conjunto, esta evolución ha sido impulsada por el incremento sostenido en el consumo de medicamentos y la necesidad de detectar, gestionar y prevenir RAM de forma más eficiente. Como afirman estudios como los de (Gauffin et al., 2023; Kiryu, 2021; 2023; Ménard et al., 2019), la consolidación de fuentes estructuradas, interoperables y escalables es hoy uno de los pilares fundamentales para una farmacovigilancia moderna basada en ciencia de datos.

### ***¿Qué Técnicas de Análisis se Han Aplicado en Farmacovigilancia y Cómo Ha Sido Su Transición del Enfoque Estadístico Tradicional Hacia el Aprendizaje Automático?***

A lo largo del tiempo, las técnicas utilizadas para el análisis de datos en farmacovigilancia han evolucionado desde enfoques estadísticos tradicionales hacia metodologías más sofisticadas, basadas en el aprendizaje automático (machine learning, ML) y

la inteligencia artificial (IA). Esta transición no ha sido abrupta, sino gradual y multifacética, acompañada de nuevos desafíos y oportunidades metodológicas.

En la etapa inicial, los estudios aplicaban técnicas estadísticas convencionales como los modelos de regresión, los análisis de frecuencia, los cálculos de tasas de incidencia y los métodos de desproporcionalidad. Estas herramientas permitieron estructurar las primeras aproximaciones cuantitativas a la detección de señales en farmacovigilancia, destacándose por su facilidad de interpretación y su compatibilidad con los sistemas de notificación espontánea. Por ejemplo, (Koneswarakantha et al., 2020; Zou et al., 2021) emplearon modelos estadísticos clásicos para gestionar riesgos en auditorías y calidad en ensayos clínicos.

Sin embargo, con el crecimiento exponencial de los datos disponibles y la creciente complejidad de las variables implicadas en los eventos adversos, se hizo evidente la necesidad de adoptar herramientas más potentes. En este contexto emergieron los algoritmos de aprendizaje automático, que permiten identificar patrones complejos, realizar predicciones con mayor precisión y adaptarse a contextos dinámicos.

Según la revisión de los 24 artículos empíricos incluidos, se observó una marcada preferencia por el uso de modelos supervisados clásicos como SVM, árboles de decisión y bosques aleatorios, los cuales representaron el 29 % de los enfoques empleados. Estas técnicas han sido utilizadas en estudios como los de (Kiryu, 2022, 2023; Ménard et al., 2019), con el propósito de predecir subregistros de eventos adversos o construir sistemas de apoyo a la gestión farmacéutica clínica.

El NLP también ha ganado relevancia, representando el 21 % de los enfoques identificados. Su valor radica en la posibilidad de extraer información valiosa desde fuentes no estructuradas como historiales clínicos o redes sociales. Estudios como los de (Allabun &

Soufiene, 2023; Pilipiec et al., 2022) han demostrado el potencial de esta técnica para enriquecer la farmacovigilancia tradicional.

Los métodos de desproporcionalidad, aún vigentes, fueron identificados en el 17 % de los estudios. Aunque más clásicos, siguen siendo herramientas clave en la detección de señales tempranas, especialmente cuando se emplean de forma complementaria a algoritmos más complejos. Por su parte, los modelos no supervisados (8 %) y los enfoques híbridos o combinados (8 %) también han sido explorados, permitiendo nuevas formas de segmentación y análisis exploratorio.

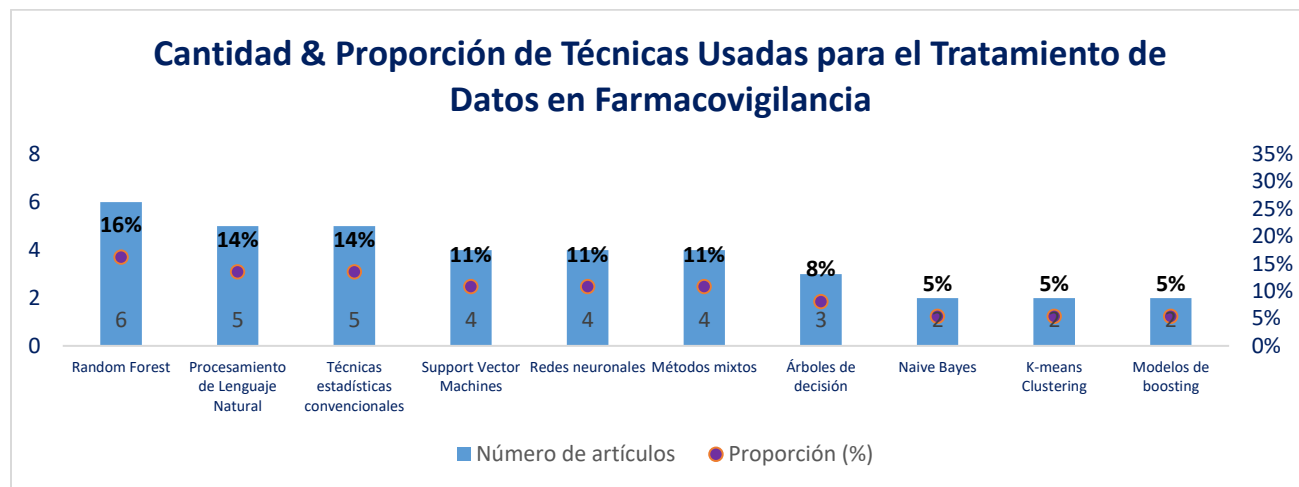
Finalmente, los enfoques estadísticos tradicionales por sí solos se identificaron en el 8 % restante, mientras que un 9 % de los estudios aplicaron otras técnicas diversas o no especificaron claramente la metodología empleada.

Para facilitar el análisis, en el desarrollo narrativo de esta sección se agruparon las técnicas utilizadas en farmacovigilancia según familias funcionales, considerando su propósito analítico y tipo de aprendizaje (por ejemplo, modelos supervisados clásicos, procesamiento de lenguaje natural, métodos estadísticos convencionales). Esta agrupación permite identificar tendencias generales y enfoques predominantes en el campo.

No obstante, en la figura 5 Cantidad & Proporción de Técnicas Usadas para el Tratamiento de Datos en Farmacovigilancia, se presentan las técnicas según su denominación específica (por ejemplo, Random Forest, Support Vector Machines, NLP, etc.), con el fin de ofrecer una visualización más precisa de la frecuencia con que cada técnica fue mencionada en los estudios revisados. Ambas perspectivas son complementarias y permiten una comprensión más profunda del panorama metodológico actual en farmacovigilancia.

**Figura 5**

*Cantidad & Proporción de Técnicas Usadas para el Tratamiento de Datos en Farmacovigilancia*



***¿En Qué Ámbitos se Ha Aplicado la Ciencia de Datos Dentro de la Farmacovigilancia y Cómo Ha Influido en la Personalización del Seguimiento Clínico?***

La aplicación de técnicas de ciencia de datos en farmacovigilancia ha evolucionado no solo en cuanto a métodos, sino también en los contextos en los que se implementa. Inicialmente, el foco estuvo puesto en el análisis pasivo de reportes espontáneos provenientes de bases de datos nacionales e internacionales, como FAERS o EudraVigilance, con el objetivo principal de detectar señales de seguridad.

Estudios como el de (Harpaz et al., 2013) reflejan esta primera etapa, centrada en el análisis retrospectivo de grandes volúmenes de reportes, sin intervención directa en los procesos clínicos. En esta fase, la farmacovigilancia era fundamentalmente reactiva, y las herramientas de análisis ayudaban a encontrar correlaciones entre medicamentos y reacciones adversas reportadas.

Con el tiempo, la incorporación de datos más estructurados y en tiempo real, como los provenientes de historiales médicos electrónicos (EHR), permitió ampliar los escenarios de aplicación. Por ejemplo, (Allabun & Soufiene, 2023) demostraron cómo el procesamiento de lenguaje natural (NLP) aplicado a EHR permite extraer información relevante sobre RAM contenida en notas clínicas, incluso cuando no está codificada de forma estandarizada. Este enfoque representa una transición hacia una farmacovigilancia más activa, apoyada por tecnologías que permiten capturar eventos adversos directamente desde la atención clínica.

Otro ámbito relevante ha sido el de los ensayos clínicos. Estudios como los de (Ménard et al., 2019; Olivier et al., 2016) han explorado cómo los algoritmos de ML pueden mejorar la calidad del reporte de eventos adversos en estudios clínicos, y cómo estas señales pueden integrarse al sistema de farmacovigilancia de manera más sistemática y automatizada. Este enfoque permite no solo mejorar la eficiencia, sino también la consistencia de los reportes en entornos regulados.

Adicionalmente, se han desarrollado aplicaciones específicas para poblaciones o condiciones clínicas concretas. Por ejemplo, (Lasky et al., 2020) utilizaron datos globales de farmacovigilancia para evaluar la seguridad de nintedanib en pacientes con fibrosis pulmonar idiopática, lo que demuestra la aplicabilidad de estos métodos a fármacos especializados. En la misma línea, (Sharif et al., 2023) aplicaron big data y ML para identificar diferencias sexo-específicas en la seguridad de bloqueadores IL4/IL13, abriendo camino a una farmacovigilancia más personalizada y centrada en subgrupos poblacionales.

En los últimos años también se ha extendido el uso de estas técnicas al monitoreo de medicamentos biotecnológicos e inmunoterapias, como lo ilustra el estudio de (Raschi et al., 2020) sobre inhibidores de puntos de control inmunitario. En estos casos, el desafío radica en

monitorear RAM complejas y multiorgánicas, cuya detección temprana puede significar una diferencia crítica en la atención del paciente.

Además, algunos estudios como el de (Kiryu, 2023) apuntan a facilitar la adopción de sistemas de análisis de RAM en farmacias comunitarias u hospitales, promoviendo una gestión local más eficiente de la seguridad del paciente. Esto sugiere que las aplicaciones no se limitan a organismos reguladores, sino que se están trasladando cada vez más a los escenarios clínicos de primera línea.

En síntesis, la ciencia de datos ha permitido extender la farmacovigilancia más allá de la detección pasiva de señales. Actualmente, sus ámbitos de aplicación abarcan desde el análisis de reportes espontáneos y EHR, hasta el monitoreo en ensayos clínicos, el seguimiento de terapias especializadas y la personalización del análisis según características del paciente o del medicamento. Esta expansión evidencia el potencial transformador de la ciencia de datos para lograr una farmacovigilancia más predictiva, integral y orientada al paciente.

### ***¿Qué Beneficios se Han Identificado en la Integración de la Ciencia De Datos a la Farmacovigilancia?***

El avance progresivo en la adopción de herramientas de ciencia de datos ha permitido que la farmacovigilancia evolucione de una actividad reactiva basada en reportes aislados hacia un sistema predictivo, continuo y mucho más eficiente. A medida que los estudios han incorporado técnicas más avanzadas, también se han registrado beneficios crecientes, tanto en la capacidad de detección de reacciones adversas como en la optimización de los sistemas de gestión de riesgos.

En los primeros trabajos que introdujeron técnicas estadísticas y de minería de datos a las bases de datos de vigilancia, como el estudio de (Harpaz et al., 2013) sobre algoritmos aplicados a la base FAERS, uno de los beneficios más destacables fue la posibilidad de identificar señales

de RAM de manera más temprana y con mayor precisión que con los métodos convencionales basados en frecuencias simples o reportes manuales.

Posteriormente, con la inclusión del aprendizaje automático, se empezaron a observar mejoras en la predicción de eventos adversos y en la priorización de señales. Por ejemplo, (Ménard et al., 2019) reportaron que los modelos predictivos basados en ML permitieron anticipar qué ensayos clínicos tenían mayor probabilidad de subreportar eventos adversos, lo que facilitó una asignación más eficaz de recursos de auditoría y mejoró la calidad regulatoria.

Estudios más recientes, como el de (Kiryu, 2021; 2022; 2023) subrayan beneficios adicionales derivados de la incorporación de sistemas completos de análisis de datos en farmacias y hospitales. Estos sistemas permiten una gestión farmacéutica más proactiva, al brindar visualizaciones tridimensionales de relaciones entre medicamentos y RAM, mejorar la adherencia terapéutica mediante alertas tempranas, y ofrecer herramientas para la individualización del tratamiento según patrones históricos y características del paciente.

(Deimazar & Sheikhtaheri, 2023), por su parte, destacan que el uso de modelos ML en registros electrónicos permite reducir el tiempo necesario para identificar riesgos, así como minimizar errores humanos en la clasificación de eventos. En la misma línea, (Allabun & Soufiene, 2023) muestran que la implementación de NLP en notas clínicas no estructuradas de EHR amplía el espectro de RAM detectables, incluyendo eventos que no se registran explícitamente en los sistemas tradicionales de codificación.

En cuanto al monitoreo de poblaciones específicas, (Sharif et al., 2023) demostraron que el análisis con técnicas de big data permitió descubrir diferencias sexo-específicas en la incidencia de RAM asociadas a bloqueadores de IL4/IL13, un hallazgo difícilmente visible

mediante métodos clásicos. Este tipo de beneficio apunta hacia una farmacovigilancia más equitativa y personalizada.

A nivel estratégico, (Kompa et al., 2022) enfatizan que la ciencia de datos permite una integración más holística entre los sistemas regulatorios, clínicos y de investigación. Esta sinergia favorece la reutilización de datos de múltiples fuentes y permite que los hallazgos generados tengan mayor impacto en la práctica clínica y en la toma de decisiones políticas.

En conjunto, los beneficios documentados en la literatura abarcaron mejoras en la sensibilidad, especificidad, rapidez y aplicabilidad de los sistemas de farmacovigilancia. Estas ventajas no solo optimizan el uso de recursos, sino que también contribuyen a una mayor seguridad del paciente, una mejor planificación terapéutica y una toma de decisiones más fundamentada en evidencia.

### ***¿Qué Limitaciones Enfrenta Actualmente el Uso de Ciencia De Datos En Farmacovigilancia y Cómo Afectan su Implementación?***

A pesar del notable progreso en la aplicación de técnicas de ciencia de datos en farmacovigilancia, la literatura revisada dejó en evidencia varias limitaciones que persisten y que obstaculizan su implementación plena y efectiva en entornos reales. Estas barreras, identificadas a lo largo del tiempo, abarcan aspectos técnicos, metodológicos, estructurales y éticos.

Una de las principales limitaciones señaladas (Deimazar & Sheikhtaheri, 2023), fue la falta de interoperabilidad entre los sistemas de información y la limitada estandarización de datos. Esta fragmentación impide una consolidación adecuada de la información necesaria para realizar análisis robustos y confiables, dificultando la integración entre instituciones y jurisdicciones.

Otro problema reiterado es la escasez de datos estructurados y de alta calidad. (Allabun & Soufiene, 2023) señalan que gran parte de los datos clínicos aún se encuentra en formato libre (como notas médicas), lo que representa un reto para su procesamiento automatizado. Aunque el uso de técnicas como procesamiento de lenguaje natural (NLP) ha intentado solventar esta barrera, su rendimiento sigue siendo variable dependiendo del contexto clínico y del idioma, como también advierten (Deimazar & Sheikhtaheri, 2023).

(Kaas-Hansen et al., 2023; Kompa et al., 2022) apuntan a otro obstáculo importante: la escasa transparencia de los modelos de machine learning. Muchos algoritmos funcionan como “cajas negras”, lo que dificulta su validación clínica y genera resistencia por parte del personal sanitario y de los organismos reguladores, que necesitan comprender cómo se llegó a una predicción o alerta para confiar en sus resultados.

En algunos casos, también se reportaron sesgos derivados del uso de datos históricos. Por ejemplo, si en el pasado ciertos grupos poblacionales estuvieron subrepresentados en los reportes de RAM (como mujeres embarazadas o adultos mayores), los modelos actuales entrenados con esos datos podrían reproducir y amplificar tales sesgos, como alerta (Sharif et al., 2023).

Desde el punto de vista regulatorio y ético, (Hauben, 2023) señala que aún existen vacíos en cuanto a normativas específicas que rijan el uso de inteligencia artificial en farmacovigilancia. Esto genera incertidumbre sobre la responsabilidad en caso de errores algorítmicos, la protección de datos sensibles y la trazabilidad de decisiones automatizadas.

Finalmente, estudios como el de (Kiryu, 2021; 2022; 2023; Ménard et al., 2019) advirtieron que, si bien se han logrado importantes avances en entornos de investigación y desarrollo, la implementación de estos sistemas en escenarios clínicos reales sigue siendo

limitada debido a barreras tecnológicas, costos de implementación y resistencia al cambio por parte del personal.

Estas limitaciones, si bien importantes, no niegan el valor de la ciencia de datos en farmacovigilancia, pero sí subrayan la necesidad de un enfoque más holístico que integre esfuerzos técnicos, regulatorios, formativos y colaborativos para garantizar su adopción segura, ética y eficiente.

## Conclusiones

La revisión sistemática permitió analizar de forma exhaustiva la literatura científica sobre la aplicación de la ciencia de datos en farmacovigilancia, revelando los avances técnicos y metodológicos más relevantes, así como los desafíos persistentes y las oportunidades emergentes. Este análisis confirmó que la ciencia de datos tiene el potencial de optimizar significativamente el monitoreo, la detección temprana, la recolección y análisis de datos, así como la toma de decisiones clínicas fundamentadas, contribuyendo a fortalecer la seguridad y eficacia de los medicamentos en los sistemas de salud contemporáneos.

En términos de fuentes de datos, se identificó que bases consolidadas como VigiBase, JADER, FAERS y los registros electrónicos de salud (EHR) son las más utilizadas. Estas plataformas ofrecen volúmenes masivos de datos estructurados y no estructurados, lo cual resulta clave para alimentar modelos predictivos aplicados a la detección de reacciones adversas. No obstante, se evidenció la necesidad de mejorar la calidad, estandarización e interoperabilidad de estos sistemas, ya que su fragmentación limita la posibilidad de desarrollar análisis confiables y generalizables, especialmente en contextos multijurisdiccionales.

Respecto a las técnicas de análisis utilizadas, los estudios muestran una evolución desde métodos estadísticos tradicionales hacia un repertorio más sofisticado de herramientas de ciencia de datos. Se destacan el aprendizaje automático supervisado (como random forest, SVM y redes neuronales), el NLP y los métodos mixtos. Estas herramientas han demostrado ser eficaces en tareas como la predicción de RAM, la clasificación de reportes, la extracción de señales en tiempo real y el análisis de texto libre, aunque su desempeño depende en gran medida de la calidad y estructura de los datos disponibles.

En paralelo, se identificaron varios desafíos que obstaculizan la implementación efectiva de la ciencia de datos en farmacovigilancia. Entre ellos destacan la heterogeneidad en las fuentes de datos, la presencia de sesgos inherentes a los reportes espontáneos, y la limitada validación clínica de los modelos desarrollados. Estos factores afectan la precisión, la confiabilidad y la aplicabilidad práctica de las soluciones analíticas propuestas, generando la necesidad de protocolos más rigurosos de validación externa y estandarización metodológica.

Finalmente, la literatura revisada propone una serie de recomendaciones orientadas a superar estas barreras, entre ellas: promover la integración de sistemas interoperables, fomentar la colaboración interdisciplinaria entre expertos en salud y ciencia de datos, garantizar el cumplimiento de principios éticos y de privacidad, y fortalecer la formación técnica en análisis de datos dentro del ámbito clínico. De este modo, se espera que la aplicación de técnicas avanzadas de ciencia de datos permita consolidar una farmacovigilancia más ágil, predictiva y centrada en el paciente.

## Recomendaciones

El presente trabajo evidencia que la farmacovigilancia es un proceso complejo, donde la detección de problemas relacionados con medicamentos representa solo el primer paso dentro de una cadena orientada a prevenir daños en los pacientes. En este contexto, la ciencia de datos se perfila como una herramienta con gran potencial para apoyar la detección temprana de riesgos y contribuir a la implementación de estrategias de seguridad. Por ello, se sugiere avanzar en la integración progresiva de la ciencia de datos en los distintos niveles donde se ejerce la farmacovigilancia. Este proceso debería comenzar con la captación de datos estructurados que faciliten análisis más rigurosos, o, en su defecto, mediante el uso de técnicas de NLP que permitan extraer información útil de fuentes no estructuradas. A partir de esta base, la aplicación de métodos estadísticos tradicionales y algoritmos de machine learning puede aportar valor en la estimación de riesgos y la priorización de casos que merecen seguimiento más estrecho. No obstante, cualquier acción derivada como la generación de alertas o la implementación de barreras de seguridad deberá estar respaldada por validación clínica y/o estadística convencional, asegurando así la confiabilidad de las decisiones adoptadas.

### Referencias Bibliográficas

- Allabun, S., & Soufiene, B. O. (2023). Study of the Drug-related Adverse Events with the Help of Electronic Health Records and Natural Language Processing. *International Journal of Advanced Computer Science and Applications*, 14(6), 2023.  
<https://doi.org/10.14569/IJACSA.2023.01406148>
- Busfield, J. (2010). ‘A pill for every ill’: Explaining the expansion in medicine use. *Social Science & Medicine*, 70(6), 934–941.  
<https://doi.org/10.1016/J.SOCSCIMED.2009.10.068>
- Cortes-Serra, N., Saravia, R., Grágeda, R. M., Apaza, A., González, J. A., Ríos, B., Gascón, J., Torrico, F., & Pinazo, M. J. (2020). Strengthening the Bolivian pharmacovigilance system: New surveillance strategies to improve care for Chagas disease and tuberculosis. *PLoS Neglected Tropical Diseases*, 14(9), 1–12.  
<https://doi.org/10.1371/JOURNAL.PNTD.0008370>
- Deimazar, G., & Sheikhtaheri, A. (2023). Machine learning models to detect and predict patient safety events using electronic health records: A systematic review. *International Journal of Medical Informatics*, 180.  
<https://doi.org/10.1016/j.ijmedinf.2023.105246>
- Dokumaci, A. D., & Karataş, y. (2023). Farmakovijilansın Dijitalleşmesi: Yapay Zeka ve Veri Analitiğinin Rolü. *Arşiv Kaynak Tarama Dergisi*, 32(4), 200–205.  
<https://doi.org/10.17827/AKTD.1333721>
- Zhengwu Lu (2009). Information technology in pharmacovigilance: Benefits, challenges, and future directions from industry perspectives. *Drug, Healthcare and Patient Safety*, 1(1), 35. <https://doi.org/10.2147/DHPS.S7180>

- Fornasier, G., Francescon, S., Leone, R., & Baldo, P. (2018). An historical overview over Pharmacovigilance. *International Journal of Clinical Pharmacy*, 40(4), 744–747. <https://doi.org/10.1007/S11096-018-0657-1>
- Gauffin, O., Brand, J. S., Vidlin, S. H., Sartori, D., Asikainen, S., Català, M., Chalabi, E., Dedman, D., Danilovic, A., Duarte-Salles, T., García Morales, M. T., Hiltunen, S., Jödicke, A. M., Lazarevic, M., Mayer, M. A., Miladinovic, J., Mitchell, J., Pistillo, A., Ramírez-Anguita, J. M., ... Norén, G. N. (2023). Supporting Pharmacovigilance Signal Validation and Prioritization with Analyses of Routinely Collected Health Data: Lessons Learned from an EHDEN Network Study. *Drug Safety*, 46(12), 1335–1352. <https://doi.org/10.1007/S40264-023-01353-W>
- Ghafari, A. T., Zakaria, Y., Jahidin, A. H., Nazari, Q. A., Majeed, A. B. A., & Hasan, M. H. (2023). Navigating the Polypharmacy Landscape: A Bibliometric Analysis of Computational Approaches for Predicting Adverse Drug Reactions. *Malaysian Journal of Medicine and Health Sciences*, 19, 109–114. <https://doi.org/10.47836/MJMHS.19.S12.13>
- Hamid, A. A. A., Rahim, R., & Teo, S. P. (2022). Pharmacovigilance and Its Importance for Primary Health Care Professionals. *Korean Journal of Family Medicine*, 43(5), 290–295. <https://doi.org/10.4082/KJFM.21.0193>
- Harpaz, R., Dumouchel, W., Lependu, P., Bauer-Mehren, A., Ryan, P., & Shah, N. H. (2013). Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system. *Clinical Pharmacology and Therapeutics*, 93(6), 539–546. <https://doi.org/10.1038/CLPT.2013.24>

- Hauben, M. (2023). Artificial Intelligence and Data Mining for the Pharmacovigilance of Drug–Drug Interactions. In *Clinical Therapeutics* (Vol. 45, Issue 2).  
<https://doi.org/10.1016/j.clinthera.2023.01.002>
- Jnaneswar, A., Hazarika, S., Jha, K., Vinay, S., & Kumar, G. (2020). Knowledge, attitude, practices, and barriers regarding pharmacovigilance and adverse drug reaction reporting among medical and dental faculties of the teaching hospitals in Bhubaneswar City. *Journal of Education and Health Promotion*, 9(1).  
[https://doi.org/10.4103/JEHP.JEHP\\_250\\_20](https://doi.org/10.4103/JEHP.JEHP_250_20)
- Kaas-Hansen, B. S., Gentile, S., Caioli, A., & Andersen, S. E. (2023). Exploratory pharmacovigilance with machine learning in big patient data: A focused scoping review. *Basic & Clinical Pharmacology & Toxicology*, 132(3), 233–241.  
<https://doi.org/10.1111/BCPT.13828>
- Kabanywany, A. M., Mulure, N., Migoha, C., Malila, A., Lengeler, C., Schlienger, R., & Genton, B. (2010). Experience of safety monitoring in the context of a prospective observational study of artemether-lumefantrine in rural Tanzania: lessons learned for pharmacovigilance reporting. *Malaria Journal* 2010 9:2, 9(2), 1–2.  
<https://doi.org/10.1186/1475-2875-9-S2-P24>
- Kiryu, Y. (2021). Potential for big data analysis using AI in the field of clinical pharmacy. *Yakugaku Zasshi*, 141(2). <https://doi.org/10.1248/YAKUSHI.20-00196-4>
- Kiryu, Y. (2022). Medical Big Data Analysis Using Machine Learning Algorithms in the Field of Clinical Pharmacy. *Yakugaku Zasshi*, 142(4).  
<https://doi.org/10.1248/yakushi.21-00178-1>

- Kiryu, Y. (2023). Development of a Medical Big Data Analysis System Utilizing Artificial Intelligence Analytics in Clinical Pharmacy. In *Yakugaku Zasshi* (Vol. 143, Issue 6). <https://doi.org/10.1248/YAKUSHI.22-00179-4>
- Kompa, B., Hakim, J. B., Palepu, A., Kompa, K. G., Smith, M., Bain, P. A., Woloszynek, S., Painter, J. L., Bate, A., & Beam, A. L. (2022). Artificial Intelligence Based on Machine Learning in Pharmacovigilance: A Scoping Review. *Drug Safety*, 45(5), 477–491. <https://doi.org/10.1007/s40264-022-01176-1>
- Koneswarakantha, B., Ménard, T., Rolo, D., Barmaz, Y., & Bowling, R. (2020). Harnessing the Power of Quality Assurance Data: Can We Use Statistical Modeling for Quality Risk Assessment of Clinical Trials? *Therapeutic Innovation & Regulatory Science*, 54(5), 1227–1235. <https://doi.org/10.1007/S43441-020-00147-X>
- Lasky, J. A., Criner, G. J., Lazarus, H. M., Kohlbrenner, V., Bender, S., & Richeldi, L. (2020). Safety of Nintedanib in Patients with Idiopathic Pulmonary Fibrosis: Global Pharmacovigilance Data. *Advances in Therapy*, 37(10), 4209–4219. <https://doi.org/10.1007/S12325-020-01452-5>
- Litvinova, O., Yeung, A. W. K., Hammerle, F. P., Mickael, M. E., Matin, M., Kletecka-Pulker, M., Atanasov, A. G., & Willschke, H. (2024). Digital Technology Applications in the Management of Adverse Drug Reactions: Bibliometric Analysis. *Pharmaceuticals*, Vol. 17, Page 395, 17(3), 395. <https://doi.org/10.3390/PH17030395>
- Ménard, T., Barmaz, Y., Koneswarakantha, B., Bowling, R., & Popko, L. (2019). Enabling Data-Driven Clinical Quality Assurance: Predicting Adverse Event Reporting in

Clinical Trials Using Machine Learning. *Drug Safety*, 42(9), 1045–1053.

<https://doi.org/10.1007/S40264-019-00831-4/TABLES/4>

Nasir, M., Zahan, T., Farha, N., & Chowdhury, A. S. M. S. (2020). Acquaintance, approach and application of pharmacovigilance: questionnaire based study at a tertiary care teaching hospital in Dhaka. *International Journal of Basic & Clinical Pharmacology*, 9(10), 1497–1502. <https://doi.org/10.18203/2319-2003.IJBCP20203994>

Olivier, P., Gimbert, A., Colin, A. L., Salvo, F., Becker, M., Marty, V., Montastruc, J. L., & Petitpain, N. (2016). Serious adverse drug events related to non-investigational drugs in academic clinical trials: another source of safety data for risk assessment? *British Journal of Clinical Pharmacology*, 82(4), 1069–1075.

<https://doi.org/10.1111/BCP.13035>

Pilipiec, P., Liwicki, M., & Bota, A. (2022). Using Machine Learning for Pharmacovigilance: A Systematic Review. In *Pharmaceutics* (Vol. 14, Issue 2).

<https://doi.org/10.3390/pharmaceutics14020266>

Pino, A. F. S., Ruiz, P. H., Mon, A., & Collazos, C. A. (2024). Mechanisms for measuring technology maturity on the Internet of Things in enterprises: A systematic literature mapping. *Internet of Things*, 25, 101100. <https://doi.org/10.1016/J.IOT.2024.101100>

Raschi, E., Gatti, M., Gelsomino, F., Ardizzoni, A., Poluzzi, E., & de Ponti, F. (2020).

Lessons to be Learnt from Real-World Studies on Immune-Related Adverse Events with Checkpoint Inhibitors: A Clinical Perspective from Pharmacovigilance.

*Targeted Oncology*, 15(4), 449–466. <https://doi.org/10.1007/S11523-020-00738-6>

- Riordan, D. O., Kinane, M., Walsh, K. A., Shiely, F., Eustace, J., & Bermingham, M. (2020). Stakeholders' knowledge, attitudes and practices to pharmacovigilance and adverse drug reaction reporting in clinical trials: a mixed methods study. *European Journal of Clinical Pharmacology*, 76(10), 1363–1372. <https://doi.org/10.1007/S00228-020-02921-0>
- Shahid, U. (2023). Leveraging Fine-Tuned Language Models in Bioinformatics: A Research Perspective. *Qeios*. <https://doi.org/10.32388/WE7UMN>
- Sharif, K., Omar, M., Lahat, A., Patt, Y. S., Amital, H., Zoabi, G., Bragazzi, N. L., & Watad, A. (2023). Big data- and machine learning-based analysis of a global pharmacovigilance database enables the discovery of sex-specific differences in the safety profile of dual IL4/IL13 blockade. *Frontiers in Pharmacology*, 14. <https://doi.org/10.3389/fphar.2023.1271309>
- Tillman, E. M., Suppes, S. L., Feldman, K., & Goldman, J. L. (2021). Enhancing Pediatric Adverse Drug Reaction Documentation in the Electronic Medical Record. *Journal of Clinical Pharmacology*, 61(2), 181–186. <https://doi.org/10.1002/JCPH.1717>
- Zou, M., Barmaz, Y., Preovolos, M., Popko, L., & Ménard, T. (2021). Using Statistical Modeling for Enhanced and Flexible Pharmacovigilance Audit Risk Assessment and Planning. *Therapeutic Innovation & Regulatory Science*, 55(1), 190–196. <https://doi.org/10.1007/S43441-020-00205-4>