

Detección de señales en FAERS

Notas introductorias

Nota importante

FAERS (FDA Adverse Event Reporting System) es una base de *reportes espontáneos*. Una **señal** en FAERS **no prueba causalidad**: sugiere una *asociación de reporte* que merece investigación adicional. Las métricas vistas aquí (PRR, ROR) miden **desproporcionalidad en reportes**, no el riesgo real en la población.

1. ¿Qué es una “señal” en farmacovigilancia?

Definición

Una **señal de seguridad** es evidencia estadística (y/o clínica) que sugiere que un **evento adverso** se reporta con **frecuencia inusualmente alta** en presencia de un **fármaco** (o clase de fármacos), comparado con lo esperable bajo un patrón “normal” de reportes.

En un enfoque cuantitativo básico, diremos que hay señal cuando los datos sugieren que el par (D, A) (fármaco D , evento A) aparece *más de lo esperado* dentro de la base.

2. Antes de la estadística: decisiones de modelado (muy importantes)

En FAERS, antes de construir la tabla 2×2 , hay decisiones que cambian los conteos.

2.1. Unidad de análisis: ¿reporte o caso? (duplicados y seguimientos)

Un mismo caso puede tener **seguimientos** (versiones actualizadas). Si cuentas todo sin control, puedes inflar señales.

Procedimiento

Regla práctica: deduplicar por **caseid** conservando la versión más reciente (o trabajar consistentemente con **primaryid** si tu flujo lo requiere). Lo importante es **ser consistente** y declarar la decisión.

2.2. ¿Qué significa “el fármaco está presente”? (rol del fármaco)

Un reporte puede incluir varios fármacos con distintos roles, por ejemplo:

- **PS:** *Primary Suspect* (principal sospechoso),

- **SS:** *Secondary Suspect* (sospechoso secundario),
- **Concomitante:** fármaco presente pero no sospechoso.

Nota importante

Para detección de señales, es común hacer el análisis principal con **PS** (más específico) y un análisis de sensibilidad con **PS+SS** (más sensible). Mezclar PS/SS con concomitantes puede introducir confusión por polifarmacia.

2.3. Multi-fármaco y multi-evento

Un reporte puede tener:

- **múltiples fármacos**, y
- **múltiples eventos** (reacciones) codificados (p. ej. MedDRA/PT).

Para el par (D, A) , un reporte cuenta como “ D y A ” si contiene D (según tu definición de exposición) y contiene el evento A .

3. La tabla 2×2 (el núcleo del método)

Para un fármaco D y un evento A , definimos los conteos:

	Evento A	No A
Fármaco D	n_{11}	n_{10}
No D	n_{01}	n_{00}

Interpretación:

- n_{11} : reportes con D y A .
- n_{10} : reportes con D y sin A .
- n_{01} : reportes sin D y con A .
- n_{00} : reportes sin D y sin A .

3.1. Intuición: conteo esperado bajo independencia

Sea $N = n_{11} + n_{10} + n_{01} + n_{00}$, $n_{1+} = n_{11} + n_{10}$ y $n_{+1} = n_{11} + n_{01}$. Si D y A fueran independientes, el conteo esperado en la celda (D, A) es

$$E[n_{11}] = \frac{n_{1+} n_{+1}}{N}.$$

Nota importante

Comparar n_{11} vs $E[n_{11}]$ es una intuición muy útil antes de calcular PRR/ROR.

4. Medidas básicas de señal: PRR y ROR

4.1. PRR (Proportional Reporting Ratio)

$$PRR = \frac{\frac{n_{11}}{n_{11}+n_{10}}}{\frac{n_{01}}{n_{01}+n_{00}}}.$$

Lectura: “proporción de reportes con A entre los que contienen D , comparada con la proporción de A en el resto de reportes”.

4.2. ROR (Reporting Odds Ratio)

$$ROR = \frac{n_{11}/n_{10}}{n_{01}/n_{00}} = \frac{n_{11}n_{00}}{n_{10}n_{01}}.$$

Lectura: razón de *odds* de reporte del evento A en presencia de D frente a ausencia de D .

Nota importante

Si $PRR > 1$ o $ROR > 1$, el evento aparece relativamente más asociado a D (en reportes). Pero para hablar de señal se necesita también estabilidad estadística (IC) y mínimo conteo.

5. Intervalos de confianza (IC) para ROR (y por qué importan)

5.1. Error estándar en escala log

Usamos que, aproximadamente,

$$SE(\log(ROR)) \approx \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{10}} + \frac{1}{n_{01}} + \frac{1}{n_{00}}}.$$

Luego, para un 95 % de confianza:

$$IC_{95\%}(ROR) = \exp(\log(ROR) \pm 1,96 SE(\log(ROR))).$$

Nota importante

IC ancho suele indicar **pocos datos** o celdas pequeñas. Un ROR grande con IC muy ancho es menos confiable que un ROR moderado con IC estrecho.

6. Criterios prácticos para priorizar señales (nivel básico)

No existe una regla universal, pero una regla didáctica muy usada para *priorizar* pares (D, A) es:

- $n_{11} \geq 3$ (o ≥ 5 , según el curso y el objetivo),
- $PRR \geq 2$ (opcional, como umbral de fuerza),
- Límite inferior de $IC_{95\%}(ROR) > 1$ (significancia aproximada),
- persistencia temporal (opcional): la señal aparece en ≥ 2 trimestres.

Nota importante

Estos umbrales son **convenciones** para balancear sensibilidad vs especificidad. En farmacovigilancia real, la priorización también depende de la gravedad del evento, plausibilidad clínica, subpoblaciones, etc.

7. Corrección de Haldane-Anscombe (cuando hay ceros)

7.1. El problema

Si alguna celda es 0, el ROR puede volverse 0 o infinito, y $\log(ROR)$ puede no existir.

7.2. La solución

Se ajusta la tabla sumando 0,5 a cada celda:

$$n'_{11} = n_{11} + 0,5, \quad n'_{10} = n_{10} + 0,5, \quad n'_{01} = n_{01} + 0,5, \quad n'_{00} = n_{00} + 0,5.$$

Y se calcula ROR e IC con n'_{ij} .

Nota importante

La corrección **evita extremos** (0 o ∞) y estabiliza el IC, pero introduce un ligero “empuje” hacia 1 cuando los conteos son muy pequeños.

8. Sesgos típicos en FAERS (para interpretar resultados)

- **Subreporte:** no todos los eventos se reportan.
- **Notoriedad:** noticias/alertas pueden aumentar reportes de un evento.
- **Confusión por indicación:** la enfermedad tratada puede explicar el evento.
- **Polifarmacia y comorbilidades:** múltiples fármacos y condiciones por caso.
- **Duplicados/seguimientos:** el mismo caso puede contarse varias veces si no se controla.

Nota importante

Por estas razones, PRR/ROR son herramientas de **tamizaje** (screening). Las señales deben evaluarse con análisis adicionales (estratificación, temporalidad, plausibilidad clínica, estudios observacionales, etc.).

9. Ejemplo guiado (cálculo completo)

Ejemplo guiado

Supón que para el par (D, A) obtienes:

$$n_{11} = 12, \quad n_{10} = 88, \quad n_{01} = 30, \quad n_{00} = 870.$$

1) PRR

$$\frac{n_{11}}{n_{11} + n_{10}} = \frac{12}{100} = 0,12, \quad \frac{n_{01}}{n_{01} + n_{00}} = \frac{30}{900} = 0,0333.$$

$$PRR = \frac{0,12}{0,0333} \approx 3,6.$$

2) ROR

$$ROR = \frac{12 \cdot 870}{88 \cdot 30} = \frac{10440}{2640} \approx 3,95.$$

3) IC95 % del ROR

$$SE(\log ROR) \approx \sqrt{\frac{1}{12} + \frac{1}{88} + \frac{1}{30} + \frac{1}{870}} \approx 0,359.$$

$$IC_{95\%}(ROR) = \exp(\log(3,95) \pm 1,96(0,359)) \approx [1,96, 8,00].$$

Conclusión: es compatible con señal (fuerte y estadísticamente consistente), siempre recordando que no es causalidad.

10. Mini-guía: de datos crudos a $n_{11}, n_{10}, n_{01}, n_{00}$

Procedimiento

1. Elige el fármaco D (normalizado) y el evento A (por ejemplo, PT de MedDRA).
2. Decide la **unidad** (reporte vs caso deduplicado) y el **rol** del fármaco (PS o PS+SS).
3. Para cada reporte, define indicadores:

$$I_D = \mathbf{1}(\text{reporte contiene } D), \quad I_A = \mathbf{1}(\text{reporte contiene } A).$$

4. Cuenta:

$$n_{11} = \#\{I_D = 1, I_A = 1\}, \quad n_{10} = \#\{I_D = 1, I_A = 0\}, \quad n_{01} = \#\{I_D = 0, I_A = 1\}, \quad n_{00} = \#\{I_D = 0, I_A = 0\}.$$

5. Calcula PRR/ROR e IC y aplica criterios de priorización.

11. Ejercicios

Ejercicio 1

Dados $n_{11} = 4, n_{10} = 196, n_{01} = 40, n_{00} = 760$, calcula PRR, ROR e interpreta.

Ejercicio 2

Dados $n_{11} = 2, n_{10} = 18, n_{01} = 5, n_{00} = 975$, calcula PRR y ROR. ¿Lo priorizarías si exiges $n_{11} \geq 3$?

Ejercicio 3 (Haldane)

Dados $n_{11} = 0, n_{10} = 25, n_{01} = 10, n_{00} = 965$, aplica corrección de Haldane y calcula *ROR*.

Ejercicio 4 (intuición)

Para el Ejercicio 1, calcula $E[n_{11}]$ bajo independencia y compáralo con n_{11} .