

Teoría y Estado del Arte

Contenido

1. Mortalidad materna (MM).....	1
2. Mortalidad infantil (MI)	3
3. Morbilidad materna extrema (MME)	4
4. Tasa de suicidios (SR)	4
5. Tasa de homicidios (HR).....	5
6. Tasa de violencia intrafamiliar (DVR).....	6
7. Propuesta para MM, MI y MME	8
8. Propuesta para SR, HR y DVR.....	8

1. Mortalidad materna (MM)

La mortalidad materna o muerte materna es un término estadístico que describe la muerte de una mujer durante el embarazo, el parto o el posparto. La Organización Mundial de la Salud (OMS) define la defunción materna como «la muerte de una mujer mientras está embarazada o dentro de los 42 días siguientes a la terminación del embarazo, independiente de la duración y el sitio del embarazo, debida a cualquier causa relacionada con o agravada por el embarazo mismo o su atención, pero no por causas accidentales o incidentales».

La reducción anual esperada en el mundo es del 5.5%, pero la real fue de 4.1% entre 2000 y 2010. El 99% de estas muertes ocurriendo en países de bajos ingresos o en vías de desarrollo. Según la OMS, en América latina la tasa de mortalidad materna promedio es de 270 por 100 mil nacidos vivos. En Colombia para 1990 era de 107 y en el 2013 era de 60 por cada 100 mil. El 24% de esas muertes, son en mujeres de comunidades indígenas.

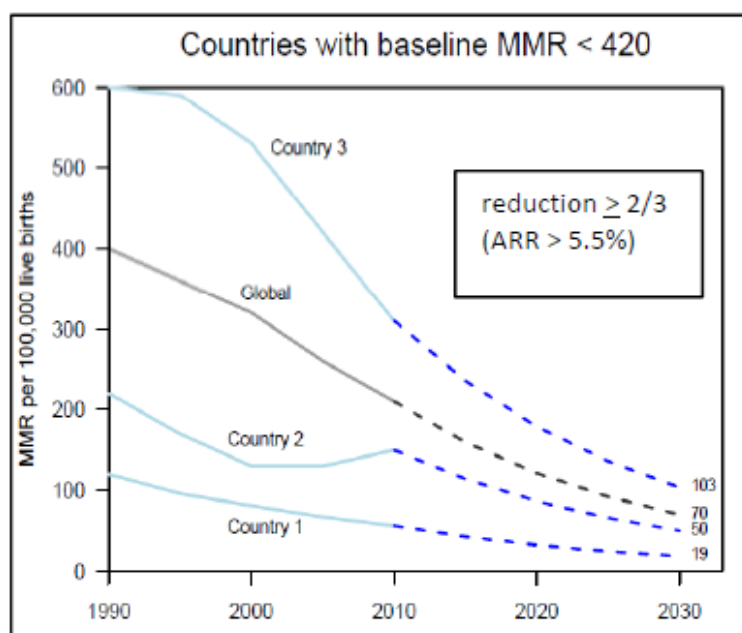
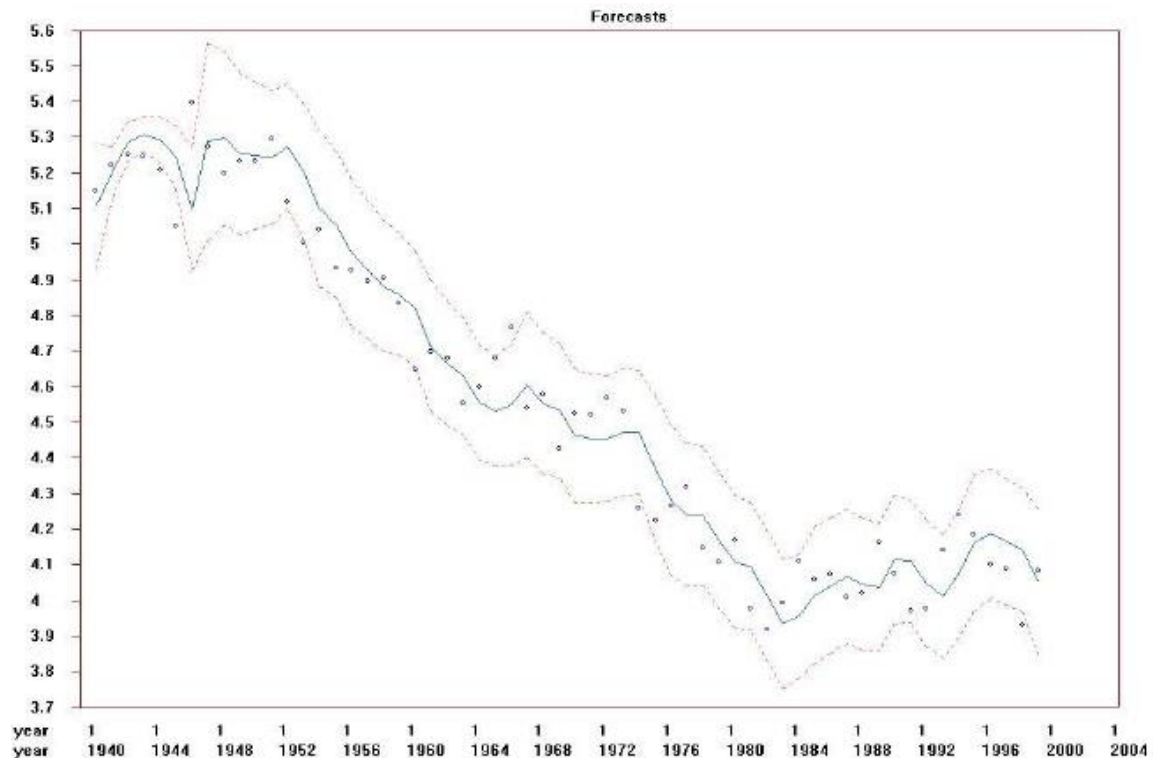


Figura 1 - Fuente: epmm@who.int

Trabajos previos

[1] Salas Maronsky, Hans. (2005). Modelo Lineal Dinámico Bayesiano en la Predicción de la Mortalidad Infantil y Materna en Venezuela (Bayesian Dynamic Linear Model for Forecasting of Maternal and Child Mortality in Venezuela). 10.13140/RG.2.2.22918.83526.



- Se creó un modelo lineal dinámico bayesiano para el comportamiento de las tasas de mortalidad infantil y materna durante el lapso 1940-1999. Periodo por años.
- Usaron las siguientes métricas de validación: Error medio cuadrático (MSE), Desviación absoluta promedio (MAD) y Log verosimilitud (LOGLIK).
- Existe correlación lineal directa alta entre las series de mortalidad infantil y la mortalidad materna (Correlación= 0.943). Tendencias a la baja.

[2] Int J Public Health. (2017). Maternal mortality and morbidity burden in the Eastern Mediterranean Region: findings from the Global Burden of Disease 2015 study.

- Usaron la técnica least-squares spline regression para modelar la tasa de mortalidad materna.
- La tasa de mortalidad materna se correlacionó fuertemente con el estado sociodemográfico, donde los países de ingresos más bajos contribuyeron más a la carga de mortalidad materna en la región.

[3] Trends in maternal mortality: 2000 to 2017. Estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division.

- Enfoque bayesiano (regresores) con intervalos de confianza de 80%.
- Estimación por subregiones del mundo y el periodo es por años.

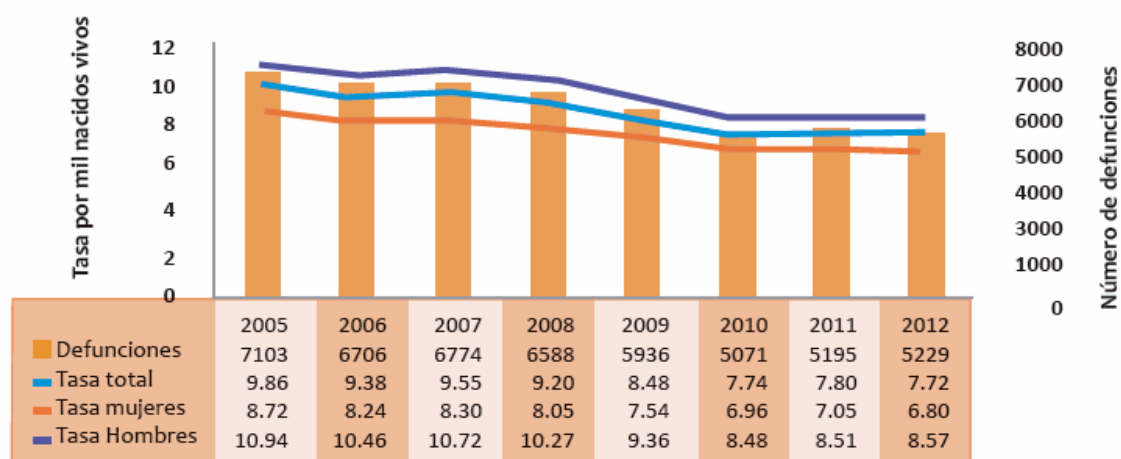
- Tasa de mortalidad materna (estimación modelada, por 100,000 nacimientos vivos) a nivel mundial y por países, entre el 2000 y el 2017:

<https://data.worldbank.org/indicator/SH.STA.MMRT>

2. Mortalidad infantil (MI)

La mortalidad infantil es el indicador demográfico que señala el número de defunciones de niños en una población de cada mil nacimientos vivos registrados, durante el primer año de su vida.

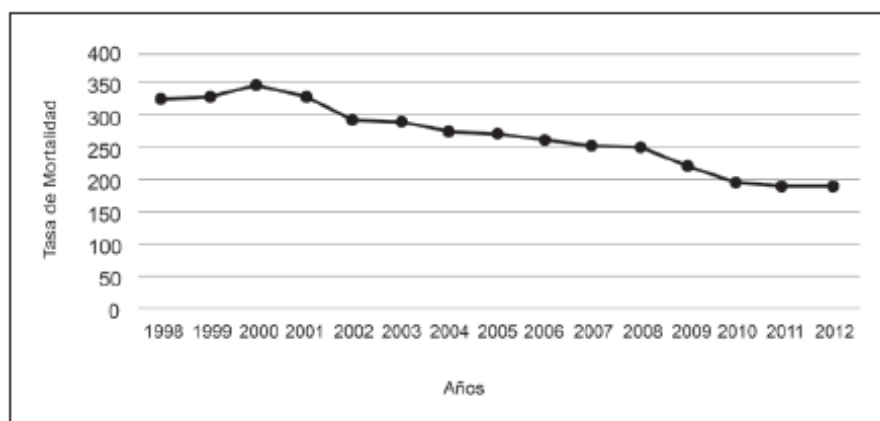
La mortalidad neonatal (<28 días) es el principal componente de la mortalidad infantil (<1 año) y participa aproximadamente del 50% de la mortalidad perinatal. Dado el progresivo y sostenido descenso de la mortalidad postneonatal (28 a 365 días de vida), actualmente los mayores esfuerzos se concentran en el período perinatal para aproximarnos a los objetivos del MDG 4.



Fuente: Análisis de situación de Salud. Colombia 2014 (Ministerio de Salud y Protección Social, 2014)

Trabajos previos

[4] Giraldo, Diana, Atehortúa, Angélica, García-Arteaga, Juan D., Romero, Eduardo, Rodríguez, Jesús. Modelo para el análisis de la mortalidad en Colombia 2000-2012. Revista de Salud Pública [en línea]. 2017, 19(2), 241-248 [fecha de Consulta 9 de junio de 2020]. ISSN: 0124-0064.



- El modelo propuesto para el análisis es similar al modelo de Murray tomando en cuenta las variables correspondientes al PIB per cápita y año: $\ln(m_{R,e,s,i}) = C_{R,e,s,i} + \beta_1 \ln(P_R) + \beta_2 \ln(P_R)^2 + \beta_3 T$. Predicción por año y departamento.

- Para la implementación del modelo propuesto es necesario incorporar la información del Producto Interno Bruto (PIB) per cápita por departamento.
- Los coeficientes del modelo: $C_{R,e,s,i}$, β_1 , β_2 y β_3 se estiman usando una regresión lineal con los datos disponibles usando el método de mínimos cuadrados.
- Usaron las librerías de R, para generar los modelos.

3. Morbilidad materna extrema (MME)

La Organización Mundial de la Salud (OMS) define la morbilidad materna extrema como un estado en el cual una mujer casi muere, pero sobrevivió a una complicación ocurrida durante el embarazo, el parto o dentro de los 42 días de la terminación del embarazo.

Las metodologías propuestas son muy similares a las usadas para la mortalidad materna y mortalidad infantil. La mayoría de los modelos predictivos son para clasificar si un paciente pudiera o no fallecer, y usan SVM o Regresiones Logísticas.

Trabajos previos

[5] Ngianga-Bakwin Kandala, Gebrenegus Ghilagaber. (2014). Advanced Techniques for Modelling Maternal and Child Health in Africa. Editorial Springer.

- Enfoque de modelado clásico: modelos multinivel y regresiones log-lineal.
- Enfoque de modelado bayesiano: Markov chain Monte Carlo (MCMC) y Bayesian Structured Additive Regression.

4. Tasa de suicidios (SR)

El suicidio supone quitarse voluntariamente la vida. Diversas estadísticas ubican el suicidio como la cuarta causa de muerte más frecuente en todo el mundo, con más de 9.000 intentos diarios.

Trabajos previos

[6] Antonio Preti & Gianluca Lentini (2016) Forecast models for suicide: Time-series analysis with data from Italy, Chronobiology International, 33:9, 1235-1246, DOI: 10.1080/07420528.2016.1211669

- Enfoque simple: regresiones lineales. Las cuadráticas generaron muy buenos resultados, pero no reproducen la estacionalidad de la serie.
- Enfoque complejo: usaron 4 métodos de series de tiempo: ARIMA, Holt-Winters, modelo ETS y el modelo TBATS.
- Para los datos de hombres, el TBATS fue el mejor modelo basado en el MAPE. Para los datos de mujeres, los enfoques simples ofrecieron mejores resultados.
- Datos mensuales, de 1969 al 2003 (35 años) en Italia. Los datos se basan en investigaciones policiales e informes forenses.
- Validación con MAE, RMSE, MAPE y MASE.
- Los datos fueron divididos en 2 datasets, uno de entrenamiento con el 80% de los datos (336 meses) y uno de pruebas con el 20% (84 meses).
- Outliers fueron removidos y se interpoló con ARIMA para completar los datos.
- Se observó y se reprodujo el comportamiento estacional del suicidio.
- El margen de error obtenido fue alrededor del 10%.

[7] Joo-Young Lee (2019). Search trends preceding increases in suicide: A cross-correlation study of monthly Google search volume and suicide rate using transfer function models.

- Se usaron series de tiempo ARIMA y la metodología Box-Jenkins para modelar las series de tiempo.
- Los términos relacionados con la ansiedad, el sueño y el desempleo mostraron una correlación significativa.
- Los datos fueron mensuales, entre el 2004 y el 2017 de Estados Unidos.
- Limitaciones: con respecto a las búsquedas en Google se pierde información asociada al sexo y la edad de la persona que la realiza.

[8] Emad Y, Hadianfard H. [Forecasting Suicide Based on Sexuality, Marital Status, Coping Strategies, Religious Orientation, and Depression Rate (Persian)]. *Iranian Journal of Psychiatry and Clinical Psychology*. 2019; 25(2):178-193. <http://dx.doi.org/10.32598/ijpcp.25.2.178>

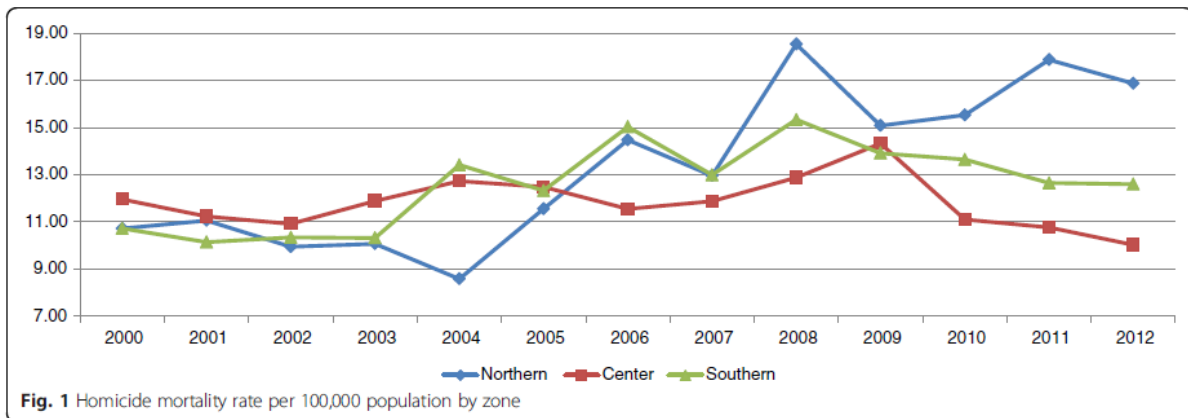
- La tasa de suicidio parece aumentar en los grupos de minorías. Hay más suicidios en hombres que en mujeres (relación 4-1).
- Los resultados mostraron que las variables anteriores en la ecuación predictiva más fuerte tenían el 78% de la varianza del riesgo de suicidio.

5. Tasa de homicidios (HR)

Número de defunciones, por causas relacionadas con homicidios/asesinatos, por cada 100.000 habitantes, en un período determinado.

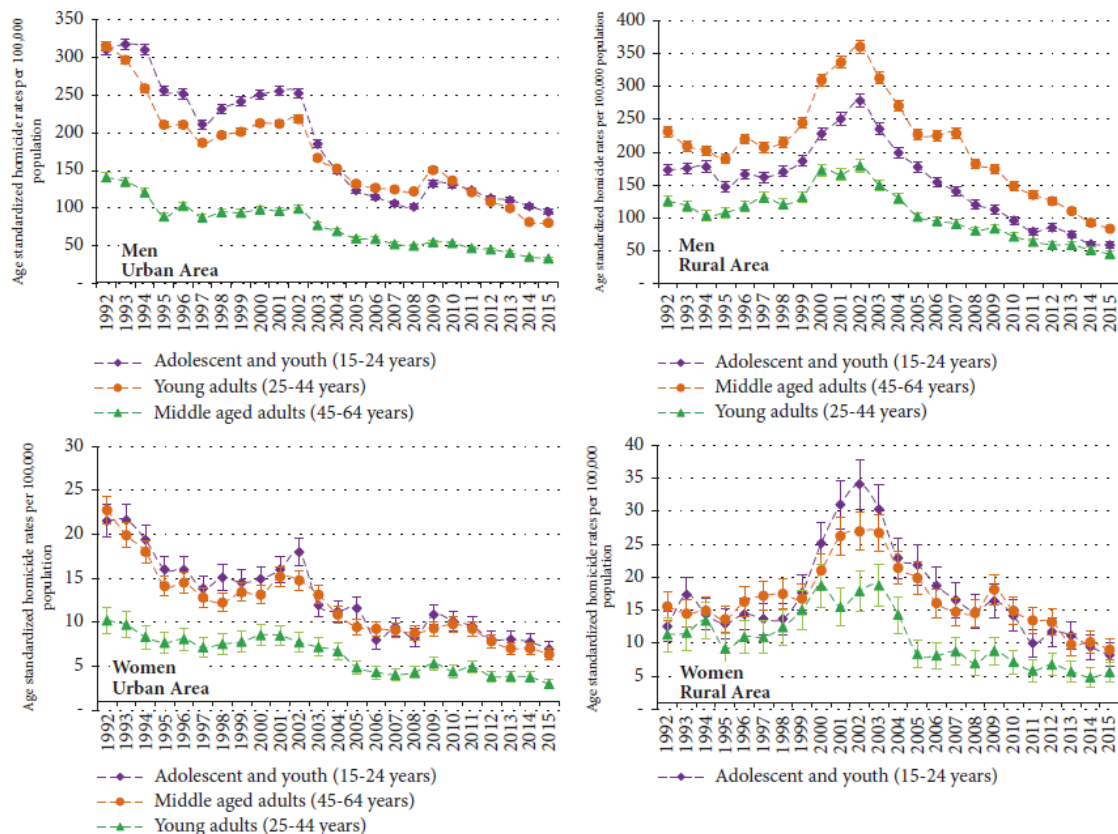
Trabajos previos

[9] Otzen, T., Sanhueza, A., Manterola, C. et al. Homicide in Chile: Trends 2000 – 2012. *BMC Psychiatry* 15, 312 (2015). <https://doi.org/10.1186/s12888-015-0632-5>



- Se analizó el cambio porcentual anual (APC) de las tasas, y se ajustó un logaritmo de las tasas por año y región mediante la aplicación de modelos de regresión lineal.
- Se usaron intervalos de confianza al 95% para las predicciones.
- Rango de datos entre el 2000 y 2012 en Chile. Los datos fueron por años.
- Variables independientes usadas: género, grupo de edad, distribución geográfica (por zona y por región) y tipo de homicidio.

[10] Katherine Vallejo, Jose Tapias, and Ivan Arroyave (2018). Trends of Rural/Urban Homicide in Colombia, 1992-2015: Internal Armed Conflict and Hints for Postconflict.



- La agrupación de los datos fue por: Sexo, Zona (Urbana y Rural) y Rango de Edad.
- Algoritmo: modelo de regresión de Poisson y el software de dominio público Joinpoint Trend Analysis Software.
- Se usaron 475640 datos de homicidios para el análisis.

6. Tasa de violencia intrafamiliar (DVR)

La violencia intrafamiliar o violencia doméstica, puede incluir distintas formas de maltrato, desde intimidación hasta golpes pasando por el acoso o los insultos. La violencia doméstica se refiere a cualquier acto de ejercicio de infracción, mental o física, entre miembros de la familia, incluidos los cónyuges (como cónyuges, excónyuges, convivientes, novios y novias, socios de del mismo sexo), hijos, hermanos, todos parientes de sangre, etc.

Trabajos previos

[11] Richard A. Berk, Susan B. Sorenson, and Geoffrey Barnes (2016). Forecasting Domestic Violence: A Machine Learning Approach to Help Inform Arraignment Decisions.

- El algoritmo usado fue Random Forest Regression
- Los datos que usaron fueron por mes, desde enero del 2007 hasta octubre del 2011.
- El entrenamiento de los modelos fue con datasets con training y test.

[12] Jin-Ding Lin, Lan-Ping Lin, Pei-Ying Lin, Jia-Lin Wu, Chien-De Li, Fang-Yu Kuo (2010). Domestic violence against people with disabilities: Prevalence and trend analyses

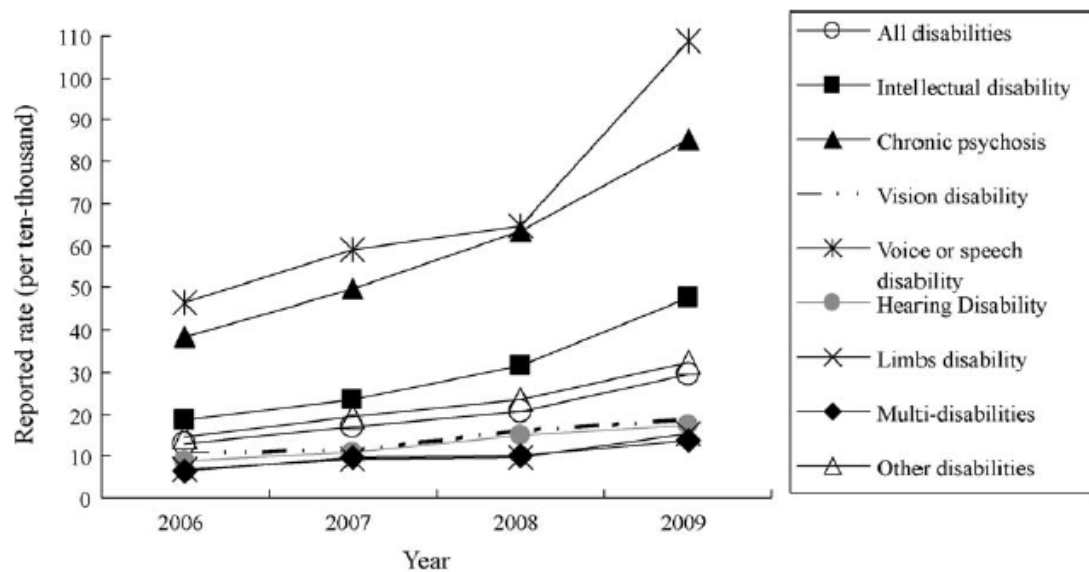
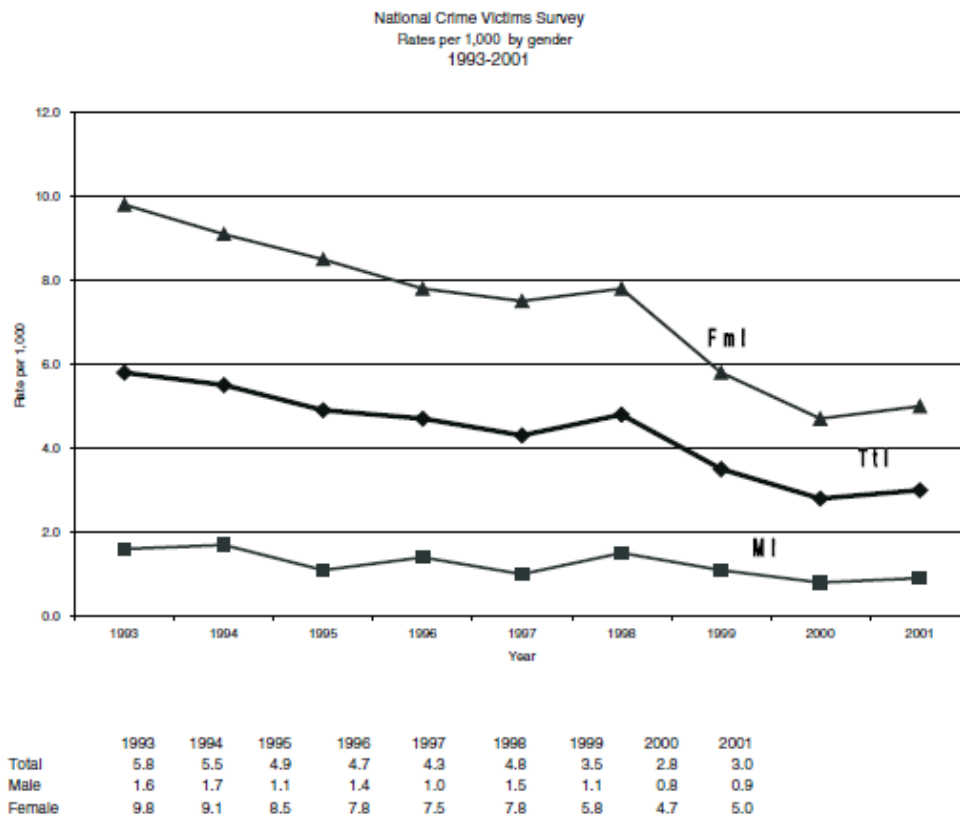


Fig. 1. Annual domestic violence reported rate by disability type in different years.

- Datos de Taiwan, del 2006 hasta el 2009 (4 años). Datos reportados por años.
- Modelo de estimación lineal (LR)
- Validación con R-cuadrado y p-value.
- Insights: las personas con discapacidad de voz o del habla, psicosis crónica y discapacidad intelectual fueron la mayor prevalencia de violencia doméstica entre las discapacidades en el estudio.

[13] Richard A. Berk, Yan He, Susan B. Sorenson (2005). Developing a practical forecasting screener for domestic violence incidents.



- Casos por años, desde 1993 hasta el 2001.

7. Propuesta para MM, MI y MME

Debido a la alta correlación que existe entre las enfermedades MM, MI y MME, se puede usar el mismo enfoque para estimarlas, más no el mismo modelo, ya que cada enfermedad tiene diferente escala.

La granularidad puede ser por mes (los estudios mencionados son por año), departamento y sexo. Se puede usar ANOVA para confirmar si la varianza del grupo es distinta a la de los intergrupos.

Se recomienda probar los siguientes algoritmos, puede ser con el siguiente orden:

- Regresiones lineales, polinomiales y logarítmicas
- Regresión spline de mínimos cuadrados
- Regresión bayesiana y variantes
- Series de tiempo

Dependiendo de los resultados, podemos explorar otros enfoques más robustos, pero menos descriptivos, como Redes Neuronales o Deep Learning. Por ahora, no se han encontrado trabajos previos con Series de Tiempo, probablemente, porque estas enfermedades no tienden a ser periódicas y menos en Colombia.

Como medidas de calibración, se pueden usar RMSE, MAPE y LOGLIK. El entrenamiento puede ser con 80-20% de los datos, respectivamente. La salida puede ser expresada en intervalos de confianza, de 80 o 90%.

Se puede evaluar el incluir otras variables independientes, tales como el PIB por departamento, la cual está fuertemente correlacionada con las políticas de salud de cada región.

Con respecto a la precisión de los modelos, se puede establecer como meta el valor usado en el intervalo de confianza. Algunos estudios hablan de errores menores al 10%, sin embargo, esto es por año y por país.

8. Propuesta para SR, HR y DVR

Debido a que para la tasa de homicidios y de suicidios el género con mayor número de casos es el masculino, y para violencia intrafamiliar el femenino, es difícil crear un solo modelos que agrupe a las 3 enfermedades. Además, la tasa de suicidios parece ser totalmente estacionaria en cambio los homicidios en Colombia están más influenciados por la región (rural >> urbana).

La granularidad puede ser por mes (se encontraron estudios tanto por mes como por año), departamento, región y sexo. Se puede usar ANOVA para confirmar si la varianza del grupo es distinta a la de los intergrupos.

Se recomienda probar los siguientes algoritmos, puede ser con el siguiente orden:

- Regresiones lineales, polinomiales y logarítmicas
- Series de tiempo ARIMA
- Regresión bayesiana y variantes

Como medidas de calibración, se pueden usar MAE, RMSE, MAPE y MASE. El entrenamiento puede ser con 80-20% de los datos, respectivamente. La salida puede ser expresada en intervalos de confianza, de 80 o 90%.