

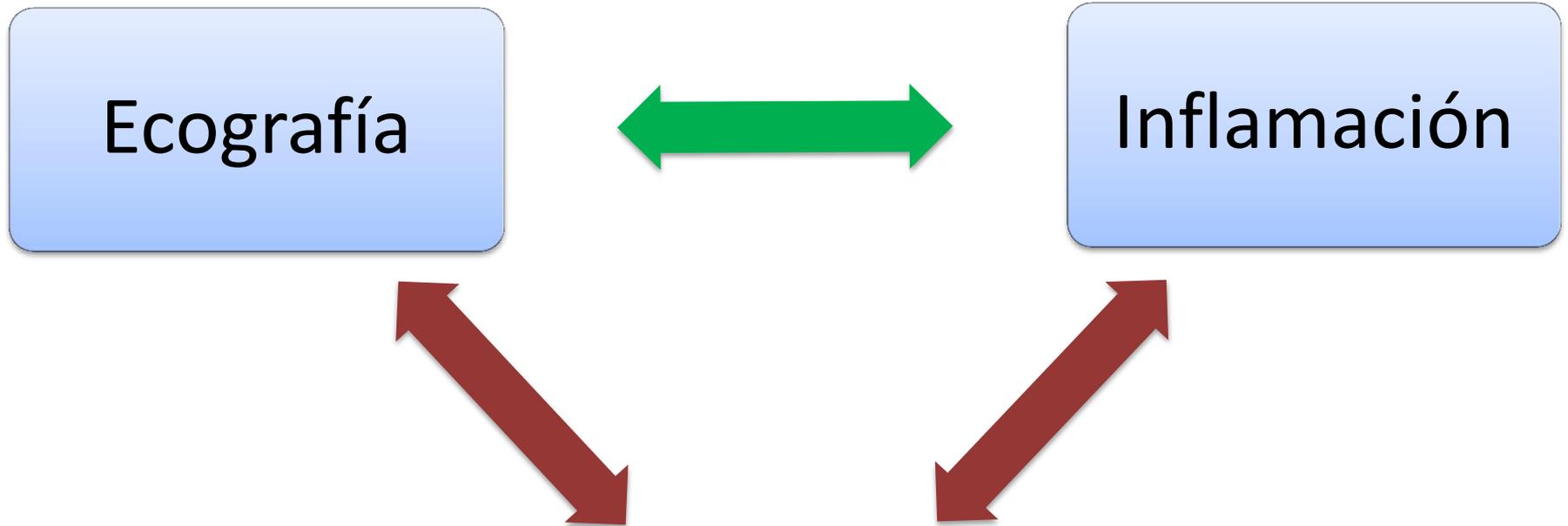
Ultrasound colour Doppler is associated with synovial pathology in biopsies from hand joints in rheumatoid arthritis patients: a cross-sectional study. **Clinical and epidemiological research 2013**

Discusión

Marco de referencia. Situación actual, bibliografía.

- **Gold standard: procedimiento por artroscopia.**
- La ecografía ha demostrado ser superior que la clínica para la evaluación, pronóstico y tratamiento.
- Existen estudios de correlación previos?
Resultados controvertidos.
- Se utiliza el gold standard habitualmente?. No es un método invasivo y costoso.
- **Es razonable tener un estudio menos costoso e invasivo.**

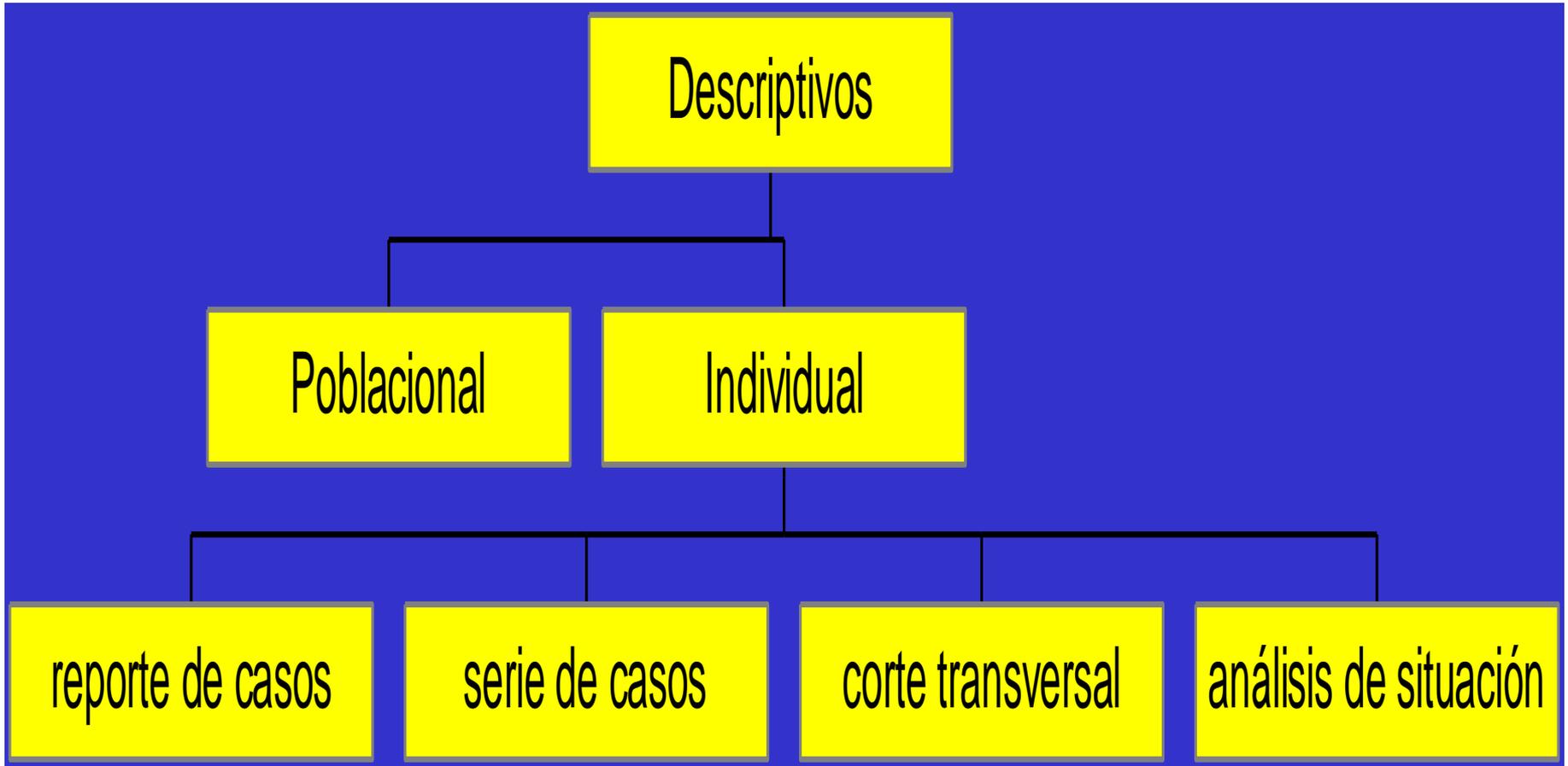
Objetivos. Causalidad.



Operador dependiente: ecografía, cirugía, patólogo.
El traumatólogo tomó de biopsia no ciego de la ecografía.
Que pasa con los falsos negativos?

Métodos. Diseño del estudio.

- Estudio de corte transversal???



Ojo!!! principal problema de los estudios transversales es la temporalidad en los criterios de causalidad, **es descriptivo de observación**

Diseño: Estudio de intervención

1. Se define el procedimiento artroscópico (gold standard), ya que no es una indicación habitual (biopsia).
2. Se define como se mide la variable (profesionales experimentados en ecografía, reumatólogos y cirujanos) esto también tiene impacto en la validez externa del estudio.
3. Que tipo de ecógrafo se utilizará y cual es el programa (ventaja: todos los pacientes están medidos de la misma forma). La ecografía se realizó en articulaciones pre-determinadas, en una posición pre-definida y con un programa específico.

Intervención.

4. Técnica de toma de biopsia se realiza en un lugar definido con una técnica pre-determinada.
5. El operador (cirujano) no era ciego, que pasa en este caso con los falsos negativos. El costo de esta estrategia favorece la correlación (sesgo?) y como se comportan los falsos negativos (son menores?)
6. Son independientes los dos estudios(biopsia y ecografía?).
- 7. Es un estudio de intervención con una sola visita.**

Muestreo/Tamaño de muestra

Criterios de elegibilidad.

- como se determinó el tamaño de muestra?



No conocemos con que poder y error alfa se calculó la muestra. En general en este tipo de estudio el tamaño de muestra es menor a otros tipos de estudios.

- La población fuente proviene de un solo centro.



La inferencia a otras poblaciones con diferentes características esta limitada.

- Elegibilidad: fundamentalmente a las limitaciones técnicas quirúrgicas.



Evita sesgos y confusores vinculados a la técnicas (ej: toma dificultosa de la muestra).

- En este caso las dos exposiciones se llevan a cabo en el mismo individuo.



La ventaja de esta estrategia es que disminuye la **variabilidad** interindividuos.

Estadística. Correlación. Modelo determinístico vs modelo de regresión

- Si uno conoce el valor de las variables, determina completamente el valor de la variable respuesta.

Variable respuesta \longrightarrow $Y = f(X_1, X_2, \dots, X_k)$

Variables \swarrow

ej: Presión=Fuerza /superficie

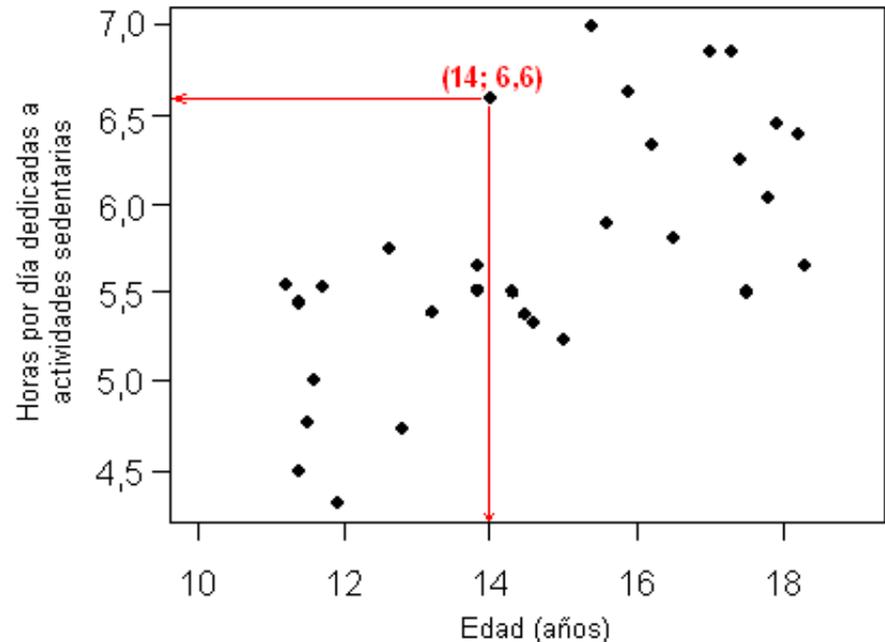
- Si uno conoce el valor de las variables, NO determina completamente el valor de la variable respuesta.

$$Y = f(X_1, X_2, \dots, X_k) + \epsilon$$

Variable error \longleftarrow

Diagrama de dispersión

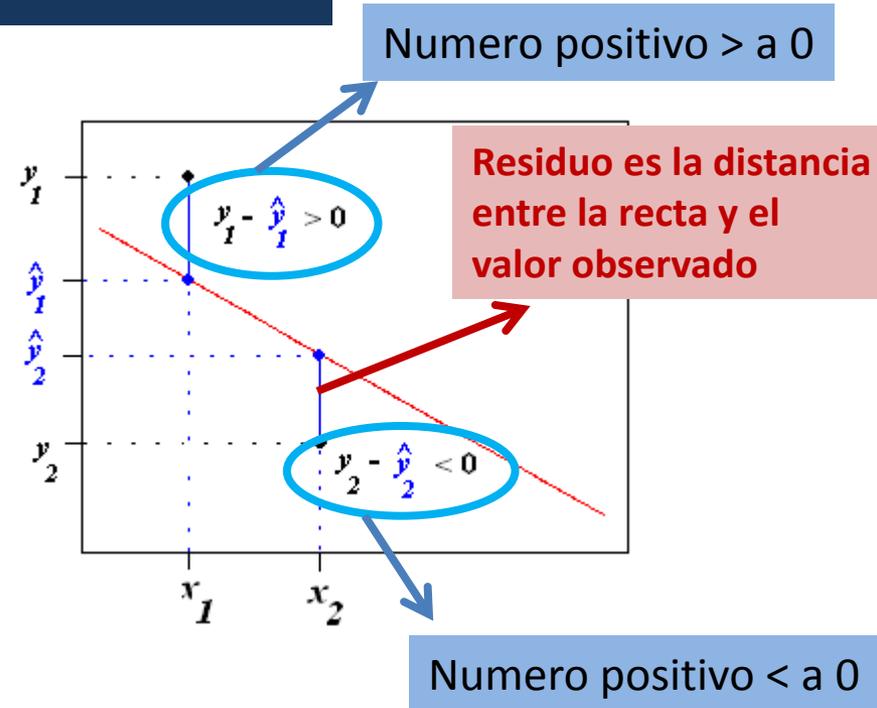
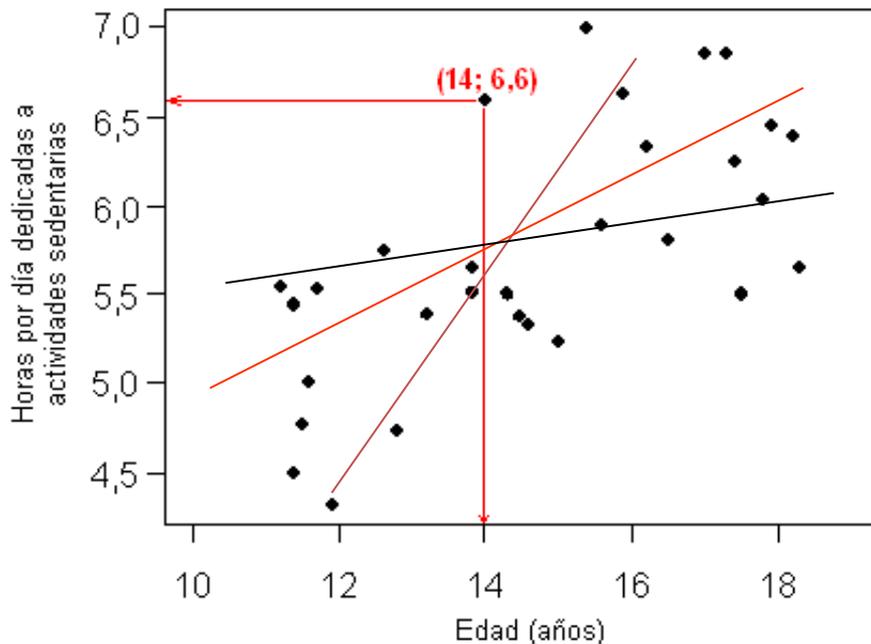
- Relación entre dos variables cuantitativas. Cada punto corresponde a un par de valores medidos sobre el mismo individuo.
- Puede mostrar una relación positiva o negativa.
- Muestra forma, dirección y grado de asociación entre las dos variables cuantitativas.



Regresión simple. Cuadrados mínimos

Que recta elijo?

$$y = a + b x$$



- Me conviene utilizar aquella recta que se encuentre mas cerca de todos los puntos. Utilizo el método de cuadrados mínimos

Modelo de Regresión Simple

- el valor de y esta determinado por x
- error permite describe la variabilidad de las observaciones.
- El valor de la pendiente es el cambio de y cuando la variable explicativa aumenta una unidad.
- La respuesta es una estimación de la media poblacional
- Los residuos proveen una medida de la calidad del ajuste

$$\overbrace{Y = \alpha + \beta x}^{\text{Recta}} + \underbrace{\varepsilon}_{\text{Otras cosas además de } x}$$

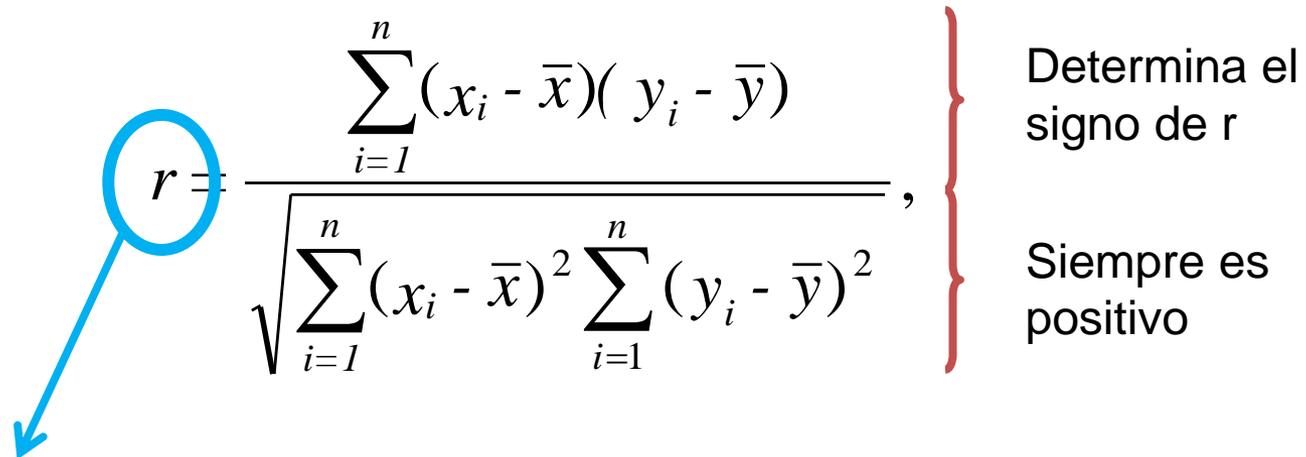
***Parámetros
poblacionales***



$y = a + bx$ Recta estimada

Coeficiente de correlación

- Mide el grado de asociación entre dos variables continuas
- Valor entre -1 y 1.
- No distingue entre variable explicativa (X) y variable respuesta (Y).
- A mayor valor absoluto de r , mayor el grado de asociación lineal.
- Cuando $r = 0$, no presenta una tendencia lineal creciente ni decreciente.
- Valores de r cercanos a 1 ó -1 indican que los puntos están cerca de una recta

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}},$$


Determina el signo de r

Siempre es positivo

Este es el valor que se informa

Interpretación/Inferencia de r

Supuestos

- Los sujetos son elegidos al azar.
- **Observaciones son independientes. Es independiente cuando el cirujano no es ciego?**
- Se miden las dos variables (x e y) en cada sujeto.
- Los valores de una variable no deben ser obtenidos mediante funciones que dependan de los valores de la otra.
- Los valores de X no están determinados por el investigador.
- **Los valores de X e Y provienen de poblaciones con distribución gaussiana. Transformaciones**
- **Se le aplica el método: Test t para ver significancia**

- El coeficiente de determinación R^2 mide la proporción de la variación total explicada por la recta ajustada.
- $R^2 = 0$ no explica nada (la SCT es igual a SCR).
- $R^2 = 1$ cuando todos los puntos están sobre la recta, la variabilidad observada de la respuesta es explicada totalmente por la regresión y la SCR es cero.
- **Un dato outliers puede cambiar la información.** (E descriptiva)

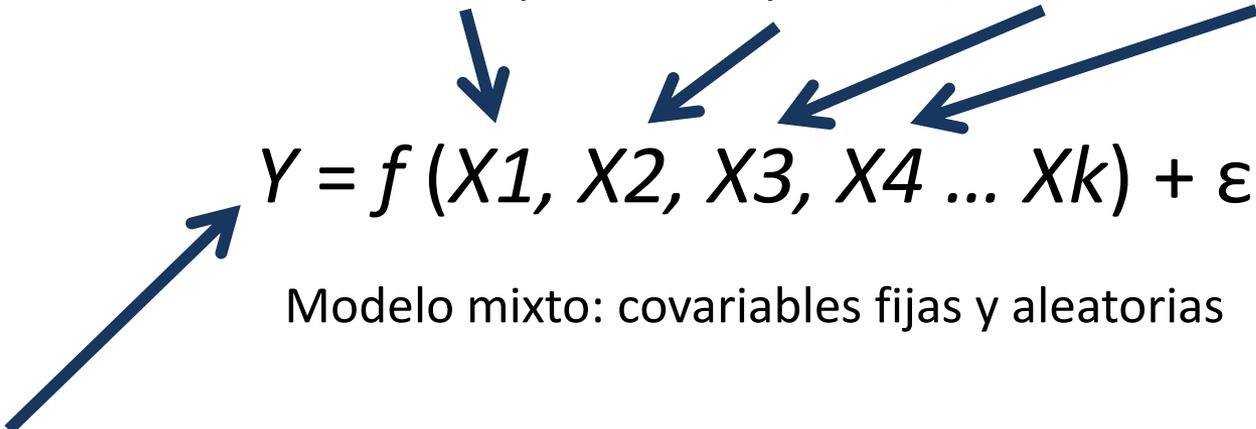
$$R^2 = \frac{SCT - SCR}{SCT}$$

Regresión lineal Múltiple

Relación entre: la variable Y y las variables explicativas X, controlando por los efectos de las otras.

Variables explicativas o independientes: X *Numérica o dicotómica.*

X1=Articulación biopsiada; x2=posición; x3=side; x4= v Cfmax.

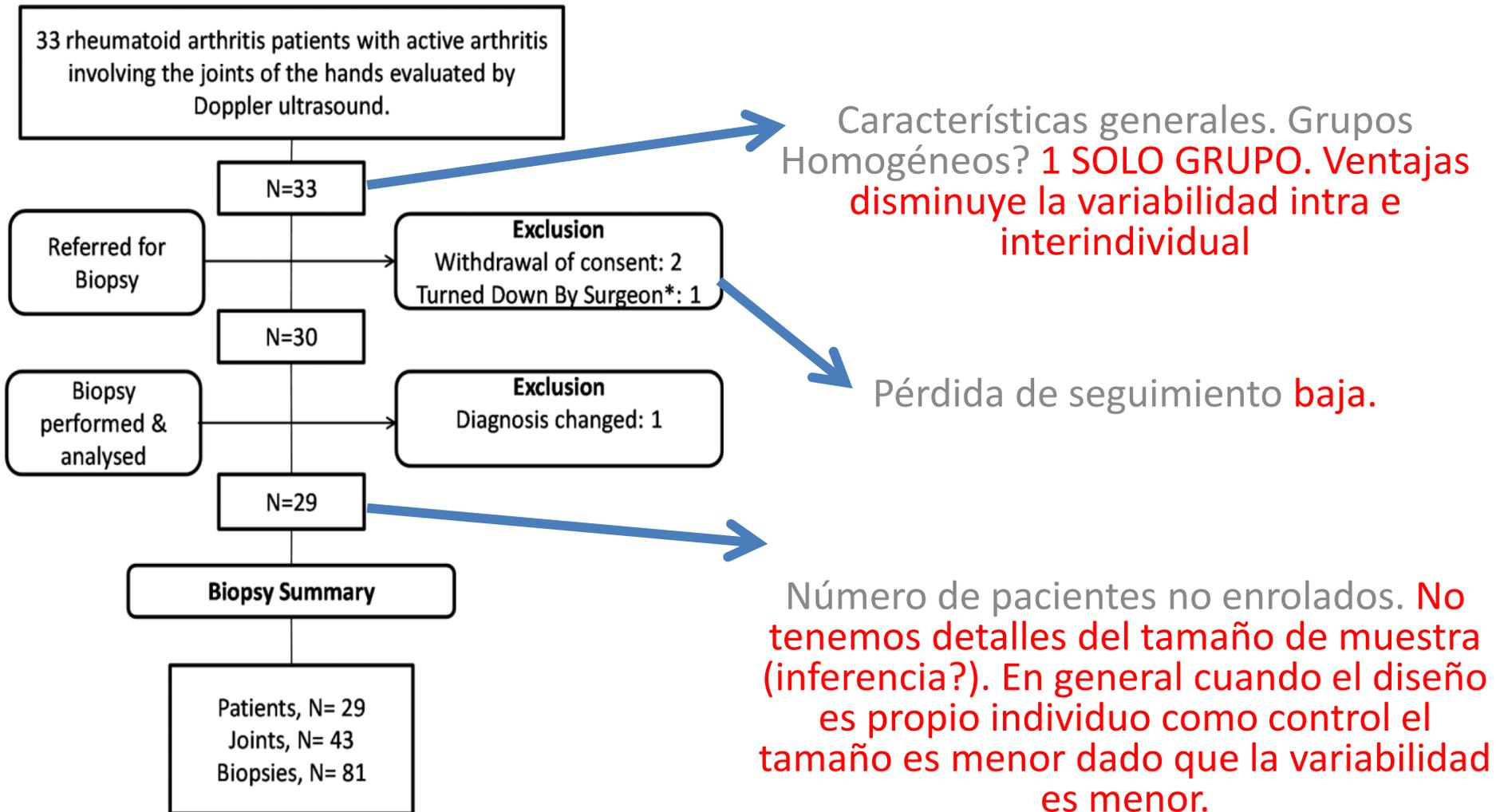


$Y = f(X_1, X_2, X_3, X_4 \dots X_k) + \varepsilon$

Modelo mixto: covariables fijas y aleatorias

Variable respuesta o dependiente: Y *debe ser Numérica.* Krenn score; factor de Von willebrand; CD8; Log10 (Ki67); Log 10 (CD3)

Resultados



Resultados grafico de dispersión

Como se interpreta?.

Dirección ?

Signo?

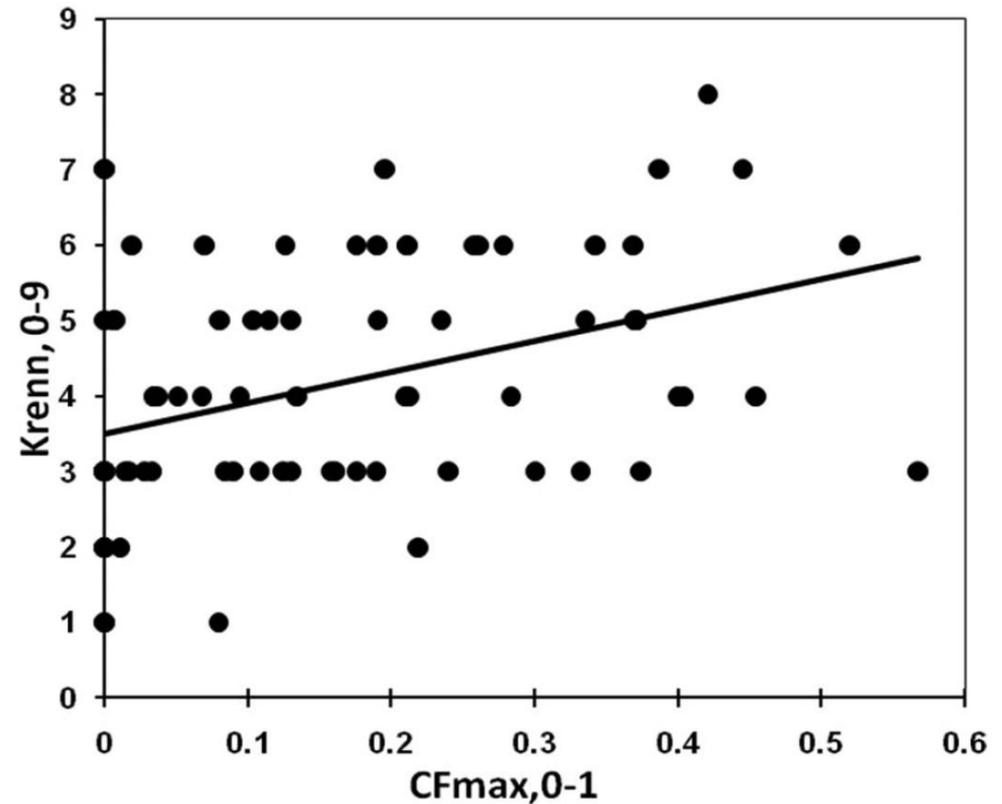


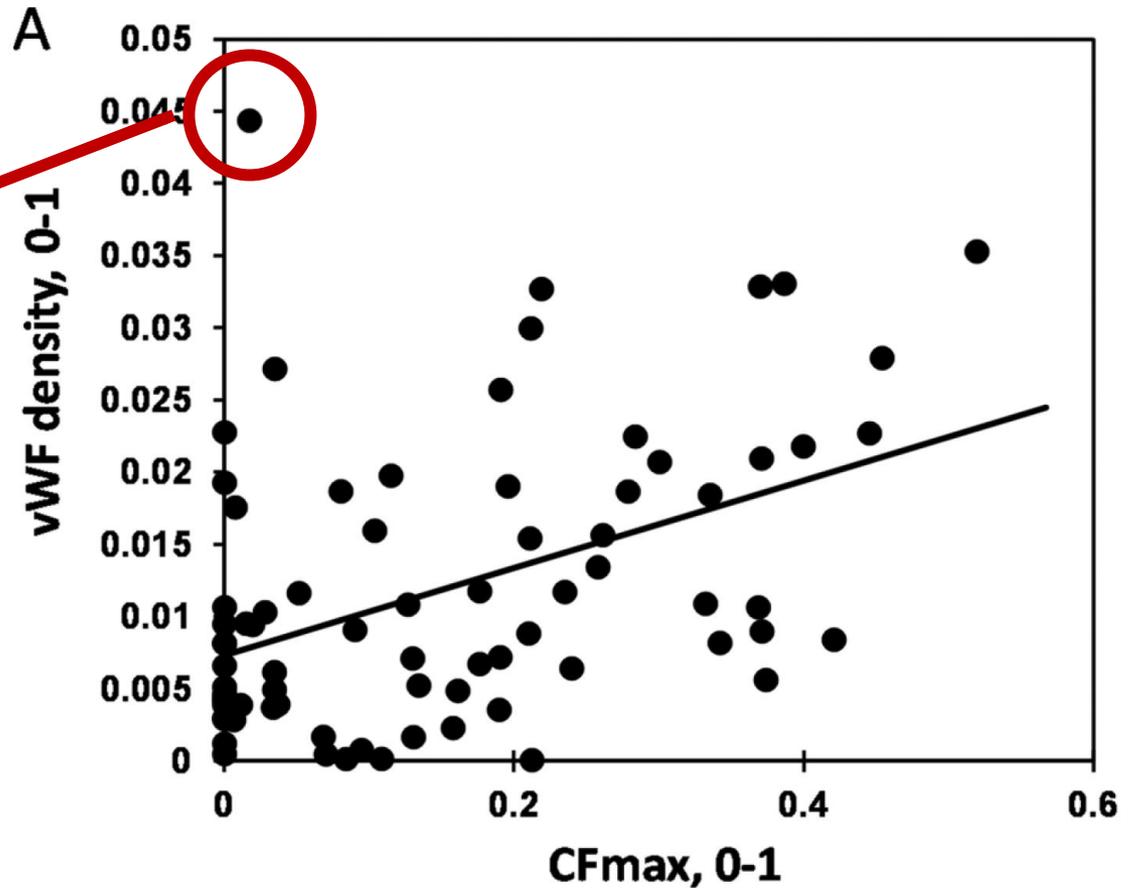
Figure 2 Scatter plot of the association between the maximum colour fraction (CF_{max}) (x-axis) and the Krenn score. Squareroot CF_{max} was used for the statistical calculations. Approximated Spearman, $r=0.44$.

Resultados grafico de dispersión 2.

Es un outliers?
Como se trato en caso
de ser así?

$r:0.44$

$r^2:?$



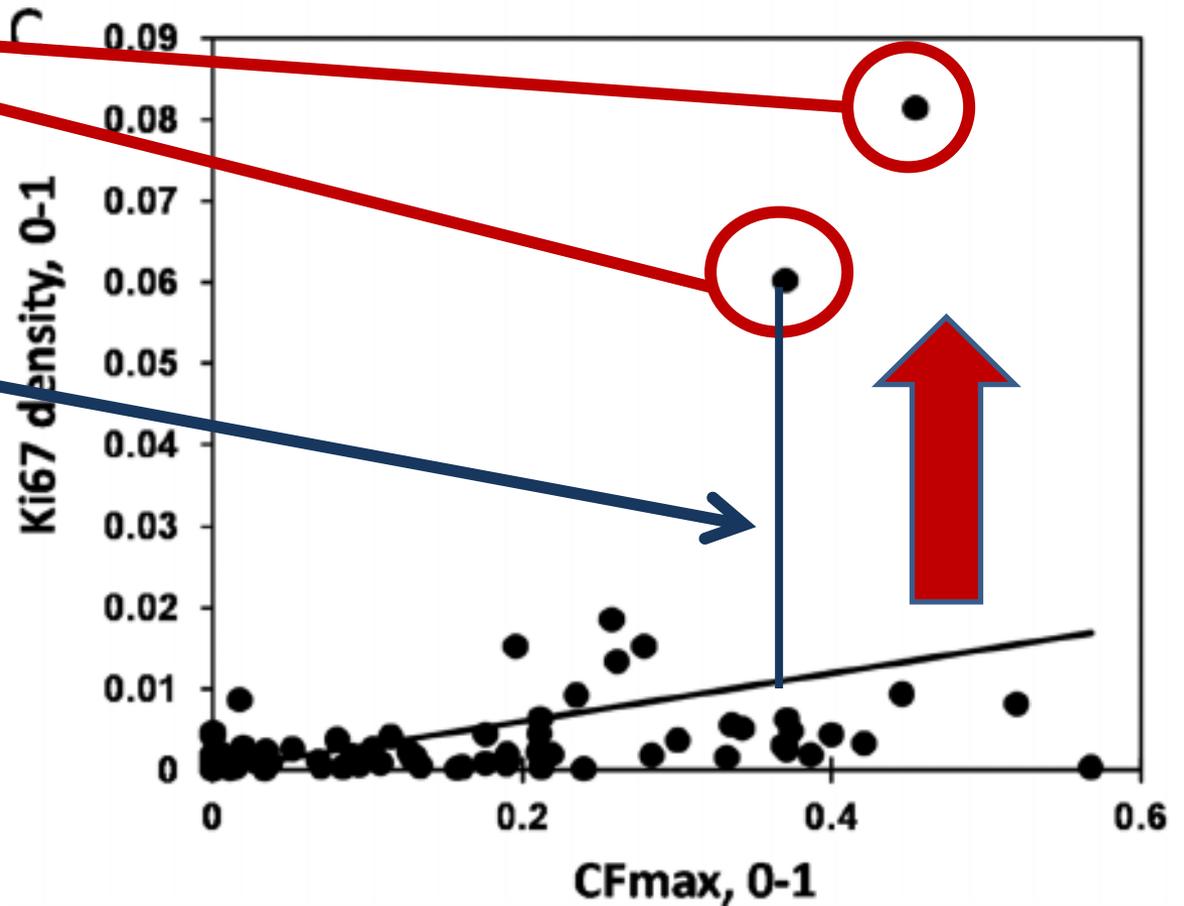
Resultados grafico de dispersión 3.

Es un outliers?

Outliers: no sigue la estructura de los demás datos, en general produce un residuo grande.

Como se trato en caso de ser así?

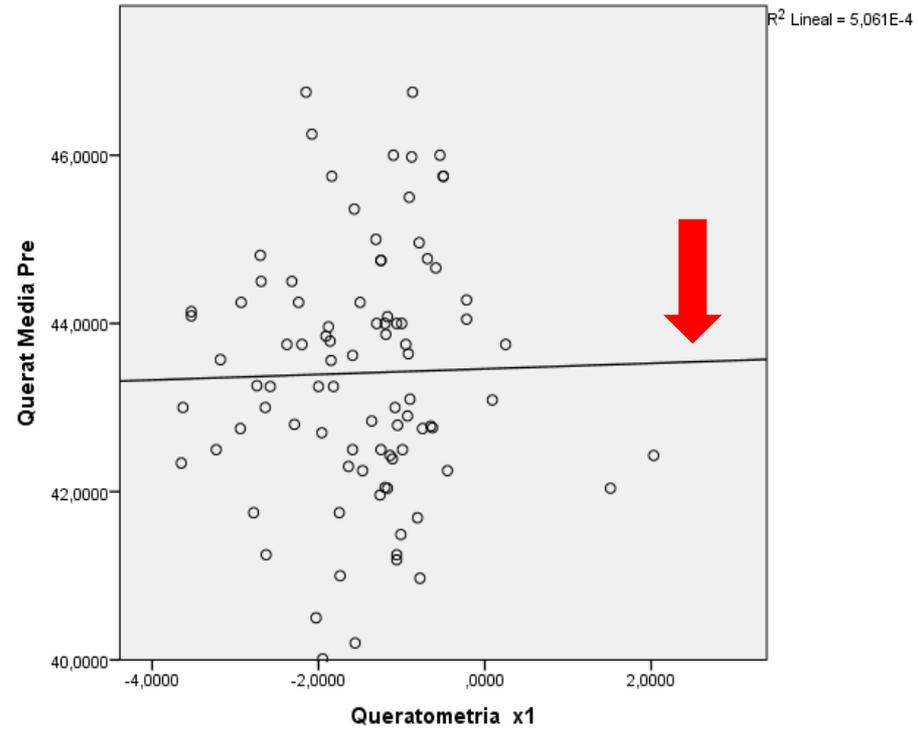
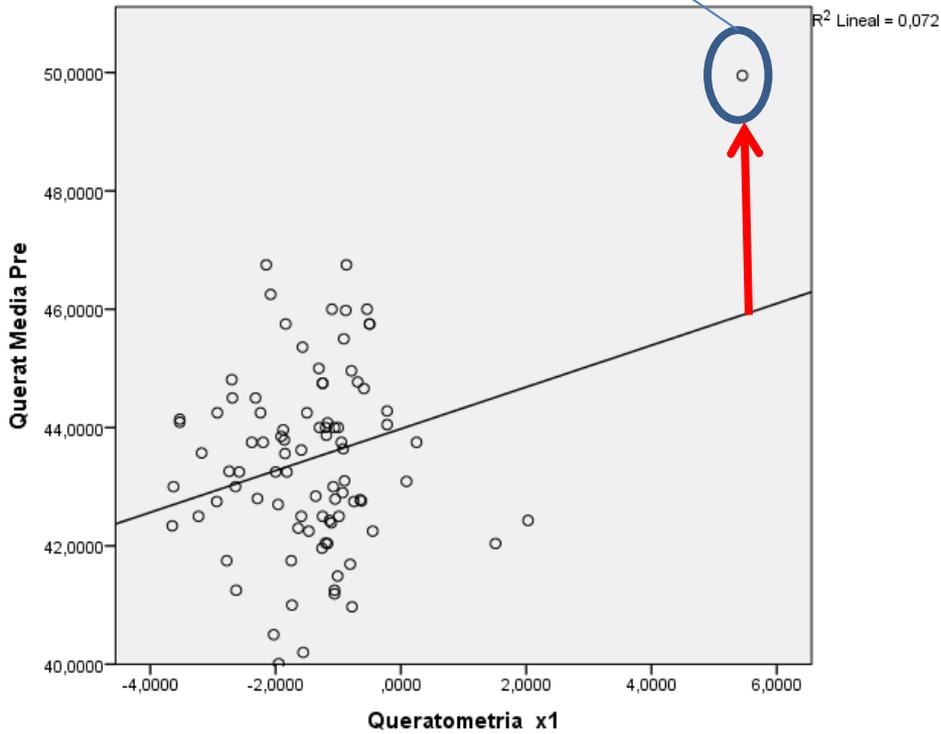
Potencialmente influyente (con alta palanca) al ser excluida del análisis podría provocar un cambio importante en la recta ajustada.



r:0.53
r²:

Ejemplo práctico

Outlier

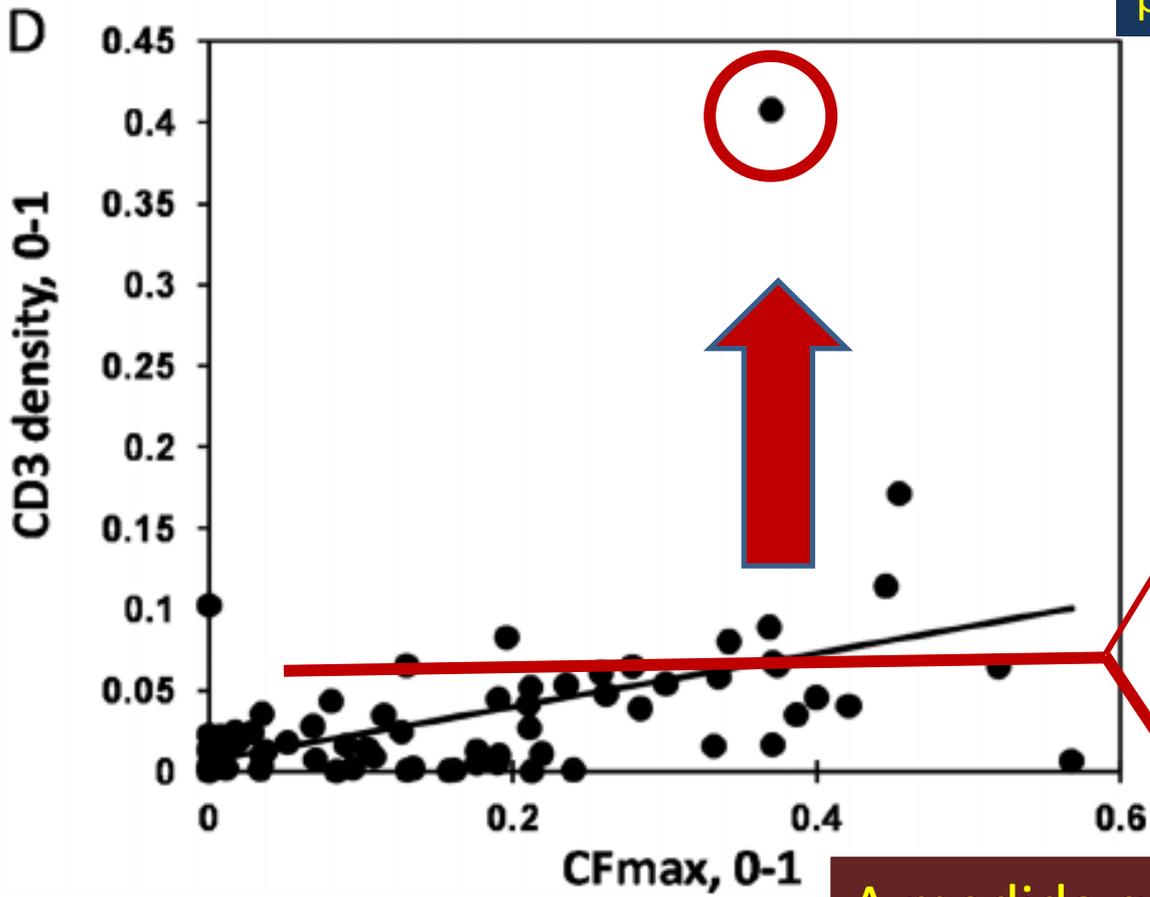


Resultados grafico de dispersión 4.

Estadística

1. Distancia de Cook (COO).
2. DFFITS.
3. DFBETA.

1 y 2 son útiles si el interés está en la predicción. DFBETA es útil si el interés está en la estimación de un parámetro de la regresión



$r:0.57$; r^2 :
Que pasa si elimino el outliers y me queda esta recta?

A medida que aumenta CFmax, CD3 no aumenta. Hay correlación?

Proponen 5 modelos (outcomes diferentes)

Table 3 RA histology versus $\sqrt{CF_{max}}$ stepwise covariate elimination

Dependent variable	Full model (p value)	1st reduced model (p value)	2nd reduced model (p value)	3rd reduced model (p value)
Krenn synovitis score* (Spearman, approximated $r=0.43$)	X1= joint biopsied (p=0.43)	} Biopsy position (p=0.400) Side (p=0.24) $\sqrt{CF_{max}}$ (p=0.03) Joint biopsied (p=0.13)	$Y = f(X1, X2, X3, X4) + \epsilon$	
	X2= biopsy position (p=0.35)			
	X3= side (p=0.26)			
	X4= $\sqrt{CF_{max}}$ (p=0.09)			
von Willebrand factor* (Spearman, approximated $r=0.44$)	Joint biopsied (p=0.15)	Biopsy position (p=0.96)	Side (p=0.51)	$\sqrt{CF_{max}}$ (p=0.10)
	Biopsy position (p=0.96)	Side (p=0.49)	$\sqrt{CF_{max}}$ (p=0.10)	Joint biopsied (p=0.38)
	Side (p=0.51)	Biopsy position (p=0.35)	Side (p=0.12)	$\sqrt{CF_{max}}$ (p=0.02)
	$\sqrt{CF_{max}}$ (p=0.10)	Side (p=0.04)	$\sqrt{CF_{max}}$ (p=0.01)	Joint biopsied (p=0.97)
CD68* (Spearman, approximated $r=0.53$)	Joint biopsied (p=0.38)	Biopsy position (p=0.35)	Side (p=0.09)	$\sqrt{CF_{max}}$ (p=0.01)
	Biopsy position (p=0.37)	Side (p=0.12)	Side (p=0.03)	$\sqrt{CF_{max}}$ (p=0.05)
	Side (p=0.04)	Side (p=0.03)	Side (p=0.03)	Joint biopsied (p=0.78)
	$\sqrt{CF_{max}}$ (p=0.02)	$\sqrt{CF_{max}}$ (p=0.01)	$\sqrt{CF_{max}}$ (p=0.01)	Biopsy position (p=0.24)
Log 10(Ki67)* (Spearman, $r=0.53$)	Joint biopsied (p=0.97)	Biopsy position (p=0.24)	Side (p=0.03)	$\sqrt{CF_{max}}$ (p=0.05)
	Biopsy position (p=0.24)	Side (p=0.03)	Side (p=0.03)	Joint biopsied (p=0.77)
	Side (p=0.04)	Side (p=0.03)	Side (p=0.03)	Biopsy position (p=0.78)
	$\sqrt{CF_{max}}$ (p=0.05)	$\sqrt{CF_{max}}$ (p=0.03)	$\sqrt{CF_{max}}$ (p=0.05)	Side (p=0.40)
Log 10(CD3)* (Spearman, approximated $r=0.57$)	Joint biopsied (p=0.78)	Joint biopsied (p=0.77)	Side (p=0.36)	$\sqrt{CF_{max}}$ (p=0.001)
	Biopsy position (p=0.78)	Side (p=0.40)	Side (p=0.40)	$\sqrt{CF_{max}}$ (p=0.0003)
	Side (p=0.40)	Side (p=0.40)	Side (p=0.40)	
	$\sqrt{CF_{max}}$ (p=0.03)	$\sqrt{CF_{max}}$ (p=0.03)	$\sqrt{CF_{max}}$ (p=0.001)	

El modelo reducido se refiere al orden en el que incluyen las co-variables al modelo

Transformaciones: \log_{10} , $\sqrt{}$.

Para cumplir con los supuestos?

Supuestos

1. Y (outcome) distribución Normal.
- Normalidad de los errores (cada valor de X).
2. Linealidad la media de la variable Y varía linealmente con X.
3. Homocedasticidad La variabilidad de Y debe ser la misma para cada X.
4. Independencia de los errores

	Enfermo	Sano	
Test positivo	70 (a)	40 (b)	110
Test negativo	30 (c)	460 (d)	490
	100	500	600

Sensibilidad: $a/(a+c)$: $70/(70+30)$: 0.7.

Probabilidad que el test sea + estando enfermo

Especificidad. $D/(d+b)$: $460/500$: 0.92.

Probabilidad de que el test sea – cuando esta sano

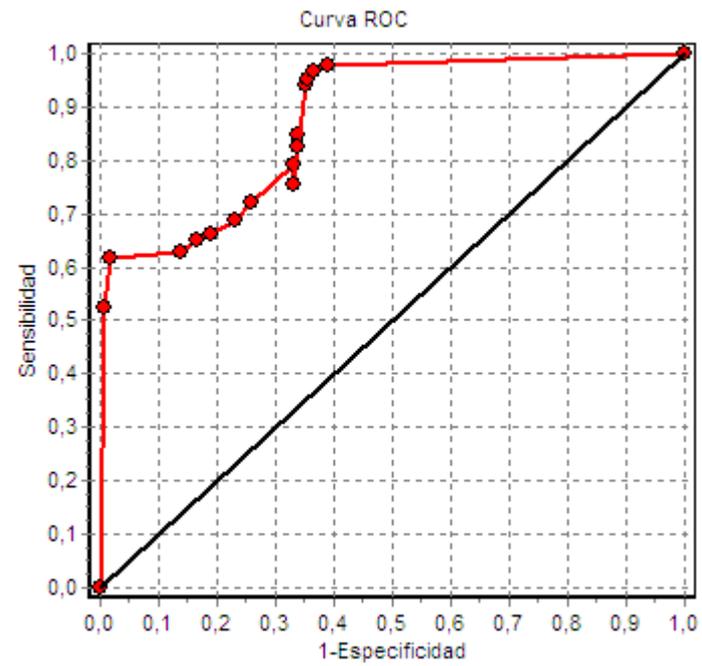
VPN: $d/d+c$.

Es la probabilidad de estar sano cuando el test es negativo

VPP: $a/a+b$.

Es la probabilidad de estar enfermo cuando el test es positivo

Curva ROC



Conclusiones

- Buena idea
- Acorde a los resultados? En general si, ver los falsos negativos
- Bibliografía previa. Consistencia con algunos trabajos (causalidad)
- Debilidades y fortalezas? No menciona, sin embargo para los estudios diagnósticos es útil tener información acerca de la sensibilidad, especificidad, VPP y VPN. No se puede sacar conclusiones.
- Excelentes datos preliminares.
- Validez externa limitada.
- Diseño y análisis poco aprovechados para el desarrollo del estudio