

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

# Análisis estadístico en pruebas diagnósticas

Cristián Garrido Inostroza




1

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Introducción

- Cruce entre el verdadero estado de la naturaleza y la percepción que tenemos de esta
  - Toma de decisiones
    - Correctas
    - Incorrectas
- En medicina
  - Falta de claridad para discernir
    - Características propias de un examen diagnóstico
    - Características propias de la población sobre la que se aplica







2

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Conceptos

- **Validez (Eficacia)**
  - Grado con que los resultados de la medición de un test de diagnóstico corresponden al verdadero estado del fenómeno que ha sido medido.
- **Niveles de medición:**
  - Niveles plasmáticos: variable continua
  - Riesgo de enfermedad: Variable ordinal
  - Presencia/Ausencia de enfermedad: Variable categórica

*Sensibilidad  
y  
Especificidad*



3

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Conceptos

- **Reproducibilidad (Precisión)**
  - Capacidad del test para ofrecer los mismos resultados cuando se repite su aplicación en circunstancias similares.
- **Depende de:**
  - Variabilidad biológica del hecho observado
  - Introducida por el propio observador
  - Introducida inter-observador
  - Derivada del test

*Consenso en  
ausencia de  
referente*



4

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Buena prueba diagnóstica

- Poder discriminatorio
  - Variabilidad de la prueba
  - Reproducibilidad de los hallazgos
  - Variabilidad de la población sana
    - Determinación de los rangos normales de la prueba en esta población
  - Marcador de la proporción de enfermos y no enfermos correctamente clasificados
- Comparación con un **estándar de referencia.**

Sensibilidad  
Especificidad  
Valores predictivos






5

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Estándar de referencia diagnóstico

- **Patrón de oro o Gold Standard**
  - Revela el verdadero estado de la naturaleza
  - Mayor veracidad disponible en un momento determinado
  - Mejor opción para identificar individuos con o sin el evento de interés
  - Utilizada para comparar con PD en estudio
    - Determinación de los rangos normales de la prueba en esta población
  - Debe ser definida
    - Hallazgos quirúrgicos
    - Hallazgos histológicos

6

## Evaluación de una prueba diagnóstica

### • Pasos:

- Definir el estándar de referencia
- Escoger dos grupos de sujetos a estudiar
  - Uno que presente la condición de interés y otro que no la presente
    - Idealmente en distintos estados evolutivos de la condición
    - Definido por el estándar de referencia
    - Aplicar prueba diagnóstica en evaluación
- Categorizar a los individuos
  - Positivos y negativos para el evento de interés
    - Según la prueba diagnóstica
- Construcción de tabla de contingencia (2x2)
  - Cálculo de S, E y VPs



7

	Patrón de oro	
	Presencia de enfermedad	Ausencia de enfermedad
Examen positivo (+)	Decisión Correcta	Decisión Incorrecta
Examen negativo (-)	Decisión Incorrecta	Decisión Correcta

	Enfermo (E)		Sano (E')		Total de sujetos con examen (+)
	a	b	c	d	
Examen (+)	a	b	c	d	a+b
Examen (-)	c	d	a+c	b+d	c+d
					$n = a+b+c+d$
		Total de enfermos	Total de sanos	Total de sujetos con examen (-)	



8

	Estado respecto a la enfermedad según el estándar de referencia	
Resultado de la prueba en estudio	Presente	Ausente
Positivo	<b>a</b> (enfermos con prueba +)	<b>b</b> (no enfermos con prueba +)
Negativo	<b>c</b> (enfermos con prueba -)	<b>d</b> (no enfermos con prueba -)

	Estado respecto a la enfermedad según el estándar de referencia	
Resultado de la prueba en estudio	Enfermo	Sano
Positivo	Verdadero positivo (VP)	Falso positivo (FP)
Negativo	Falso negativo (FN)	Verdadero negativo (VN)

9

**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Sensibilidad de la prueba

- **Fracción de verdaderos positivos (FVP)**
  - Responde a la pregunta
    - ¿Si el paciente realmente tiene la enfermedad, cuál es la probabilidad de que la prueba empleada sea positiva?
  - Capacidad del examen de detectar un enfermo

$$P(+|E) = \frac{a}{a+c} = \frac{VP}{VP+FN}$$

- Útil en el contexto de que no diagnosticar genera más problema que el exceso de diagnóstico
  - Screening: Bajo costo, fácil ejecución, mínima incomodidad

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

10

## Especificidad de la prueba

- **Fracción de verdaderos negativos (FVN)**

- Responde a la pregunta

¿Si el paciente no tiene la enfermedad, cuál es la probabilidad de que la prueba empleada sea negativa?

- Capacidad del examen de detectar un sano

$$P(-|E') = \frac{d}{b+d} = \frac{VN}{FP+VN}$$

- Útil para confirmar o **descartar** una enfermedad

- Prueba de confirmación: Más sofisticadas

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

11

- **SnNout** es el hallazgo de un signo, test o síntoma que posee una alta **Sensibilidad**

- Resultado **N**egativo (rules **out**) descarta el diagnóstico.

- **SpPin** es el hallazgo de un signo, test o síntoma que posee una alta **e**specificidad

- Resultado **P**ositivo (rules **in**) confirma el diagnóstico.

- Valores de Sensibilidad o Especificidad sobre 95%



12

## Eficiencia de la prueba

- Es la probabilidad de ser VP o VN
  - Capacidad del examen para clasificar correctamente a un paciente

$$P((E \cap +) \cup (E' \cap -)) = \frac{a + d}{a + b + c + d} = \frac{VP + VN}{n}$$

Por lo tanto, indica la proporción de resultados válidos entre el total de resultados

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

13

- Sensibilidad, especificidad y eficiencia son **características propias del examen y no dependen** de la población donde se apliquen, ni de la prevalencia de la enfermedad
- Estrecha relación entre S y E
  - Generalmente de tipo inversa
    - Cuando resultado de una PD usa escala continua
    - Sensibilidad se incrementa sólo a expensas de la Especificidad
    - Curvas ROC



14

**RED**  
HOSPITAL CLINICO UNIVERSIDAD DE CHILE

## Ejemplo numérico

	Enfermo (E)	Sano (E')	
Examen (+)	85	100	185
Examen (-)	15	800	815
	100	900	1.000

$$\text{Sensibilidad} = P(+|E) = \frac{a}{a+c} = \frac{VP}{VP+FN} = \frac{85}{100} = 0,85 = 85\%$$

$$\text{Especificidad} = P(-|E') = \frac{d}{b+d} = \frac{VN}{FP+VN} = \frac{800}{900} = 0,88 = 88\%$$

$$\text{Eficiencia} = P((E \cap +) \cup (E' \cap -)) = \frac{a+d}{a+b+c+d} = \frac{VP+VN}{n} = \frac{85+800}{1000} = 0,885 = 88,5\%$$

Por lo tanto, este examen tiene más capacidad para descartar la enfermedad

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

15

**RED**  
HOSPITAL CLINICO UNIVERSIDAD DE CHILE

## Sensibilidad y Especificidad

- Test de diagnósticos son imperfectos
- Individuos sanos se pueden clasificar como enfermos (Falsos positivos)
- Otros enfermos pueden no ser detectados (falsos negativos)

Se llaman errores de mala-clasificación

16



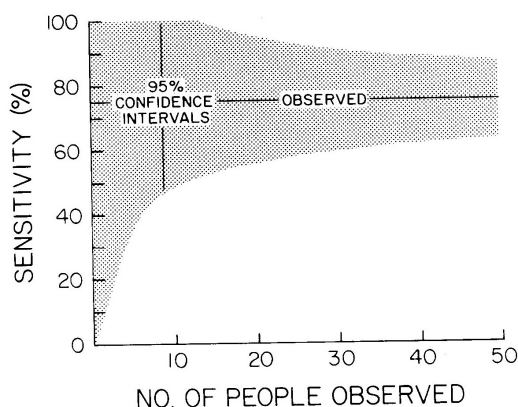
## Debilidades de la Sensibilidad y Especificidad

- Hablan de la evaluación del test con respecto a una sola enfermedad o exposición
- Positivo/negativo y no sirven a la hora de evaluar el grado de estenosis de un vaso o coeficiente de atenuación de una lesión focal hepática.
- No sirven a la hora de tomar decisiones clínicas.



17

## Precisión de la Sensibilidad de acuerdo al tamaño de la muestra



**Figure 3.7.** The precision of an estimate of sensitivity. The 95% confidence interval for an observed sensitivity of 75%, according to the number of people observed.



18

## S y E como herramientas de marketing

- Comparación artificial de distintos test que no están en competición.
- Destacan solo una alta Sensibilidad o Especificidad y no si son “modestas”
- Ambos “con intervalos de confianza”
- Cifras derivadas de población no representativa de la población en que el test finalmente se usara.



19

## ¿Debe preferirse alta Sensibilidad o alta Especificidad?

- Si se estudia enfermedad grave o hay mayor probabilidad de enfermedad
  - Prefiera test con alta especificidad si desea baja tasa de falsos positivos
- Si esta monitoreando para prevención de enfermedad
  - Prefiera test con alta sensibilidad si desea baja tasa de falsos negativos



20

## La “cuestión clínica”

- Sensibilidad, especificidad y eficiencia carecen de utilidad en la práctica clínica
  - Clínico no tiene información “a priori” sobre el verdadero diagnóstico del paciente.
- Surge la pregunta
  - Ante un resultado (+) o (-) en la prueba ¿Cuál es la probabilidad de que el paciente esté realmente enfermo o “no enfermo”?

Valores predictivos



21

## Valor predictivo positivo (VPP)

- Responde a la pregunta
  - ¿Si el resultado de la prueba es positivo, cuál es la probabilidad de que el paciente tenga la enfermedad en estudio?
- Proporción de VP entre los que tienen resultados (+)

$$P(E|+) = \frac{a}{a+b} = \frac{VP}{VP+FP}$$

- También se denomina probabilidad a posteriori o probabilidad “post-test”

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

22

## Valor predictivo negativo (VPN)

- Responde a la pregunta  
¿Si el resultado de la prueba es negativo, cuál es la probabilidad de que el paciente no tenga la enfermedad en estudio?
- Proporción de VN entre los que tienen resultados (-)

$$P(E'|-) = \frac{d}{c+d} = \frac{VN}{VN+FN}$$

- VPP y VPN dependen de la prevalencia de la enfermedad en la población en estudio

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

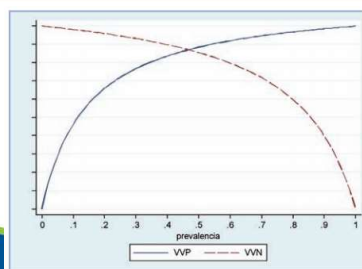
23

## Prevalencia

- Proporción de individuos de una población que presentan una determinada condición en un momento dado

$$Prevalencia = P = \frac{a+c}{a+b+c+d} = \frac{VP+FN}{n}$$

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d





24

**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Estimación de la prevalencia

- Revisión literatura médica (ACL)
- Acceso a bases de datos on-line
- Consenso de expertos
- Juicio Clínico

25

**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Ejemplo numérico

	Enfermo (E)	Sano (E')	
Examen (+)	85	100	185
Examen (-)	15	800	815
	100	900	1.000



$$VPP = P(E|+) = \frac{a}{a+b} = \frac{VP}{VP+FP} = \frac{85}{185} = 0,459 = 45,9\%$$

$$VPN = P(E'|-) = \frac{d}{c+d} = \frac{VN}{VN+FN} = \frac{800}{815} = 0,982 = 98,2\%$$

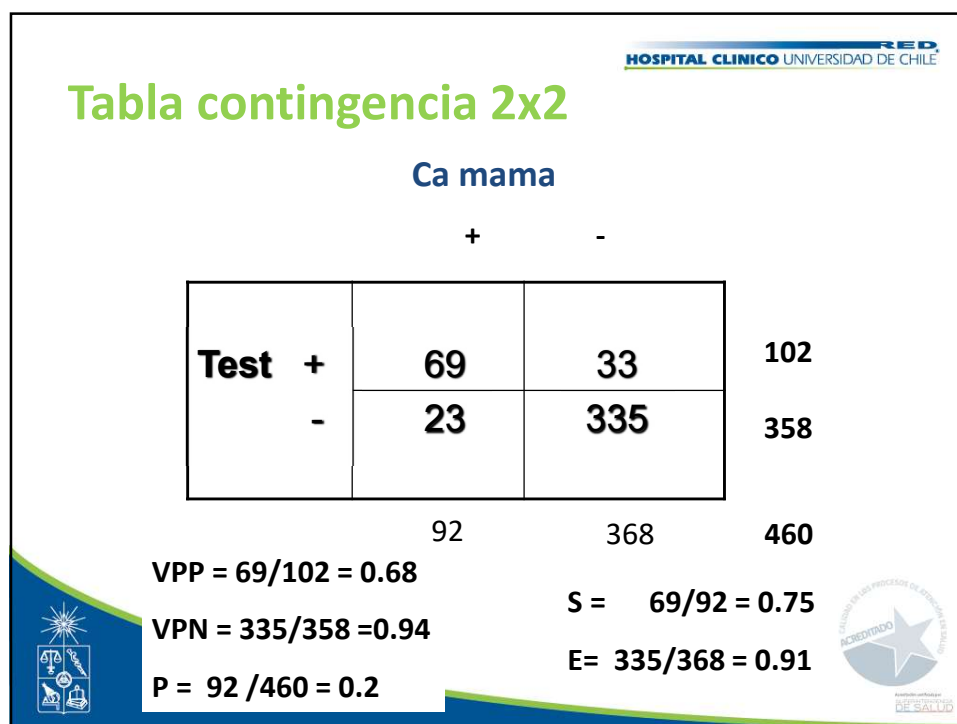
$$Prevalencia = P = \frac{a+c}{a+b+c+d} = \frac{VP+FN}{n} = \frac{100}{1000} = 0,1 = 10\%$$

Por lo tanto, el valor del pronóstico de “no enfermedad” es muy superior en el caso de obtener un examen con resultado negativo. Si la prevalencia es baja, un resultado negativo permite descartar la enfermedad con mayor seguridad, por lo que el VPN será mayor. Al contrario, un resultado positivo no permitirá confirmar el diagnóstico, por lo que el VPP será bajo

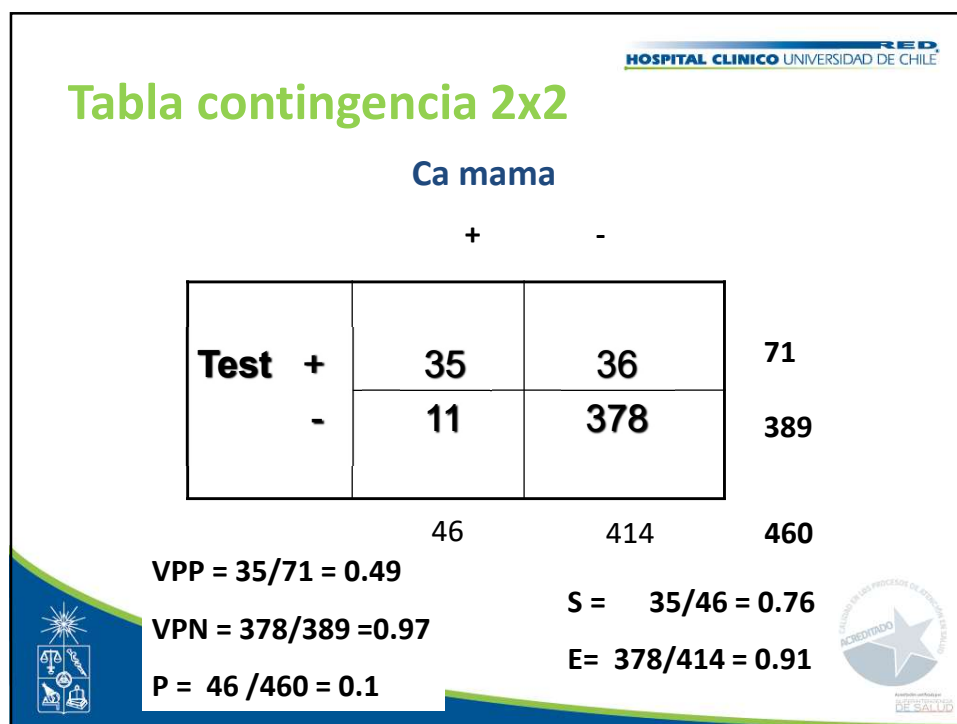
	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

26



27



28

### Otra “cuestión clínica”

- Prevalencia influye en los valores predictivos
  - VPs no permiten comparar dos métodos diagnósticos
  - VPs no permiten extrapolar resultados de otros estudios en datos propios o entre centros de salud
- Surge la pregunta
  - ¿Cuánto es más probable es un resultado positivo o negativo, según la presencia o ausencia de la condición?

*Razones de probabilidad*



29

### Prevalencia: Probabilidad Pre-Test

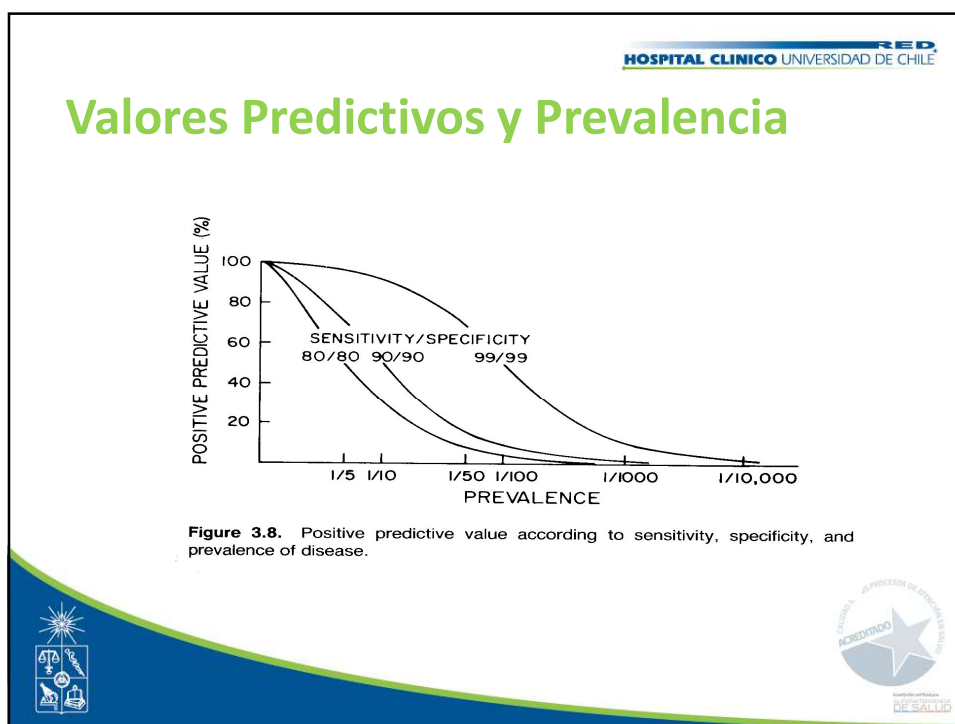
- Prob. Pre-test %    Chance de enfermedad %

	Test +	Test –
1	7	1
50	88	15
99	99	95

*Razones de probabilidad*



30



31

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Razones de probabilidad (RP)

- **Combinación de Sensibilidad y Especificidad**
  - Resultados más confiables sobre la validez
  - Más precisión en la información de una prueba diagnóstica
  - Puede ser positiva (RPP) y negativa (RPN)
- **Conocidas como Likelihood ratios (LR+ y LR-)**
- **Indica cuanto más probable es un resultado determinado en una PD en un paciente con una condición dada, comparado con uno sin la condición de interés**

32



## Razón de probabilidad positiva (RPP)

- Probabilidad de tener la enfermedad en oposición a no tenerla, teniendo un resultado (+) en la PD
- Cuociente entre:
  - $P(E|+) = \text{sensibilidad}$  y  $P(+|E') = 1 - \text{especificidad}$

$$LR+ = \frac{\frac{a}{a+c}}{\frac{b}{b+d}} = \frac{\text{Sensibilidad}}{1 - \text{Especificidad}}$$

- Cuociente entre la fracción de VP y la de FP
- Recorrido  $[1, \infty[$

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

33

## Razón de probabilidad negativa (RPN)

- Probabilidad de no tener la enfermedad en oposición a tenerla, teniendo un resultado (-) en la PD
- Cuociente entre:
  - $P(E|-) = 1 - \text{sensibilidad}$  y  $P(-|E') = \text{especificidad}$

$$LR- = \frac{\frac{c}{a+c}}{\frac{d}{b+d}} = \frac{1 - \text{Sensibilidad}}{\text{Especificidad}}$$

- Cuociente entre la fracción de FN y la de VN
- Recorrido  $[0, 1]$

	Enfermo (E)	Sano (E')	
Examen (+)	a <b>VP</b>	b <b>FP</b>	a+b
Examen (-)	c <b>FN</b>	d <b>VN</b>	c+d
	a+c	b+d	n = a+b+c+d

34

## Ventajas de las Razones de probabilidad

- Relaciona S y E en un solo índice
- Nueva medida es información no dicotómica (N, NoN, +, -)
  - Más precisión en la información de una prueba diagnóstica
  - Puede ser positiva (RPP) y negativa (RPN)
- No varía con la prevalencia
  - Índices de comparación entre diferentes PD para el mismo Dg
  - Son constantes aunque varíe la prevalencia en los sujetos a los que se aplica la PD
- Clínicamente más útil  $\Rightarrow$  Mejor entendimiento de resultados
  - Con que fuerza un resultado (+) indica presencia real de la enfermedad
  - Con que fuerza un resultado (-) descarta la presencia de enfermedad

RPs Indican como el resultado de una PD cambia la probabilidad pre-test a la probabilidad post-test de la enfermedad

35

Tabla II. Ejemplo 1, con datos obtenidos de un estudio del año 2001 <sup>(13)</sup> , sobre la exactitud diagnóstica de la tomografía computada para predicción de daño cerebral irreversible luego de accidente cerebrovascular isquémico.			
TC en infarto agudo	Gold standard		TOTAL
	Positivo	Negativo	
Test positivo	433	16	449
Test negativo	246	91	337
<b>Totales marginales</b>	<b>679</b>	<b>107</b>	<b>786</b>
IC 95%			
Sensibilidad	63,8%	60,2 - 67,4	
Especificidad	85,0%	78,3 - 91,8	
VPP	96,4%	94,7 - 98,2	
VPN	27,0%	22,3 - 31,7	
LR+	4,3	2,7 - 6,7	
LR-	0,4	0,37 - 0,48	
<b>Prevalencia</b>	<b>86,4%</b>		

Tabla III. Ejemplo 2 con datos modificados, con una prevalencia menor de la enfermedad y un tamaño muestral ligeramente menor.			
TC en infarto agudo -2	Gold standard		TOTAL
	Positivo	Negativo	
Test positivo	289	48	337
Test negativo	164	273	437
<b>Totales marginales</b>	<b>453</b>	<b>321</b>	<b>774</b>
IC 95%			
Sensibilidad	63,8%	59,4 - 68,2	
Especificidad	85,0%	81,1 - 88,9	
VPP	85,8%	82,0 - 89,4	
VPN	62,5%	57,9 - 67,0	
LR+	4,3	3,3 - 5,6	
LR-	0,4	0,37 - 0,49	
<b>Prevalencia</b>	<b>58,5%</b>		

36

## Relación de probabilidades condicionales

- Uso de la probabilidad pre-test para estimar la probabilidad post-test utilizando Odds ratios

$$\text{Odds} = \frac{\text{Probabilidad}}{1 - \text{Probabilidad}} = \frac{P(x < X)}{1 - P(x < X)}$$

- Probabilidad pre-test: Prevalencia informada

$$\text{Probabilidad pre - test} = \text{Prevalencia}$$

$$\text{Odds pre - test} = \frac{\text{Prevalencia}}{1 - \text{Prevalencia}}$$

$$\text{Odds post - test} = \text{Odds pre - test} \times \text{LR -}$$

$$\text{Probabilidad post - test} = \frac{\text{Odds post - test}}{1 + \text{Odds post - test}}$$



37

## Relación de probabilidades condicionales

TC en infarto agudo	Gold standard		TOTAL
	Positivo	Negativo	
Test positivo	433	16	449
Test negativo	246	91	337
Totales marginales	679	107	786
IC 95%			
Sensibilidad	63,8%	60,2 - 67,4	
Especificidad	85,0%	78,3 - 91,8	
VPP	96,4%	94,7 - 98,2	
VPN	27,0%	22,3 - 31,7	
LR+	4,3	2,7 - 6,7	
LR-	0,4	0,37 - 0,48	
Prevalencia	86,4%		

Prevalencia informada: 70%

$$\text{Probabilidad pre - test} = 0,7$$

$$\text{Odds pre - test} = \frac{\text{Prevalencia}}{1 - \text{Prevalencia}} = \frac{0,7}{0,3} = 2,33$$

$$\text{Odds post - test} = \text{Odds pre - test} \times \text{LR}(-) = 2,33 \times 0,4 = 0,932$$

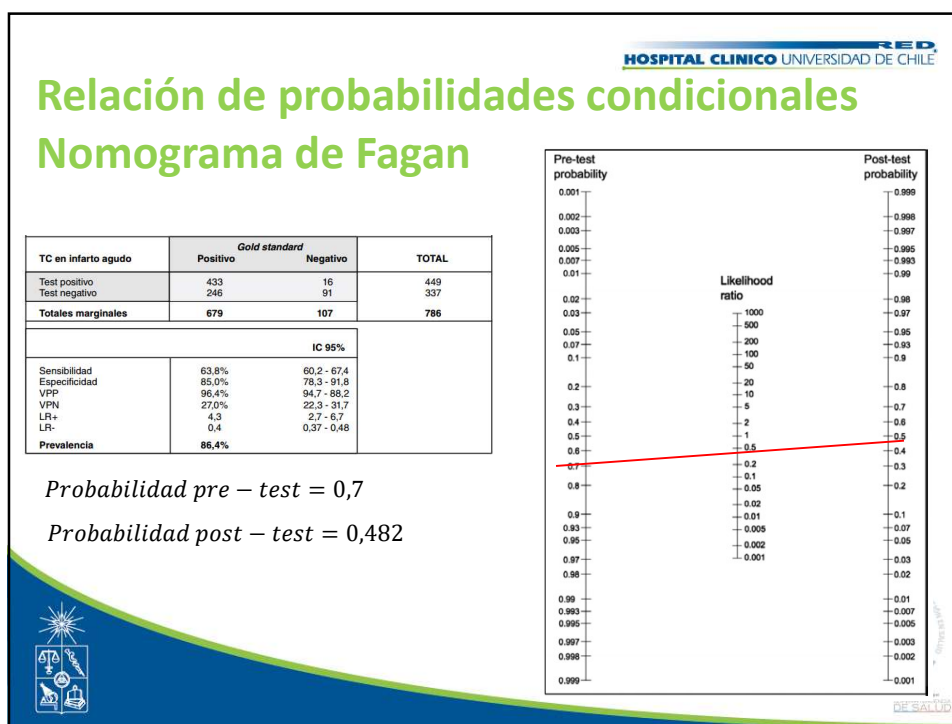
$$\text{Probabilidad post - test} = \frac{\text{Odds post - test}}{1 + \text{Odds post - test}} = \frac{0,932}{1,932} = 0,482 = 48,2\%$$



Como P(post-test) sigue siendo alta, una TC normal no descarta ACV



38



39

**HOSPITAL CLINICO UNIVERSIDAD DE CHILE**

## Resumen

Parámetro	Fórmula	Definición
Sensibilidad	$a/(a+c)$	Proporción de pacientes con la enfermedad que tendrán test positivo
Especificidad	$d/(b+d)$	Proporción de pacientes sin la enfermedad que tendrán test negativo
Valor predictivo positivo	$a/(a+b)$	Probabilidad de que el paciente tenga la enfermedad dado que el test es positivo
Valor predictivo negativo	$d/(c+d)$	Probabilidad de que el paciente no tenga la enfermedad dado que el test es negativo
Likelihood ratio (+)	sensibilidad/(1-especificidad)	Describe cuántas veces es más probable que reciba un resultado determinado una persona con la enfermedad que una persona sin la enfermedad
Likelihood ratio (-)	(1-sensibilidad)/especificidad	
Exactitud	$(a+d) / (a+b+c+d)$	La probabilidad de que el resultado del test prediga correctamente la presencia o ausencia de la enfermedad

40

**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Curvas ROC (Receiver Operating Characteristics)

- Enfoque expuesto hasta ahora es dicotómico
- Mayoría de test en escala continua (o discreta ordinal)
  - Evaluación de función renal: Creatinina sérica
  - Dg de DM con glicemia (mg/dl)
- Elección de un punto de corte para dicotomizar
  - Resultado positivo para hacer el diagnóstico
  - Determinará la Sensibilidad y Especificidad
  - Ejemplo: Glicemia con punto de corte 70 mg/dl y 140 mg/dl
- Gráfico de todos los pares sensibilidad/especificidad con distintos niveles de decisión

**Curvas ROC proporcionan una representación global de la exactitud diagnóstica**

41

**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Curvas ROC (Receiver Operating Characteristics)

- Concepto de Área bajo la Curva (AUC)
- Ideal: Sensibilidad y Especificidad de 100%  $\Rightarrow$  Mayor exactitud global
- Sin discriminación: Igual distribución de resultados en los subgrupos
  - Discriminación por azar:  $S=1-E$



Área bajo la curva ROC	Discriminación
0,5	por azar
0,7 a 0,8	aceptable
0,8 a 0,9	muy buena
0,9 a 1	excelente

42

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Ventajas de las Curvas ROC

- Fácil comprensión de la capacidad de discriminación del test en todo el rango de los puntos de corte
- Simples, gráficas y fácil de interpretar visualmente
- No requieren de nivel de decisión particular
  - Includo todo el espectro de los puntos de corte
- Independientes de la prevalencia
  - No es necesario obtener prevalencias “representativas”
- Comparación visual directa entre pruebas con escala común
- Sensibilidad y Especificidad son accesibles en el mismo gráfico






43

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Conceptos sobre pruebas diagnósticas

- **Reproducibilidad**
  - Intraobservador: Mismo observador, varias mediciones con el mismo instrumento
  - Interobservador: Varios observadores, misma muestra, personas diferentes
- **Consistencia**
  - Mismo ranking de las observaciones al medirlas varias veces
- **Acuerdo o concordancia**
  - Valores en las distintas mediciones son idénticos
  - Siempre que hay acuerdo hay consistencia

44



**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Concordancia interobservador

- **Se evalúa a través de:**
  - Kappa de Cohen (cualitativas)
    - Dos observadores: Kappa de Cohen
    - Tres o más observadores: Kappa de Fleiss
  - CCI (cuantitativas)
- **Acuerdo simple**
  - Incluye el azar

Acuerdo simple:  $(35+55)/107 = 0.8411$  (**84.1%**)

		Entrevista 1: ¿Usted consume suplementos vitamínicos?		
		NO	SI	Total
Entrevista 2: Responda si consume vitaminas sin contar sus aportes alimentarios	NO	<b>35</b>	12	47
	SI	5	<b>55</b>	60
		40	67	107

45



**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

## Concordancia interobservador

- **Acuerdo + Azar = Concordancia simple u observada**
- **Concordancia esperada**
  - % de acuerdo exclusivamente por azar

Concordancia esperada por azar:  $(17.57+37.37)/107 = 0.5153$  (**51.5%**)

		Entrevista 1: ¿Usted consume suplementos vitamínicos?		
		NO	SI	Total
Entrevista 2: Responda si consume vitaminas sin contar sus aportes alimentarios	NO	$(47 \times 40)/107$ 17.57	$(47 \times 67)/107$ 29.43	47
	SI	$(60 \times 40)/107$ 22.43	$(60 \times 67)/107$ 37.57	60
		40	67	107

46



**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Concordancia interobservador

Acuerdo simple o Concordancia observada:  $(35+55)/107 = 0.8411$  (**84.1%**)

Concordancia esperada por azar:  $(17.57+37.37)/107 = 0.5153$  (**51.5%**)

$$Kappa = \frac{\text{Concordancia observada} - \text{concordancia esperada}}{1 - \text{concordancia esperada}}$$

$$Kappa = \frac{0.8411 - 0.5153}{1 - 0.5153} = 0.672$$



47


**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Concordancia interobservador

- **Kappa varía entre -1 y 1:**
  - -1 : Mayor grado de discordancia interobservador
  - 0: Concordancia exclusivamente por azar
  - +1: Mayor grado de concordancia interobservador
- **Problemas**
  - Paradoja del sesgo: Si un observador tiende a diagnosticar mas que el otro el kappa tenderá a ser más bajo
  - Paradoja de la prevalencia: A mayor prevalencia de la enfermedad el kappa tiende a ser más alto (especialistas)

**Tabla 3. Valoración del coeficiente kappa (Landis y Koch, 1977)<sup>a</sup>**

Coeficiente kappa	Fuerza de la concordancia
0,00	Pobre ( <i>Poor</i> )
0,01 - 0,20	Leve ( <i>Slight</i> )
0,21 - 0,40	Aceptable ( <i>Fair</i> )
0,41 - 0,60	Moderada ( <i>Moderate</i> )
0,61 - 0,80	Considerable ( <i>Substantial</i> )
0,81 - 1,00	Casi perfecta ( <i>Almost perfect</i> )



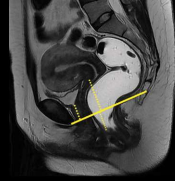
48



**HOSPITAL CLINICO UNIVERSIDAD DE CHILE**

### TABLAS DE MEDICION PROLAPSO

**PCL**




GRADO	DISTANCIA DESDE PCL
Leve	1-3 cm
Moderado	3-6 cm
Severo	> 6 cm

### MEDICION PROLAPSO

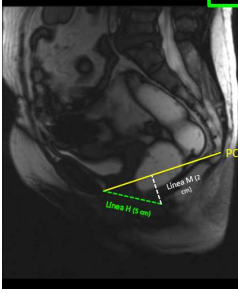
**MPL**



GRADO	DISTANCIA DESDE MPL
0	> 3 cm sobre
1	Entre 1 y 3 cm sobre
2	≤ 1 cm sobre o bajo
3	≥ 1 cm bajo
4	Completa eversion del organo

### MEDICION PROLAPSO

**LINEA H**



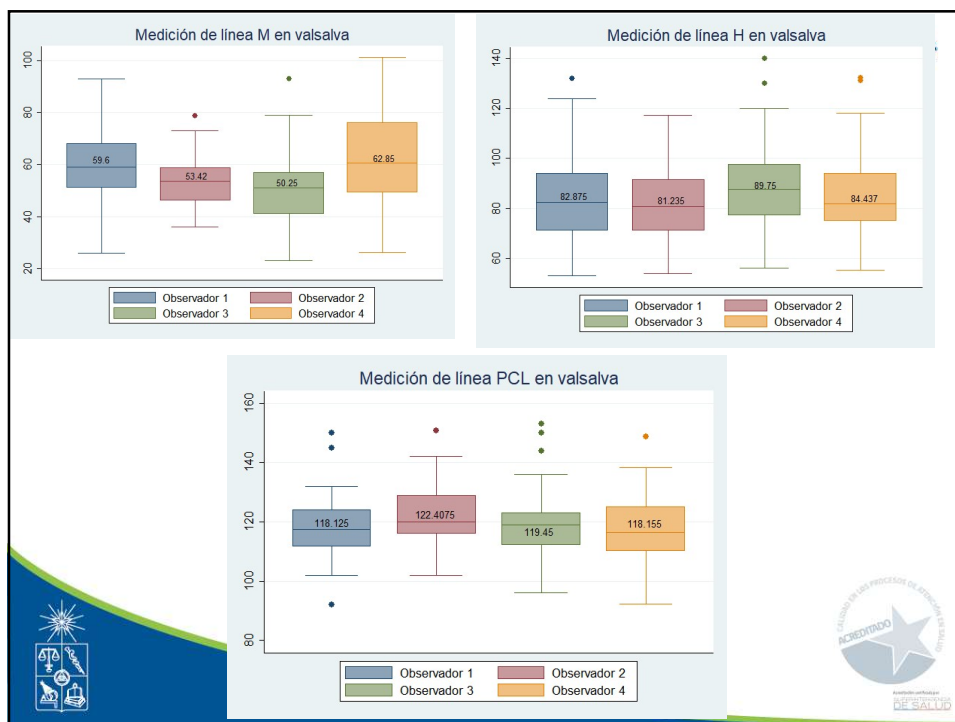
GRADO	ENSANCHAMIENTO HIATAL	DESCENSO DEL PISO PELVICICO
0 (no hay prolapso)	< 6 cm	0-2 cm
1 (leve)	6-8 cm	2-4 cm
2 (moderado)	8-10 cm	4-6 cm
3 (severo)	≥ 10 cm	≥ 6 cm

GRADO	DISTANCIA DESDE LINEA H
0 (no hay prolapso)	Por encima
1 (leve)	0-2 cm bajo
2 (moderado)	2-4 cm bajo
3 (severo)	≥ 4 cm bajo

**ACREDITADO**  
Asociación Chilena de Acreditación de Procesos de Atención en Salud

49



50

**Tabla A**  
Coeficiente de correlación de Pearson para la línea PCL entre observadores

Línea PCL	Observador	Observador 1	Observador 2	Observador 3	Observador 4
Observador 1	1				
Observador 2		0.8242			
Observador 3			0.8818		
Observador 4				0.8416	
					0.8104
					1

Significancia del coeficiente de correlación (Rho) de Pearson con p valor < 0.001

**Tabla B**  
Porcentaje de acuerdo entre observadores para las líneas H y M

Línea H	Observador	Observador 1	Observador 2	Observador 3	Observador 4
Observador 1	96.67%				
Observador 2		95.83%			
Observador 3			95.83%		
Observador 4				94.44%	
Promedio del % de acuerdo Línea H: 95.83% (IC 95%: 94.87% - 96.79%)					
Línea M	Observador	Observador 1	Observador 2	Observador 3	Observador 4
Observador 1	89.38%				
Observador 2		83.75%			
Observador 3			91.88%		
Observador 4				82.50%	
Promedio del % de acuerdo Línea M: 88.02% (IC 95%: 83.34% - 92.7%)					

**Tabla C**  
Nivel de concordancia (Índice kappa de Cohen con ponderación cuadrática) entre observadores para las líneas H y M

Línea H	Observador	Observador 1	Observador 2	Observador 3	Observador 4
Observador 1	1				
Observador 2		0.7744 (0.652 - 0.806)			
Observador 3			0.717 (0.575 - 0.838)		
Observador 4				0.783 (0.676 - 0.810)	
Promedio del Kappa ponderado Línea H: 0.714 (IC 95%: 0.645 - 0.783)					
Línea M	Observador	Observador 1	Observador 2	Observador 3	Observador 4
Observador 1	1				
Observador 2		0.485 (0.402 - 0.647)			
Observador 3			0.343 (0.217 - 0.486)		
Observador 4				0.527 (0.365 - 0.682)	
Promedio del Kappa ponderado Línea M: 0.451 (IC 95%: 0.292 - 0.610)					

51

## Concordancia interobservador

- Coeficiente de correlación intraclass (CCI)**
  - En variables cuantitativas → Evita la recodificación
  - Coeficiente de correlación de Pearson
    - Mide asociación entre variables
    - Solo valora consistencia

id	w1	w2
1	52	55
2	54	57
3	56	60
4	67	70
5	78	80
6	80	82
7	85	86
8	82	88
9	90	90
10	110	111

w1: Peso declarado  
w2: Peso medido en una pesa

id	rater	w
1	1	52
2	1	55
3	2	54
4	2	57
5	3	56
6	3	60
7	4	67
8	4	70
9	5	78
10	5	80
11	6	80
12	6	82
13	7	85
14	7	86
15	8	82
16	8	88
17	9	90
18	9	90
19	10	110
20	10	111

52

**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

- **ANOVA**
  - Descompone la variabilidad total de la muestra
    - Sujetos (id)
    - Valoraciones (rater)

```
. icc w id rater
```



Intraclass correlations  
Two-way random-effects model  
Absolute agreement

Random effects: **id**                      Number of targets =        10  
Random effects: **rater**                    Number of raters =        2

	w	ICC	[95% Conf. Interval]	
Individual		<b>.9863192</b>	.6240103	.9975845
Average		.9931125	.7684807	.9987908

F test that  
ICC=0.00: F(9.0, 9.0) = 436.83                      Prob > F = 0.000

Note: ICCs estimate correlations between individual measurements and between average measurements made on the same target.






53

**HOSPITAL CLINICO** UNIVERSIDAD DE CHILE

## Validez aceptable de una PD

- No existe un parámetro para todas las situaciones
  - Depende de la enfermedad y condiciones reales en el medio y población estudiada
- Si quiero detectar el mayor número posible de enfermos: ALTA S
  - Enfermedad es grave y no quiero que pase piola
  - La enfermedad es tratable
  - Los FP no traumen a los individuos examinados
- Si quiero “asegurar” el diagnóstico: Máxima E
  - La enfermedad es importante pero difícil de tratar o incurable
  - Los FP trauman a los individuos examinados
  - El tratamiento de los FP puede tener consecuencias graves

54


**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE

**A tener en cuenta**

- La validez de una prueba diagnóstica depende de su capacidad para detectar correctamente la presencia o ausencia de la enfermedad que se estudia.
- No existe un parámetro único para evaluar la validez aceptable de un método diagnóstico en todas las situaciones.
- Las curvas ROC son índices de la exactitud diagnóstica y proporcionan un criterio unificador en el proceso de evaluación de una prueba, debido a sus diversas aplicaciones.
- La sensibilidad nos indica la capacidad de una prueba diagnóstica para identificar una enfermedad.
- La especificidad valora la utilidad de una prueba para identificar a los no enfermos.
- El valor predictivo positivo valora la probabilidad de que un individuo con resultado positivo en la prueba tenga la enfermedad.
- El valor predictivo negativo valora la probabilidad de que un individuo con resultado negativo en la prueba no tenga la enfermedad.


**Errores más habituales**


- No considerar que los valores predictivos de un test son variables y dependen de la prevalencia de la enfermedad en la población.
- No considerar que la sensibilidad y la especificidad son características propias del test y no se modifican con cambios en la prevalencia.




55

**HOSPITAL CLINICO** **RED** UNIVERSIDAD DE CHILE







56