

**IX CONGRESO INTERNACIONAL
INVESTIGAR PARA INNOVAR EN SALUD:
EL FUTURO ESTÁ AQUÍ**



Como seleccionar una prueba estadística

CURSO DE INVESTIGACIÓN Y REDACCIÓN CIENTÍFICA BIOMÉDICA

Jorge L. Maguiña MSc

Universidad Peruana Cayetano Heredia
Facultad de Salud Pública y Administración
Escuela de Postgrado

Datos existen en todas partes

- **Prevalencia y causas de ceguera en el Perú - Rev Panam Salud Publica 36(5), 2014**

Resultados. *Se examinaron 4 849 personas. La prevalencia de ceguera fue 2,0% (intervalo de confianza de 95%: 1,5–2,5%). La catarata fue la causa principal de ceguera (58,0%), seguida por el glaucoma (13,7%) y la degeneración macular relacionada con la edad (11,5%).*

- **Prácticas de alimentación de lactantes en la Amazonia Peruana: implicaciones para los programas de mejora de la alimentación - Rev Panam Salud Publica. 2014, vol.36, n.3, pp. 150-157.**

Resultados. *La mediana de duración de la lactancia materna exclusiva fue 19 días. Sin embargo, en los 180 primeros días de vida los niños habían sido amamantados exclusivamente 46,1% de los días.*

Datos proveen información

- Los buenos datos pueden ser analizados y resumidos para proveer información útil
- Los malos datos pueden ser analizados y resumidos para proveer información incorrecta, engañosa, confusa y no informativa

Pasos en los proyectos de investigación

- Diseño y planteamiento del diseño
- Colección de datos
- Análisis de datos
- Presentación de resultados
- Interpretación

**El bioestadístico juega un rol importante en todos los pasos. Sin embargo, sólo lo llaman par analizar datos.*

**Empecemos con un entrenamiento en bioestadística*

Necesitamos saber bioestadística?

- Planteamiento y diseño de estudios:
 - Pregunta primaria de interés:
 - Cuantificando la información relacionada a un grupo
 - Comparando múltiples grupos
 - Tamaño de la muestra:
 - Cuantos sujetos necesitan en el estudio?
 - Cuantos sujetos son necesarios para comparar grupos
 - Seleccionando participantes
 - Muestreo de sujetos
 - Seleccionado de un pool de personas
 - Como asignar a los grupos de interés

Necesitamos saber bioestadística?

- Colecta de datos
- Análisis de datos
 - Cual es el método estadístico apropiado para los datos colectados
 - Lidar con la variabilidad de los datos
 - Modelos importantes en los datos están oscurecidos por la variabilidad
 - Distinguir los modelos reales de la variación aleatoria
 - Inferencia

Necesitamos saber bioestadística?

- Presentación
 - Que medidas de resumen son las más convenientes “mensajes principales”
 - Dar estimados adecuados y estos que implicancia tienen
- Interpretación:
 - Implicancias de los resultados en salud pública y epidemiología

Ensayo clínico

MAJOR ARTICLE

Comparison of the Immunogenicity and Reactogenicity of Cervarix and Gardasil Human Papillomavirus Vaccines in HIV-Infected Adults: A Randomized, Double-Blind Clinical Trial

Lars Toft,¹ Merete Storgaard,¹ Martin Müller,² Peter Sehr,³ Jesper Bonde,^{4,5} Martin Tolstrup,¹ Lars Østergaard,¹ and Ole S. Søgaard¹

¹Department of Infectious Diseases, Aarhus University Hospital, Aarhus, Denmark; ²Research Program Infection and Cancer, German Cancer Research Center; ³Chemical Biology Core Facility, European Molecular Biology Laboratory, Heidelberg, Germany; and ⁴Department of Pathology and ⁵Clinical Research Centre, Copenhagen University Hospital Hvidovre, Copenhagen, Denmark



Ensayo clínico: Presentación de datos

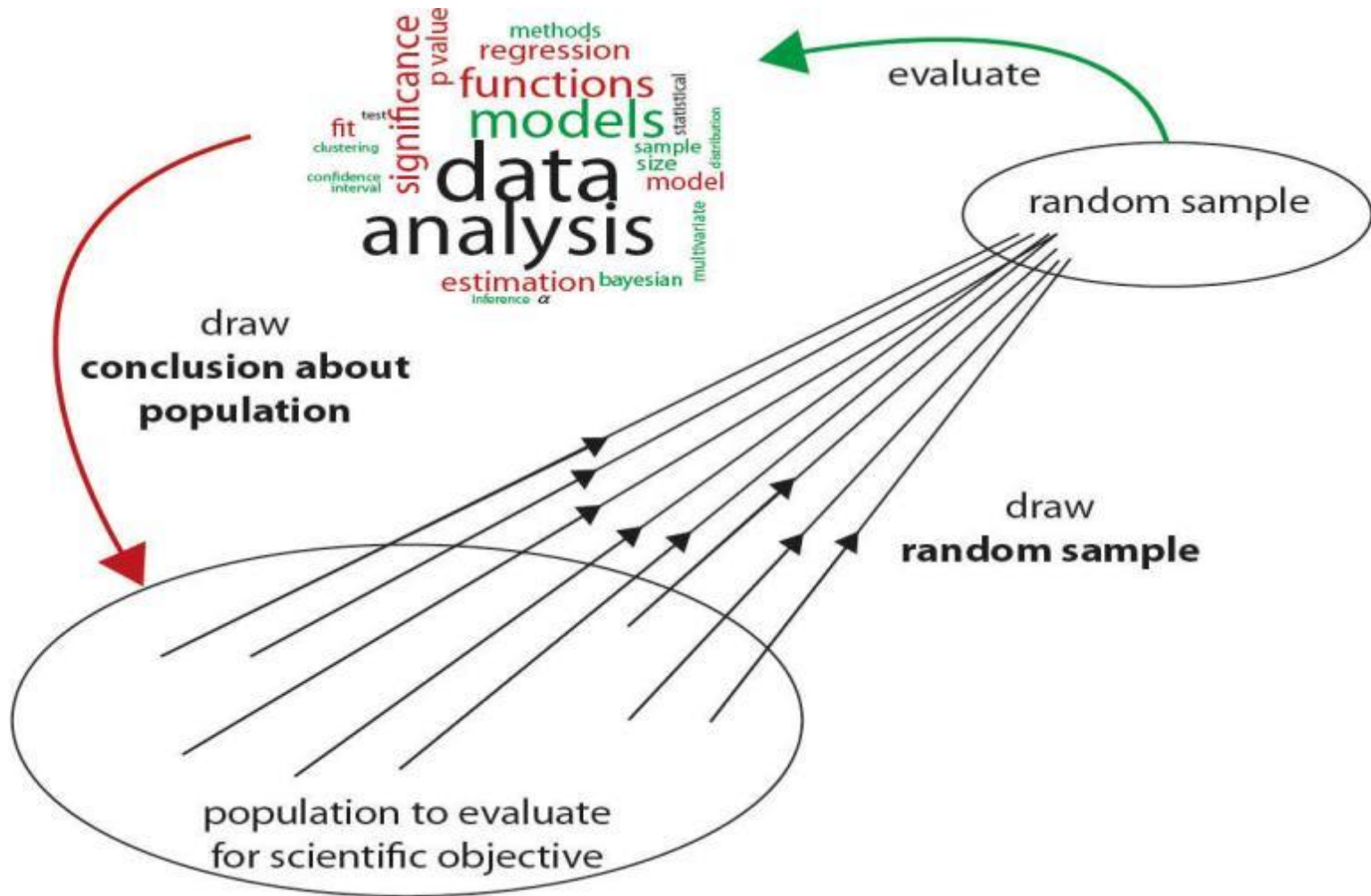
Table 1. Baseline Characteristics of the Study Population

Characteristic	Cervarix Group (n = 45)	Gardasil Group (n = 46)
Sex		
Male	30 (66.7)	31 (67.4)
Female	15 (33.3)	15 (32.6)
Race		
White	38 (84.4)	37 (80.4)
Other	7 (15.6)	9 (19.6)
Age, y, median (IQR)	47.0 (38.6–54.2)	44.5 (38.2–51.9)
Body mass index, kg/m ² , median (IQR)	23.8 (21.8–26.7)	24.1 (21.0–26.4)
Current smoker	6 (13.3)	16 (34.8)
CD4 ⁺ cell count, cells/ μ L, median (IQR)	600 (470–750)	585 (440–760)
Undergoing HAART		
Yes	40 (88.9)	40 (87.0)
No	5 (11.1)	6 (13.0)
HIV RNA level, log ₁₀ copies/mL, median (IQR)		
In patients receiving HAART	1.28	1.28
In patients not receiving HAART	4.97 (4.81–5.32)	4.58 (3.85–5.22)

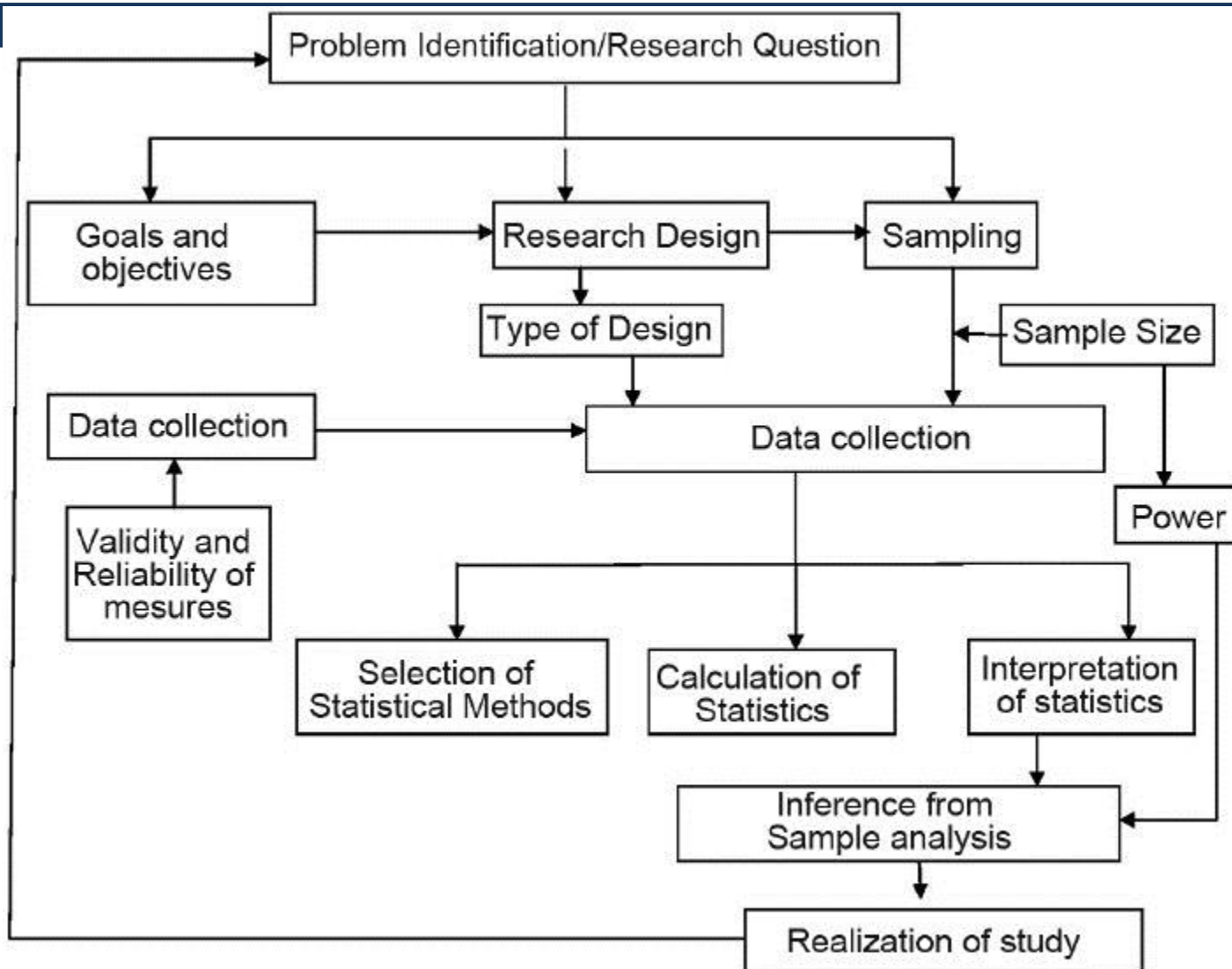
Como seleccionar una prueba estadística?

Necesito llamar a un bioestadístico?

Diseño muestral



Marco de investigación



Datos binarios:

- Datos binarios (Dicotómicos):
 - Si/No
 - Enfermedad: (Si/No)
 - Cura: (Si/no)
 - Sexo: (Masculino/Femenino)

Datos categóricos:

- Datos categóricos nominales:
 - Raza/Etnicidad
 - País de nacimiento
 - Religión
- Datos categóricos ordinales:
 - Estatus económico
 - Grado de acuerdo

- Datos continuos:
 - Presión sanguínea
 - Peso
 - Talla
 - Edad
 - Ingreso económico

- Datos tiempo para evento:
 - Datos que con una combinación de datos continuos y binarios
 - Momento en que un evento ocurre desde que se empezó a seguir a un sujeto.

Diferentes métodos para diferentes datos

Cual es la pregunta de investigación?

Cual es la pregunta de investigación?

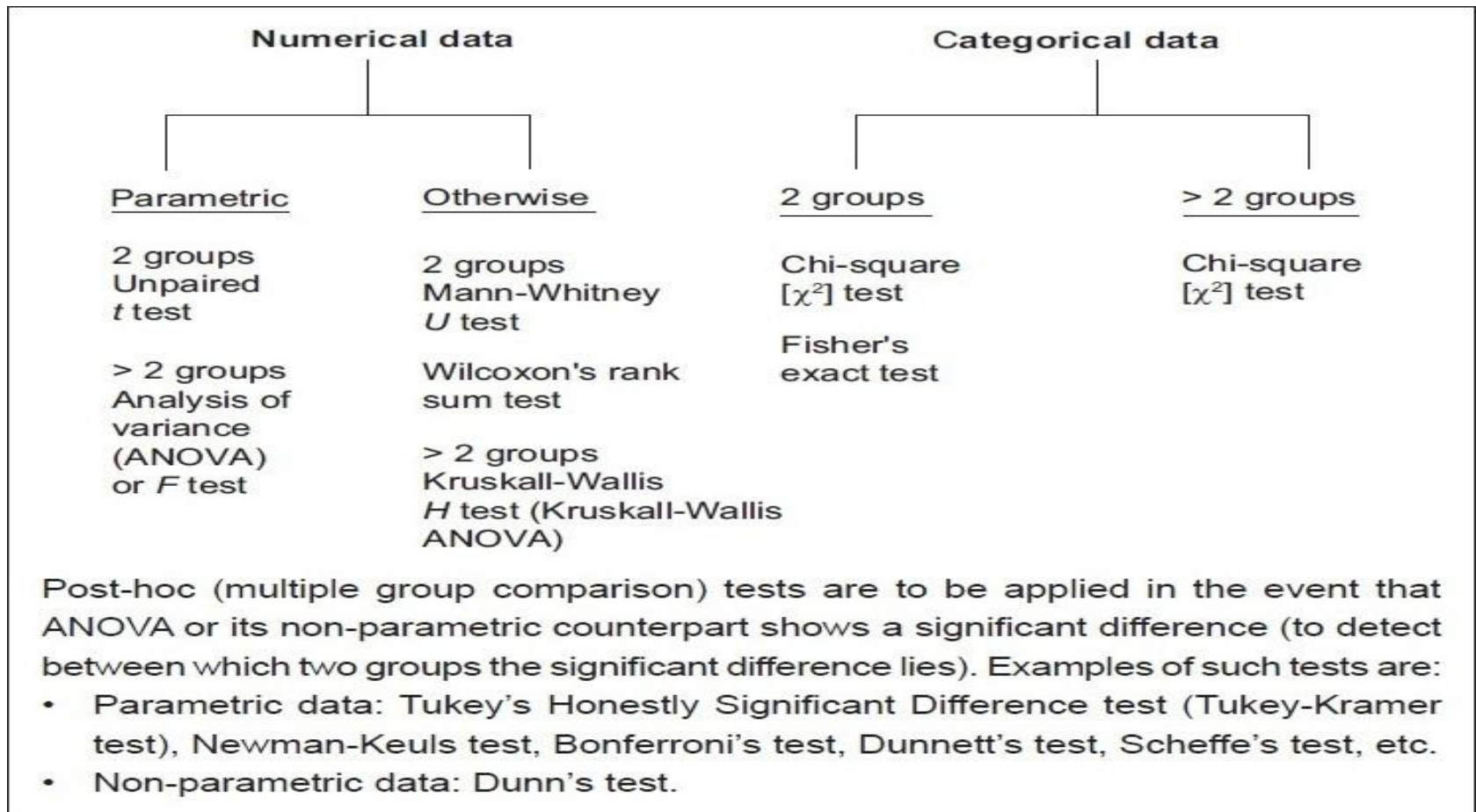
- Variable Independiente
- Variable Dependiente



Cual es la pregunta de investigación?

- Hay alguna diferencia de la variable X entre los grupos ?
- Existe alguna dependencia o relación entre X e Y?
- Hay algún cambio de la variable X en el tiempo?
- Cual es el nivel de asociación entre las variables?

Variables y pruebas estadísticas



Numerical data

Categorical data

Parametric

2 groups
Paired *t* test

> 2 groups
Repeated
measures
ANOVA

Otherwise

2 groups
Wilcoxon's
matched
pairs signed
rank test

> 2 groups
Friedman's ANOVA

2 groups

McNemar's
test

McNemar's
test
exact variants

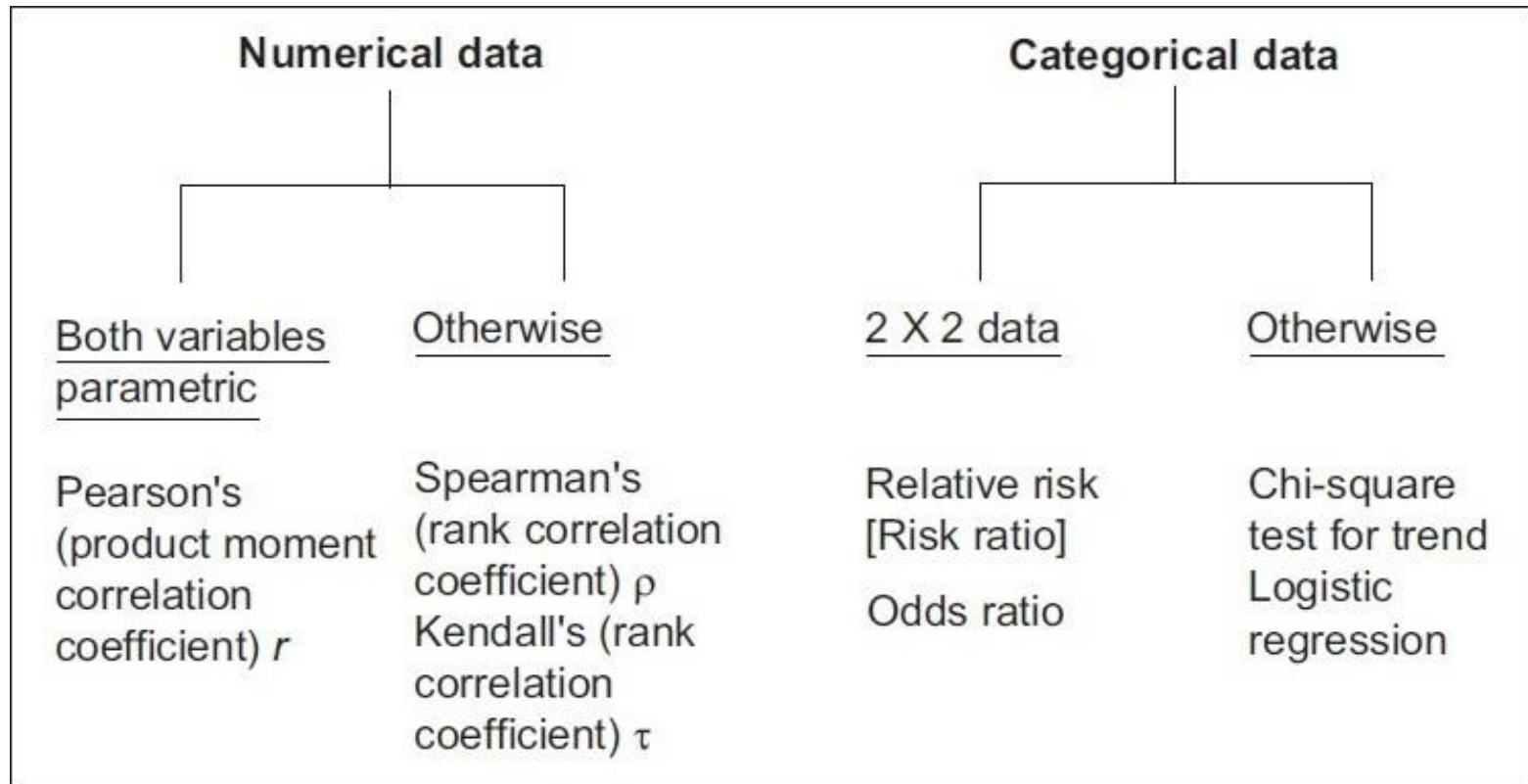
> 2 groups

Cochran's
Q test

Post-hoc (multiple group comparison) tests to be applied in the event that Repeated measures ANOVA or its non-parametric counterpart shows a significant difference (to detect between which two data sets the significant difference lies) include:

- Parametric data: Wilcoxon's matched pairs signed rank test.
- Non-parametric data: Dunn's test

Variables y pruebas estadísticas



Variables y pruebas estadísticas

Diseño de estudio	Medida de asociación	Modelo de regresión
Transversal	RP	GLM binomial, log
Caso - Control	OR	Regresión logística
Cohortes	RR o IRR	Regresión Poisson
Sobrevivencia	HR	Regresión Cox
Ensayo clínico	RR	Regresión de Poisson

Variables y pruebas estadísticas

Numerical data

Intraclass correlation coefficient
(quantitative method) Bland-Altman
plot (graphical method)

Categorical data

Cohen's Kappa statistic

Gracias!!

- Contacto:

Jorge L. Maguiña

Jorge.luis.maguina@gmail.com