

# Plan de análisis estadístico

Edith Alarcón Matutti  
César Gutiérrez Villafuerte

# PROCESO DE INVESTIGACION CIENTIFICA

**Paso 1**  
Concebir la idea a investigar

**Paso 2**  
Plantear el problema de investigación:  
•Objetivos  
•Preguntas  
•Justificación y viabilidad

**Paso 3**  
Elaborar el MARCO TEORICO  
•Revisión de la Literatura  
•Construcción del Marco teórico

**Paso 4**  
Definir el Tipo de investigación:  
Exploratoria, Descriptiva, Correlacional, Explicativa, y hasta que nivel llegará

**Paso 5**  
Establecer hipótesis Y definir variables conceptual y operacionalmente

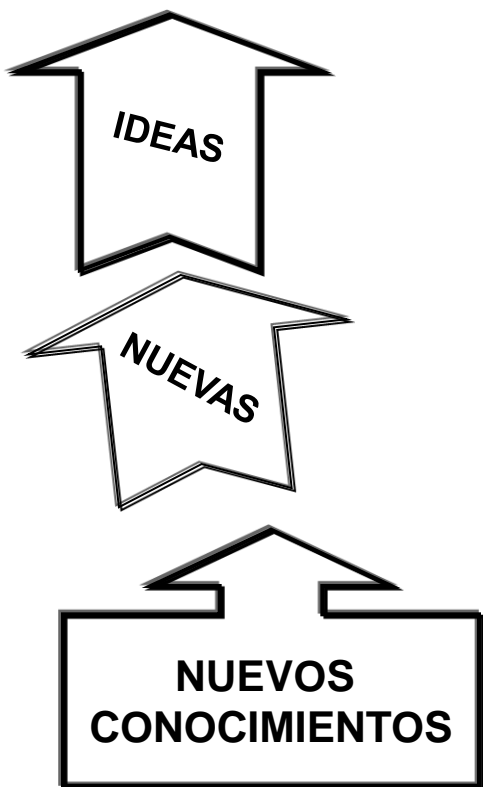
**Paso 6**  
Seleccionar el diseño apropiado De investigación

**Paso 7**  
Selección de la Muestra : Universo

**Paso 8**  
Recolección de datos:  
•Instrumento de medición  
•Validación del instrumento  
•Codificación y archivo de datos

**Paso 9**  
Análisis de datos:  
• Selección de pruebas estadísticas  
•Análisis de resultados

**Paso 10**  
Presentación de resultados:  
•Elaborar el reporte  
•Presentar el reporte de investigación



# Plan de análisis estadístico

## Definición

Se expone y detalla

- Las medidas de resumen de las variables y como serán presentadas, indicando los modelos y técnicas estadísticas a usar.
- El software a usar en el procesamiento
- El nivel de significancia fijado para las pruebas estadística.

# Uniform Requirements for Manuscripts Submitted to Biomedical Journals

## *Statistics*

*Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as  $P$  values, which fail to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used.*

# Proceso de selección de la PRUEBA

- Conocer las características de cada variable
- Verificar si nuestros datos son independientes o pareados.
- Verificación del cumplimiento de los supuestos para pruebas paramétricas.

# Estrategia de análisis

- Análisis exploratorio de datos
- Análisis univariado
- Análisis bivariado
- Análisis multivariado
- Incorporación de variables de estratificación

# Análisis exploratorio de datos

- El análisis exploratorio de datos consiste en la obtención de medidas de resumen y gráficos para un conjunto de datos. El hacer un análisis exploratorio de datos es importante para:
  - Evaluación general de los datos
  - Identificación de datos atípicos, extremos, perdidos
  - Descripción
  - Chequeo de supuestos para análisis posteriores
  - Identificar diferencias entre sub-grupos
- El análisis exploratorio ayuda en la determinación de las técnicas estadísticas a emplear luego, además puede indicar si es necesaria la transformación de alguna variable.

# Análisis univariado

- Variables cualitativas:
  - Distribuciones de frecuencias
  - Proporciones
  - Razones
  - Tasas
- Variables cuantitativas:
  - Medidas de tendencia central
  - Medidas de dispersión
  - Medidas de localización
  - Distribuciones de frecuencia (datos agrupados)



# Análisis bivariado

Situación	Variables cuantitativas con distribución normal	Variables cuantitativas sin distribución normal y cualitativas
	Prueba Paramétrica	Prueba No Paramétrica
Comparación de medias, 2 grupos independientes	t de Student – prueba Z	U de Mann Withney
Comparación de medias, 2 grupos pareados	t de Student	Wilcoxon
Comparación de medias, 3 ó más grupos	ANOVA	Kruskal Wallis
Comparación de variables cualitativas	-----	X <sup>2</sup> , prueba exacta de Fisher
Asociación entre dos variables cuantitativas	Coeficiente de correlación de Pearson	Coeficiente de correlación de Spearman

# Análisis multivariado

<b>Variable dependiente</b>	<b>Variables independientes</b>	<b>Modelo</b>	<b>Medición del resultado</b>
Cuantitativa	Cuantitativas o cualitativas	Regresión lineal múltiple	Coefficientes estandarizados
Cualitativa dicotómica	Cuantitativas o cualitativas	Regresión logística	Odds ratio (OR) ajustado
Tiempo de ocurrencia de un evento	Cuantitativas o cualitativas	Regresión de Cox	Hazard ratio (HR)

# Conclusiones

- El plan de análisis estadístico es una etapa definida antes de la recolección de datos.
- Es consecuencia de un buen marco teórico del tema a investigar.
- El diseño del formulario y el tamaño de muestra dependen de la complejidad del análisis estadístico establecido.

# Presentación de resultados

César Gutiérrez Villafuerte

# TABLAS ESTADÍSTICAS

Según el ICMJE, se recomienda la presentación de tablas con tres líneas horizontales y sin líneas verticales.

Las tablas deben de tener notas explicativas al pie de ser necesario, p.e. para precisar la prueba de hipótesis empleada.

**Tabla 2.** Prevalencia de vida de violencia psicológica en mujeres por su pareja por año. Perú, 2004–2007.

Año	Sí sufrió de violencia psicológica		No sufrió de violencia psicológica	
	n	%	n	%
2004	2 264	74.1	792	25.9
2005	2 526	75.2	835	24.8
2006	2 717	73.1	998	26.9
2007	2 661	74.1	928	25.9
Total	10 168	74.1	3 553	25.9

**Tabla 3:** Percepción de mujeres en edad fértil del riesgo de transmisión del VIH/SIDA de acuerdo a características sociodemográficas. Perú, 2004-2006.

<i>Característica sociodemográfica</i>	<b>Percepción del riesgo¶</b>					
	Alto o moderado		Mínimo o ninguno		No sabe	
	n	%	n	%	n	%
<i>Edad</i>						
▪ 15-19*	84	2.5	2988	90.5	228	7.0
▪ 20-24*	138	4.9	2466	87.9	203	7.2
▪ 25-29*	158	6.4	2143	86.3	183	7.3
▪ 30-34*	147	6.2	2041	85.9	189	7.9
▪ 35-39*	180	8.1	1857	83.4	189	8.5
▪ 40-44*	127	6.6	1621	83.7	189	9.7
▪ 45-49†	107	6.7	1325	82.5	174	10.8
<i>Grado de instrucción</i>						
▪ No estudia*	42	5.5	240	89.0	447	34.5
~~~~~						
<i>Alfabetización</i>						
▪ No puede leer el párrafo u oración*	25	3.3	507	66.1	235	30.6
▪ Lo lee parcialmente*	32	5.6	395	69.3	143	25.1
▪ Lo lee completamente†	883	5.8	13511	87.9	971	6.3
▪ No hay tarjeta para el idioma de la entrevistado§	0	0	8	88.9	1	11.1
▪ Ceguera/problemas visuales§	1	8.3	8	66.7	3	25.0

\* Regresión logística multinomial, diferencia significativa ( $p < 0.05$ ) con categoría de referencia.

† Categoría de referencia para las comparaciones.

‡ Estas categorías fueron recodificadas como "lengua distinta al castellano" para el análisis multivariado.

§ Categoría excluida del análisis.

¶ Frecuencias y porcentajes obtenidos de respuestas válidas.

TABLA 5

Ingesta de energía y nutrientes y adecuación a los estándares de referencia según índice de masa corporal (IMC).

IMC	Energía			Hierro			Calcio		
	X	DS	% Adec.	X	DS	% Adec	X	DS	% Adec
NORMAL Nº= 24	1900	661.0	92	12.6	5.7	70.0	586.6	353.1	58.6
SOBREPESO Nº= 25	1580	526.6	76	10.9	3.7	60.5	471.4	306.2	47.1
OBESIDAD Nº= 5	1353	242.8	68	9.6	2.2	53.3	224.8	86.9	22.4
TOTAL Nº 54	1701	596.5	82	11.5	4.6	63.9	499.8	328.7	49.9
P*	.065			.288			.066		

\* Anova

***¿Está bien elaborada esta tabla?***



**Tabla 1.** Distribución de personas encuestadas según ciudad y tipo de establecimiento de salud (EESS).

	Lima N=444 (69%)	Pucallpa N=141 (9%)	Ayacucho N=59 (22%)	Total N=644 (100%)
<b><i>EESS Público</i></b>				<b>N=522 (81%)</b>
Hospital Nacional Cayetano Heredia	N=103	--	--	--
Hospital Nacional Dos de Mayo	N=86	--	--	--
Hospital Nacional Hipólito Unanue	N=133	--	--	--
Hospital Regional de Pucallpa	--	N=89	--	--
Hospital Amazónico de Yarinacocha	--	N=52	--	--
Hospital Regional de Ayacucho	--	--	N=59	--
<b><i>EESS Privado</i></b>				<b>N=122 (19%)</b>
Clínica Montefiori	N=81	--	--	--
Clínica Maison de Santé Este	N=41	--	--	--

***¿Está bien elaborada esta tabla?***

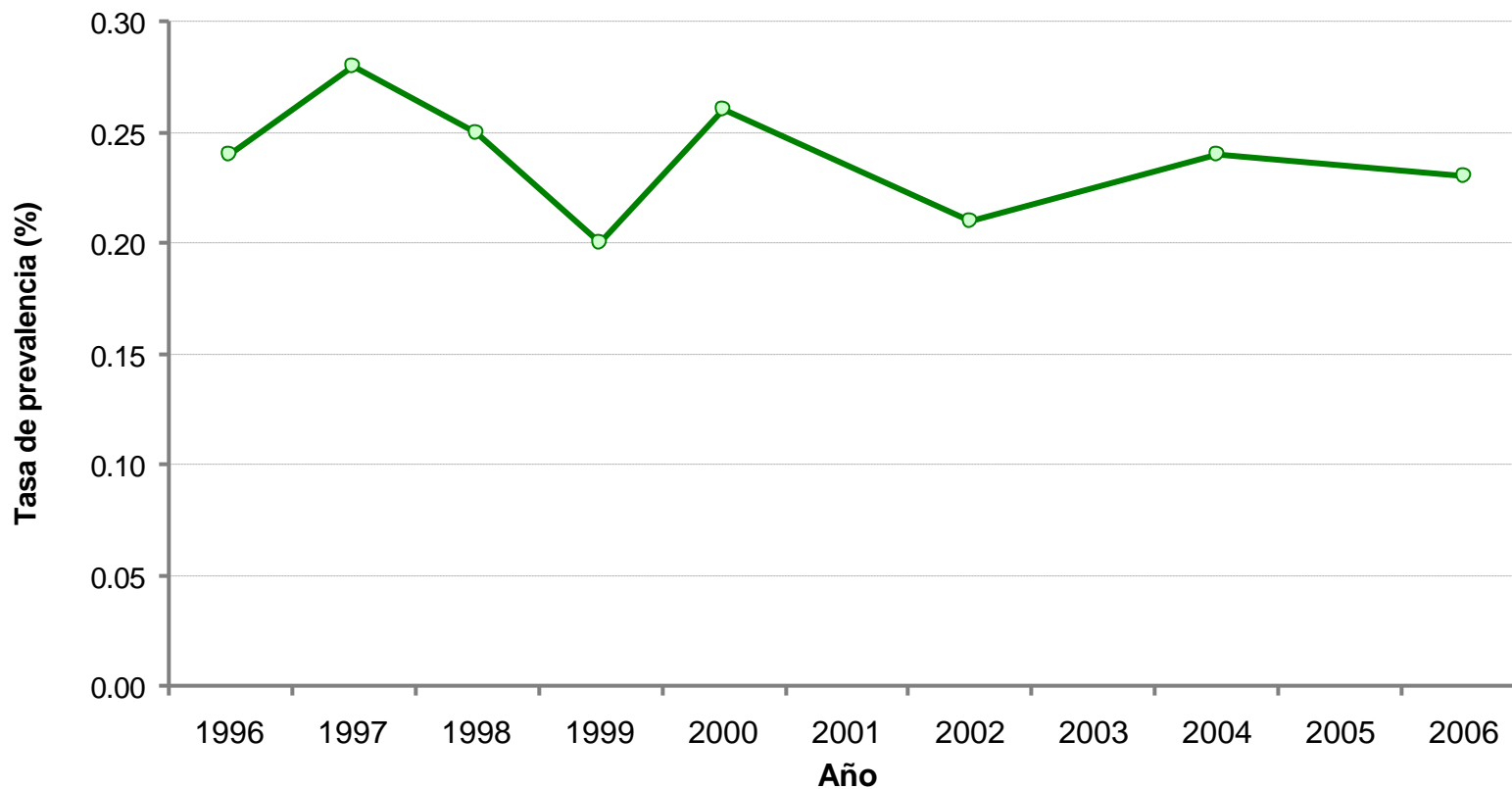
# GRÁFICOS

Los gráficos deben de ser seleccionados adecuadamente según la naturaleza de las variables analizadas.

Deben guardar también una adecuada proporción.

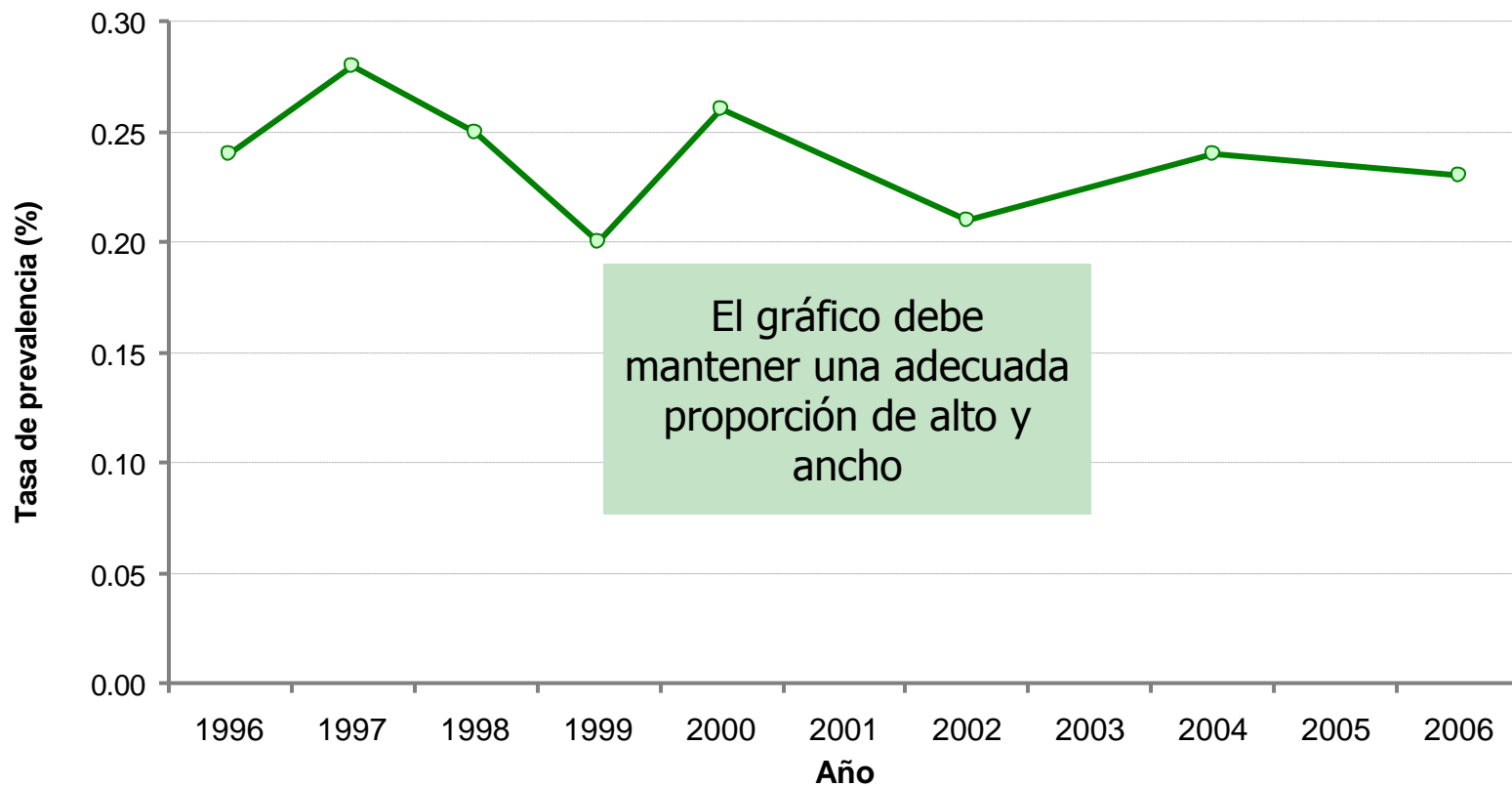
No deben inducir a conclusiones diferentes a las que los hechos muestran.

### Prevalencia de VIH en población donante de sangre. San Juan, 1996-2006



Fuente: Programa Nacional de Bancos de Sangre y Hemoterapia. Informe Anual 2007. Ciudad de Dios, San Juan.

### Prevalencia de VIH en población donante de sangre. San Juan, 1996-2006

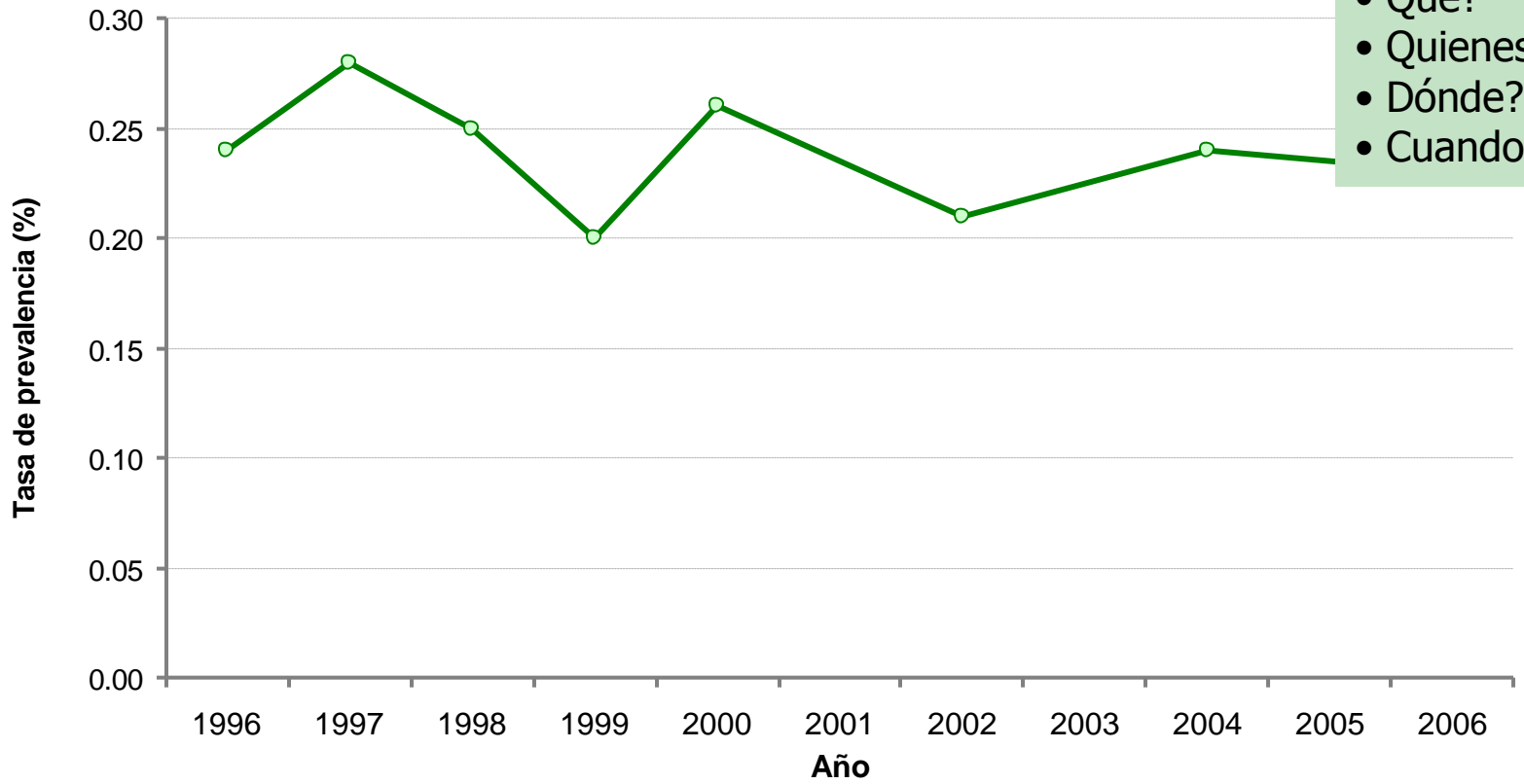


Fuente: Programa Nacional de Bancos de Sangre y Hemoterapia. Informe Anual 2007. Ciudad de Dios, San Juan.

Prevalencia de VIH en población donante de sangre.  
San Juan, 1996-2006

El título debe indicar:

- Qué?
- Quienes?
- Dónde?
- Cuando?



Fuente: Programa Nacional de Bancos de Sangre y Hemoterapia. Informe Anual 2007. Ciudad de Dios, San Juan.

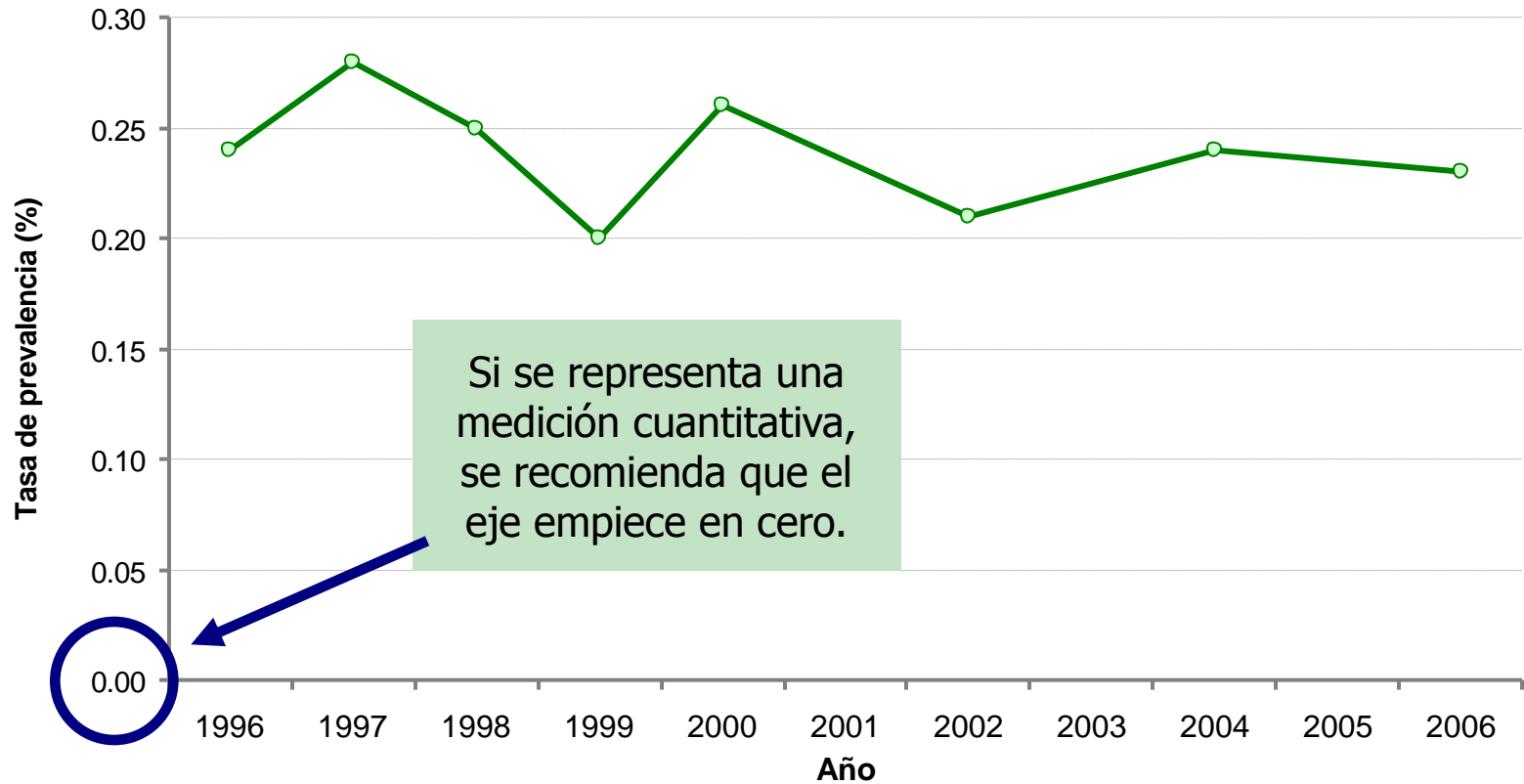
### Prevalencia de VIH en población donante de sangre. San Juan, 1996-2006



Los ejes deben llevar un título que identifique lo que están representando.

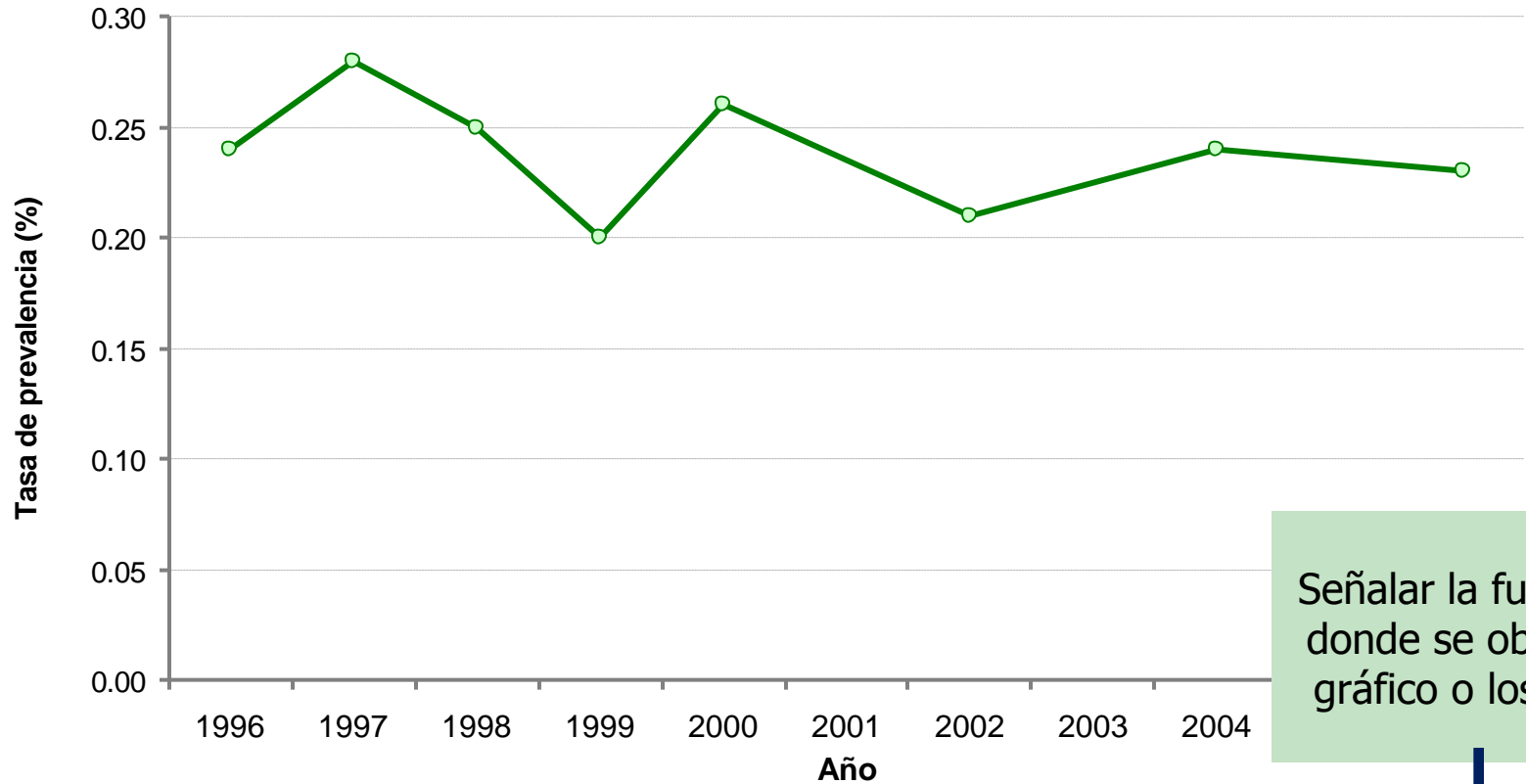
ia. Informe Anual 2007. Ciudad de Dios, San Juan.

### Prevalencia de VIH en población donante de sangre. San Juan, 1996-2006



Fuente: Programa Nacional de Bancos de Sangre y Hemoterapia. Informe Anual 2007. Ciudad de Dios, San Juan.

### Prevalencia de VIH en población donante de sangre. San Juan, 1996-2006



Señalar la fuente de donde se obtuvo el gráfico o los datos

Fuente: Programa Nacional de Bancos de Sangre y Hemoterapia. Informe Anual 2007. Ciudad de Dios, San Juan.

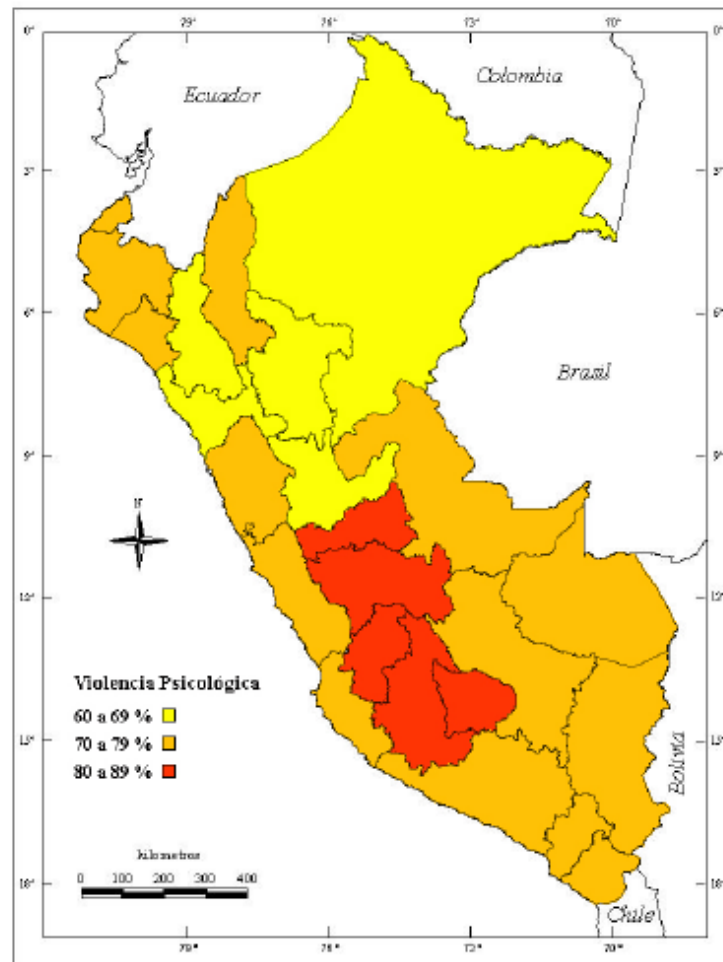


# GRÁFICOS ESTADÍSTICOS - MAPAS

1. Representan la distribución geográfica de un fenómeno.
2. Generalmente se emplean mapas con la división política de un territorio.

# GRÁFICOS ESTADÍSTICOS - MAPAS

Figura 1. Prevalencia de vida de violencia psicológica en mujeres en edad fértil por su pareja según departamento de residencia. Perú, 2004–2007.



# GRÁFICOS ESTADÍSTICOS – BARRAS

1. Se emplean para analizar una o dos variables cualitativas.
  - Si se analiza una variable – barras simples.
  - Si se analizan dos variables – barras agrupadas.
2. Generalmente representan frecuencias (absolutas o relativas) o tasas.
3. También se pueden emplear para representar una medida de resumen de una variable cuantitativa (generalmente la media).

# GRÁFICOS ESTADÍSTICOS – BARRAS

**Figura 5.** Razones de no empleo de métodos anticonceptivos. Gestantes adolescentes de la sierra rural del sur de Ayacucho, 2004-2008.

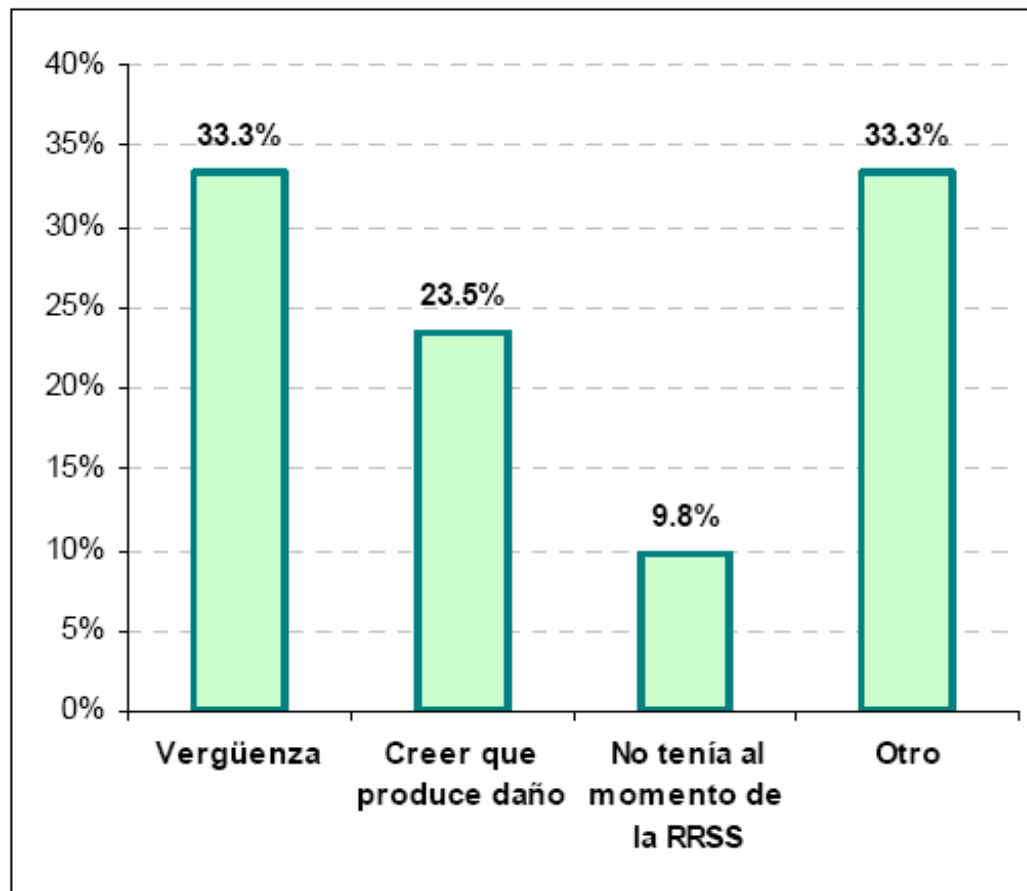
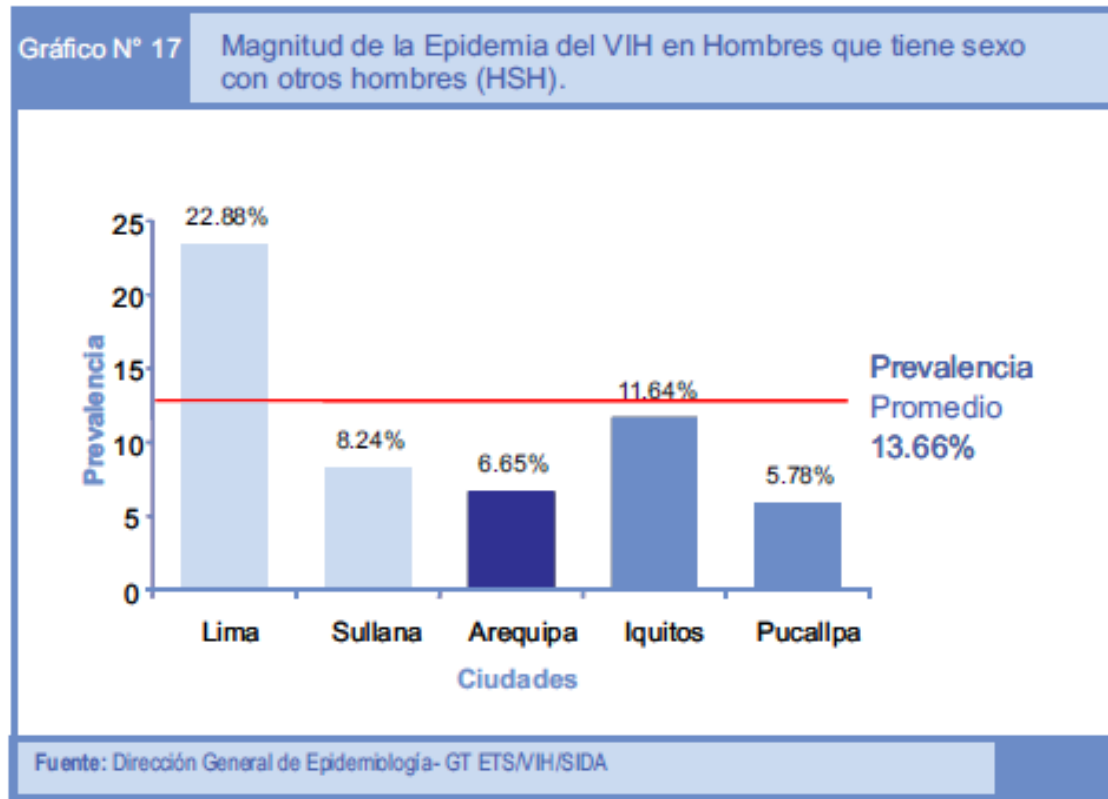


Gráfico de  
barras simples

# GRÁFICOS ESTADÍSTICOS – BARRAS

## Gráfico de barras simples



# GRÁFICOS ESTADÍSTICOS – BARRAS

**Figura 6.** Diagnóstico principal de los atendidos por accidentes de tránsito. Hospital Nacional Daniel A. Carrión y Hospital San José. Callao, 2002-2004.

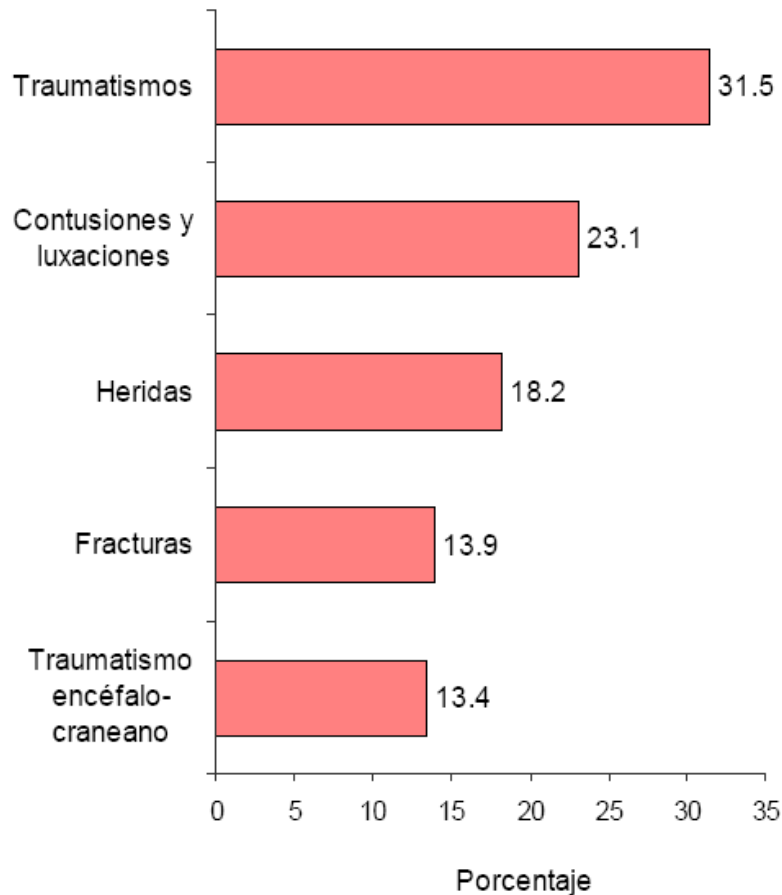


Gráfico de barras simples  
(horizontales)

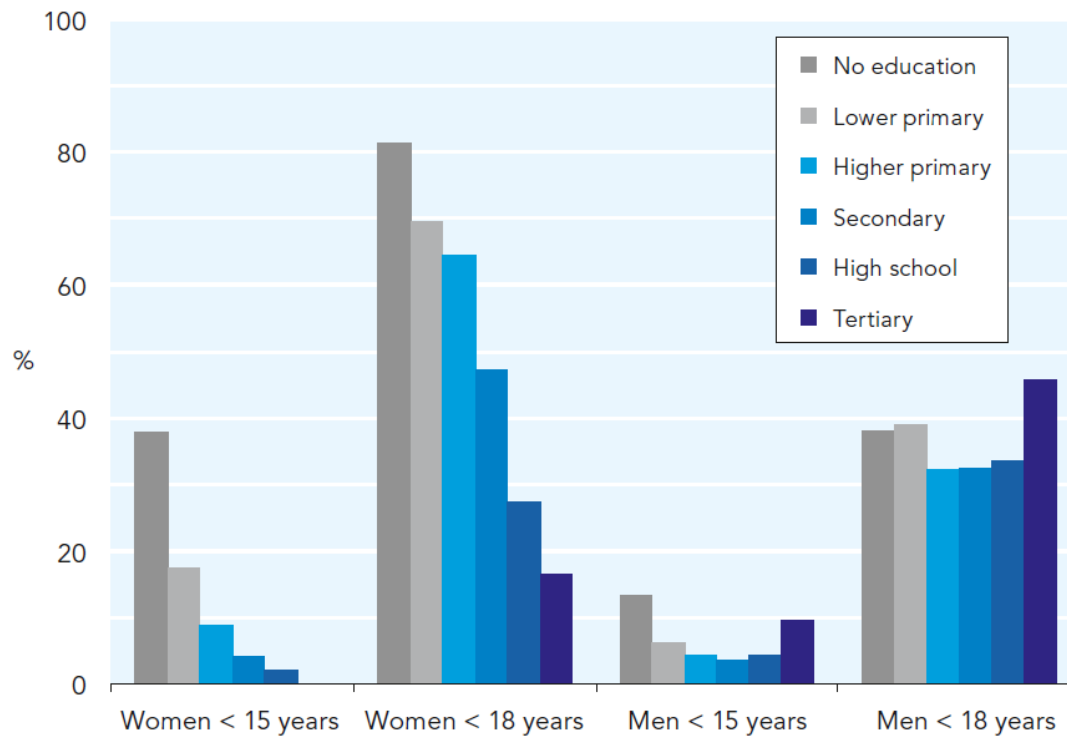
Fuente de datos: HNDAC y HSJ.

# GRÁFICOS ESTADÍSTICOS – BARRAS

## Gráfico de barras agrupadas

Figure 5

Age at first sexual intercourse by education status in Swaziland, 2007



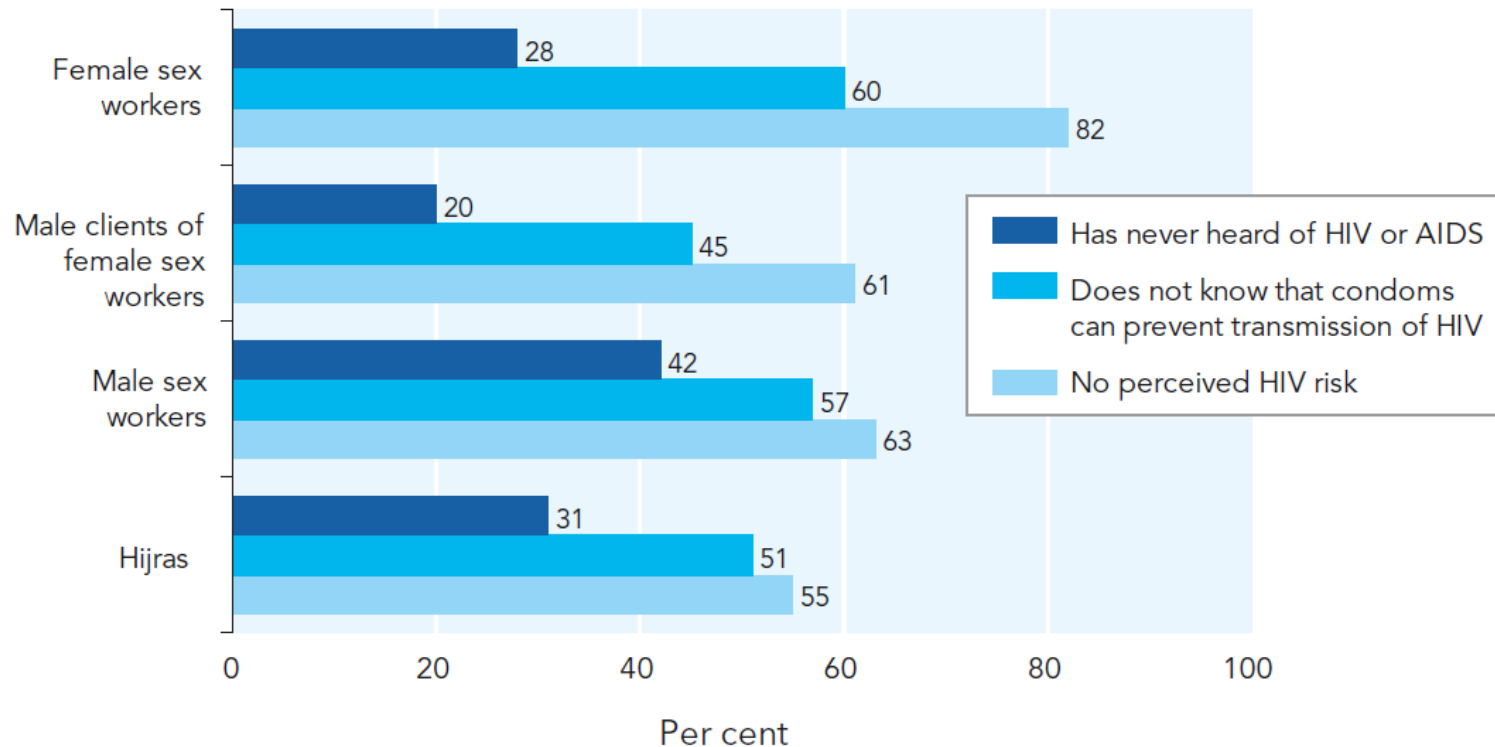
Source: Central Statistical Office & Macro International (2008).

# GRÁFICOS ESTADÍSTICOS – BARRAS

## Gráfico de barras agrupadas (horizontales)

Figure 9

Vulnerability to sexual HIV transmission in commercial sex in Karachi and Lahore, Pakistan



Source: Bokhari et al. (2007).



# GRÁFICOS ESTADÍSTICOS – BARRAS

**Figura 1.** Nivel de conocimientos según ocupación.

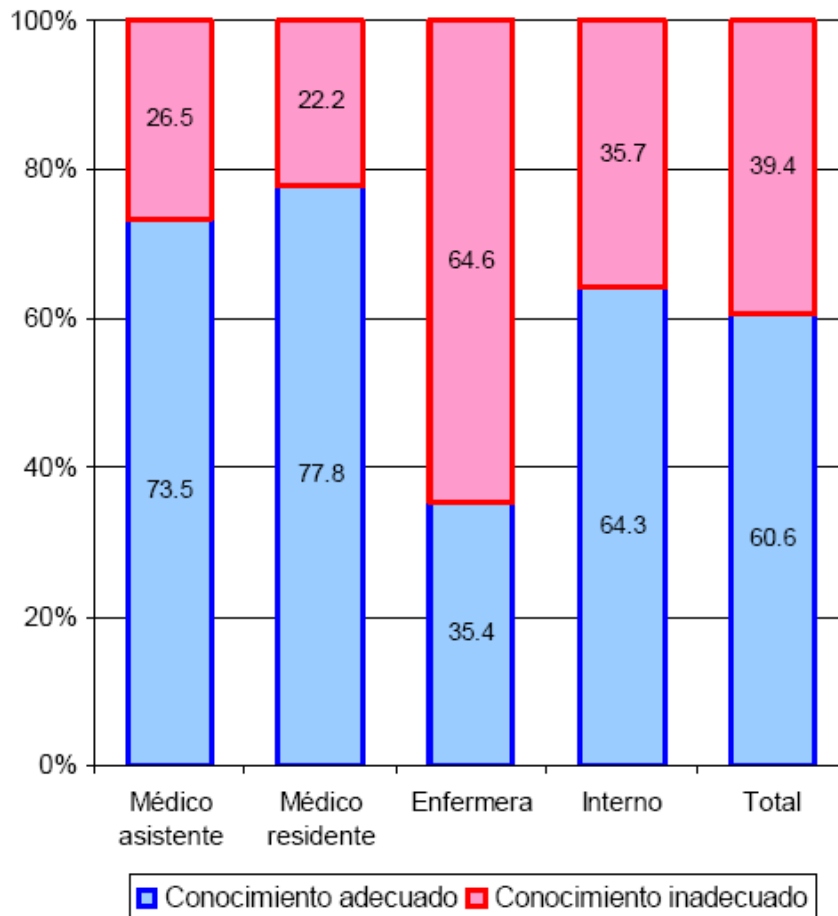


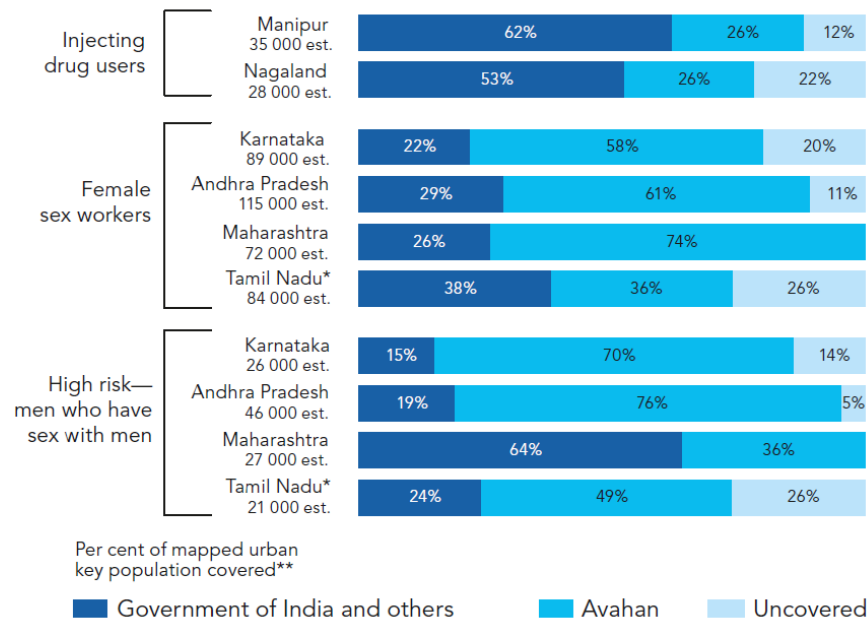
Gráfico de  
barras apiladas

# GRÁFICOS ESTADÍSTICOS – BARRAS

## Gráfico de barras apiladas (horizontales)

Figure 10

Saturating prevention coverage through complementary programming.  
Avahan has achieved a high coverage of target populations (routine programme monitoring data)



Percentages indicate intended coverage through establishment of services in specific geographic areas.

\* Includes districts with no intended coverage.

\*\* Mapping and size estimation quality varies by state.

Does not include rural areas

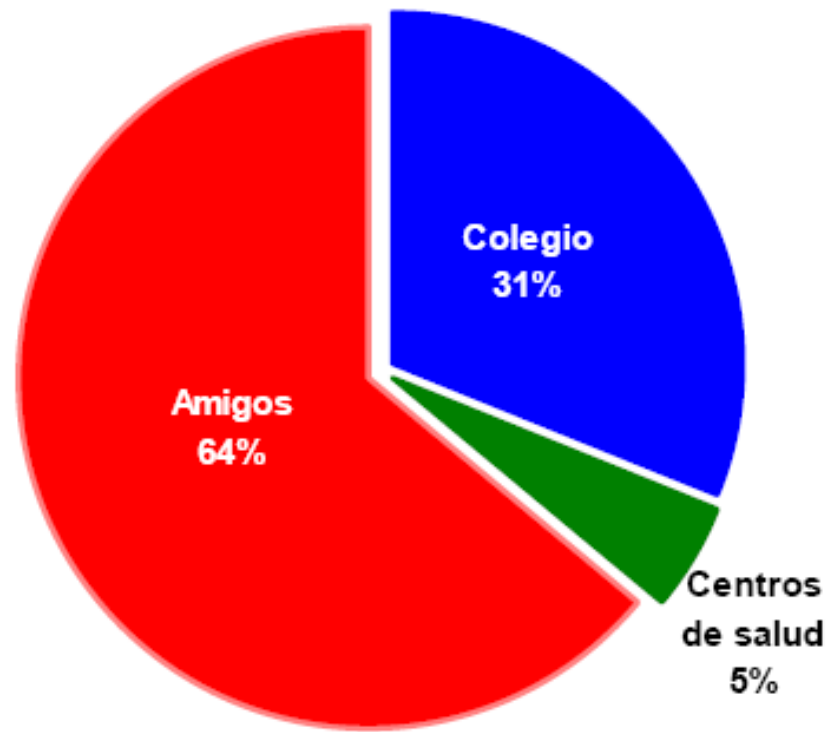
Source: Avahan and State AIDS Control Society programme data (2008).

# GRÁFICOS ESTADÍSTICOS – CIRCULAR

1. Se emplean para analizar una variable cualitativa.
2. Se recomienda que representen frecuencias relativas de valores que son mutuamente excluyentes.
3. La variable no debe tener más de cinco categorías.

# GRÁFICOS ESTADÍSTICOS – CIRCULAR

**Figura 3.** Lugar dónde las adolescentes recibieron educación sexual. Gestantes adolescentes de la sierra rural del sur de Ayacucho, 2004-2008.



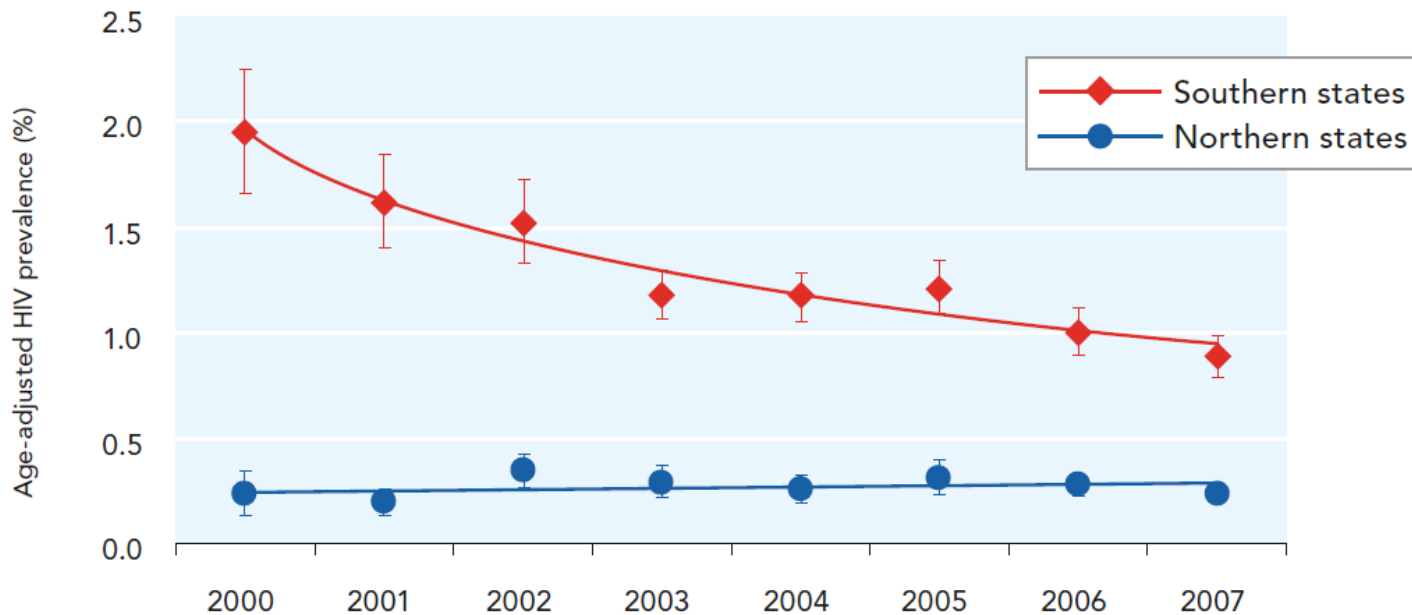
# GRÁFICOS ESTADÍSTICOS – GRÁFICO LINEAL

1. Representan la evolución o tendencia de una variable en el tiempo.

# GRÁFICOS ESTADÍSTICOS – GRÁFICO LINEAL

Figure 8

Age-adjusted HIV prevalence among antenatal attendees aged 15–24 from 2000 to 2007 in high-prevalence southern states (Andhra Pradesh, Karnataka, Maharashtra and Tamil Nadu) and northern states of India



Logarithmic trend line; test for trend by logistic regression, with age adjustment to the entire study population,  $n = 202\ 254$  for the south,  $n = 221\ 588$  for the north.

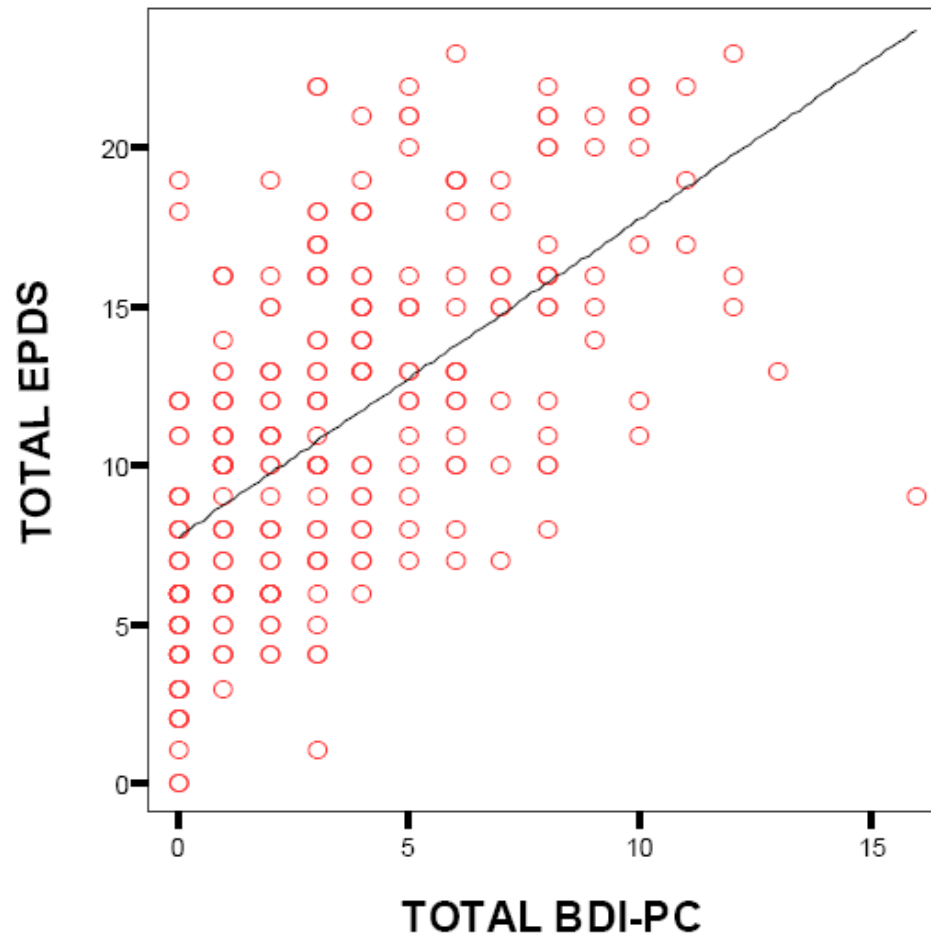
Source: Arora et al. (2008).

# GRÁFICOS ESTADÍSTICOS – DE DISPERSIÓN

1. Sirven para poder identificar que tan lineal es la relación entre dos variables numéricas.

# GRÁFICOS ESTADÍSTICOS – DE DISPERSIÓN

**Figura 2.** Grafico de dispersión entre los puntajes de BDI-PC y EPDS.

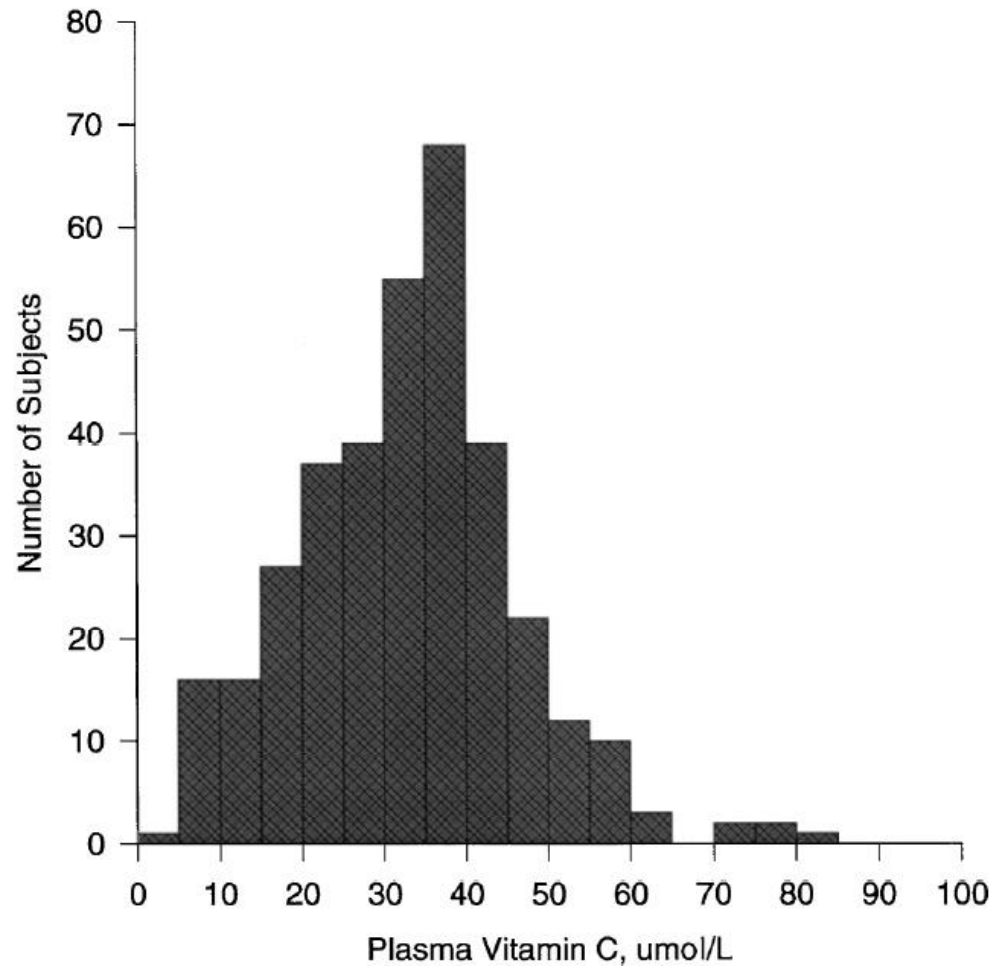




# GRÁFICOS ESTADÍSTICOS – HISTOGRAMA

1. Sirven para representar la distribución de una variable cuantitativa (continua).

# GRÁFICOS ESTADÍSTICOS – HISTOGRAMA



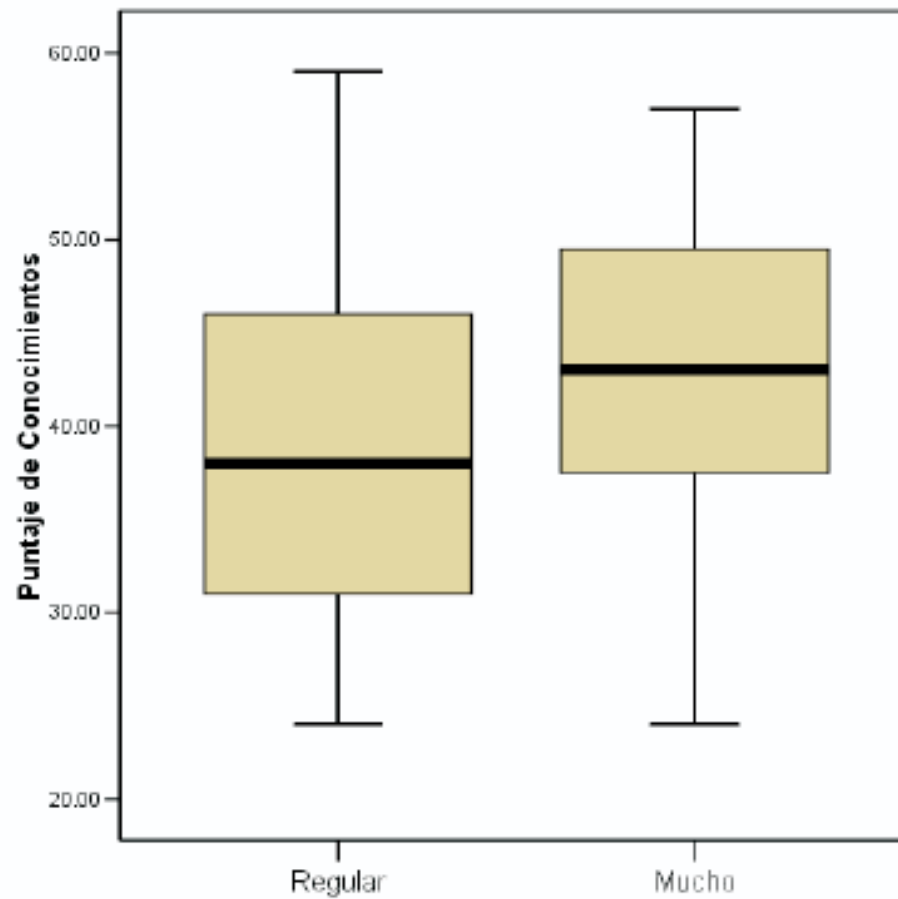
**Fig. 1.** Distribution of plasma vitamin C concentrations in females (n=350).

# GRÁFICOS ESTADÍSTICOS – CAJAS (BOX PLOT)

1. Sirven para representar la distribución de una variable cuantitativa en función a los cuartiles.

# GRÁFICOS ESTADÍSTICOS – CAJAS (BOX PLOT)

FIGURA 3. Comparación del puntaje de conocimiento obtenido según la auto-percepción del conocimiento de médicos pediatras de dos hospitales de emergencia de Lima, 2006.



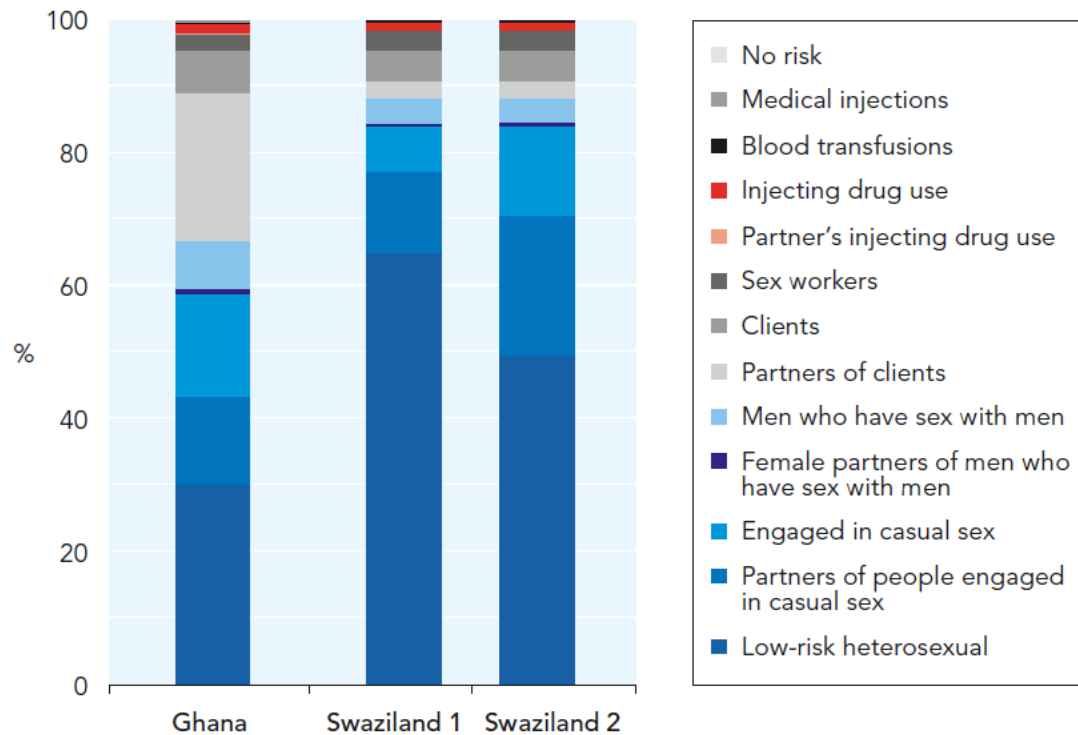
# **GRÁFICOS ESTADÍSTICOS**

## **ERRORES FRECUENTES**

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

Figure 3

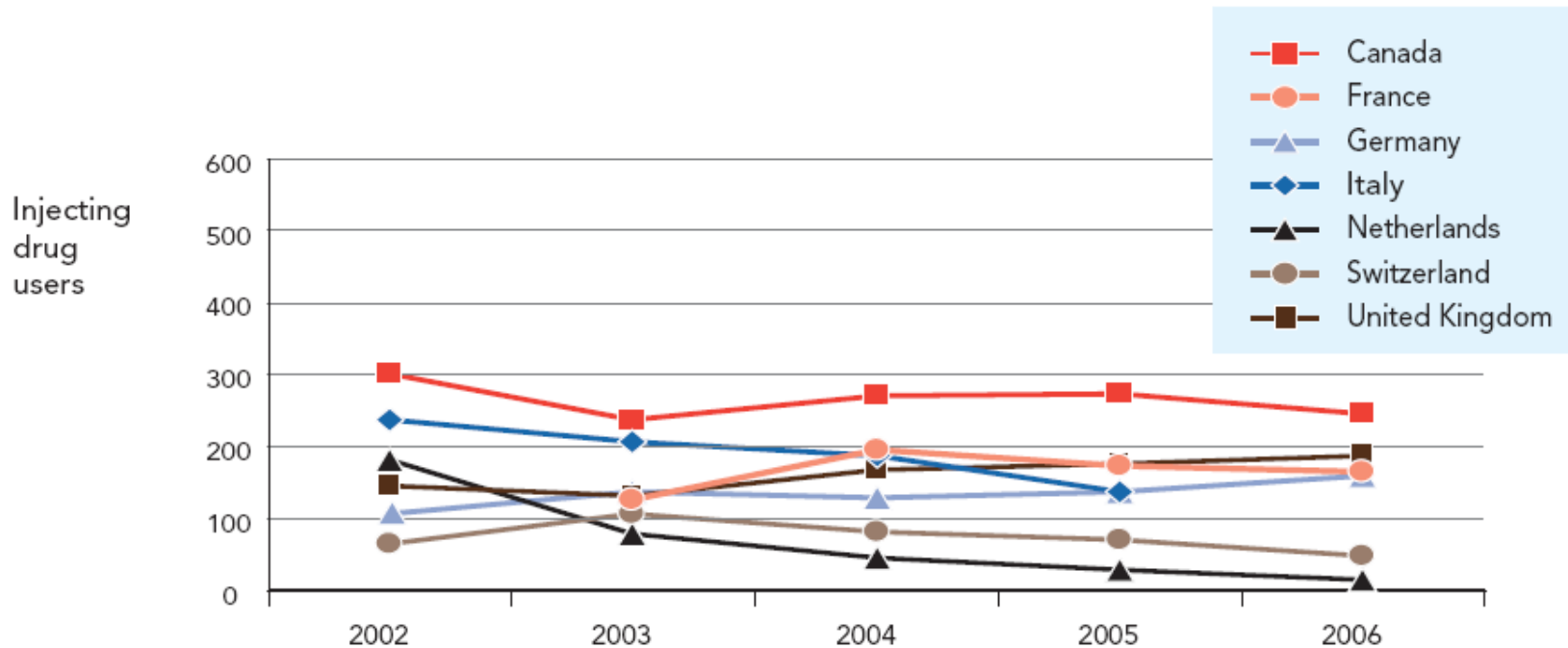
Distribution of new infections by mode of exposure in Ghana and Swaziland, 2008



Note: sensitivity analysis for Swaziland used different data sources.

Sources: Bosu et al. (2009) and Mngadi et al. (2009).

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES



Sources:

(1) Public Health Agency of Canada. HIV and AIDS in Canada. Selected Surveillance Tables to June 30, 2007. Surveillance and Risk Assessment Division, Centre for Infectious Disease Prevention and Control, Public Health Agency of Canada, 2007. (2) Epidemiologisches Bulletin (5. Oktober 2007/ Sonderausgabe B aktuelle daten und informationen zu infektionskrankheiten und public health). (3) The UK Collaborative Group for HIV and STI Surveillance. Testing Times. HIV and other Sexually Transmitted Infections in the United Kingdom: 2007. London: Health Protection Agency, Centre for Infections. November 2007. (4) EuroHIV. HIV/AIDS Surveillance in Europe. End-year report 2006. Saint-Maurice: Institut de veille sanitaire, 2007. No. 75. (5) EuroHIV. HIV/AIDS Surveillance in Europe. Mid-year report 2007. Saint-Maurice: Institut de Veille Sanitaire, 2007. No. 76.

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

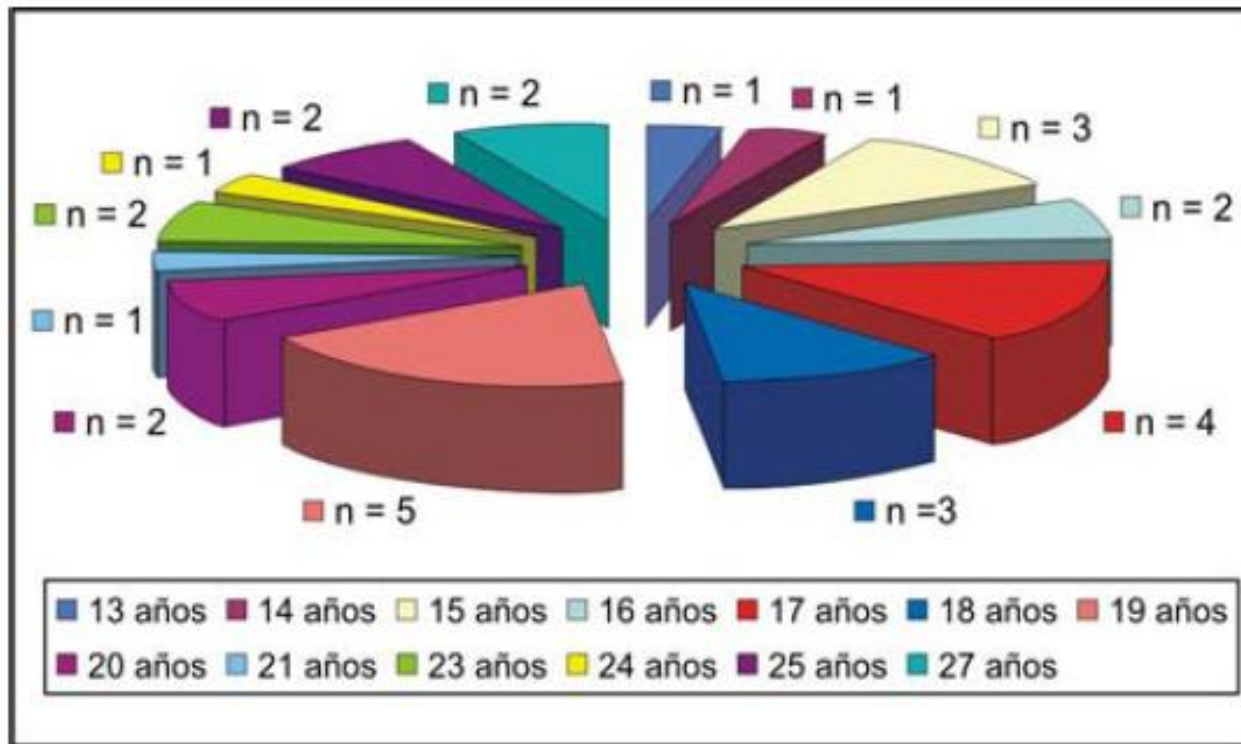
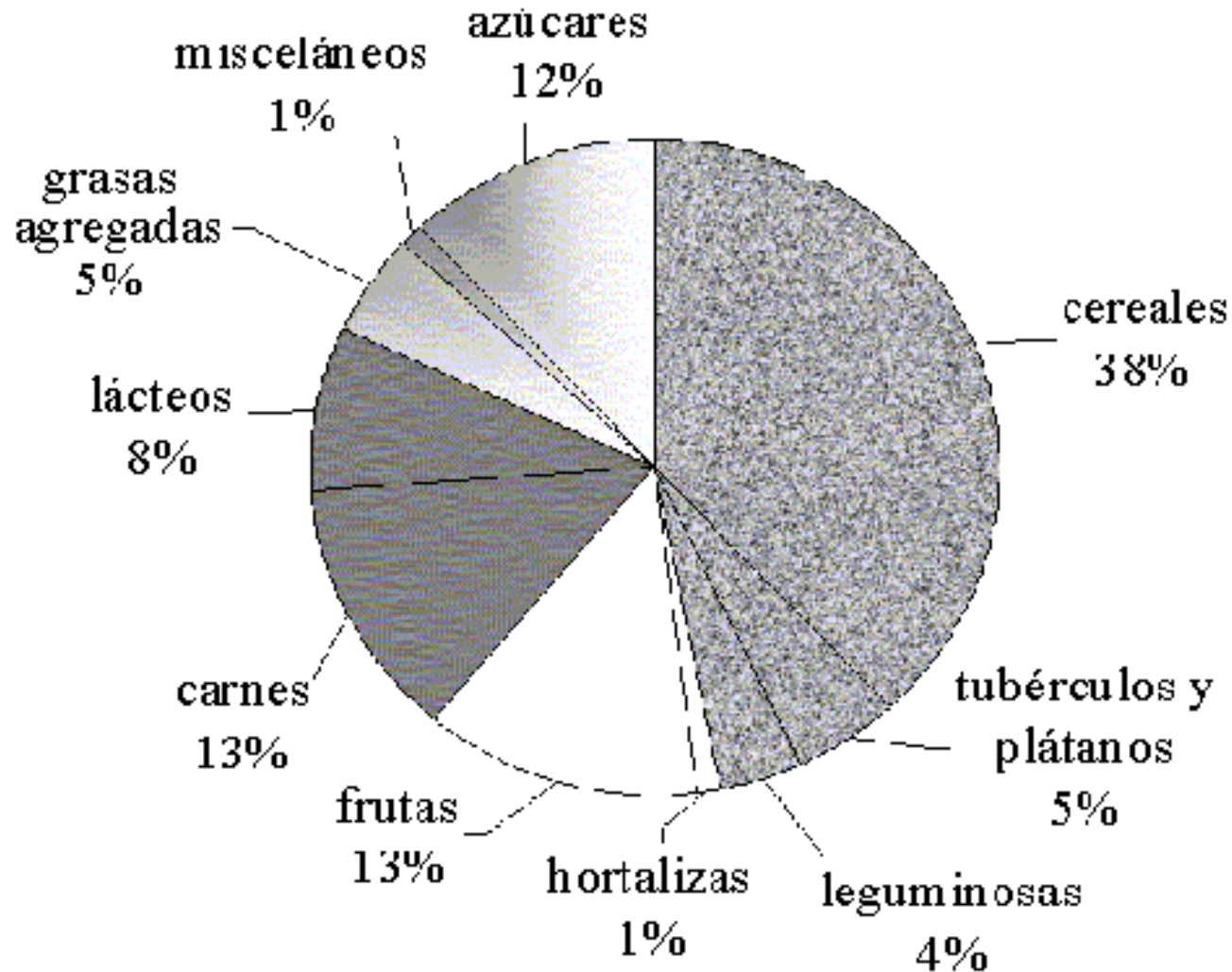


Figura 1. Casos según edades de pacientes.



# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

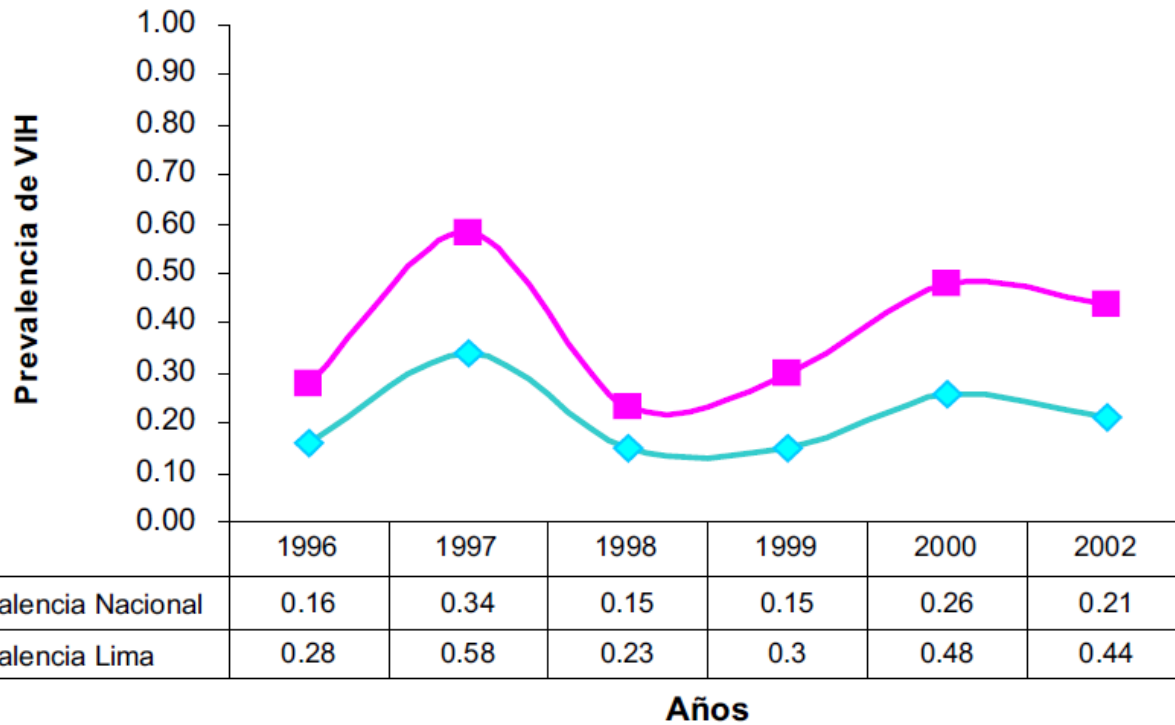
Figura 2. Aporte calórico por grupo de alimentos



# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

Gráfico N° 11

Evolución de la Prevalencia de VIH en población gestantes  
Perú 1996-2002



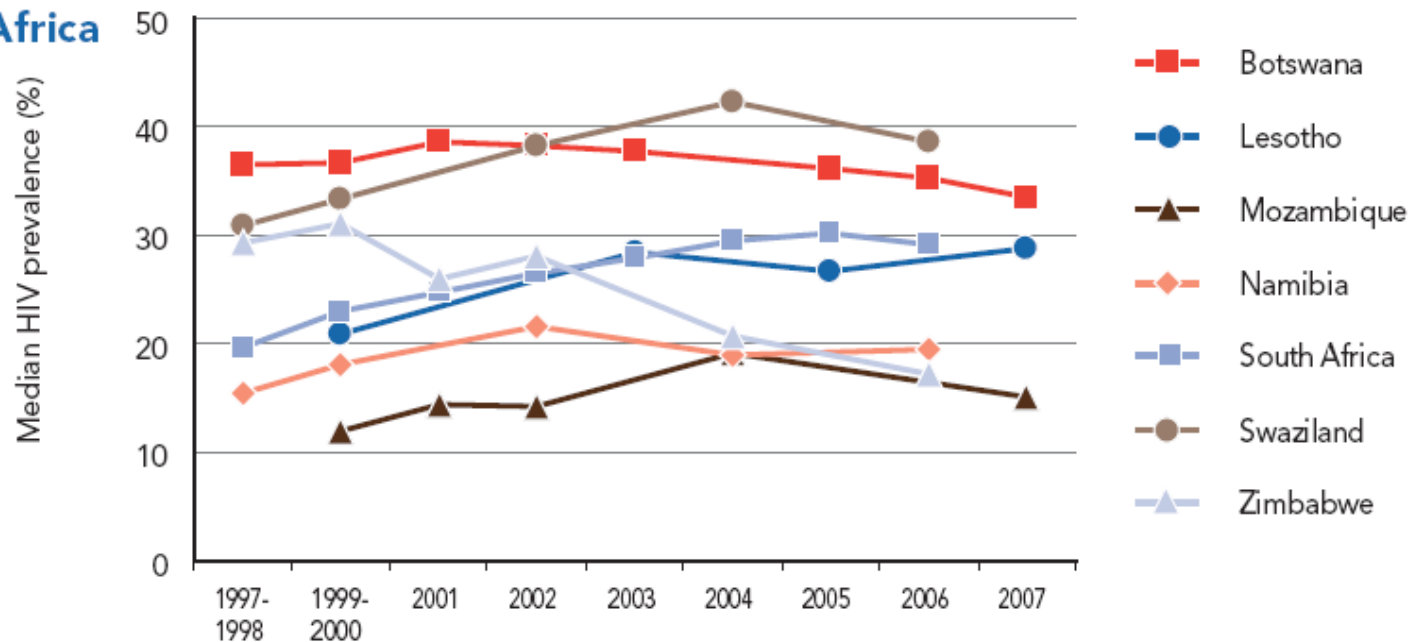
Fuente: Dirección General de Epidemiología- GT ETS/VIH/SIDA

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

FIGURE 2.9

HIV prevalence (%) among pregnant women attending antenatal clinics in sub-Saharan Africa, 1997–2007

## Southern Africa



# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

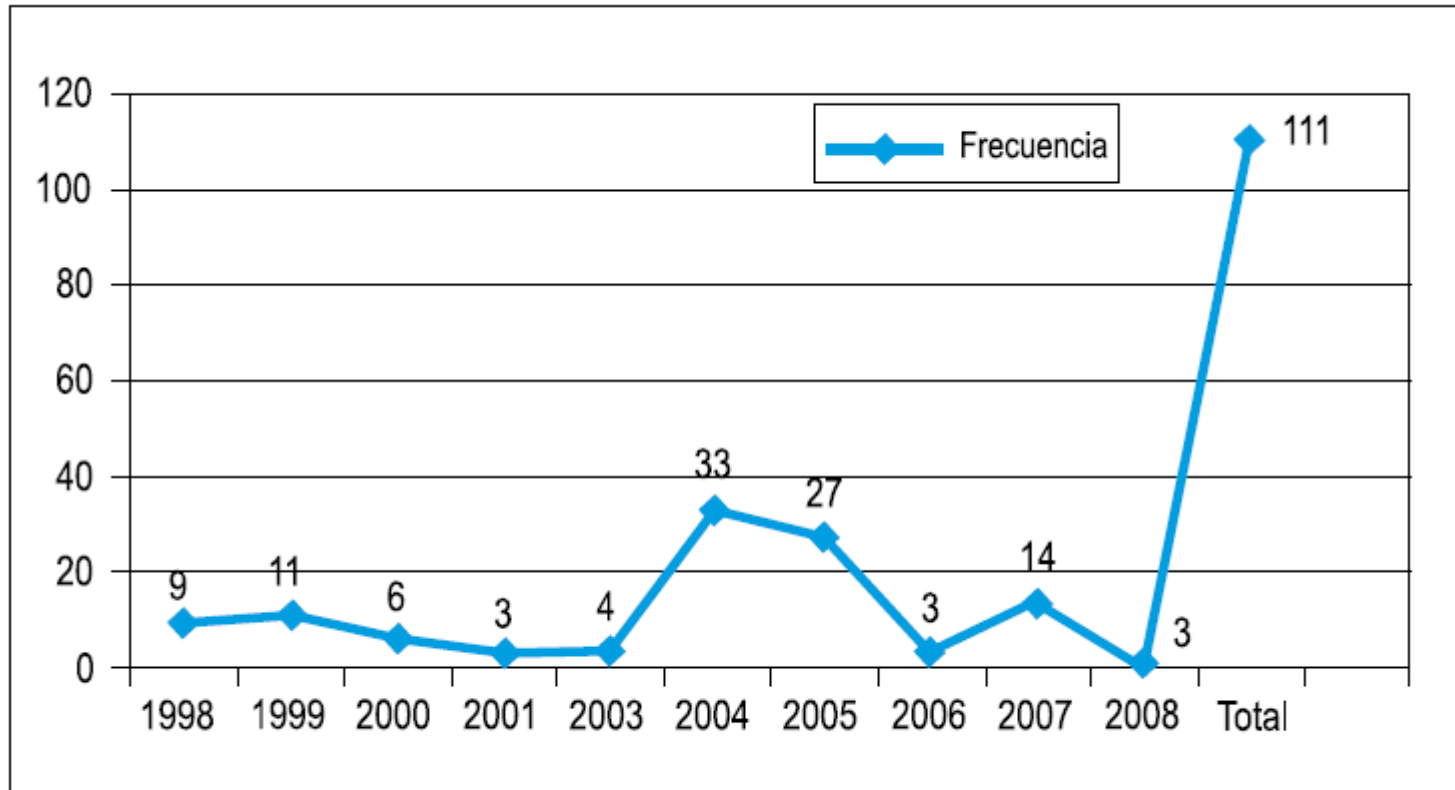
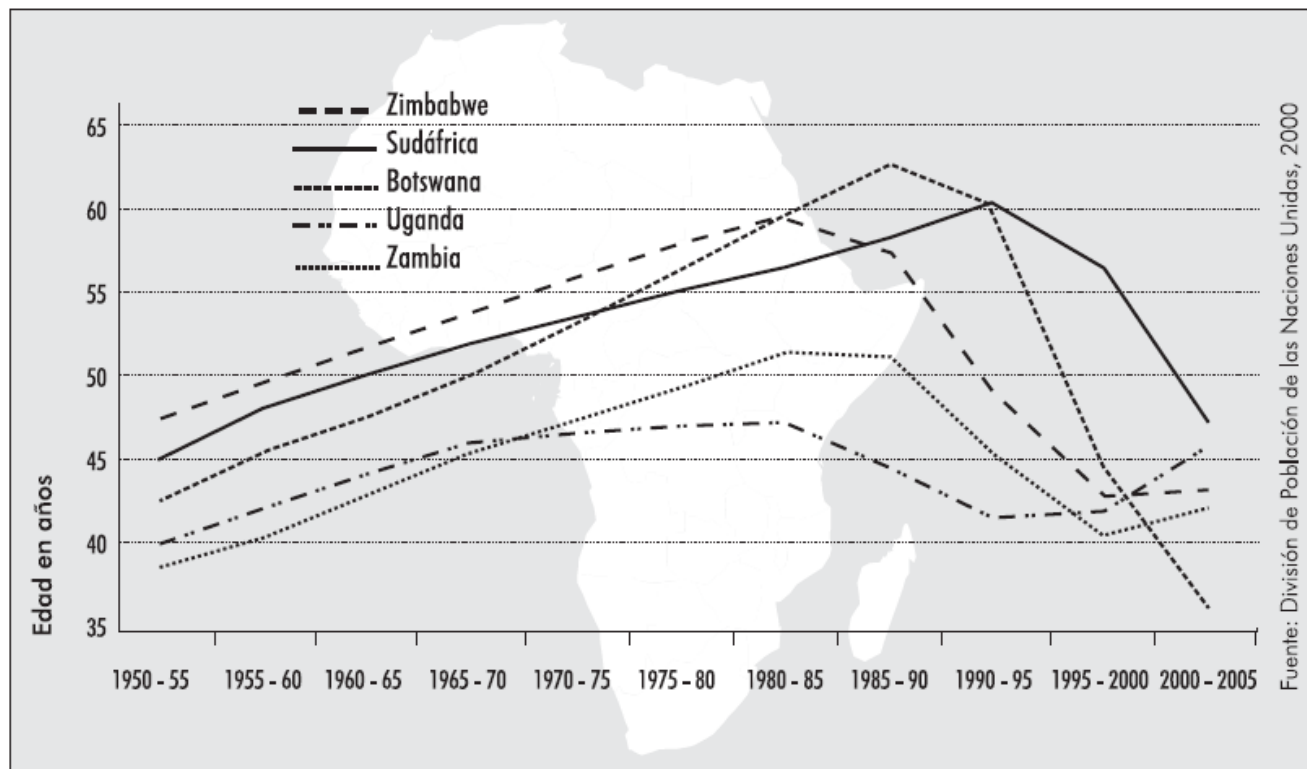


Figura 1. Frecuencia de casos de bartonellosis.

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

Figura 2. Cambios en la esperanza de vida en determinados países africanos con elevada prevalencia del VIH, 1950-2005



# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

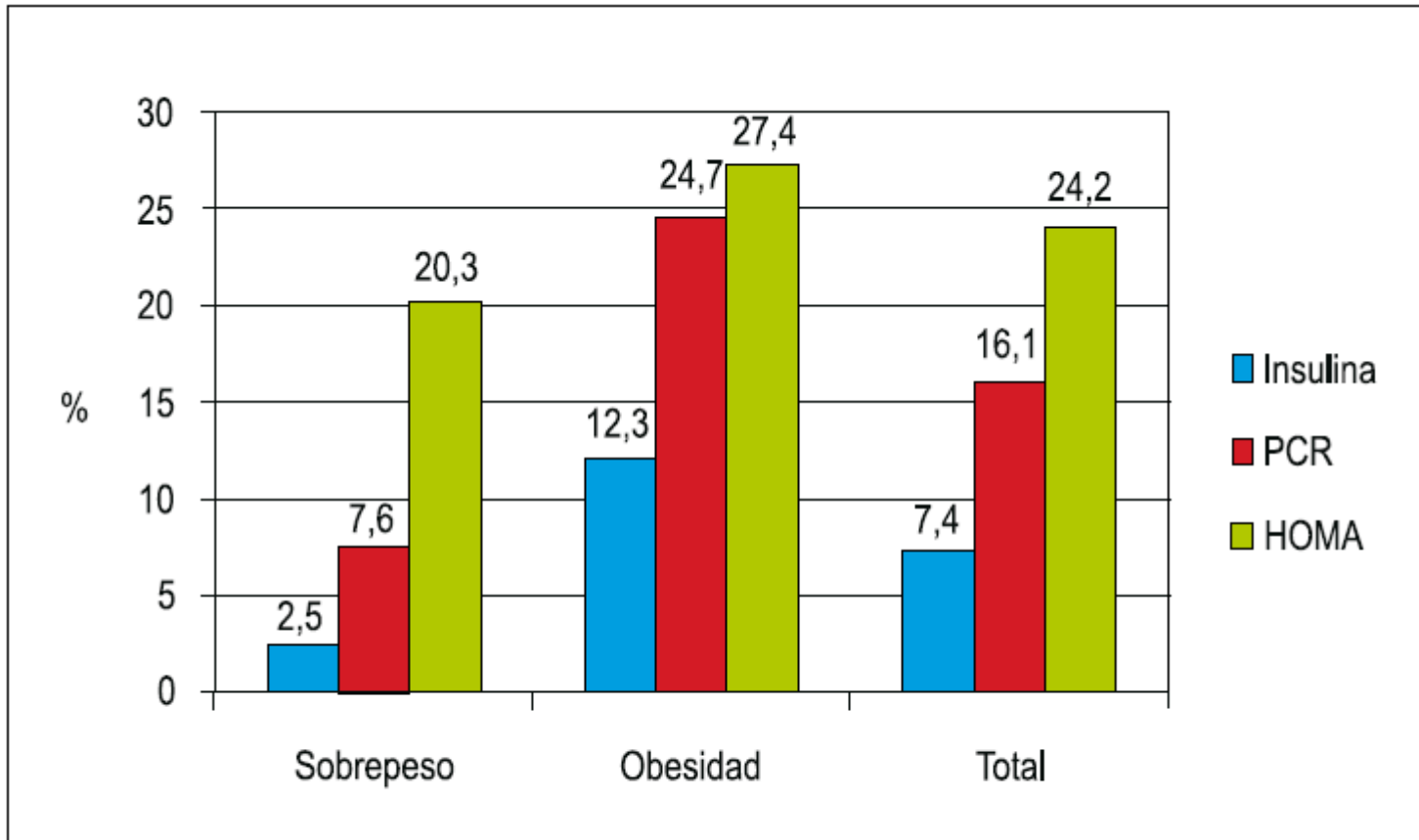
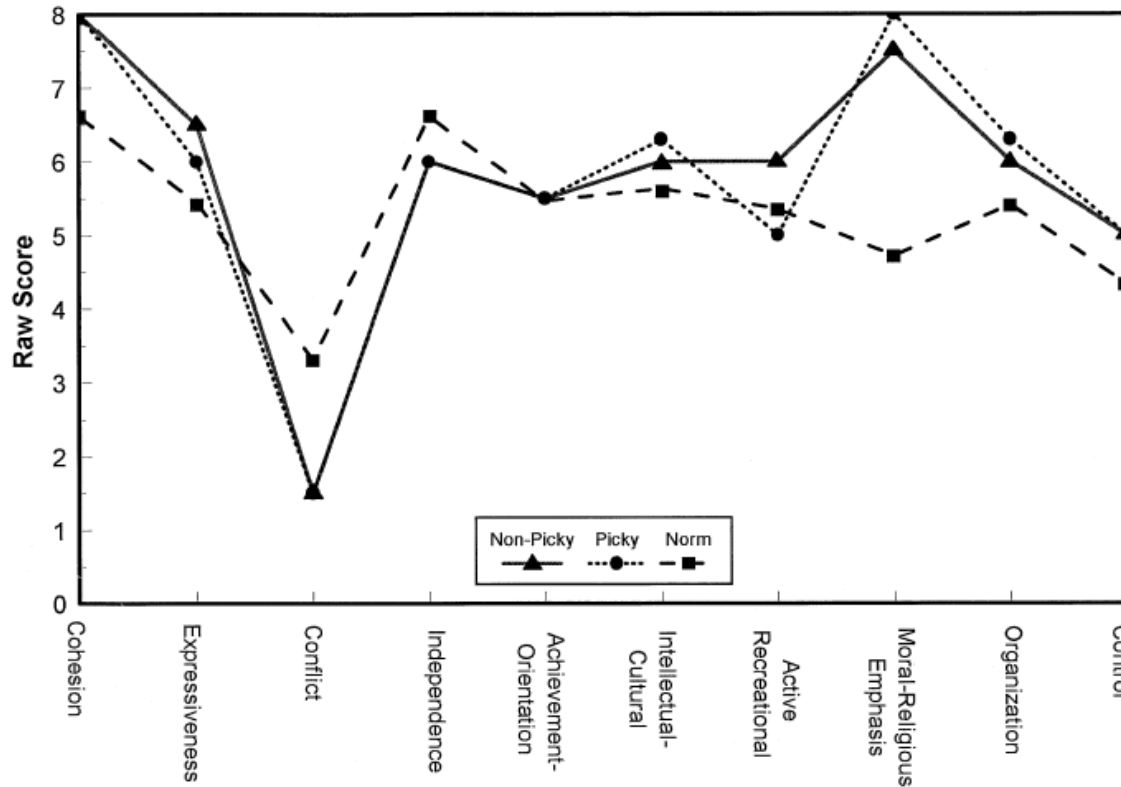


Figura 1. Porcentaje de adolescentes de acuerdo a su estado nutricional, que se encuentran por encima de lo considerado normal para las variables bioquímicas.

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES



**Fig. 1.** Comparison of normative mean scores for Family Environment Scales (FES) and mean scores for toddlers (24 to 36 months) perceived as picky eaters (n=47) and non-picky eaters (n=74). A higher raw score denotes a stronger family characteristic than a lower score (0 to 8).

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

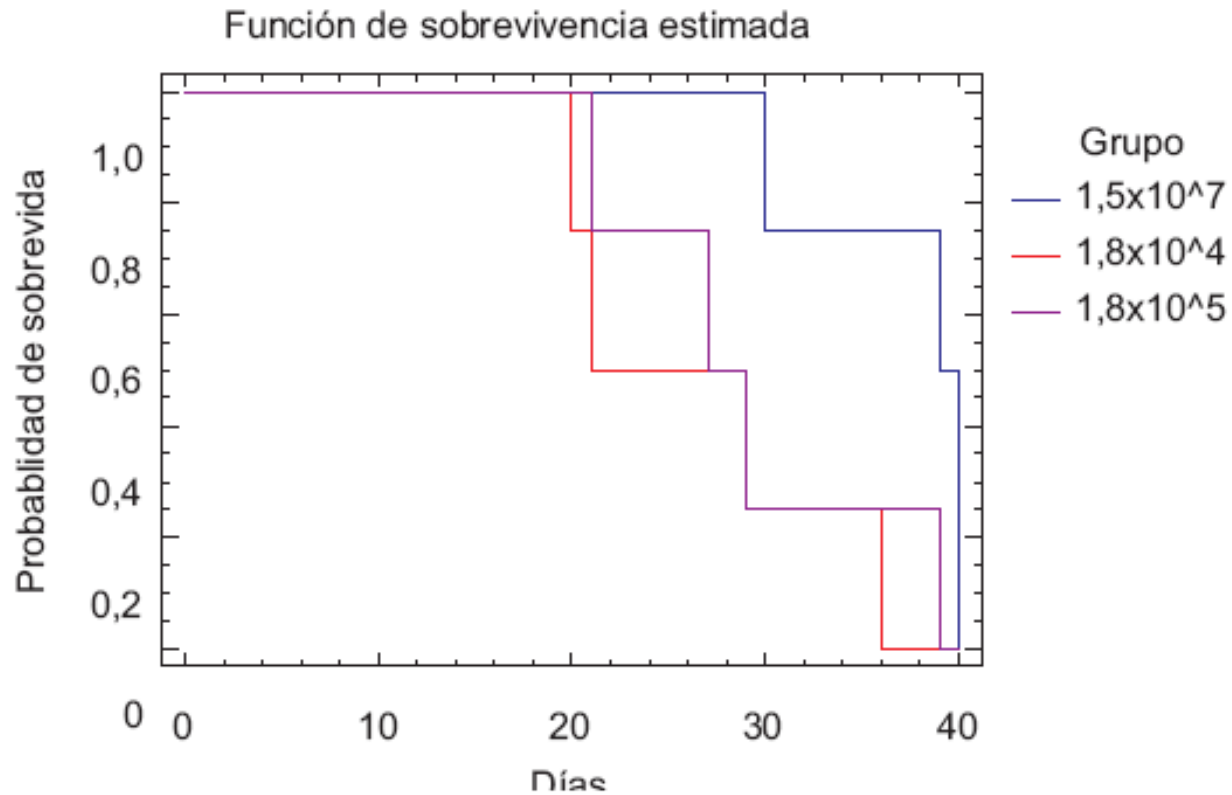


Figura 2. Supervivencia de los ratones inoculados con *T. cruzi* de acuerdo a la dosis y forma parasitaria.



# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

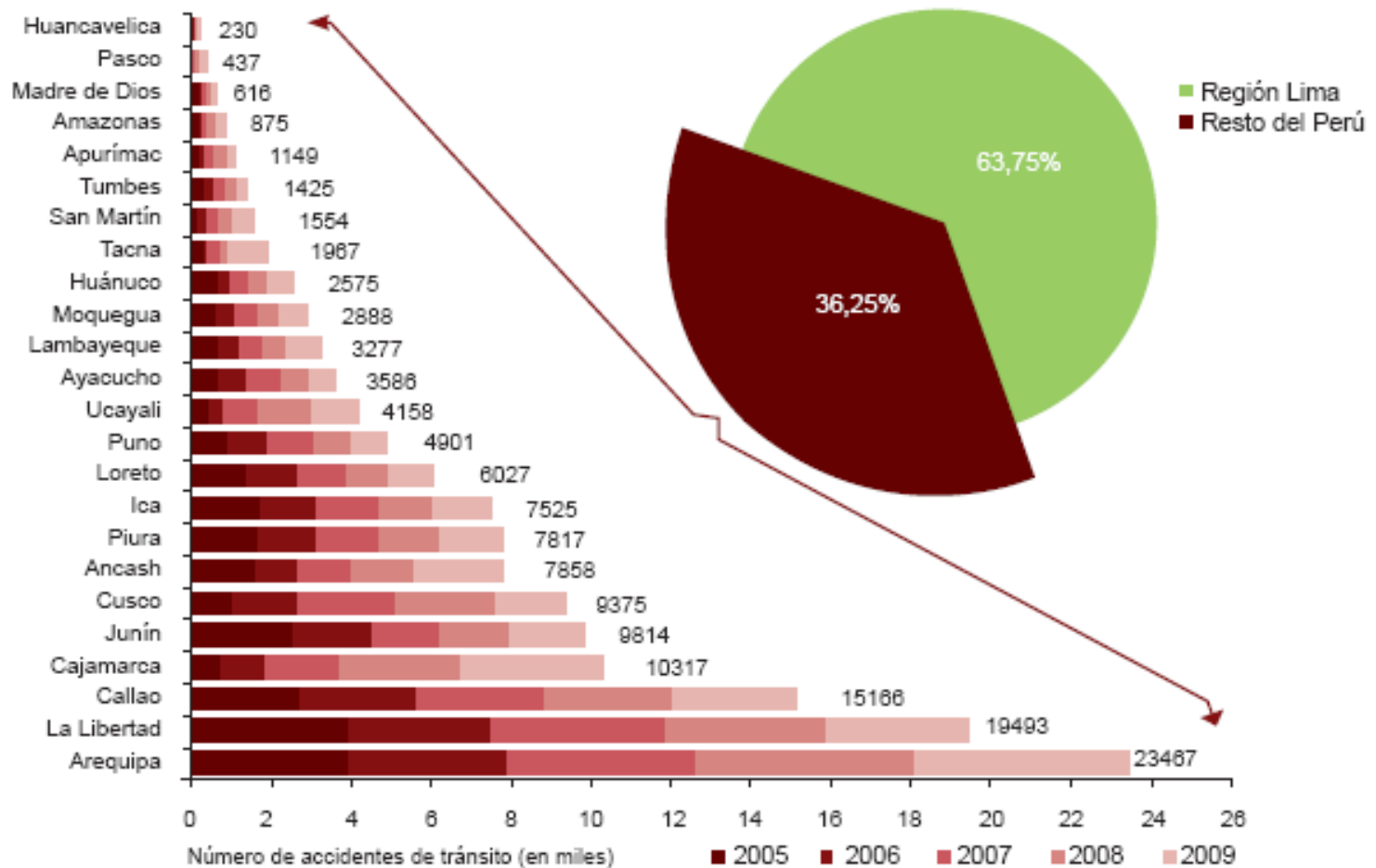


Figura 3. Accidentes de tránsito por regiones en el Perú, 2005-2009.

Fuente: Dirección de Estadística del Estado Mayor de la Policía Nacional del Perú.

# GRÁFICOS ESTADÍSTICOS – ERRORES FRECUENTES

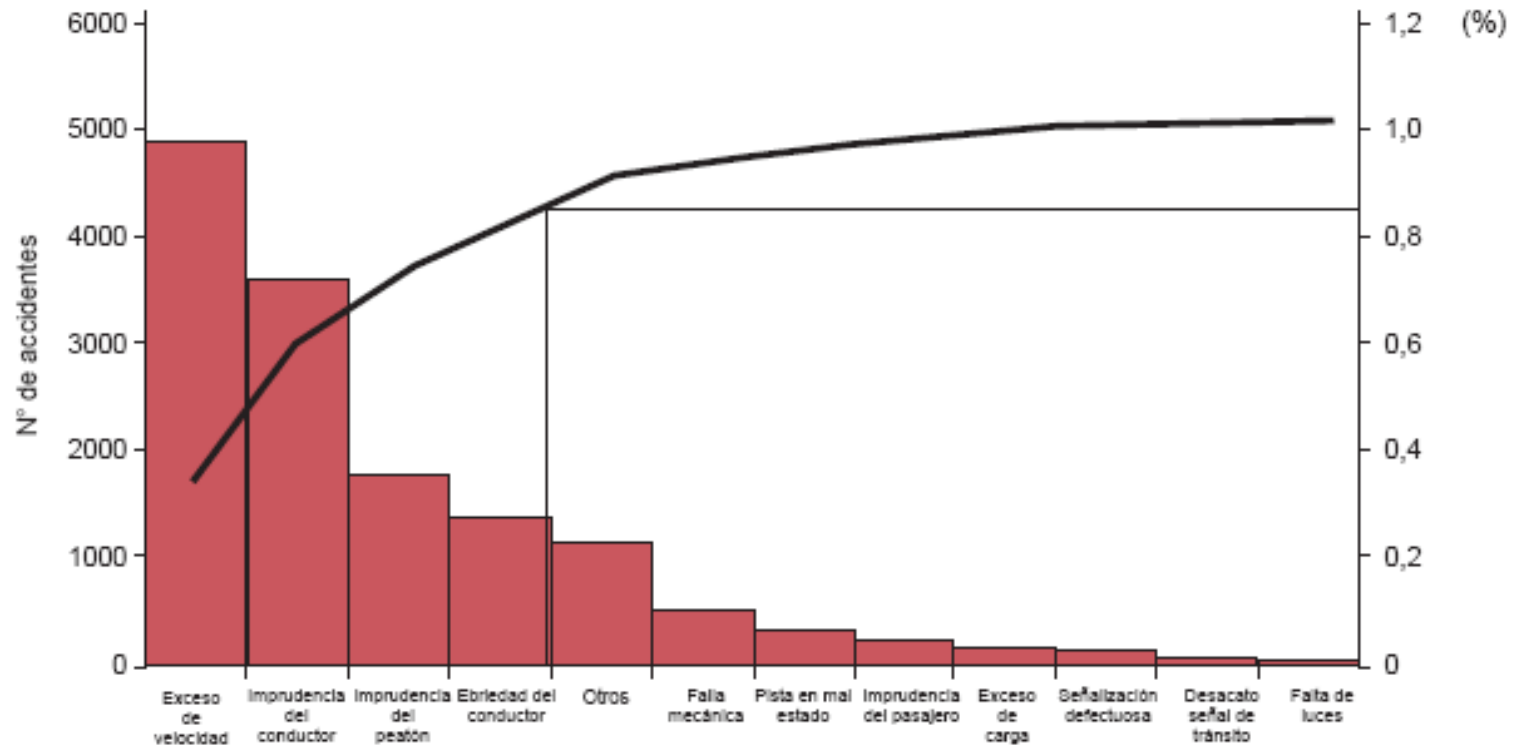
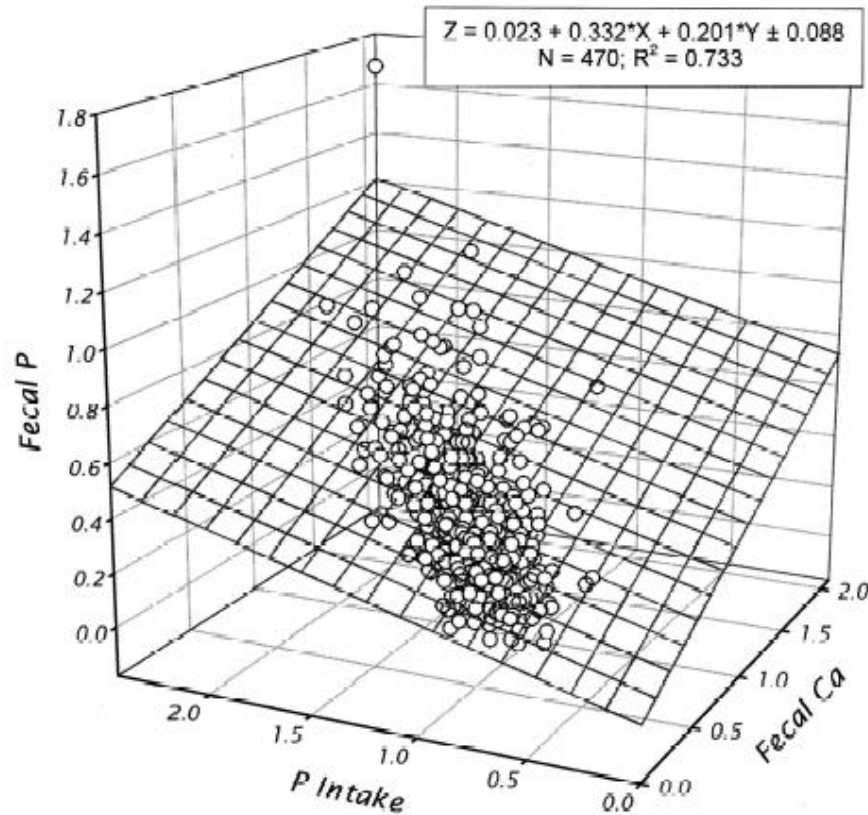


Figura 2. Causas de los accidentes de tránsito en el Perú, 2005-2009.

Fuente: Dirección de Estadística del Estado Mayor de la Policía Nacional del Perú.



**Fig. 1.** Three dimensional plot of fecal phosphorus as a function of diet phosphorus and fecal calcium for the 470 balances of data set 1. The diagonal plane through the data is the least squares fit according to the parameters set forth in Table 2. In the equation inset,  $Z = \text{fecal P}$ ,  $X = \text{fecal Ca}$ , and  $Y = \text{diet P}$ . (Copyright Robert P. Heaney, 2002. Used with permission.)

***¿Es este gráfico fácil de entender?***

