Hospital Universitario Ramón y Cajal

SaludMadrid Unidad de Bioestadística Clínica

Conceptos básicos de estadística para clínicos

Víctor Abraira

A Coruña. Noviembre 2008

Programa

- “Valor p”: ¿qué significa?
- Aleatorización
- Pruebas de hipótesis para variables continuas
- Pruebas de hipótesis para variables categóricas
- Tiempo hasta un evento

Valor *p*

Articles

Randomised trial of a perindopril-based blood-pressure-lowering regimen among 6105 individuals with previous stroke or transient ischaemic attack

*PROGRESS Collaborative Group**

Summary

Background Blood pressure is a determinant of the risk of stroke among both hypertensive and non-hypertensive individuals with cerebrovascular disease. However, there is uncertainty about the efficacy and safety of blood-pressure-lowering treatments for many such patients. The perindopril protection against recurrent stroke study (PROGRESS) was designed to determine the effects of a blood-pressure-lowering regimen in hypertensive and non-hypertensive patients with a history of stroke or transient ischaemic attack.

Methods 6105 individuals from 172 centres in Asia, Australasia, and Europe were randomly assigned active treatment (n=3051) or placebo (n=3054). Active treatment comprised a flexible regimen based on the angiotensin-converting-enzyme inhibitor perindopril (4 mg daily), with the addition of the diuretic indapamide at the discretion of treating physicians. The primary outcome was total stroke (fatal or non-fatal). Analysis was by intention to treat.

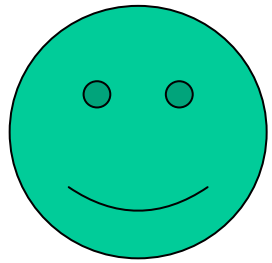
Findings Over 4 years of follow up, active treatment reduced blood pressure by 9/4 mm Hg. 307 (10%) individuals assigned active treatment suffered a stroke, compared with 420 (14%) assigned placebo (relative risk reduction 28% [95% CI 17–38], p<0·0001). Active treatment also reduced the risk of total major vascular events (26% [16–34]). There were similar reductions in the risk of stroke in hypertensive and non-hypertensive subgroups (all p<0·01). Combination therapy with perindopril plus indapamide reduced blood pressure by 12/5 mm Hg and stroke risk by 43% (30–54). Single-drug therapy reduced blood pressure by 5/3 mm Hg and produced no discernable reduction in the risk of stroke.

Introduction

stroke or transient ischaemic attack. No treatment has been proven to reduce recurrent stroke risk among

¿Qué son?

- Reducción relativa del riesgo (RRR)
- Intervalo de confianza (CI)
- Valor *p*



la probabilidad de encontrar una diferencia en las proporciones de recurrencia de ictus como la que se ha encontrado, o mayor, en la hipótesis, llamada hipótesis nula, de que el tratamiento no tenga efecto.

Ejemplo: La dama y las tazas de té

- Supóngase que una dama asegura que es capaz de distinguir entre dos formas de preparar el té
- Experimento: se preparan 8 tazas, 4 de cada tipo, la dama lo sabe, y se le presentan al azar
- H_0: la dama no distingue
- Resultado: Señala todas correctamente.
- $C_{8;4}=70$ ➔ $p=1/70=0{,}014$

Ejemplo: La dama y las tazas de té

- Si se hubieran preparado 4 tazas; 2 y 2
- $C_{4;2} = 6$ ➔ $p = 1/6 = 0{,}17$
- No sería un experimento que aportara suficiente carga contra H_0 aunque acertara todas
- Notar que el cálculo de *p* depende del diseño

Ejemplo: La dama y las tazas de té

- Si la dama no supiera cuantas tazas se preparan de cada tipo, sólo que hay, o puede haber, de los dos tipos.
- La variable número de aciertos en H_0 es $Bin(n=8, p=0,5)$

$$p = f(8) = \binom{8}{8}(0,5)^8(0,5)^0 = 0,0039$$

Críticas

A Randomized, Controlled Trial of the Effects of Remote, Intercessory Prayer on Outcomes in Patients Admitted to the Coronary Care Unit

William S. Harris, PhD; Manohar Gowda, MD; Jerry W. Kolb, MDiv; Christopher P. Strychacz, PhD; James L. Vacek, MD; Philip G. Jones, MS; Alan Forker, MD; James H. O'Keefe, MD; Ben D. McCallister, MD

Context: Intercessory prayer (praying for others) has been a common response to sickness for millennia, but it has received little scientific attention. The positive findings of a previous controlled trial of intercessory prayer have yet to be replicated.

Objective: To determine whether remote, intercessory prayer for hospitalized, cardiac patients will reduce overall adverse events and length of stay.

Design: Randomized, controlled, double-blind, prospective, parallel-group trial.

Setting: Private, university-associated hospital.

Patients: Nine hundred ninety consecutive patients who were newly admitted to the coronary care unit (CCU).

Intervention: At the time of admission, patients were randomized to receive remote, intercessory prayer (prayer group) or not (usual care group). The first names of patients in the prayer group were given to a team of outside intercessors who prayed for them daily for 4 weeks. Patients were unaware that they were being prayed for, and the intercessors did not know and never met the patients.

Main Outcome Measures: The medical course from CCU admission to hospital discharge was summarized in a CCU course score derived from blinded, retrospective chart review.

Results: Compared with the usual care group (n = 524), the prayer group (n = 466) had lower mean ± SEM weighted (6.35 ± 0.26 vs 7.13 ± 0.27; $P = .04$) and unweighted (2.7 ± 0.1 vs 3.0 ± 0.1; $P = .04$) CCU course scores. Lengths of CCU and hospital stays were not different.

Conclusions: Remote, intercessory prayer was associated with lower CCU course scores. This result suggests that prayer may be an effective adjunct to standard medical care.

Arch Intern Med. 1999;159:2273-2278

5

a

Alternativas

- Hacer estimaciones con IC
  - Misma base conceptual

    Feinstein AR. P-values and confidence intervals: two sides of the same unsatisfactory coin. J Clin Epidemiol 1998; 51:355-360.

    Abraira V. Estimación: intervalos de confianza. SEMERGEN 2002; 28:84-85.

- Alternativa Bayesiana

  Abraira V. Inferencia estadística bayesiana. SEMERGEN 2005; 31:18-20.

Aleatorización

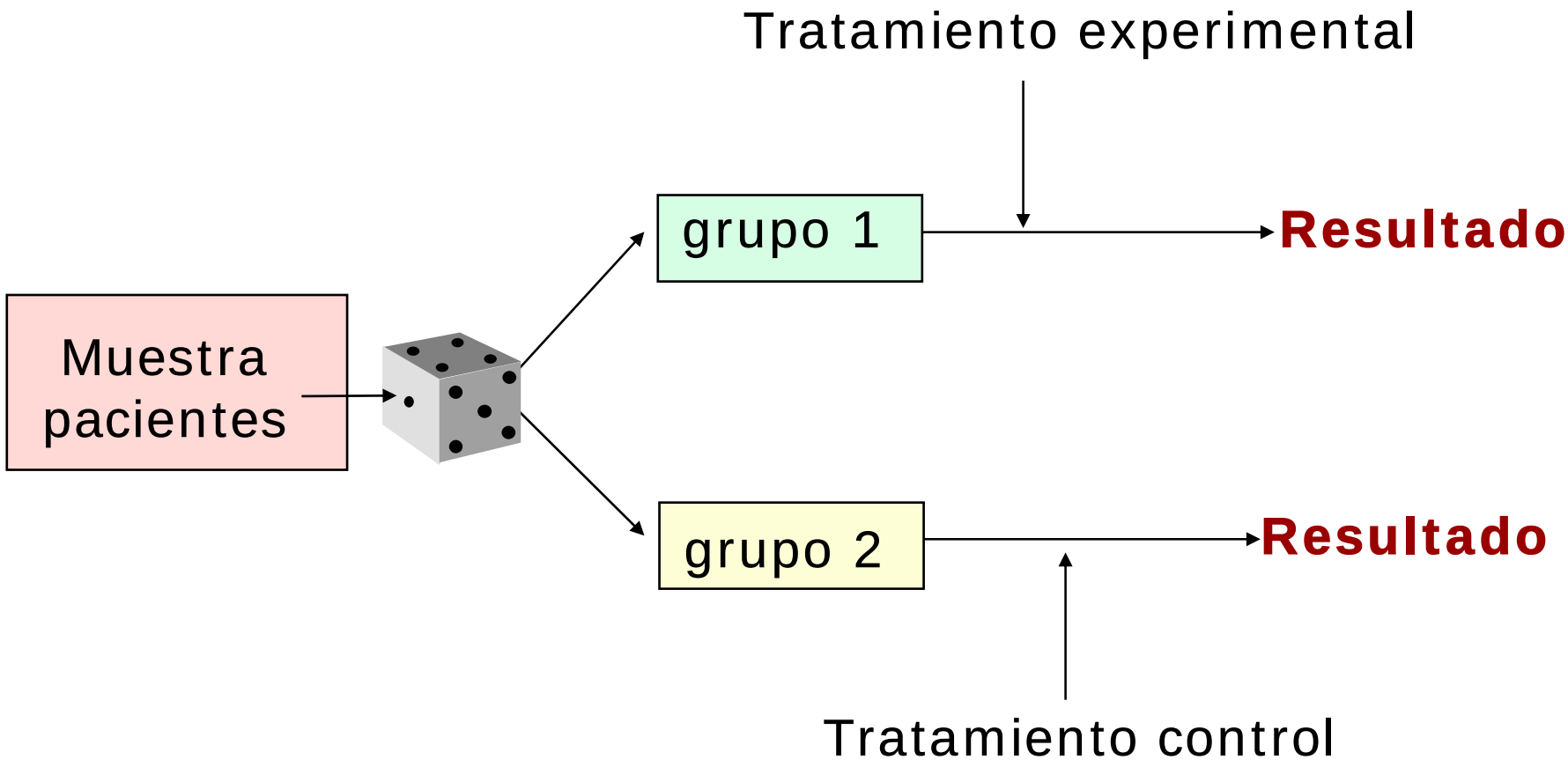
- Dos conceptos:
  - Muestra aleatoria (afecta a toda la inferencia). Es necesaria para:
    - muestra sea representativa
    - cuantificar incertidumbre muestreo
  - Asignación aleatoria de tratamientos (afecta a los EECC). Necesaria para igualar el pronóstico de los pacientes.

Muestra aleatoria

- Aquella en que todos sus elementos tienen la misma probabilidad y cada uno está independientemente de los demás.
- Cada muestra es independiente de las demás y con la misma probabilidad.
- En la práctica rara vez se dispone de muestras aleatorias, por la tanto la situación habitual es

Muestra aleatoria

Diseño ensayo clínico



¿Cómo se calcula la p?

Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety

Chris T Bolliger, Jean-Pierre Zellweger, Tobias Danielsson, Xandra van Biljon, Annik Robidou, Åke Westin, André P Perruchoud, Urbain Säwe

Abstract

Objectives To determine whether use of an oral nicotine inhaler can result in long term reduction in smoking and whether concomitant use of nicotine replacement and smoking is safe.
Design Double blind, randomised, placebo controlled trial. Four month trial with a two year follow up.
Setting Two university hospital pulmonary clinics in Switzerland.
Participants 400 healthy volunteers, recruited through newspaper advertisements, willing to reduce their smoking but unable or unwilling to stop smoking immediately.
Intervention Active or placebo inhaler as needed for up to 18 months, with participants encouraged to limit their smoking as much as possible.

Results At four months sustained reduction of smoking was achieved in 52 (26%) participants in the active group and 18 (9%) in the placebo group ($P<0.001$; Fisher's test). Corresponding figures after two years were 19 (9.5%) and 6 (3.0%) ($P=0.012$).
Conclusion Nicotine inhalers effectively and safely

time, plus the fact that many smokers try to quit several times before succeeding, new treatment approaches are clearly needed. One such strategy could be to reduce tobacco consumption substantially in smokers who are unwilling or unable to quit right away. For such smokers, sustained reduction might reduce the

University Medical Policlinic, CHUV, 1000 Lausanne, Switzerland
Jean-Pierre Zellweger
lecturer

Cumulative event curves were estimated with the Kaplan-Meier procedure, and the effects of treatment on the primary and secondary endpoints were estimated from unadjusted Cox's proportional hazards models. Among participants who had more than one outcome

PROGRESS

¿Cómo se calcula la p?

- Gran variedad de métodos (los llamados *tests* estadísticos) dependiendo de la hipótesis, del diseño, del tipo de variables, etc.
- Un amplio grupo son aquellos en que la hipótesis es la asociación entre dos variables:
  - Ambas categóricas
  - Una categórica y otra continuas
    - Tiempo a un evento
  - Ambas continuas

Dos variables categóricas

Results At four months sustained reduction of smoking was achieved in 52 (26%) participants in the active group and 18 (9%) in the placebo group ($P<0.001$; Fisher's test). Corresponding figures after two years were 19 (9.5%) and 6 (3.0%) ($P=0.012$). **Conclusion** Nicotine inhalers effectively and safely

BMJ 2000;**321**:329-333

Findings Over 4 years of follow up, active treatment reduced blood pressure by 9/4 mm Hg. 307 (10%) individuals assigned active treatment suffered a stroke, compared with 420 (14%) assigned placebo (relative risk reduction 28% [95% CI 17–38], p<0·0001). Active treatment also reduced the risk of total major vascular events (26% [16–34]). There were similar reductions in the risk of stroke in hypertensive and non-hypertensive subgroups (all p<0·01). Combination therapy with perindopril plus indapamide reduced blood pressure by 12/5 mm Hg and stroke risk by 43% (30–54). Single-drug therapy reduced blood pressure by 5/3 mm Hg and produced no discernable reduction in the risk of stroke.

Lancet 2001;**358**:1033-1041

Contrastes sobre diferencia de proporciones

- H_0: $p_1 = p_2 = p_0$
- El estadístico para el contraste es (si se puede aproximar a la normal)

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} : \quad Nor(0,1)$$

- O, equivalentemente, la prueba ji-cuadrado
- O la prueba de Fisher si no se puede aproximar a la normal

Ejemplo

- Resultados del ensayo del BMJ:
  - 200 participantes en cada grupo; 52 con éxito en el grupo tratado y 18 en el placebo

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \hat{p}_1 = \frac{52}{200}; \quad \hat{p}_1 = \frac{18}{200}; \quad \hat{p}_0 = ??$$

$$\hat{p}_0 = \frac{70}{400} \qquad Z = \frac{0,26 - 0,09}{\sqrt{0,175 \cdot 0,825\left(\frac{1}{200} + \frac{1}{200}\right)}} = 4,474$$

$$p = 0,0000038$$

Una categórica y otra continua

A Randomized, Controlled Trial of the Effects of Remote, Intercessory Prayer on Outcomes in Patients Admitted to the Coronary Care Unit

William S. Harris, PhD; Manohar Gowda, MD; Jerry W. Kolb, MDiv; Christopher P. Strychacz, PhD; James L. Vacek, MD; Philip G. Jones, MS; Alan Forker, MD; James H. O'Keefe, MD; Ben D. McCallister, MD

Context: Intercessory prayer (praying for others) has been a common response to sickness for millennia, but it has received little scientific attention. The positive findings of a previous controlled trial of intercessory prayer have yet to be replicated.

Objective: To determine whether remote, intercessory prayer for hospitalized, cardiac patients will reduce overall adverse events and length of stay.

Design: Randomized, controlled, double-blind, prospective, parallel-group trial.

Setting: Private, university-associated hospital.

Patients: Nine hundred ninety consecutive patients who were newly admitted to the coronary care unit (CCU).

Intervention: At the time of admission, patients were randomized to receive remote, intercessory prayer (prayer group) or not (usual care group). The first names of patients in the prayer group were given to a team of outside intercessors who prayed for them daily for 4 weeks. Patients were unaware that they were being prayed for, and the intercessors did not know and never met the patients.

Main Outcome Measures: The medical course from CCU admission to hospital discharge was summarized in a CCU course score derived from blinded, retrospective chart review.

Results: Compared with the usual care group (n = 524), the prayer group (n = 466) had lower mean ± SEM weighted (6.35 ± 0.26 vs 7.13 ± 0.27; $P = .04$) and unweighted (2.7 ± 0.1 vs 3.0 ± 0.1; $P = .04$) CCU course scores. Lengths of CCU and hospital stays were not different.

Conclusions: Remote, intercessory prayer was associated with lower CCU course scores. This result suggests that prayer may be an effective adjunct to standard medical care.

Arch Intern Med. 1999;159:2273-2278

Estadísticos para los contrastes sobre diferencias de medias

- H_0: $\mu_1 = \mu_2$
- Los estadísticos para el contraste son (si las variables son normales)

Varianzas conocidas

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} : \; Nor(0,1)$$

Varianzas desconocidas

- desiguales

$$T = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} : \; t_n$$

- iguales

$$T = \frac{(\bar{X}_1 - \bar{X}_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} : \; t_{n_1+n_2-2}$$

En el ejemplo

$$T = \frac{2,7 - 3}{\sqrt{0,1^2 + 0,1^2}} = -2,12 \qquad p = 0,017$$

Tiempo al evento

A predictive model for mortality of bloodstream infections: Bedside analysis with the Weibull function

Emilio Casariego Vales[a,*], Victor Abraira[b], Juan Carlos Corredoira Sánchez[a], María Pilar Alonso García[c], Asunción Rodríguez Feijoo[c], María José López Alvarez[a], José Varela Otero[a], Amparo Coira Nieto[c], Ramón Rabuñal Rey[a], María Teresa Rigueiro Veloso[a]

[a]*Department of Internal Medicine, Hospital Xeral-Calde, C/ Severo Ochoa, s/n. 27004, Lugo, Spain*
[b]*Biostatistics Unit, Hospital Ramón y Cajal, Crta. de Colmenar, Km. 9,100, 28034 Madrid, Spain*
[c]*Microbiology Unit, Hospital Xeral-Calde, C/ Severo Ochoa, s/n. 27004, Lugo, Spain*

Received 29 October 1999; received in revised form 5 December 2001; accepted 12 December 2001

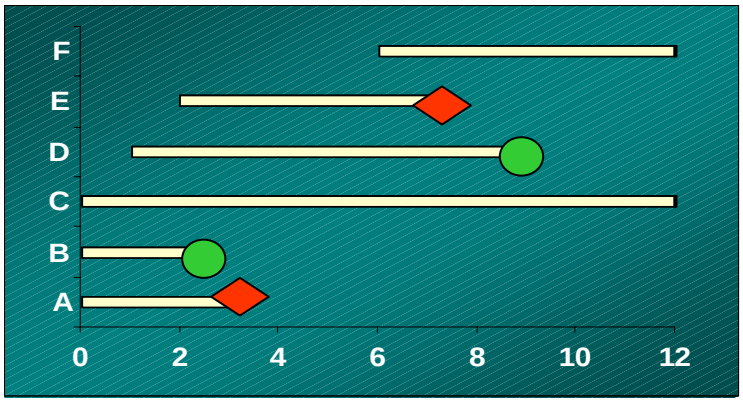
Abstract

This paper describes the construction and validation of a prognostic model for predicting post-bloodstream infection survival up to Day 21. A Weibull multiple regression model was adopted in a prospective cohort study of all patients diagnosed with true bacteremia or fungemia in a teaching hospital between 1991 and 1994 (training set, 1,577 patients). The final model included six variables easily detected in any institution: source of infection, underlying neoplasm, septic shock, community-acquired, age over 65, and polymicrobial bacteremia. Using this model, it is possible to obtain a graphic representation of survival probability for any combination of these risk factors. The model was tested on a second set of patients diagnosed in the same hospital between 1996 and 1997 (validation set, 952 patients), confirming its reliability in predicting survival. In conclusion, the Weibull function, together with variables easily identified at bedside, enables a precise prediction of the short-term, post-bloodstream infection mortality of a given patient.

Keywords: Bloodstream infection; Septicemia; Survival analysis; Mortality; Prognosis, Weibull function

¿Por qué distinto?

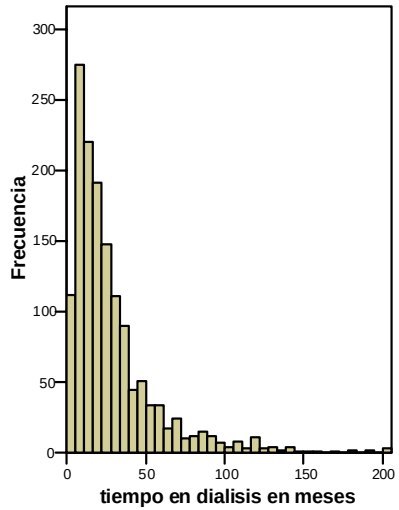
- Problemas con la medición:



Desconocido en Pérdidas

No eventos (evento competitivo)

} Censuras

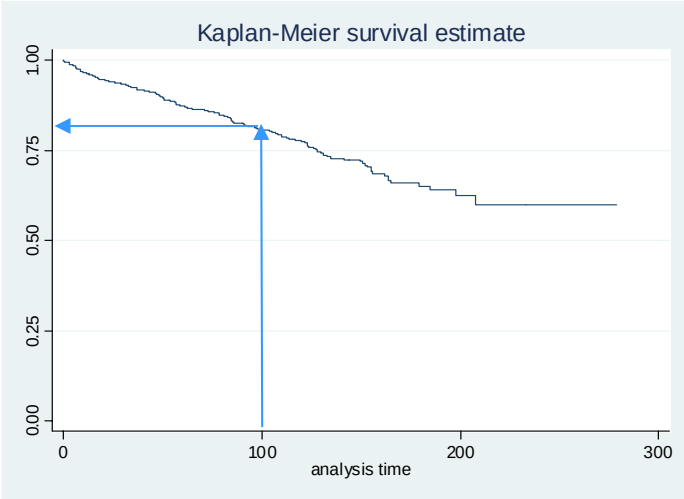


En general, no normal

¿Cómo se resume?

- La variable tiempo de espera, usando la información parcial de las censuras, se resume mediante:
  - la función de supervivencia *S(t)*

Función de supervivencia



- *S(t)* es la probabilidad de que, en un individuo, el evento ocurra en un tiempo igual o mayor que *t* (si el evento es muerte, sobreviva al menos *t*).
- P.e. *S(*100*)* es la probabilidad de que un individuo sobreviva 100 ó más meses.

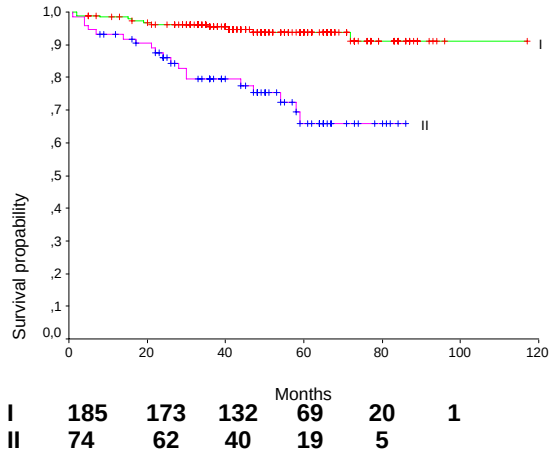
Puntos críticos

- El método asume que las “pérdidas” son al azar (no tienen información).
- Tiempo de seguimiento. Lo mide:
  - Mediana de seguimiento (más habitual)
  - Mediana de “pérdidas”

Estabilidad de la gráfica

- Las gráficas de K-M deberían incluir algún índice de la precisión de las estimaciones:
  - Intervalos de confianza
  - Tablas de individuos en r

1A



Eje tiempo hasta el 10% de los individuos

Pocock, Clayton, Altman. *Lancet* 2002: **359**:1686-89

A modo de resumen

<table>
<tr><td rowspan="2">VARIABLE INDEPENDIENTE (X)</td><td colspan="5">VARIABLE DEPENDIENTE (Y)</td></tr>
<tr><td>Categórica
(2 categorías)</td><td>Categórica
(más de
2 categorías)</td><td>Cuantitativa
(No Normal)</td><td>Cuantitativa
(Normal)</td><td>Tiempo al evento</td></tr>
<tr><td>Categórica
(2 categorías)</td><td>Comparación 2 proporciones o Chi-cuadrado o Fisher</td><td rowspan="2">Chi-cuadrado</td><td>Mann-Whitney</td><td>Comparación 2 medias - t de Student/</td><td rowspan="2">Kaplan-Meier y log rank</td></tr>
<tr><td>Categórica
(más de
2 categorías)</td><td>Chi-cuadrado</td><td>Kruskall-Wallis</td><td>ANOVA</td></tr>
<tr><td>Cuantitativa</td><td colspan="2">Regresión logística</td><td>Correlación Spearman</td><td>Correlación Pearson/
Regresión lineal</td><td>Regresión Cox</td></tr>
</table>