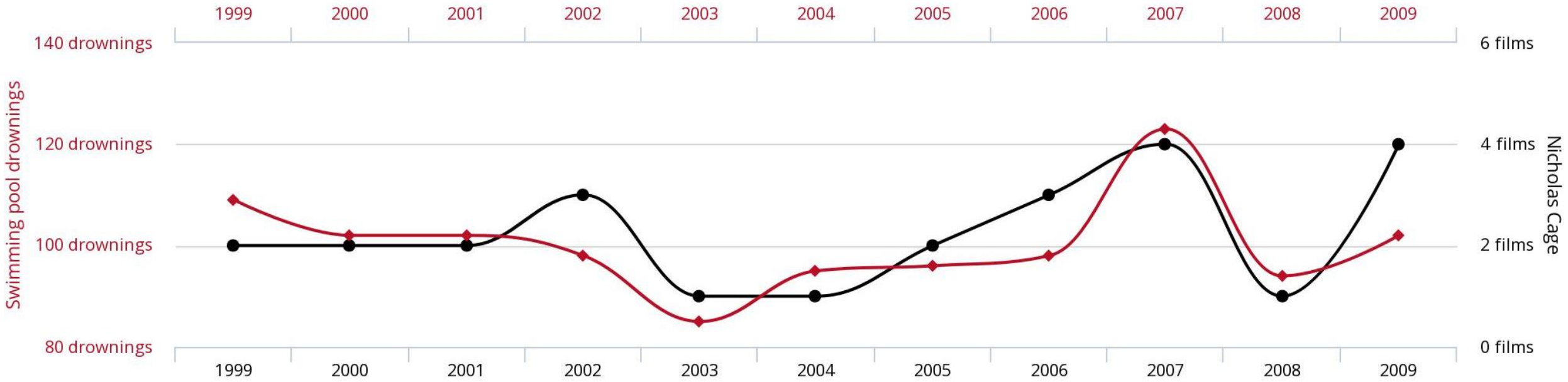
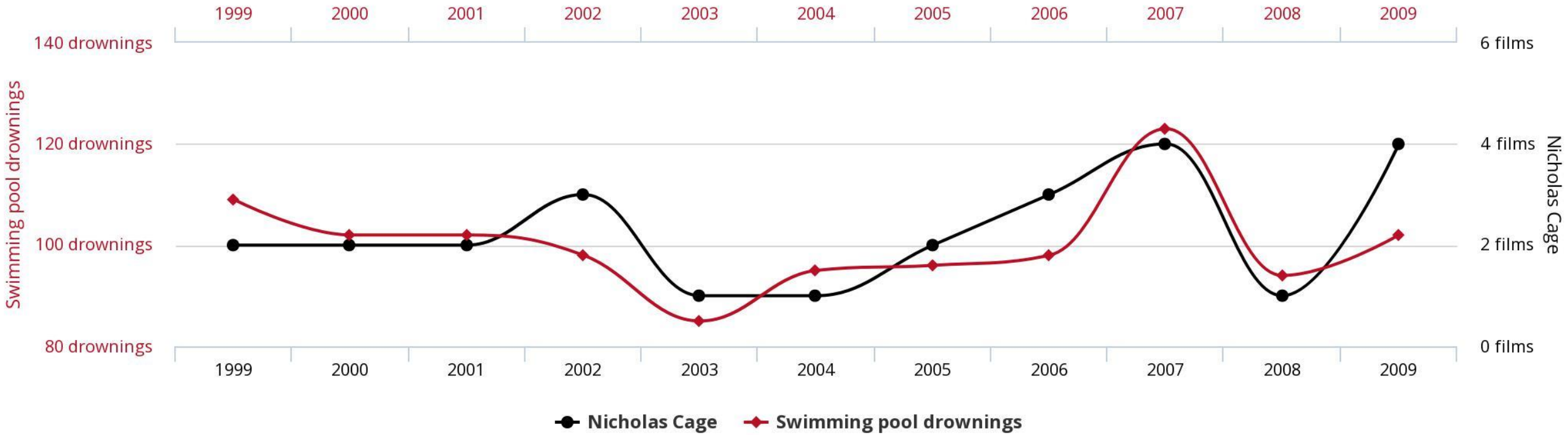
y "Dios creó los números" …¿qué te preocupa?

Martín Lázaro

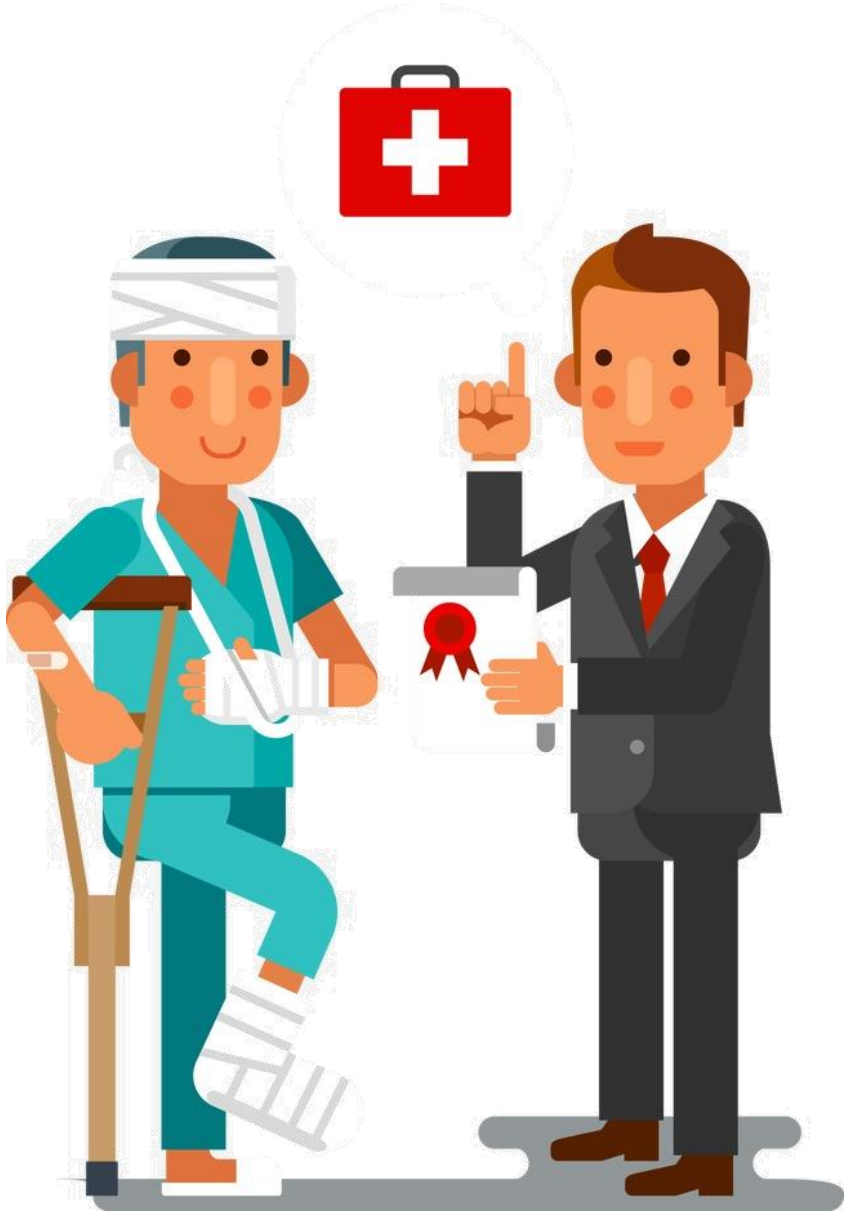
Oncología Médica, Complexo Hospitalario Universitario de Vigo


1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
140 drownings
120 drownings
100 drownings
80 drownings
Swimming pool drownings
6 films
4 films
2 films
0 films
Nicholas Cage

El número de personas que mueren ahogadas tras caer en una piscina
se correlaciona
con el número de películas en las que aparece Nicolas Cage



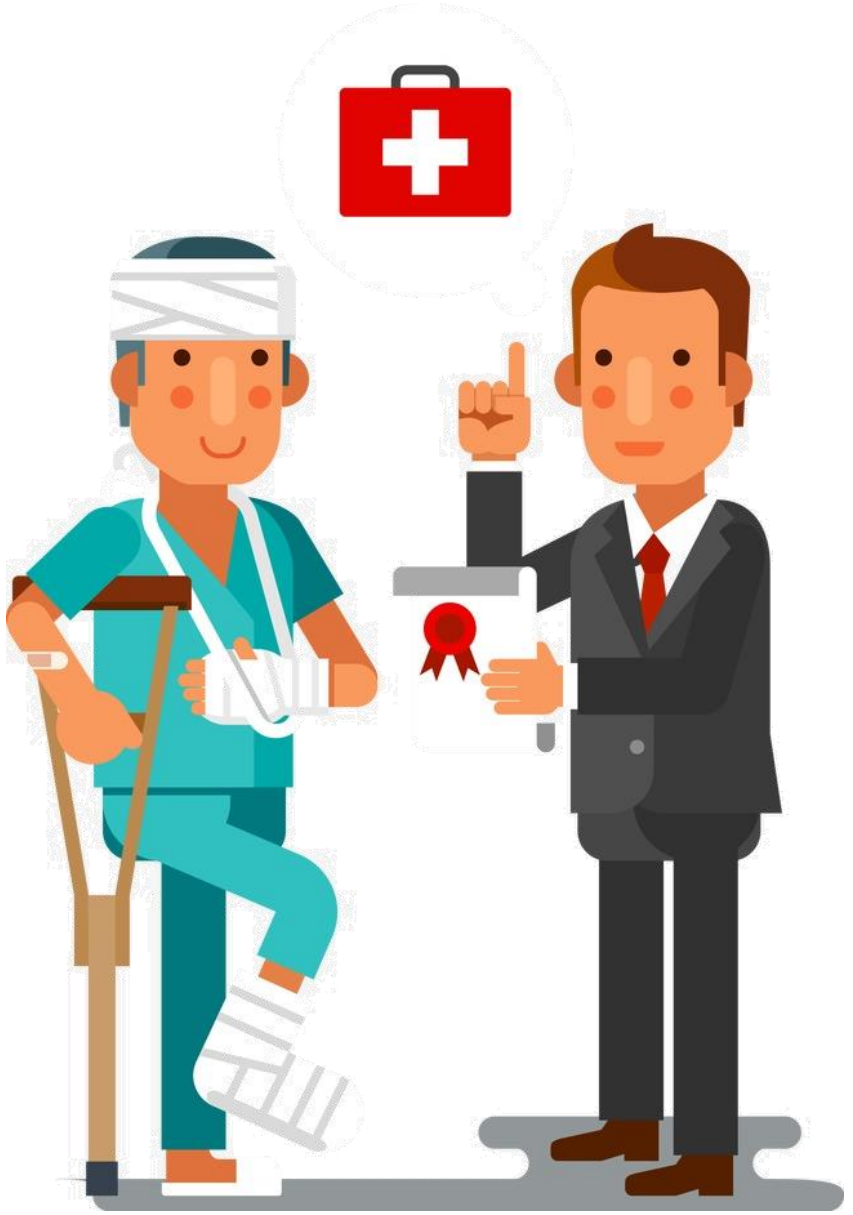
tylervigen.com



Cherny NI et al. Ann Oncol. 2015;26(8):1547-73

1

Vivan más



- Supervivencia global
- Supervivencia libre de enfermedad
- Supervivencia libre de progresión

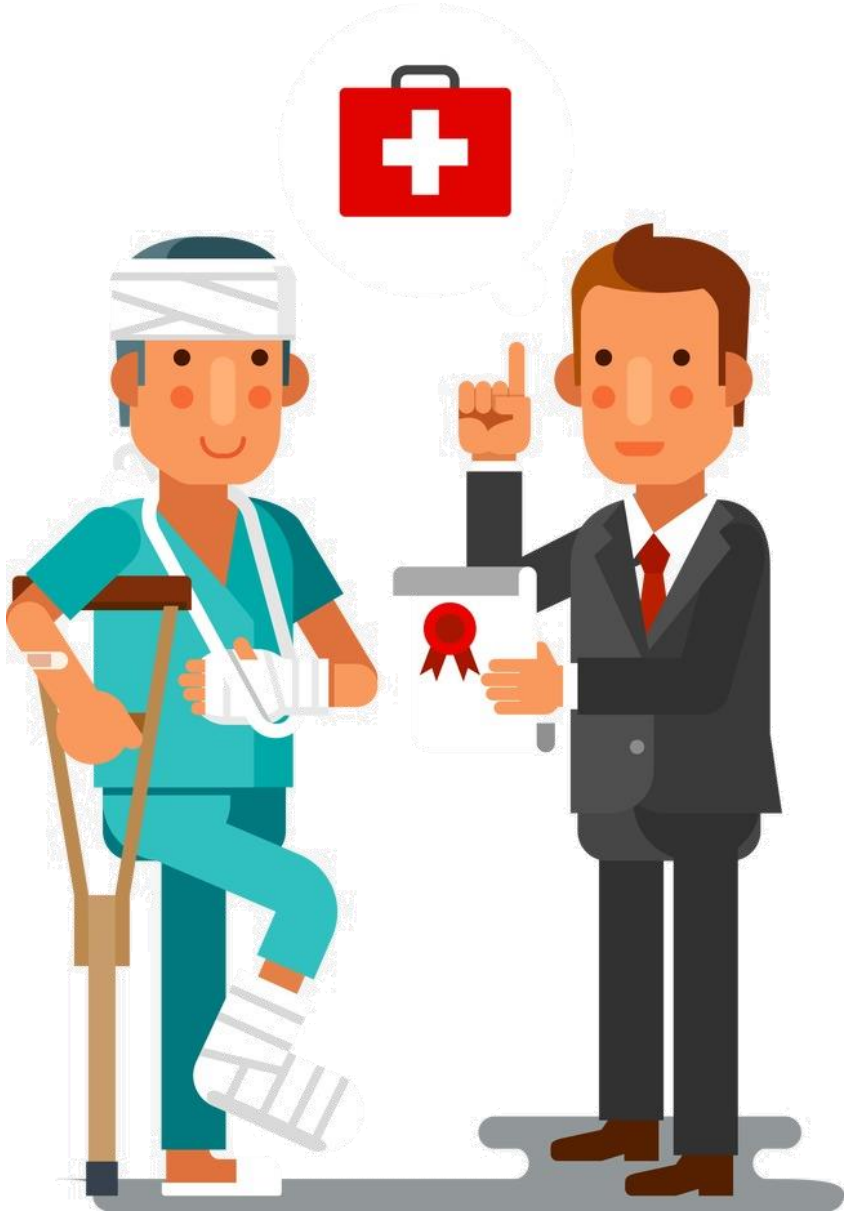
Cherny NI et al. Ann Oncol. 2015;26(8):1547-73

1 Vivan más

Supervivencia global

Supervivencia libre de enfermedad

Supervivencia libre de progresión



Supervivencia libre de progresión

Tasa de respuestas

Toxicidad

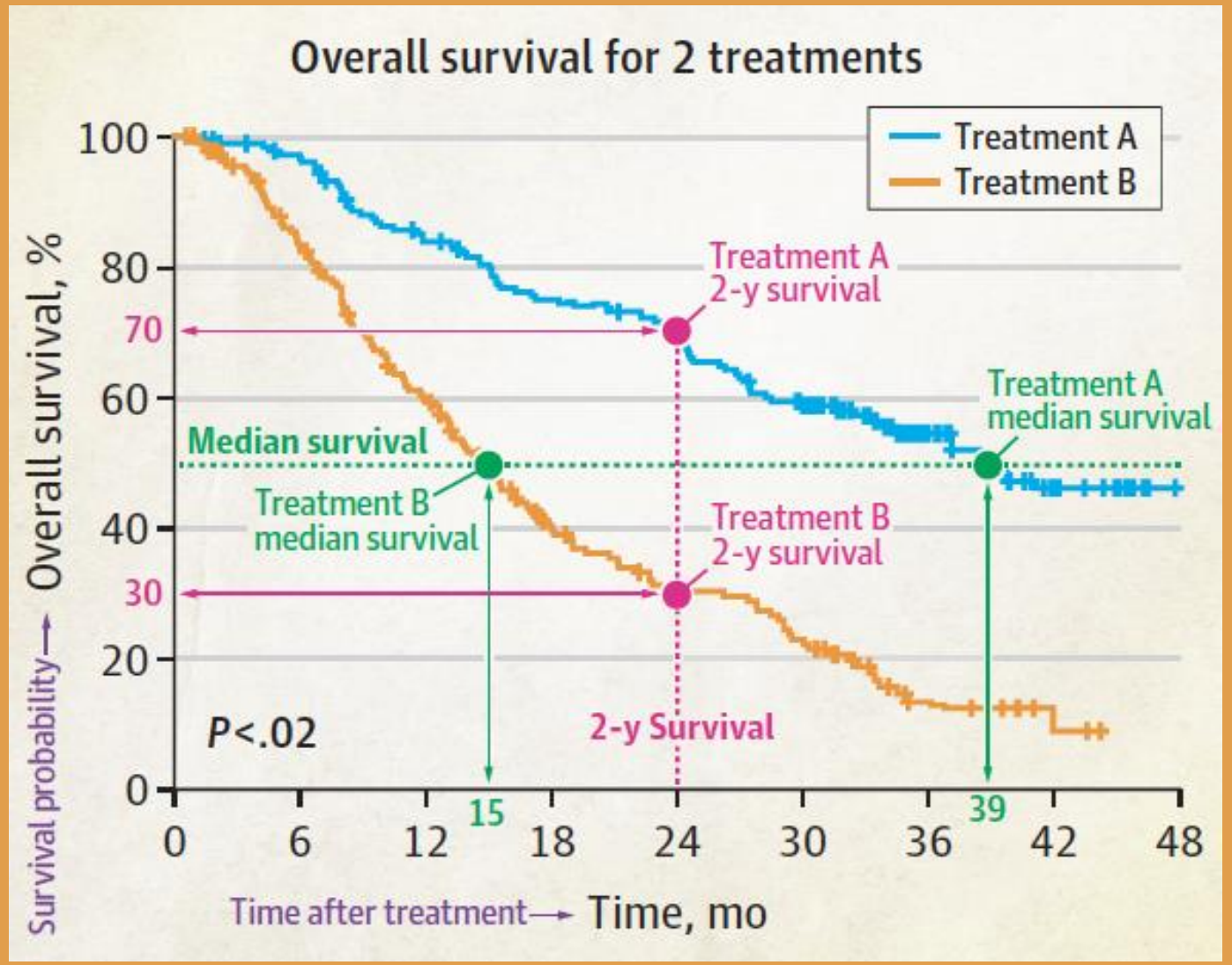
Vivan mejor 2

Cherny NI et al. Ann Oncol. 2015;26(8):1547-73

Es necesario valorar la respuesta de un tratamiento de una manera objetiva y estandarizar la forma de hacerlo para poder comparar los resultados de diferentes estudios

[Miller et al,1981]

¿Qué queremos medir?



1

SUPERVIVENCIA GLOBAL

"Periodo que transcurre desde la administración del tratamiento en estudio hasta el último control realizado o el fallecimiento del paciente"

1

SUPERVIVENCIA GLOBAL

“Periodo que transcurre desde la administración del tratamiento en estudio hasta el último control realizado o el fallecimiento del paciente”

SUPERVIVENCIA LIBRE DE PROGRESIÓN

2

“Tiempo transcurrido durante y después del tratamiento en el que el cáncer no progresa ni a nivel local ni a distancia”

1

SUPERVIVENCIA GLOBAL

"Periodo que transcurre desde la administración del tratamiento en estudio hasta el último control realizado o el fallecimiento del paciente"

SUPERVIVENCIA LIBRE DE PROGRESIÓN

2

"Tiempo transcurrido durante y después del tratamiento en el que el cáncer no progresa ni a nivel local ni a distancia"

3

TASA DE RESPUESTAS

"Relación directa entre reducción del tamaño tumoral y efecto del tratamiento. Imagen como herramienta"

VARIABLES *SUBROGADAS*

Variable correlacionada con la SG, o con cualquier otra variable, que sea clínicamente relevante para el paciente

75 % de los ensayos en oncología utilizan SLP-THP como variable final

>80% aprobaciones FDA-EMA basadas en variables intermedias

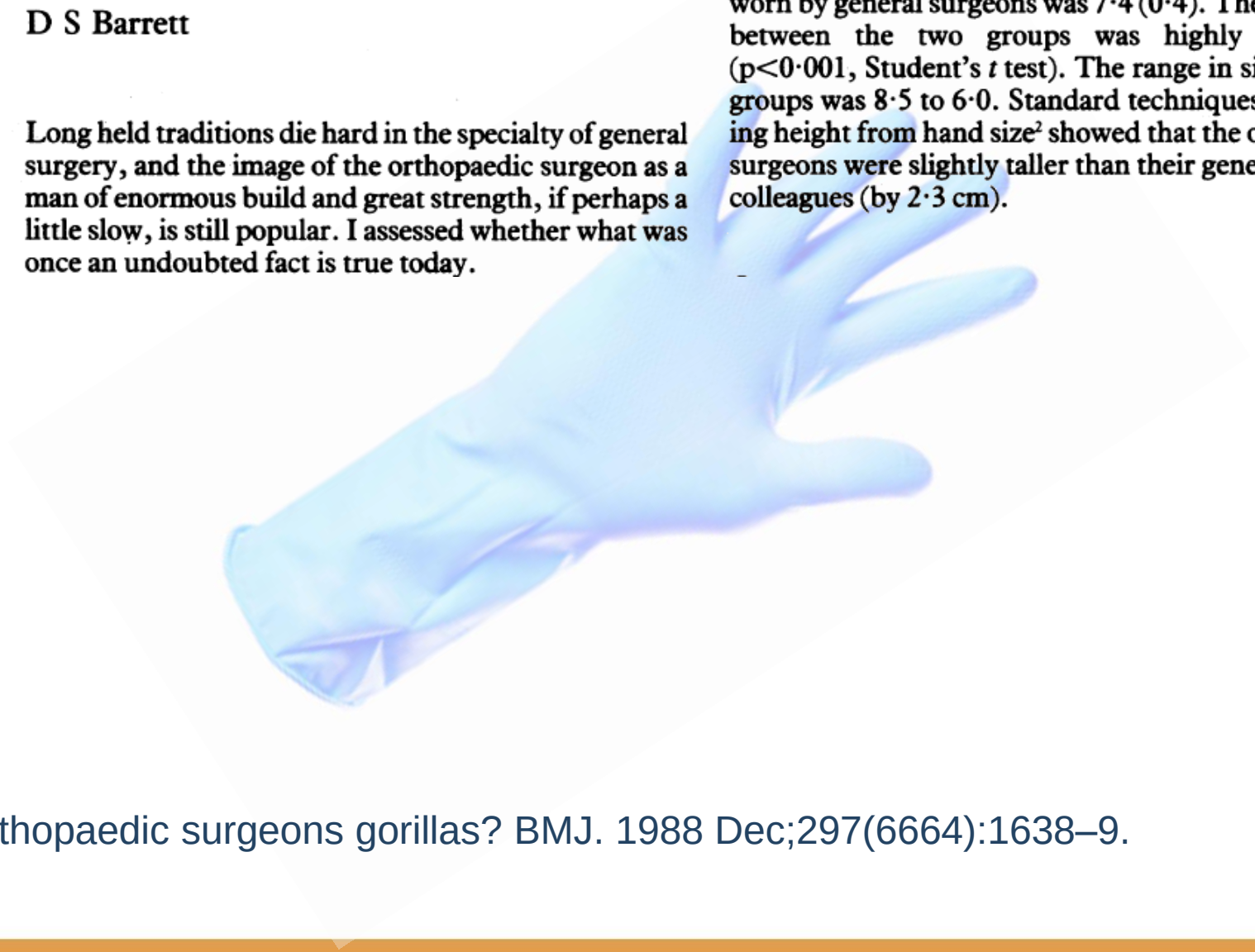
Are orthopaedic surgeons gorillas?

D S Barrett

Long held traditions die hard in the specialty of general surgery, and the image of the orthopaedic surgeon as a man of enormous build and great strength, if perhaps a little slow, is still popular. I assessed whether what was once an undoubted fact is true today.

orthopaedic surgeons. When the surgeons wore two pairs of gloves only the size of the inner glove was used.

The mean size of glove worn by orthopaedic surgeons was 7·6 (SD 0·4) while the mean size of gloves worn by general surgeons was 7·4 (0·4). The difference between the two groups was highly significant ($p<0·001$, Student's t test). The range in size for both groups was 8·5 to 6·0. Standard techniques for assessing height from hand size[2] showed that the orthopaedic surgeons were slightly taller than their general surgical colleagues (by 2·3 cm).



Barrett DS. Are orthopaedic surgeons gorillas? BMJ. 1988 Dec;297(6664):1638–9.

Are orthopaedic surgeons gorillas?

D S Barrett

Long held traditions die hard in the specialty of general surgery, and the image of the orthopaedic surgeon as a man of enormous build and great strength, if perhaps a little slow, is still popular. I assessed whether what was once an undoubted fact is true today.

orthopaedic surgeons. When the surgeons wore two pairs of gloves only the size of the inner glove was used.

The mean size of glove worn by orthopaedic surgeons was 7·6 (SD 0·4) while the mean size of gloves worn by general surgeons was 7·4 (0·4). The difference between the two groups was highly significant ($p<0·001$, Student's *t* test). The range in size for both groups was 8·5 to 6·0. Standard techniques for assessing height from hand size[2] showed that the orthopaedic surgeons were slightly taller than their general surgical colleagues (by 2·3 cm).

Are orthopaedic surgeons really gorillas?

John S Fox, Gordon R Bell, Patrick J Sweeney

Critical comparison between orthopaedic and general surgeons is woefully absent in scientific publications, leading to speculation, innuendo, and myth. This often results in derogatory remarks about one of these two categories. Previous reports in the *BMJ* have lent some credence to the long held traditions and popular myths that the orthopaedic surgeon is a man of enormous build and great strength, if perhaps a little slow[1]; that orthopaedic surgery requires brute

on the glove size of actual gorillas. In addition, the study was not prospective or double blind and did not include a crossover. There was no disclaimer that funds were not received from any interested party. At best, results could only support the contention that orthopaedic surgeons are bigger gorillas than are general surgeons. To correct for these flaws we undertook a randomised double blind study.

Barrett DS. Are orthopaedic surgeons gorillas? BMJ. 1988 Dec;297(6664):1638–9.

Fox JS, et al. Are orthopaedic surgeons really gorillas? BMJ. 1990 Dec;301(6766):1425–6.

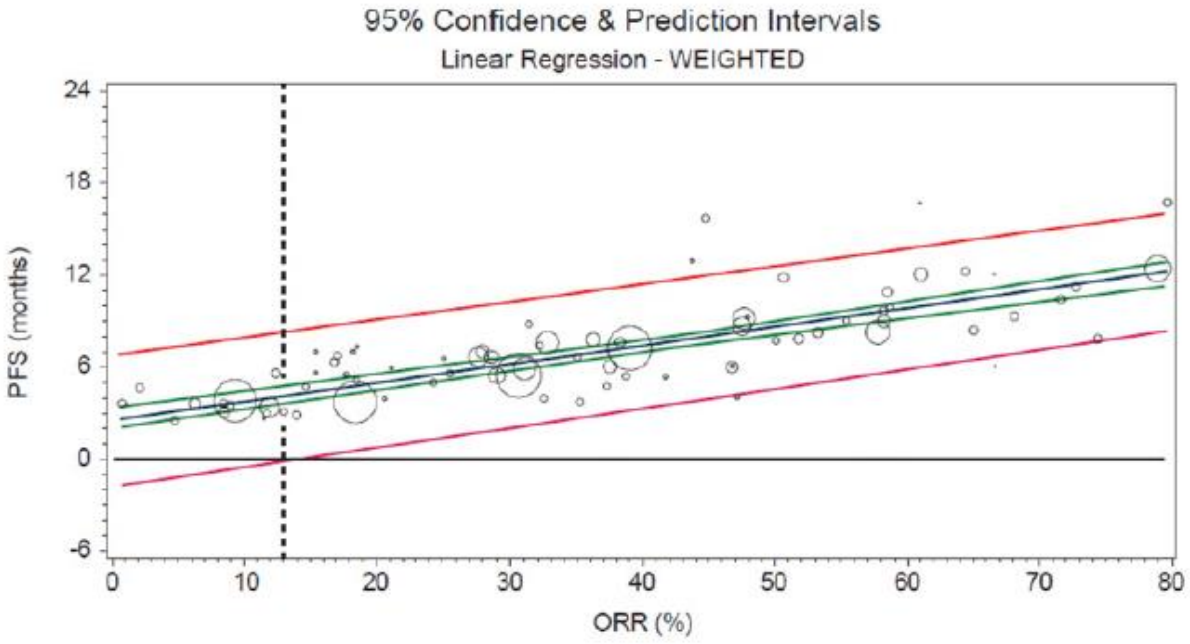
TASA DE RESPUESTAS

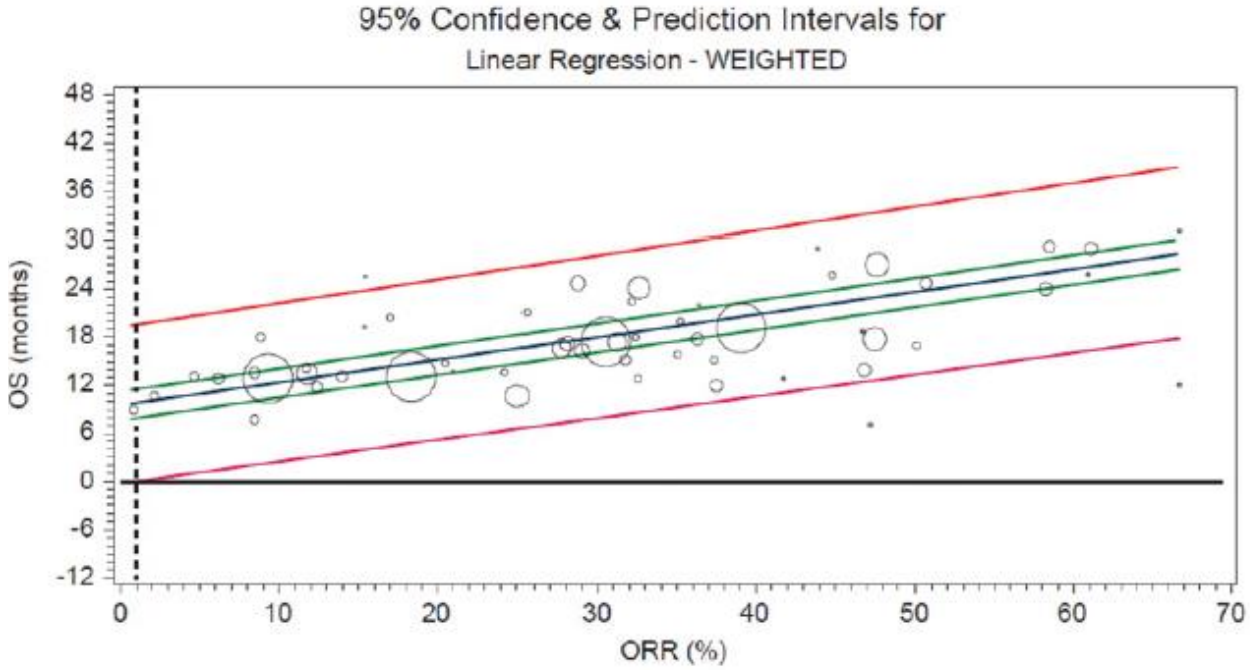
- Información puntual sobre disminución en el tamaño de las lesiones
- RECIST 1.1
- InmunoRECIST
- Muy útil en fase II y neoadyuvancia

Fácil y rápida de medir

RECIST: **R**esponse **E**valuation **C**riteria **I**n **S**olid **T**umors Eisenhauer EA et al Eur J Cancer. 2009;45(2):228-47

Cáncer de ovario



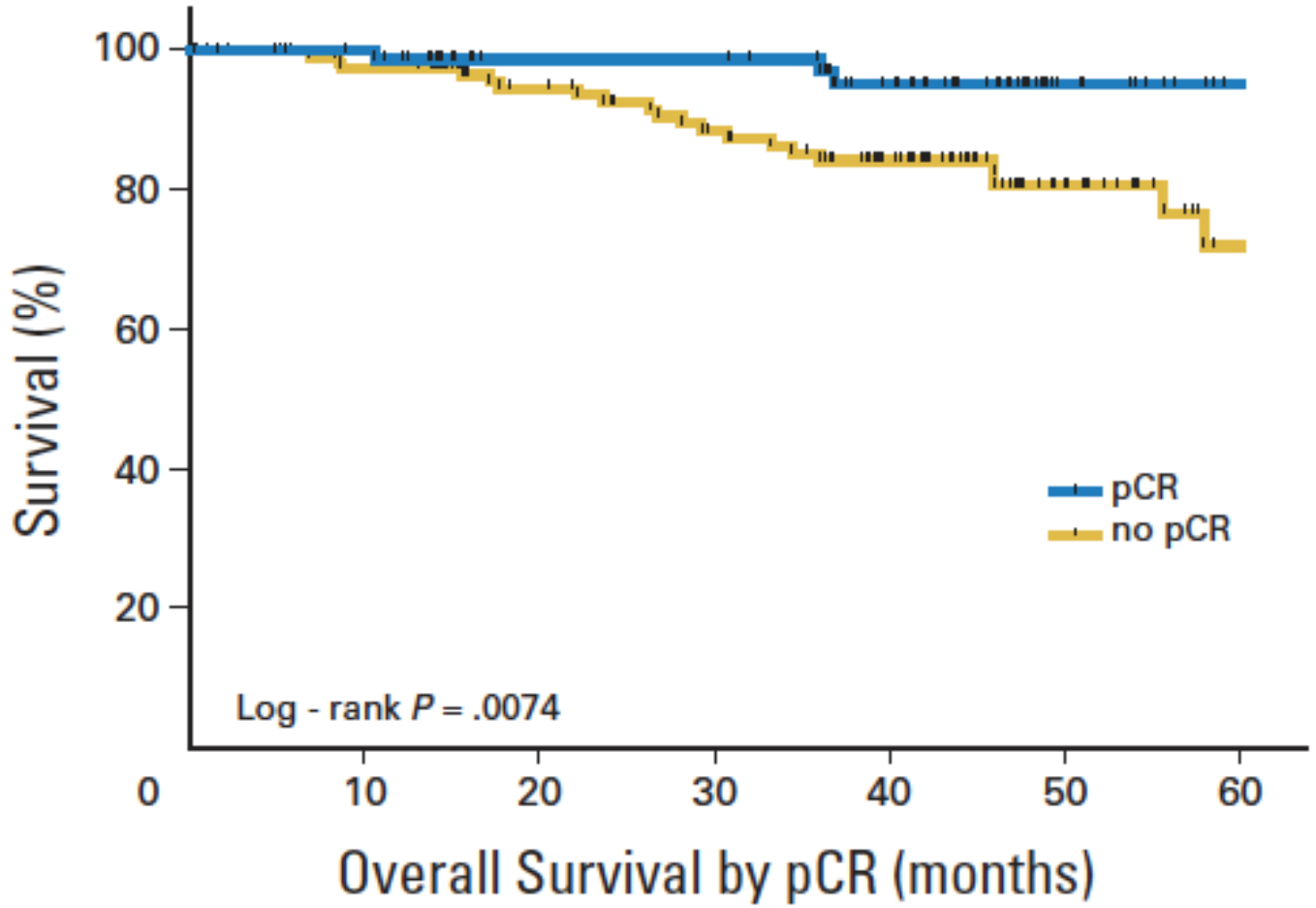


Por cada incremento del 10% en tasa de respuestas:

- incremento en 1.2 meses en SLP
- 2.83 m en SG

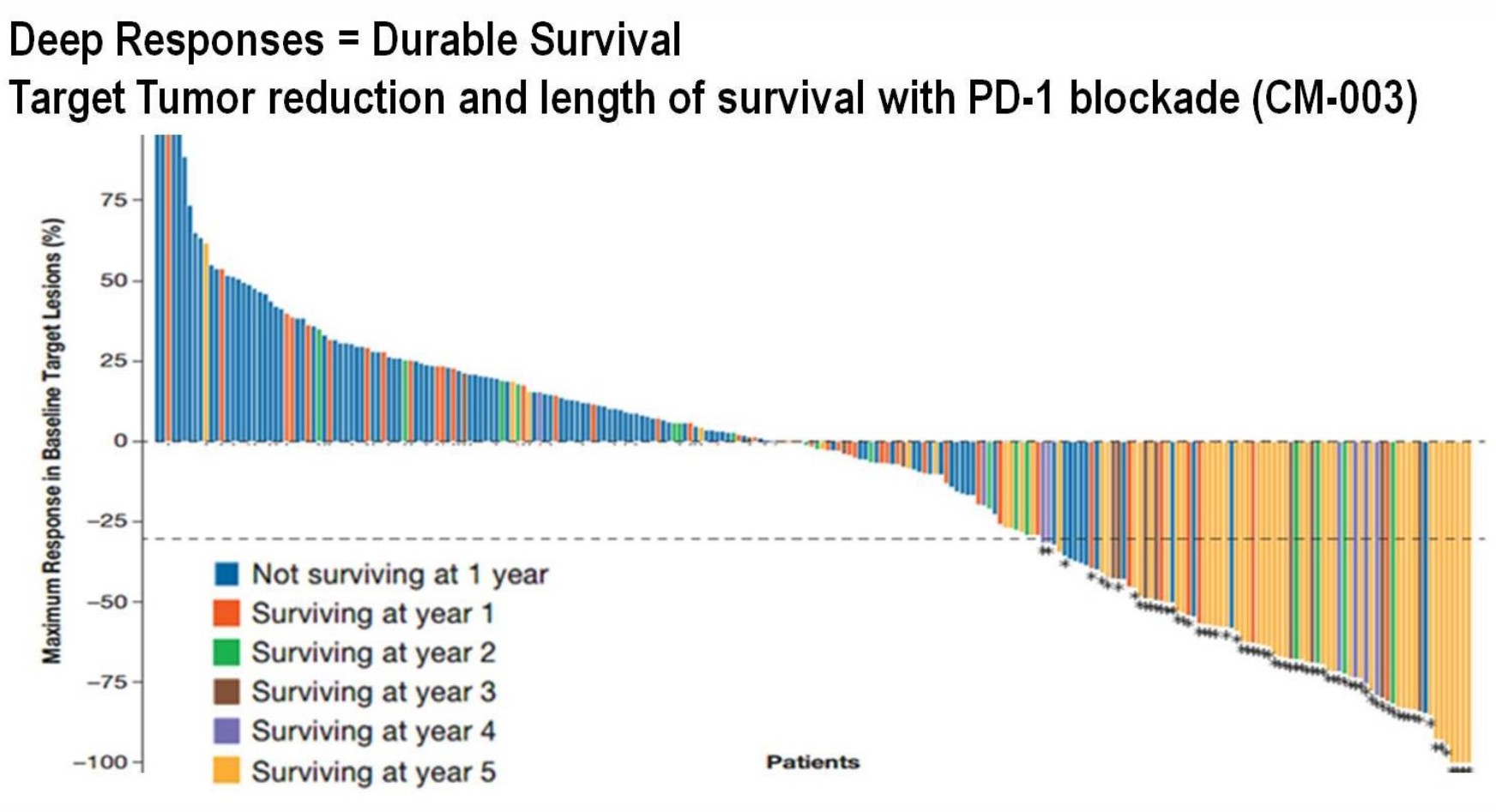
Siddiqui MK, Tyczynski J, Pahwa A, Fernandes AW. Objective response rate is a possible surrogate endpoint for survival in patients with advanced, recurrent ovarian cancer. Gynecologic Oncology; 2017 Jul 1;146(1):44–51.

Cáncer de mama



Untch M, Fasching PA, Konecny GE, Hasmüller S, Lebeau A, Kreienberg R, et al. Pathologic complete response after neoadjuvant chemotherapy plus trastuzumab predicts favorable survival in human epidermal growth factor receptor 2-overexpressing breast cancer: results from the TECHNO trial of the AGO and GBG study groups. J Clin Oncol. 2011 Sep 1;29(25):3351–7.

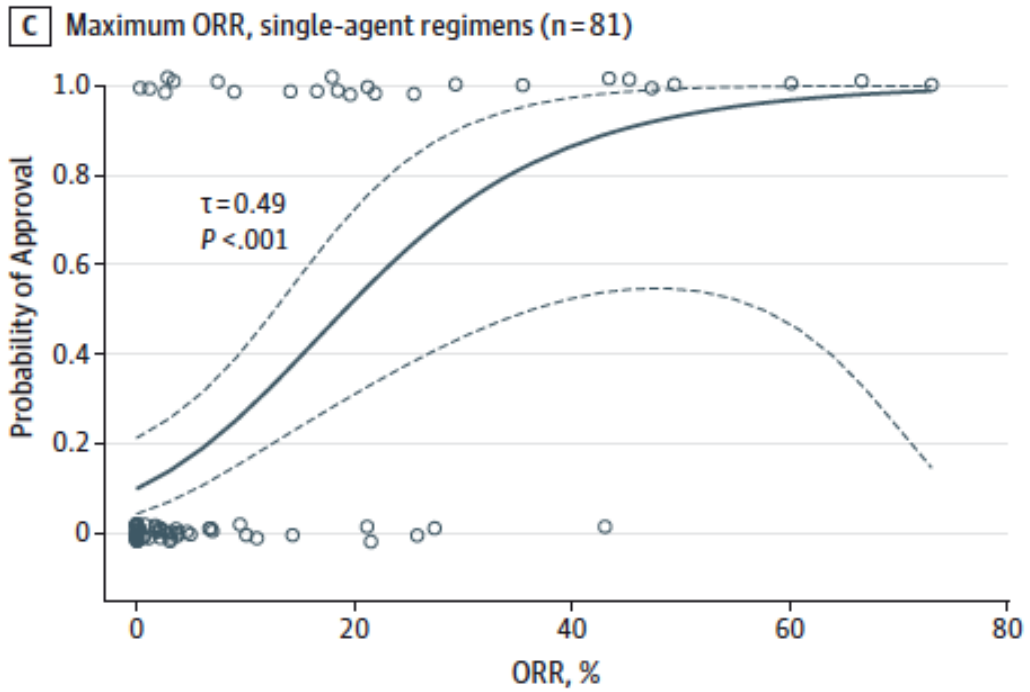
Cáncer renal: nivolumab

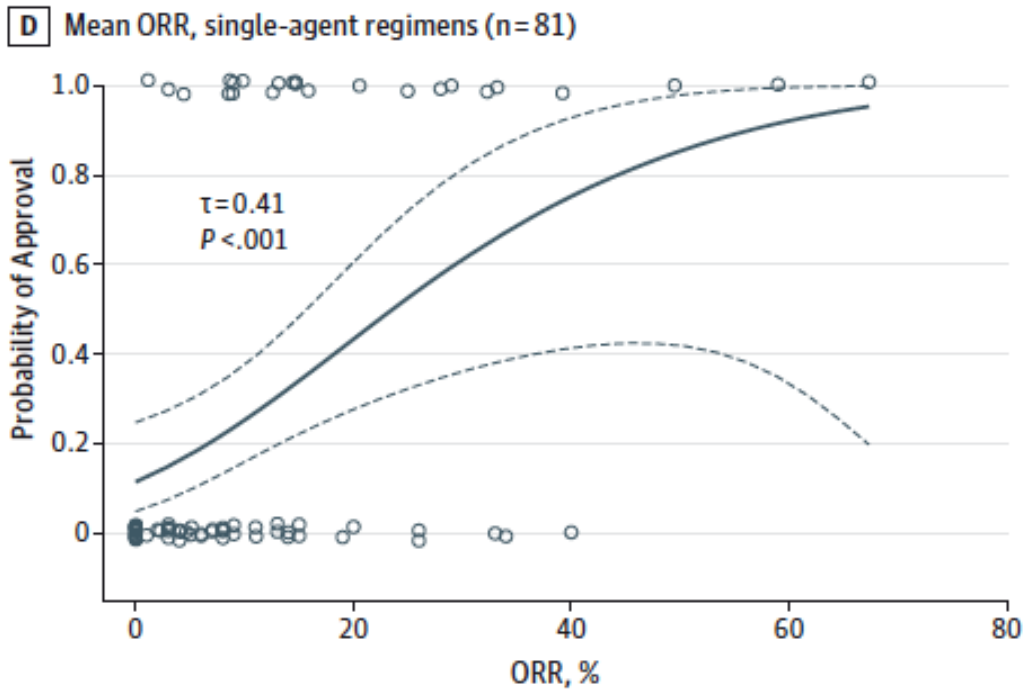


Topalian, SITC 2017; #P216

Original Investigation

Response Rate as a Regulatory End Point in Single-Arm Studies of Advanced Solid Tumors



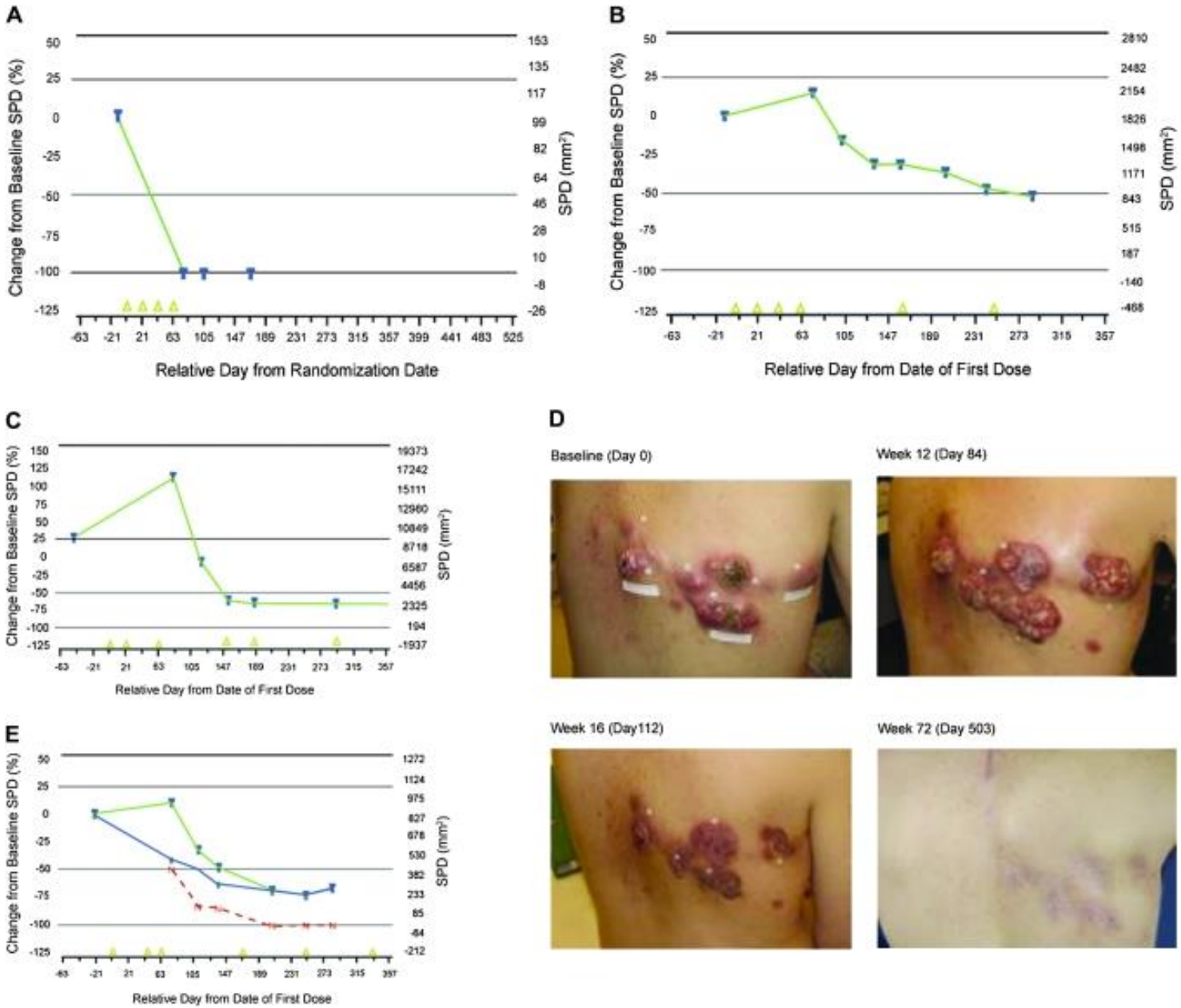


ORR statistically exceeding 30% with a single agent had 98% specificity and 89%positive predictive value for identifying regimens achieving regulatory approval.

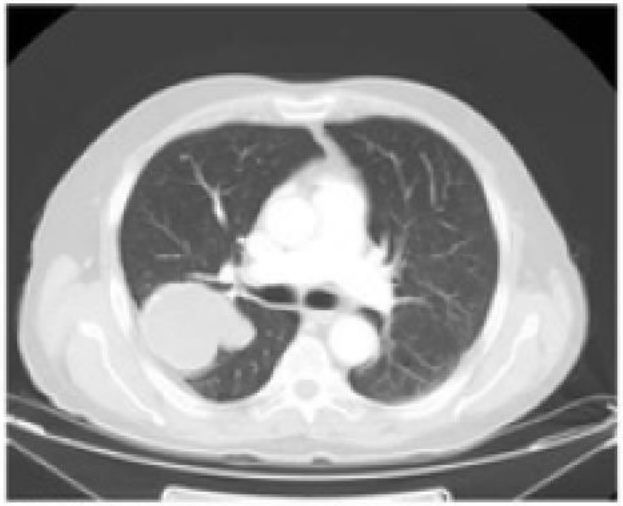
Oxnard GR. JAMA Oncol. 2016 Jun 1;2(6):772–9.

Inconvenientes

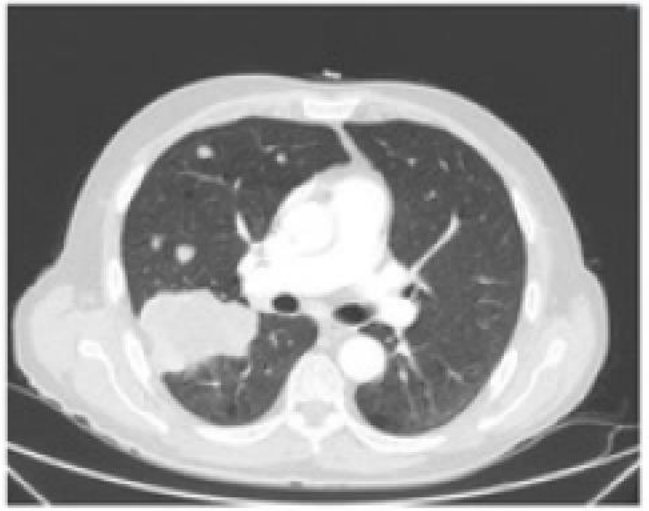
- Categorización amplia
- No indica duración ni calidad de la respuesta
- Sesgo de evaluación
- Variabilidad en ensayos fase II-III
- Subrogada?
- Lesiones no medibles
- Pseudoprogresión; necrosis
- Duración de respuesta: IT



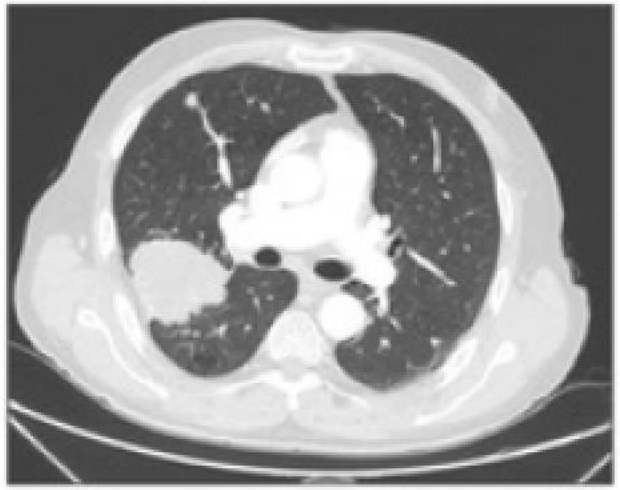
Hoos A, et al. Improved endpoints for cancer immunotherapy trials. J Natl Cancer Inst 2010, Sep 22;102(18):1388-97.



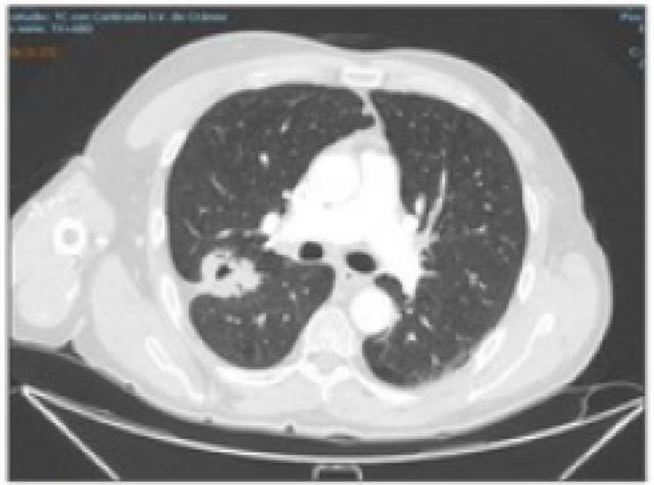
A) Inicio de nivolumab. Septiembre de 2016.



B) 4 ciclos de nivolumab. Noviembre de 2016. Progresión.



C) 6 ciclos de nivolumab. Diciembre de 2016. Pseudoprogresión irRECIST.



D) 12 ciclos de nivolumab. Marzo de 2017. Respuesta parcial irRECIST.

Adrián Sánchez Vegas; Sevilla; +MIR 2017

ANÁLISIS DE SUPERVIVENCIA

Supervivencia global: inicio de tratº hasta exitus

Supervivencia libre de progresión: desde inicio de tratamiento a progresión (o recurrencia en el caso de supervivencia libre de enfermedad)

Tiempo hasta fallo de tratamiento: desde inicio de tratamiento hasta que se suspende (progresión, toxicidad…)

SUPERVIVENCIA LIBRE DE PROGRESIÓN

- Mide desde el inicio del tratamiento hasta progresión /fallecimiento
- Variable subrogada de supervivencia global??
- No influenciada por líneas posteriores de tratamiento
- Menor tamaño muestral; menor seguimiento: ahorro de costes, acelera aprobación

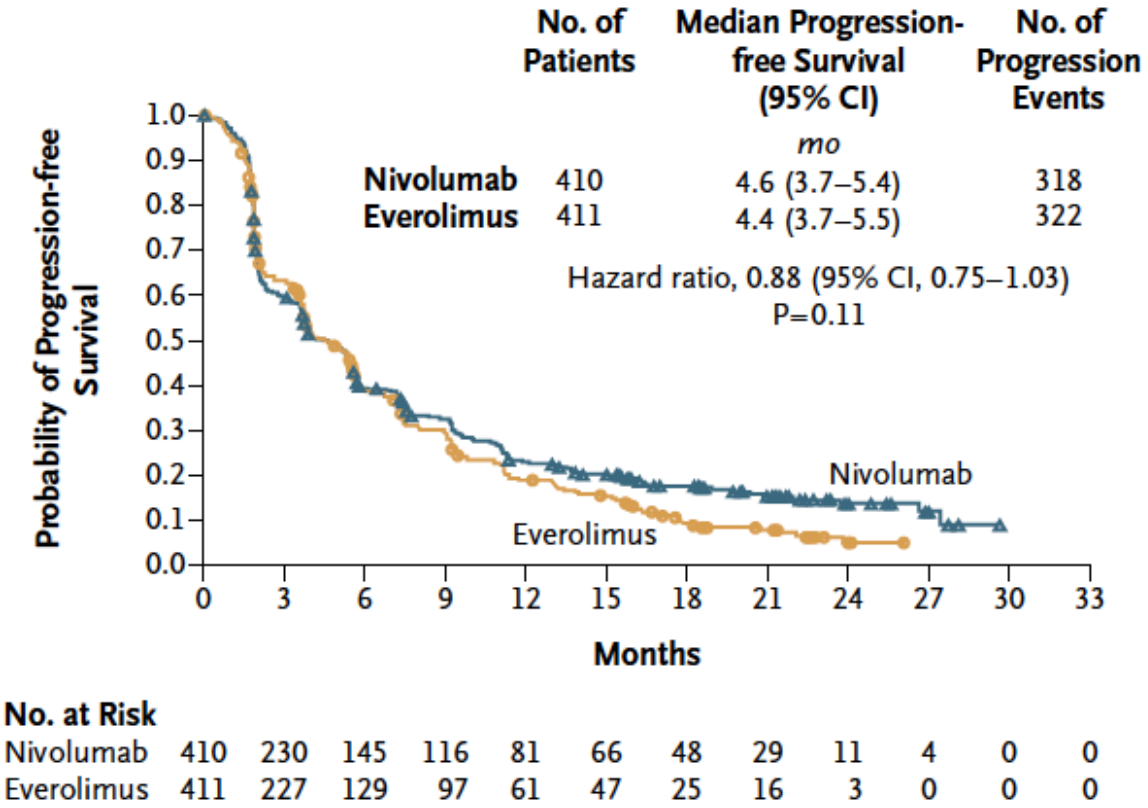
Inconvenientes

- Sesgo de evaluación
- Frecuencia de la evaluación. Método de evaluación
- Validación insuficiente
- Inmunoterapia
- Puede infraestimar el beneficio de tratº con estabilizaciones prolongadas

	Objetivo subrogado	Conclusión
Cáncer Colorrectal	DFS PFS	Válido Válido
Cáncer de ovario	PFS	Aceptado como válido en fases III en primera línea
Cáncer de mama	DFS, RR, TTP, PFS	Evidencia inadecuada
Cáncer de pulmón	PFS (SCLC) RR, TTP, PFS (NSCLC)	Válido Baja correlación
Cáncer renal	PFS	Evidencia limitada
Cáncer de próstata	PFS	Baja evidencia

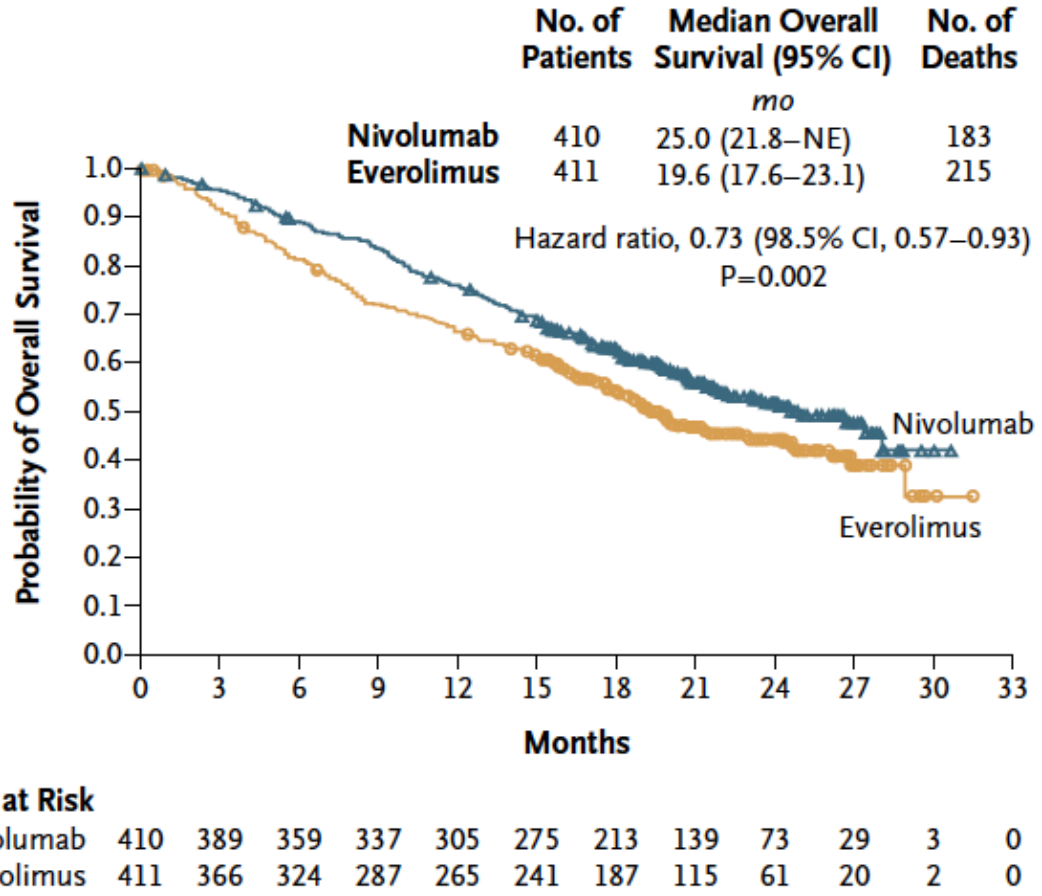
Sherrill, Onco Targets Ther 2012;5:287-296

CheckMate 025: nivolumab vs everolimus ca renal en 2ª línea



No. at Risk

Nivolumab	410	230	145	116	81	66	48	29	11	4	0	0
Everolimus	411	227	129	97	61	47	25	16	3	0	0	0



No. at Risk

Nivolumab	410	389	359	337	305	275	213	139	73	29	3	0
Everolimus	411	366	324	287	265	241	187	115	61	20	2	0

Motzer RJ. N Engl J Med. 2015 Nov 5;373(19):1803–13.

SUPERVIVENCIA GLOBAL

- Gold standard
- Importa a paciente, médico

- No es ambigua ni sujeta a sesgos

- *Seguimiento prolongado*

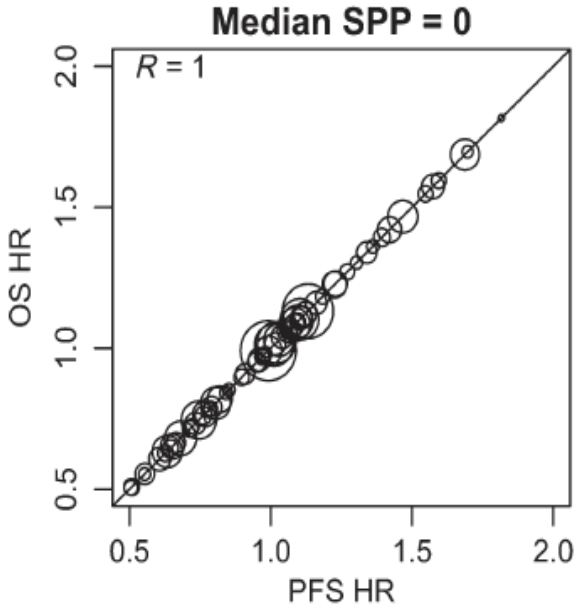
- *Tratamientos ulteriores*

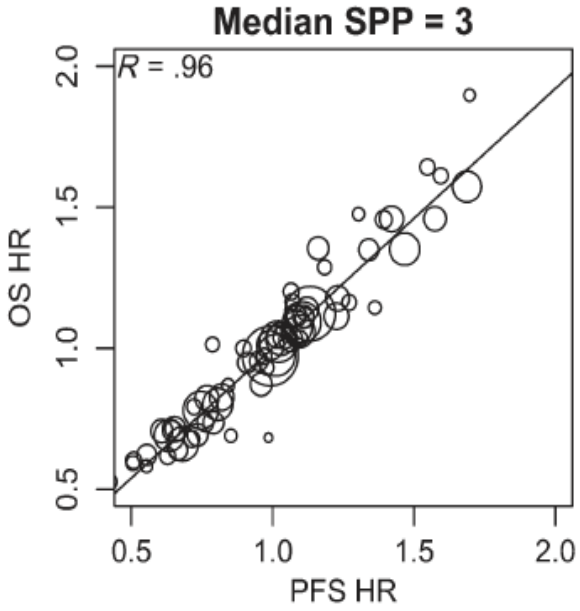
- *No útil como endpoint en tratamientos que pueden conseguir beneficios a largo plazo*

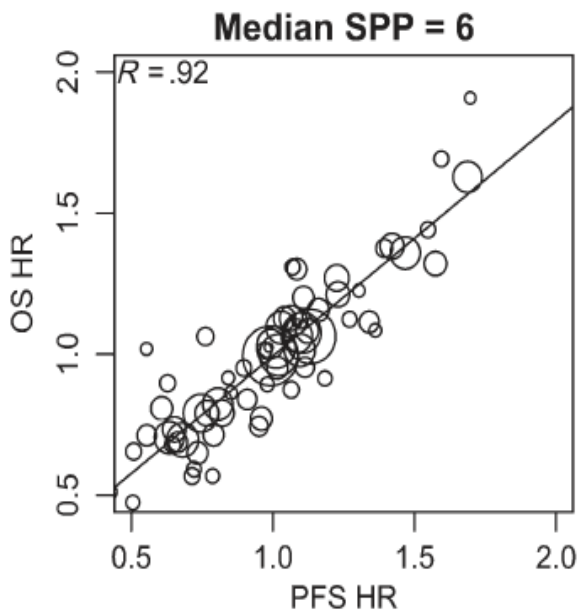
OS= PFS + Supervivencia post-progresión

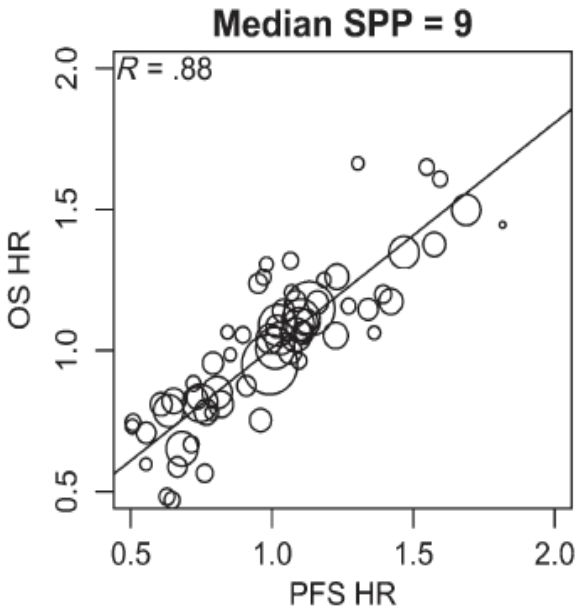
- Estrategia no aleatorizada a la progresión
- Estrategias sesgadas
- Dificultad para comparar regímenes en base a OS

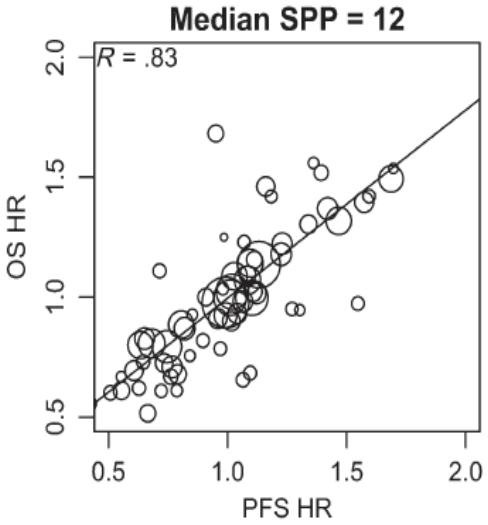
Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. Journal of the National Cancer Institute. 2009 Dec 2;101(23):1642–9.

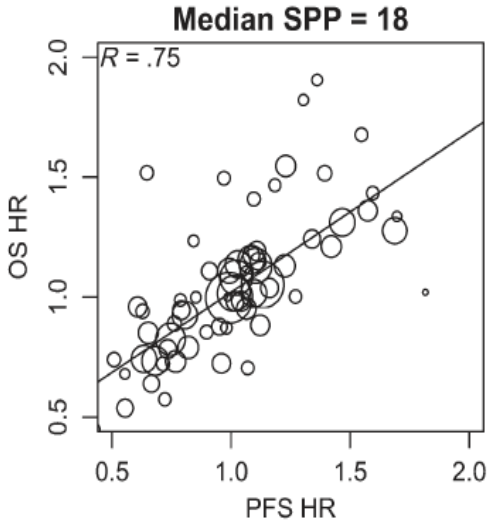


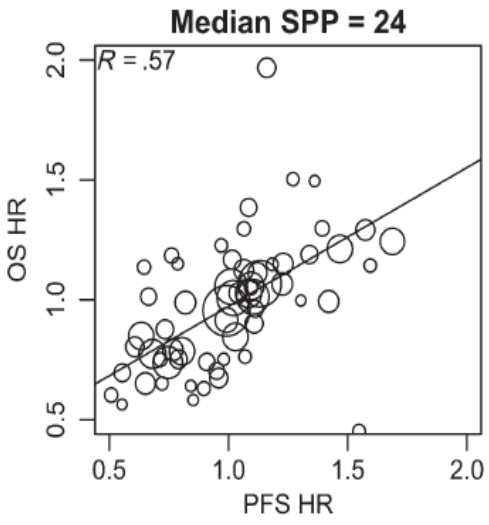




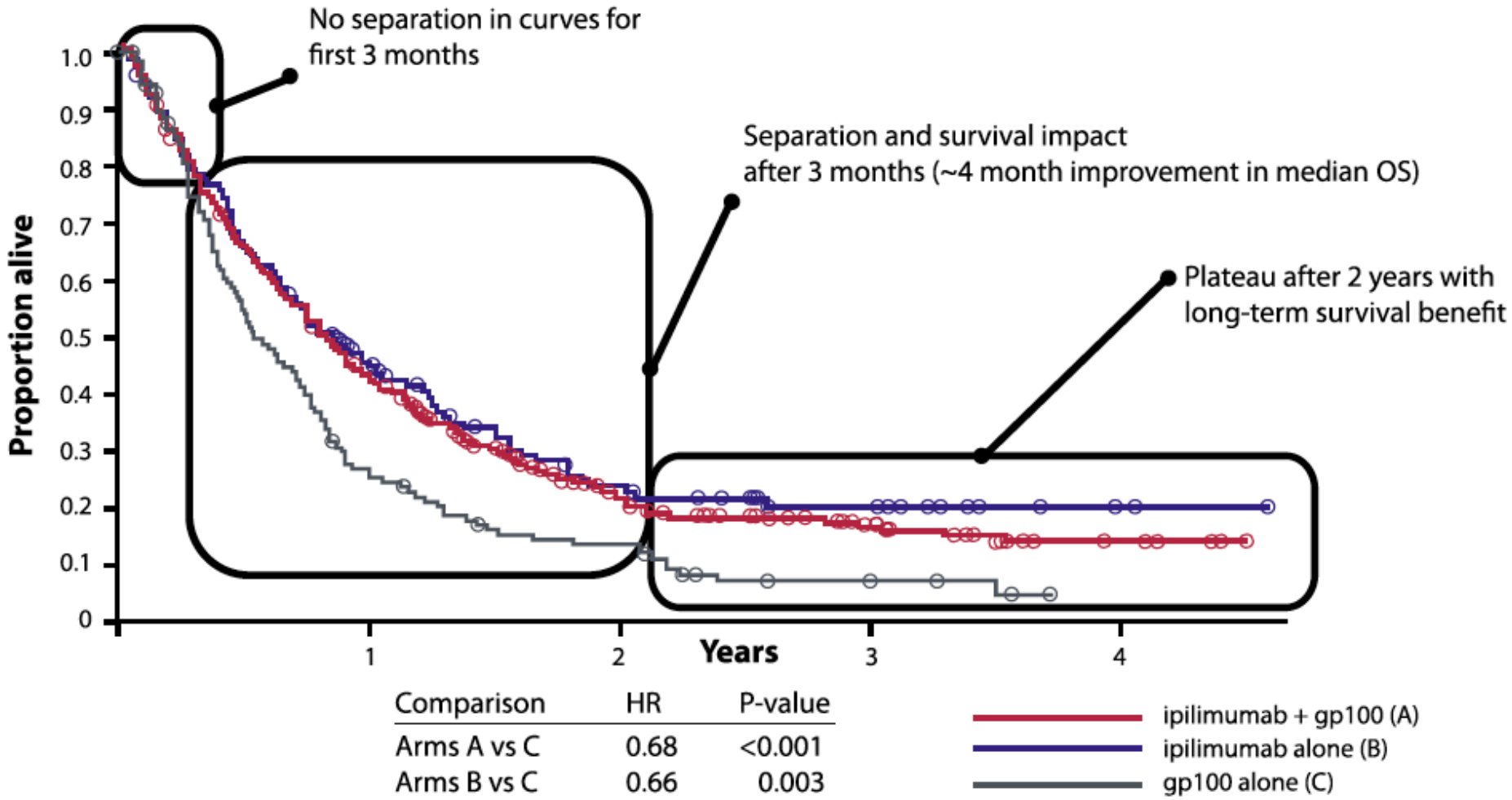






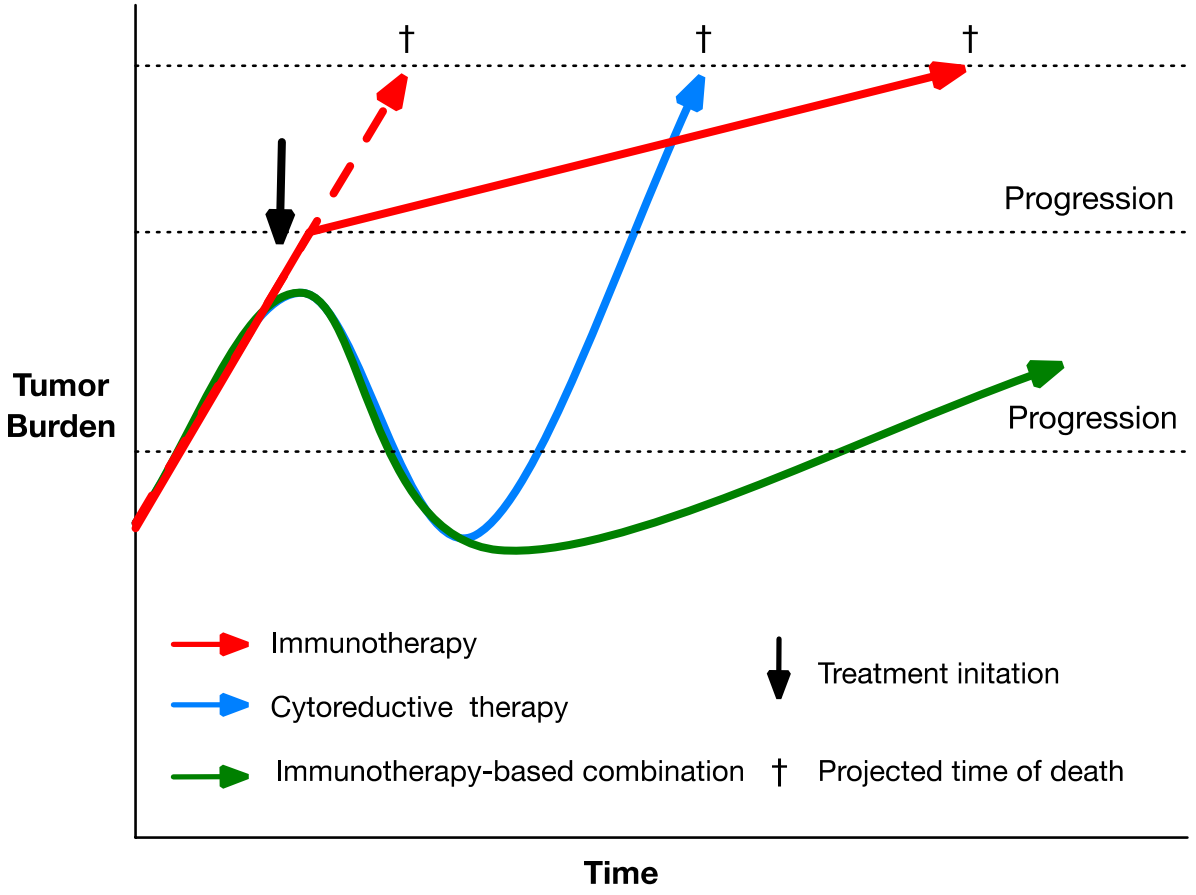


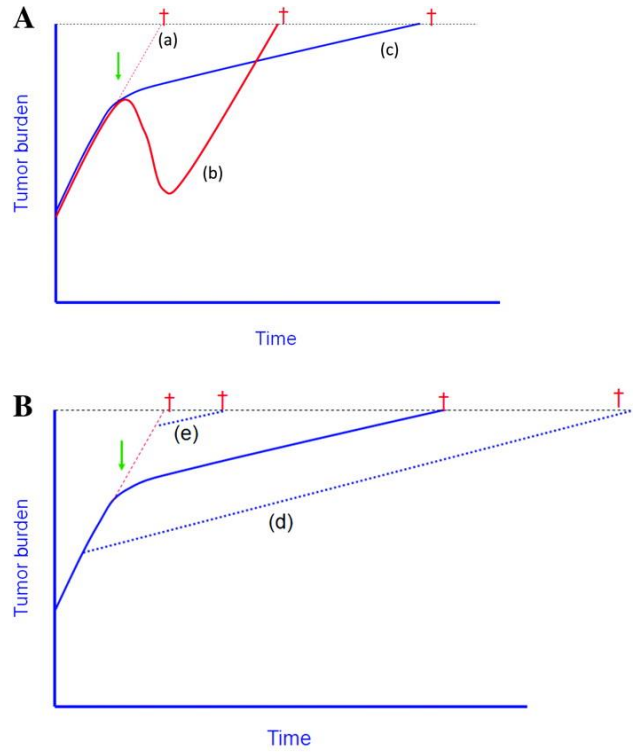
OS es un objetivo principal razonable cuando la mediana de SPP es corta, quizás menos de 6 meses, pero es un obstáculo importante cuando es superior a los 12 meses



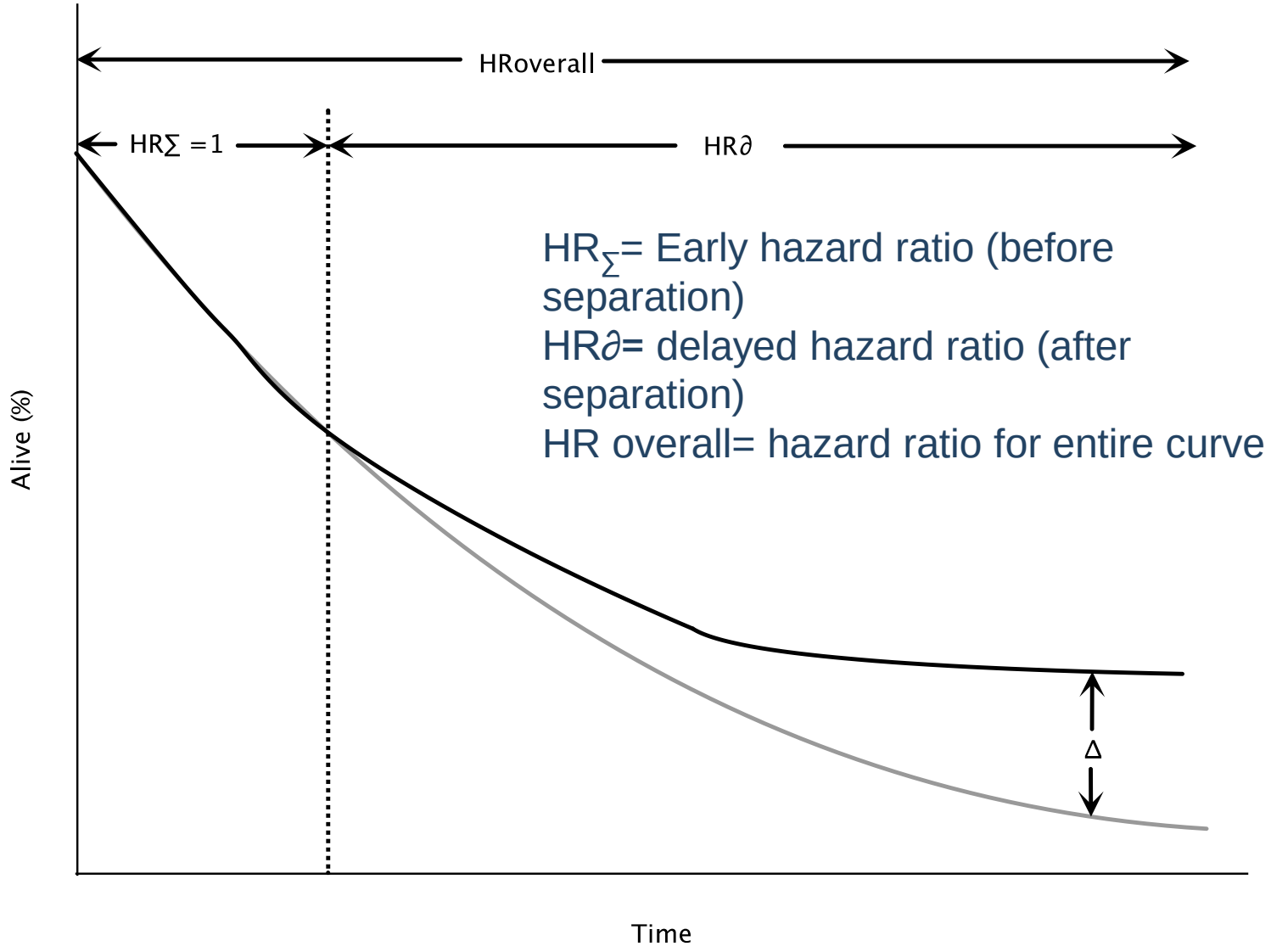
McDermott D, Lebbé C, Hodi FS, Maio M, Weber JS, Wolchok JD, et al. Durable benefit and the potential for long-term survival with immunotherapy in advanced melanoma. Cancer Treatment Reviews; 2014 Oct 1;40(9):1056–64.

PROBLEMAS AL EVALUAR





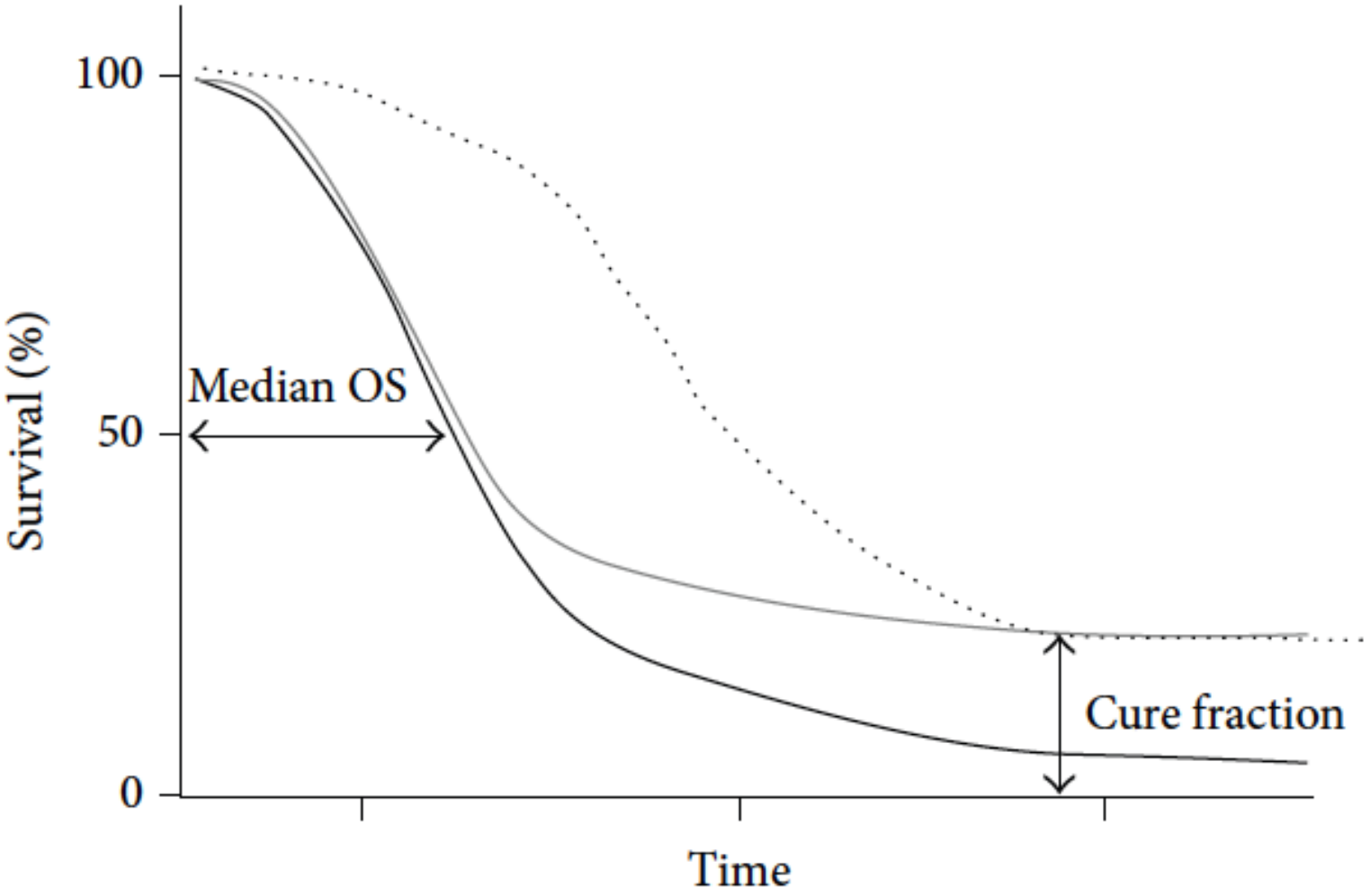
Madan RA, Schwaab T, Gulley JL. Strategies for optimizing the clinical impact of immunotherapeutic agents such as sipuleucel-t in prostate cancer. J Natl Compr Canc Netw 2012, Dec 1;10(12):1505-12.
Schlom J. Therapeutic cancer vaccines: Current status and moving forward. J Natl Cancer Inst 2012, Apr 18;104(8):599-613.



$HR_{\sum}$= Early hazard ratio (before separation)

HR∂= delayed hazard ratio (after separation)

HR overall= hazard ratio for entire curve



Johnson P, Greiner W, Al-Dakkak I, Wagner S. Which Metrics Are Appropriate to Describe the Value of New Cancer Therapies? BioMed Research International; 2015;2015(12):1–9.

When Are “Positive” Clinical Trials in Oncology Truly Positive?

Alberto Ocana, Ian F. Tannock

BENEFICIO ESTADISTICAMENTE SIGNIFICATIVO

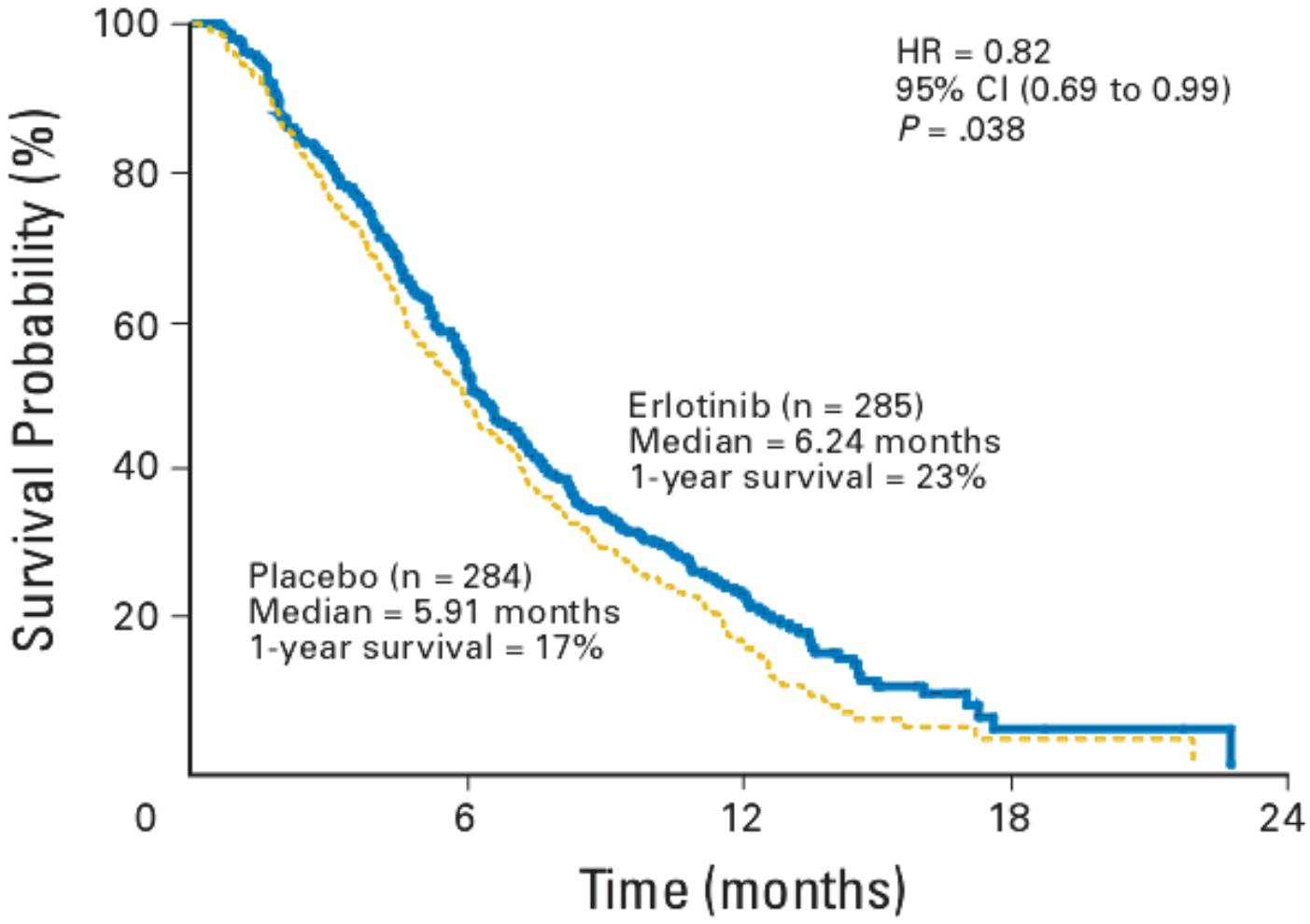
–p < 0.05 IC95% HR No atraviesa el valor 1

BENEFICIO CLINICO (RELEVANCIA CLINICA)

–**Valor δ:** Mínima diferencia observada entre el tratamiento experimental y el control considerada como clínicamente relevante

Ocana A, Tannock IF. When are “positive” clinical trials in oncology truly positive? Journal of the National Cancer Institute. 2011 Jan 5;103(1):16–20.

Erlotinib vs gemcitabina en cáncer de páncreas



El estudio fue diseñado para detectar una reducción del riesgo relativo del 25% (HR≤75)

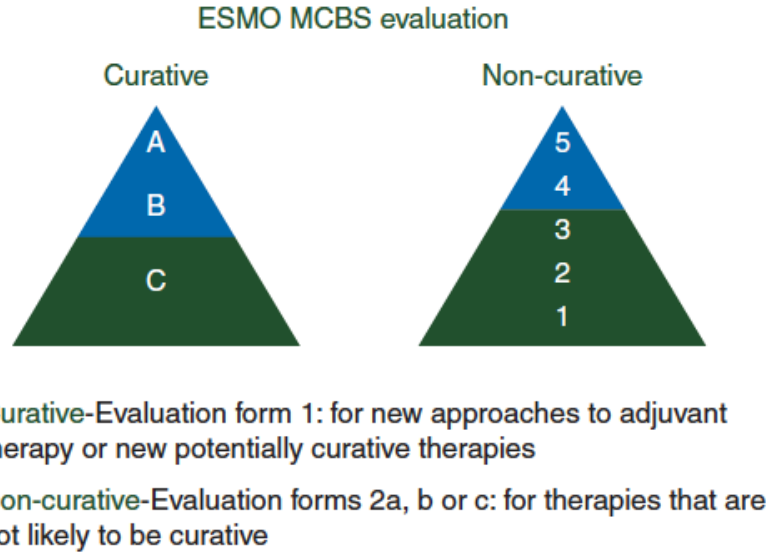
Moore, et al. J Clin Oncol 2007;25:1960-1966

SOCIEDADES: ASCO, ESMO

Table 1. Summary of Recommended Targets for Meaningful Clinical Trial Goals

			Primary End Point		Secondary End Point	
Cancer Type	Patient Population	Current Baseline Median OS (months)	Improvement Over Current OS That Would Be Clinically Meaningful (months)	Target HRs	Improvement in 1-Year Survival Rate (%)*	Improvement in PFS (months)
Pancreatic cancer	FOLFIRINOX-eligible patients	10 to 11[19]	4 to 5	0.67 to 0.69	48 → 63	4 to 5
Pancreatic cancer	Gemcitabine or gemcitabine/nab-paclitaxel–eligible patients	8 to 9[20,21]	3 to 4	0.6 to 0.75	35 → 50	3 to 4
Lung cancer	Nonsquamous cell carcinoma	13[22]	3.25 to 4	0.76 to 0.8	53 → 61	4
Lung cancer	Squamous cell carcinoma	10[23,23a]	2.5 to 3	0.77 to 0.8	44 → 53	3
Breast cancer	Metastatic triple negative, previously untreated for metastatic disease	18[24,25]	4.5 to 6	0.75 to 0.8	63 → 71	4
Colon cancer	Disease progression with all prior therapies (or not a candidate for standard second- or third-line options)	4 to 6[26]	3 to 5	0.67 to 0.67	25 → 35	3 to 5

Abbreviations: FOLFIRINOX, leucovorin, fluorouracil, irinotecan, and oxaliplatin; HR, hazard ratio; OS, overall survival; PFS, progression-free survival.
*Current → target.

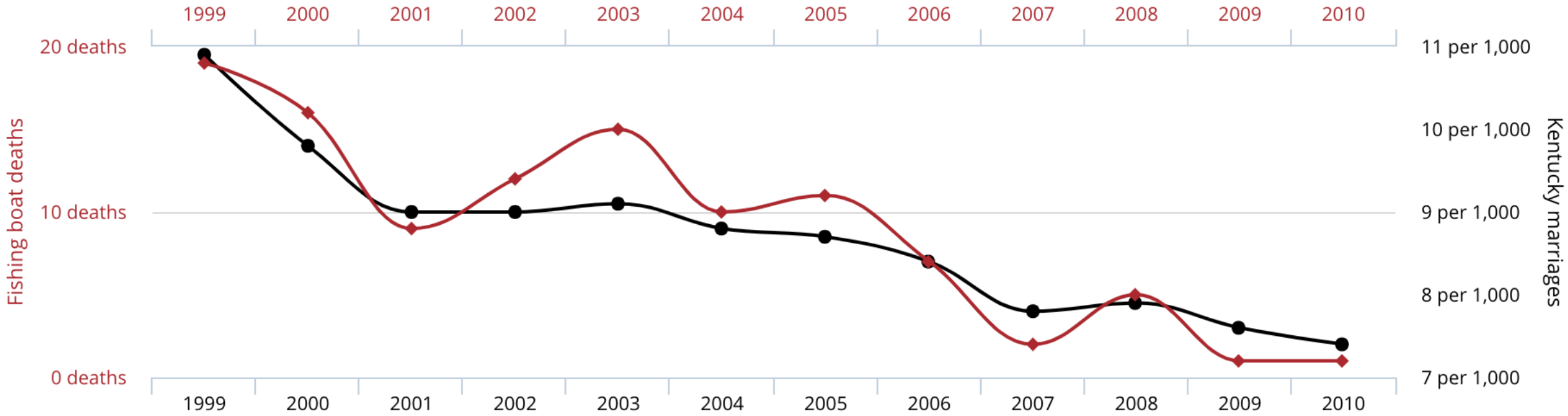


Curative-Evaluation form 1: for new approaches to adjuvant therapy or new potentially curative therapies

Non-curative-Evaluation forms 2a, b or c: for therapies that are not likely to be curative

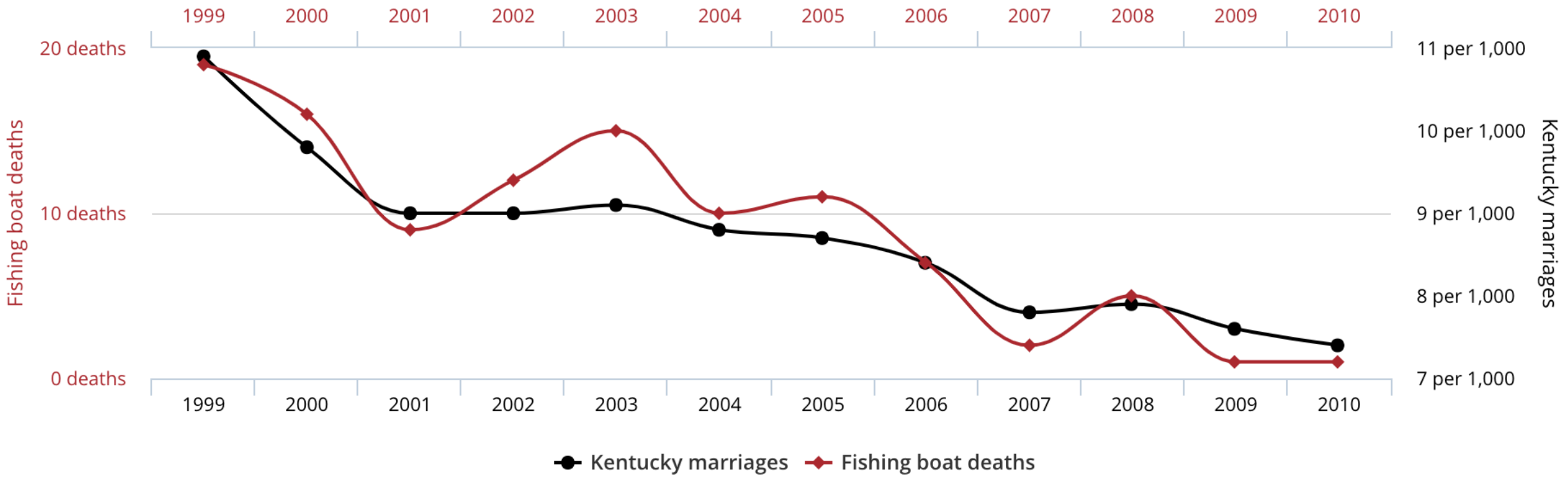
Ellis LM, Bernstein DS, Voest EE, Berlin JD, Sargent D, Cortazar P, et al. American Society of Clinical Oncology perspective: Raising the bar for clinical trials by defining clinically meaningful outcomes. Ann Oncol. American Society of Clinical Oncology; 2014 Apr 20;32(12):1277–80.

Dafni U, et al. Detailed statistical assessment of the characteristics of the ESMO Magnitude of Clinical Benefit Scale (ESMO-MCBS) threshold rules. ESMO Open. 2017 Oct 9;2(4):e000216–13.


1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
20 deaths
10 deaths
0 deaths
Fishing boat deaths
11 per 1,000
10 per 1,000
9 per 1,000
8 per 1,000
7 per 1,000
Kentucky marriages

Personas ahogadas tras caer de una barca de pesca
se correlaciona con
la tasa de matrimonios en Kentucky

Correlation: 95.24% (r=0.952407)



tylervigen.com

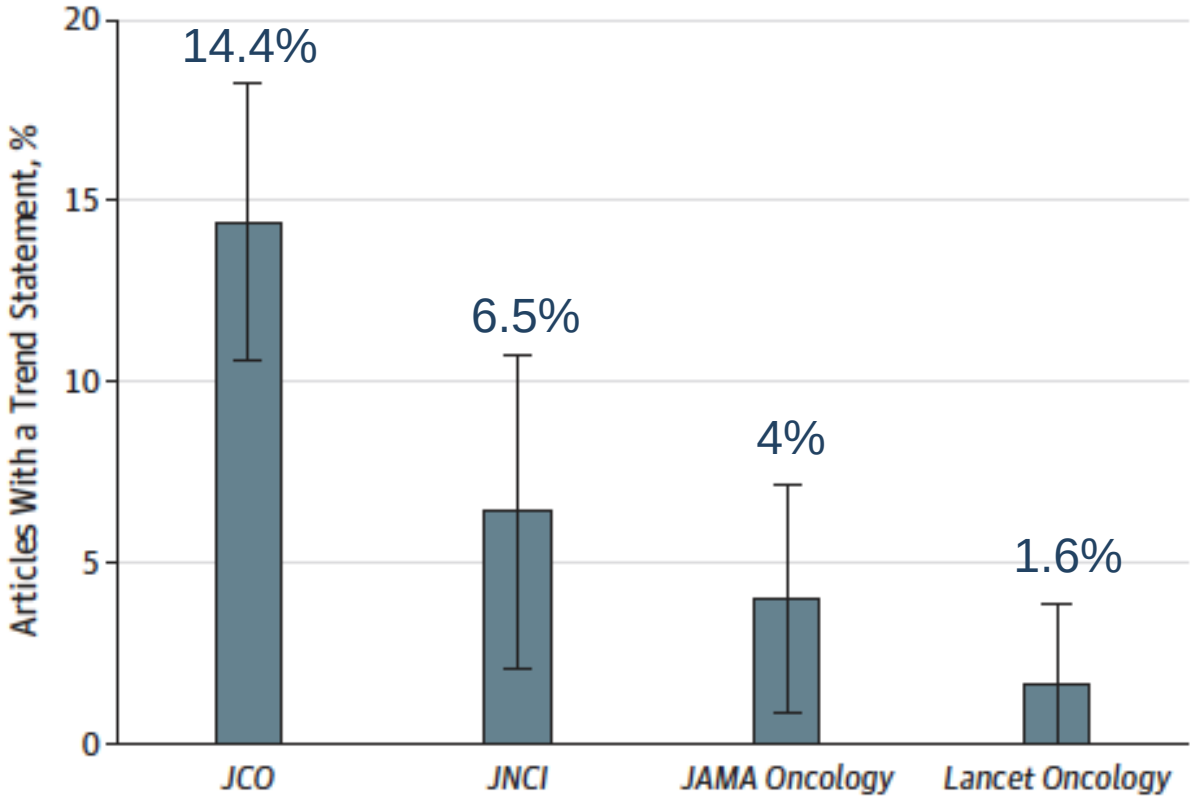
Data sources: Centers for Disease Control & Prevention and National Vital Statistics Reports

LA COMUNICACIÓN….

EL USO DE “TENDENCIA” CUANDO NO HAY SIGNIFICACIÓN ESTADÍSTICA

Artículos originales publicados entre 11/2016 y 10/2017: 722

10%



La “*p*”

- Las conclusiones no deben basarse en la *p*
- El valor de la *p* no mide la importancia del resultado
- La *p* no es una buena medida de evidencia

Nead KT, JAMA Oncol. 2018 Oct 18;:1–2
Wasserstein; ASA’s statetments on p values. Am Stat 2016;70(2)129-33

Table. Examples of Problematic Uses of the Word "Trend" to Describe Statistically Nonsignificant Results

Text	Trend Statement Location	Comparator Statistics Not Included With Trend Statement (Location in Article)	Point Illustrated
"There was a trend toward long-term survival in favor of GemErlo (estimated survival after 1, 2, and 5 y for GemErlo was 77%, 53%, and 25% vs 79%, 54%, and 20% for Gem, respectively)."[2]	Abstract	*P* = .61 (Figure 2B)	Large *P* value and trend statement in abstract without supporting comparator statistics
"BCL2 expression indicated a trend toward inferior outcome within GCB-like DLBCL, but not within ABC-like DLBCL (Appendix Figure A2, online only)."[3]	Results, main text	*P* = .547 (EFS), *P* = .351 (PFS), *P* = .125 (OS) (Appendix)	Large *P* value and comparator statistics to support trend statement found only in appendix
"Women with stage I-II PBL had overall survival superior to women with stage I-II systemic presentations of the same lymphoma subtype (Figure 4), except for ALCL-PBL where the same trend was seen, although not statistically significant at the 5% level."[4]	Results, main text	*P* = .69 (Figure 4)	Large *P* value
"In unselected patients, a trend for ramucirumab survival benefit was observed in patients with HCC in the Child-Pugh 5 disease subgroup."[5]	Key points section; abstract conclusion	HR = 0.80; 95% CI = 0.63-1.02; *P* = .06 (results text and Figure 1A)	Trend statement highlighted in Key Points section and as primary study conclusion (abstract) without mention of potentially clinically relevant effect size

Abbreviations: ABC, activated B cell; ALCL, anaplastic large cell lymphoma; DLBCL, diffuse large B-cell lymphoma; EFS, event-free survival; GCB, germinal center B cell; Gem, gemcitabine; GemErlo, gemcitabine and erlotinib; HCC, hepatocellular carcinoma; HR, hazard ratio; PBL, primary breast lymphoma; PFS, progression-free survival; OS, overall survival.

Nead KT, JAMA Oncol. 2018 Oct 18;:1–2

Reporting harms more transparently in trials of cancer drugs

Studies of cancer drugs often use terms that downplay the seriousness of adverse events. **Bishal Gyawali and colleagues** call for greater clarity and transparency

Bishal Gyawali *postgraduate trainee*[1], Tomoya Shimokata *assistant professor*[1], Kazunori Honda *medical oncologist*[2], Yuichi Ando *professor*[1]

Box 1: Terms used to downplay the harms of cancer drugs and reasons for avoiding them

Acceptable—Acceptable to whom? Were the patients asked if the toxicities were acceptable to them?

Manageable—Serious events and deaths can never be considered manageable. Even manageable toxicities incur burden and decrease patients' quality of life

Feasible—What is the threshold for feasibility of a treatment? Will the mention of "the treatment is feasible" be enough to obtain patient's consent to a treatment?

Favourable toxicity profile—Favourable compared with what? Threshold of enduring toxicities and thus favourability is different from patient to patient

Tolerable or well tolerated—Only the patient can decide whether any side effect is tolerable

Safe—Any cancer treatment that has resulted in a treatment related death cannot be considered safe

2016:
New England Journal of Medicine, *Lancet*, *Lancet Oncology*, *Journal of the American Medical Association,* and *Journal of Clinical Oncology*

122 estudios

43% infravaloran toxicidad

La toxicidad se ve incrementada en el brazo experimental en el 77% de los estudios; muertes incrementadas en el 66%; efectos adversos serios en el 84%

Gyawali B, et al. Reporting harms more transparently in trials of cancer drugs. BMJ. 2018 Nov 1;363:k4383.

RIBOCICLIB + LETROZOL EN CÁNCER DE MAMA

Table 3. Adverse Events.*

Adverse Event	Ribociclib Group (N=334)			Placebo Group (N=330)†		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
	number of patients (percent)					
Any adverse event	329 (98.5)	221 (66.2)	50 (15.0)	320 (97.0)	105 (31.8)	3 (0.9)
Neutropenia‡	248 (74.3)	166 (49.7)	32 (9.6)	17 (5.2)	3 (0.9)	0
Nausea	172 (51.5)	8 (2.4)	0	94 (28.5)	2 (0.6)	0
Infections	168 (50.3)	12 (3.6)	2 (0.6)	140 (42.4)	7 (2.1)	1 (0.3)
Fatigue	122 (36.5)	7 (2.1)	1 (0.3)	99 (30.0)	3 (0.9)	0
Diarrhea	117 (35.0)	4 (1.2)	0	73 (22.1)	3 (0.9)	0
Treatment related serious adverse events		25			5	

"Most patients had an acceptable adverse-event profile with long-term administration of ribociclib plus letrozole"

Hortobagyi GN, Stemmer SM, Burris HA, Yap Y-S, Sonke GS, Paluch-Shimon S, et al. Ribociclib as First-Line Therapy for HR-Positive, Advanced Breast Cancer. N Engl J Med. 2016 Nov 3;375(18):1738–48.

Randomized, Double-Blind, Placebo-Controlled Phase III Study of Tasquinimod in Men With Metastatic Castration-Resistant Prostate Cancer

Table 3. Most Common AEs Occurring in at Least 5% of Patients in Either Treatment Group

	Tasquinimod (n = 830)				Placebo (n = 411)			
	All Grades		Grades 3 to 5		All Grades		Grades 3 to 5	
AE	No.	%	No.	%	No.	%	No.	%
All AEs	791	95.3	355	42.8	381	92.7	138	33.6

Serious adverse events: 36 vs 23%

"The tolerability was good overall"

Sternberg C, et al. Randomized, Double-Blind, Placebo-Controlled Phase III Study of Tasquinimod in Men With Metastatic Castration-Resistant Prostate Cancer. J Clin Oncol. 2016 Aug 1;34(22):2636–43.

Impower 131: atezolizumab +/- carboplatino-nabpaclitaxel or paclitaxel

Safety Summary

	Arm B: Atezo + CnP (N = 334)	Arm C (control): CnP (N = 334)
Treatment duration, median (range), mo		
Atezolizumab	6.7 (0-30)	NA
Carboplatin	2.6 (0-7)	2.4 (0-7)
Paclitaxel/nab-paclitaxel	3.0 (0-7)	2.8 (0-7)
All-cause AE, n (%)	332 (99)	324 (97)
Grade 3-4	243 (73)	220 (66)
Grade 5	31 (9)	14 (4)
Treatment-related AE, n (%)	316 (95)	303 (91)
Grade 3-4	227 (68)	190 (57)
Grade 5	4 (1)	3 (1)
Serious AE, n (%)	152 (46)	96 (29)
Treatment-related serious AE	68 (20)	35 (10)
AEs of special interest, n (%)	162 (49)	71 (21)
Grade 3-4	39 (12)	8 (2)
Grade 5	1 (< 1)	0
AE leading to any treatment withdrawal, n (%)	97 (29)	58 (17)
AE leading to any dose interruption or modification, n (%)	258 (77)	219 (66)

Data cutoff: January 22, 2018.

"Manageable safety profile"

Presented By Robert Jotte at 2018 ASCO Annual Meeting

Contents lists available at SciVerse ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Original Article

The role of facial hair in women's perceptions of men's attractiveness, health, masculinity and parenting abilities

Barnaby J. Dixson *, Robert C. Brooks

Evolution & Ecology Research Centre, School of Biological, Earth & Environmental Sciences, University of New South Wales, Kensington, Sydney 2052 NSW, Australia

Fig. 1. An example of the stimuli used in this study. Images show the same man when clean-shaven, with light stubble, heavy stubble and a full beard.

1. Dixson BJ, Brooks RC. The role of facial hair in women“s perceptions of men”s attractiveness, health, masculinity and parenting abilities. Evolution and Human Behavior. 2013 May;34(3):236–41.

¿Qué queremos medir?

¿Qué queremos medir?

¿Cómo queremos medirlo?

¿Qué queremos medir?

¿Cómo lo vamos a interpretar?

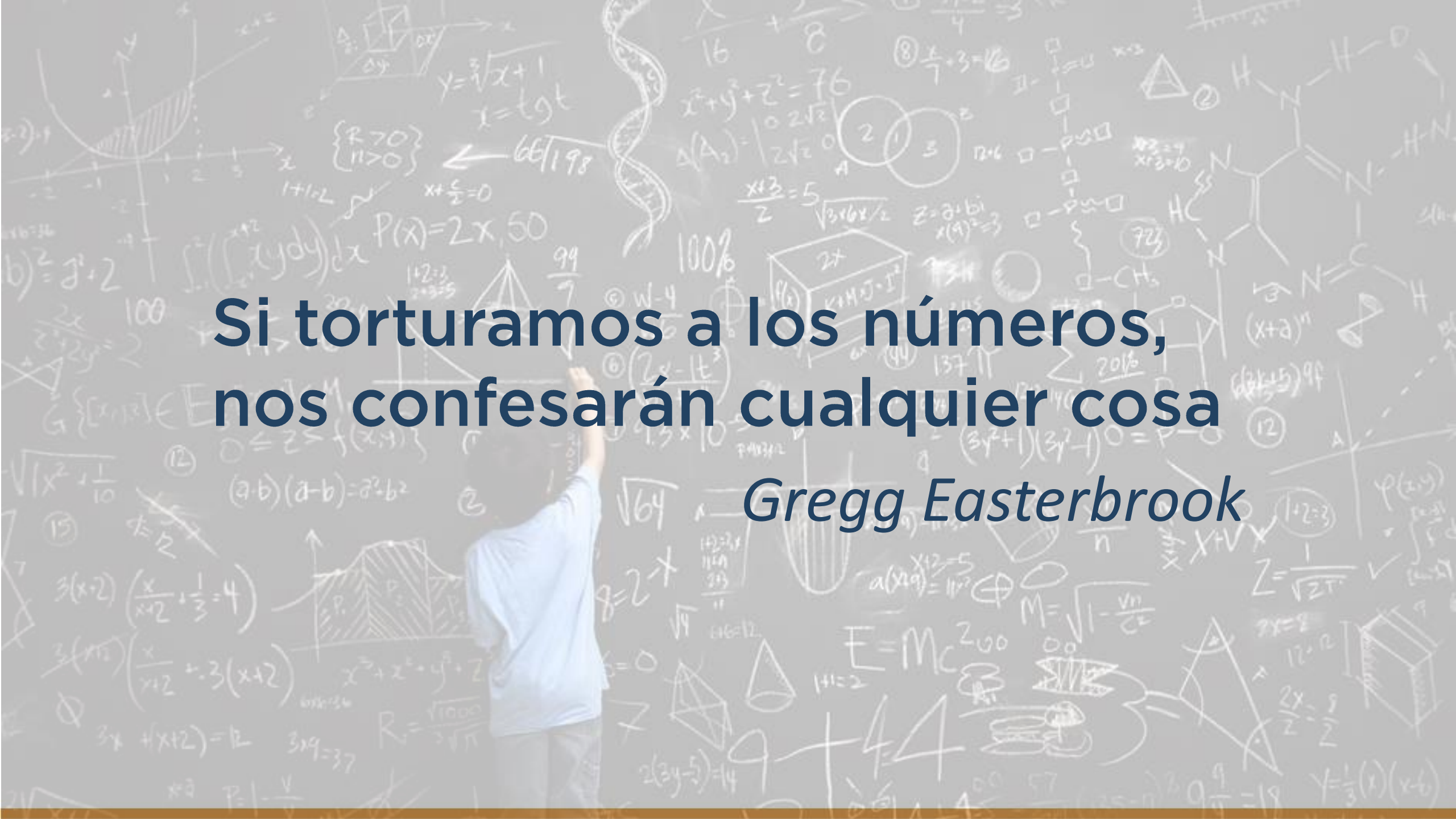
¿Cómo queremos medirlo?

¿Qué queremos medir?

¿Cómo lo vamos a interpretar?

¿Cómo queremos medirlo?

¿Cómo lo vamos a comunicar?



Si torturamos a los números, nos confesarán cualquier cosa

Gregg Easterbrook