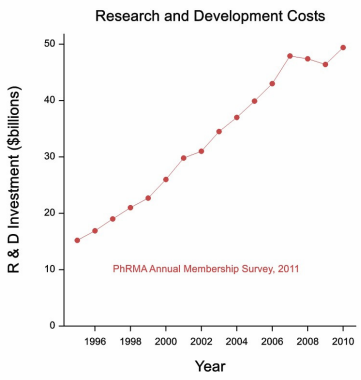
Increasing Efficiency in Clinical Trials

Stephen Senn

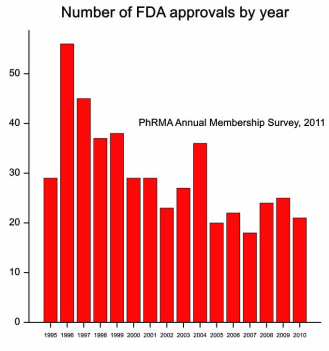
crp SANTÉ
CENTRE DE RECHERCHE PUBLIC
MÉTHODOLOGY AND STATISTICS

Why everybody is worried about costs

Input



Output



Basic Thesis

- There is an increased enthusiasm for new statistical methods in the pharmaceutical industry
- A great deal of interesting work is being done
- However many simple things that we have known about for decades are not being done
- It is time we changed this

Examples

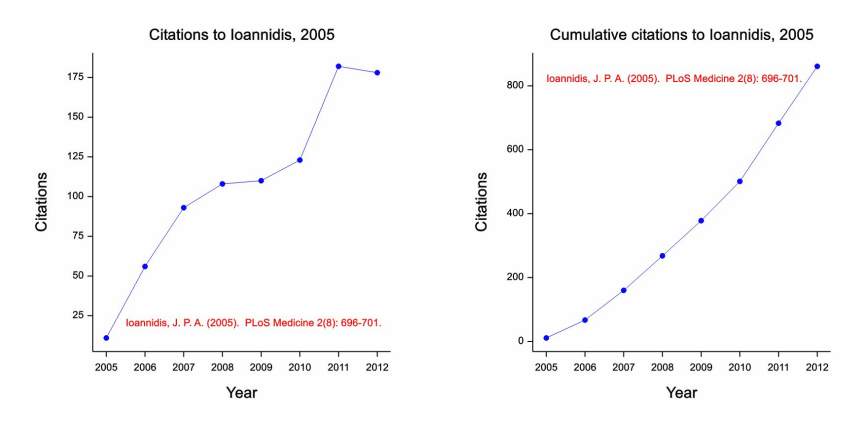
Improvements

- Mixed models replacing summary measures
- Better approaches to missing data
- More use of non-linear models
- Flexible designs

Continued bad practice

- Obsession with the log rank test
- Change from baseline
  - Percentage change from baseline
- Ordinal measures treated as categorical
- Fear of linearity
- Failure to implement decision analysis
- **Dichotomisation**
  - **Responder analysis**

Why most research findings are wrong



Source Web of Knowledge
21 Nov 2013

Information and Inference 1

A Simple Model

τ 'true' treatment effect, ϵ experimental error, Y observed effect

$$Y = \tau + \varepsilon$$

$$E[\tau] = \mu, V[\tau] = \gamma^2, E[\epsilon] = 0, V[\epsilon] = \varphi^2\ (n)$$

Here $\varphi^2(n)$ is the variance of the experimental result. It is a decreasing function of the number of observations. For example in a parallel group trial with patients allocated with equal probability to one of two groups we have $\varphi^2(n) \cong 4\frac{\sigma^2}{n}$.

The unconditional variance of Y is $V(Y) = \gamma^2 + \varphi^2(n)$ and the ratio of the variance of τ to the variance of Y is $\rho(n) = \gamma^2/\{\gamma^2 + \varphi^2(n)\}$.

Information and Inference 2
A Simple Model

Variance-Covariance Matrix

	τ	Y
τ	γ^2	γ^2
Y	γ^2	$\gamma^2 + \varphi^2(n)$

$$E[Y|\tau] = \mu + \frac{\gamma^2}{\gamma^2}(\tau - \mu) = \tau$$

$$E[\tau|Y] = \mu + \frac{\gamma^2}{\gamma^2 + \varphi^2(n)}(Y - \mu) = \mu + \rho(n)(Y - \mu) \neq Y$$

To put it another way, it does not follow from the fact that Y is an unbiased estimate of τ that on average $\tau = Y$

Consequences

- Expect to be disappointed
  - Senn's law 'you can expect to be disappointed even if you have taken account of Senn's law'
- The less information you gather the more disappointed you will be
- Reducing the cost of the information you have to gather is much more valuable than reducing the amount of information you need to gather

Flexible Designs

Good

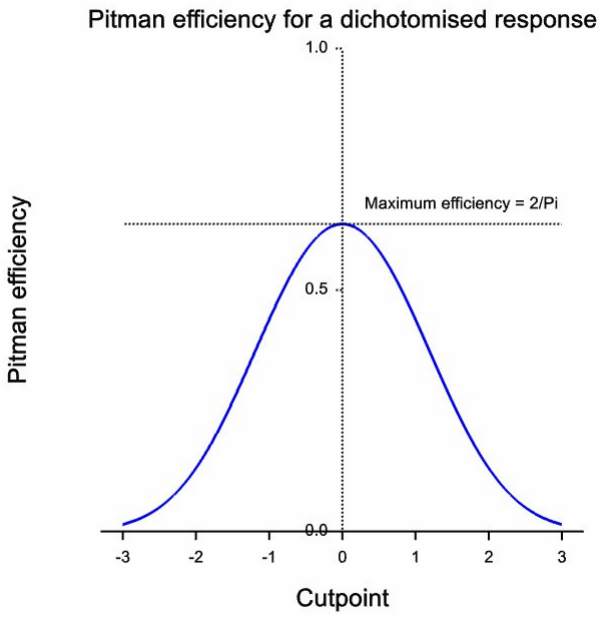
- Flexibility *per se* is always valuable
- Killing what is useless is good
  - Treatments
  - Doses
  - Measures?

Bad

- Can lead to less information on treatments eventually chosen
- What optimises power for a controlled type I error rate is bad for many other criteria
- Can lead to unnatural weighting of information

How to do better for free

- Study past trials to find out what works and what doesn't
- Build planning data-bases
- Be sensible about dose-response
- **Use original values rather than dichotomies**
- **Make more use of covariates**



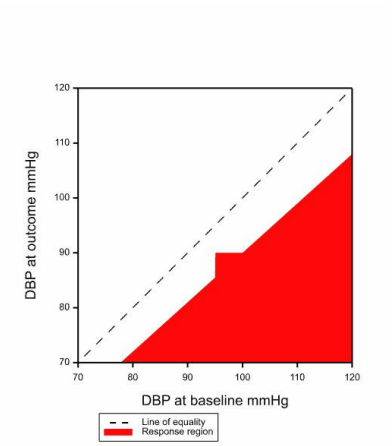
11

Typical Nonsense

The wisdom of the EMA

"Blood pressure lowering effects of anti-hypertensive therapy should be documented as the pre-/post-treatment reduction of blood pressure. As a secondary endpoint these effects can also be assessed with respect to response criteria. Arbitrarily, response criteria for antihypertensive therapy include the percentage of patients with a normalisation of blood pressure (reduction SBP < 140 mmHg and DBP < 90 mmHg) and/or reduction of SBP ≥ 20 mmHg and/or DBP ≥ 10 mmHg."

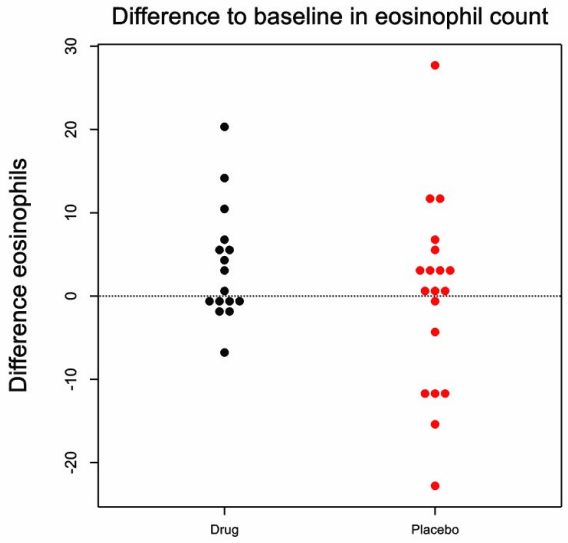
CPMP Note for guidance, 1997



12

Theory

- There is a ton of theory and a huge number of publications showing how bad this is
- Nobody pays any attention
- I am going to give you a single example
- This is not a satisfactory means of proof
- However, the proof is readily available elsewhere for anybody who wants it
- An example at least has the advantage of illustrating the problem



Wei and Zhang, *Drug Information Journal* 2001; 35: 1201-1214

Parallel group trial in asthma

Eosinophil count in sputum

Differences is baseline-outcome so high values are good

Two-sample t-test

Variate: Difference
Group fact or: Treatment

Summary

Sample	Size	Mean	Variance	Standard deviation	Standard error of mean
Drug	16	3.650	45.82	6.769	1.692
Placebo	19	-0.100	133.7	11.565	2.653

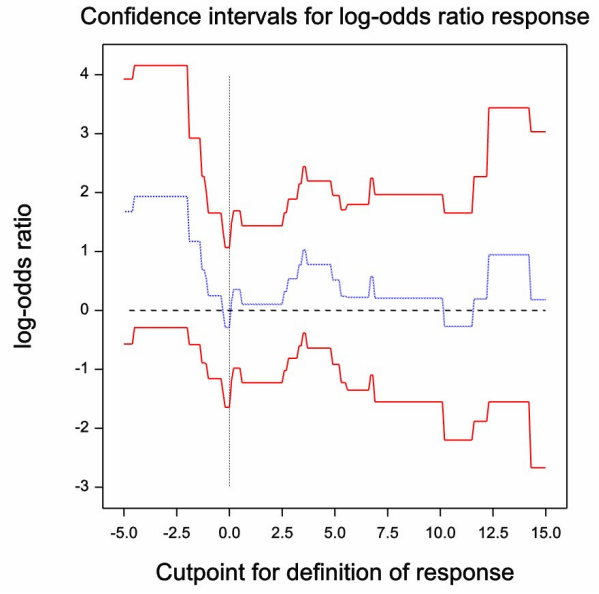
Difference of means: 3.750
Standard error of difference: 3.147

95% confidence interval for difference in means: (-2.680, 10.18)

Test of null hypothesis that mean of Difference eosinophils with Treatment = Drug is equal to mean with Treatment = Placebo

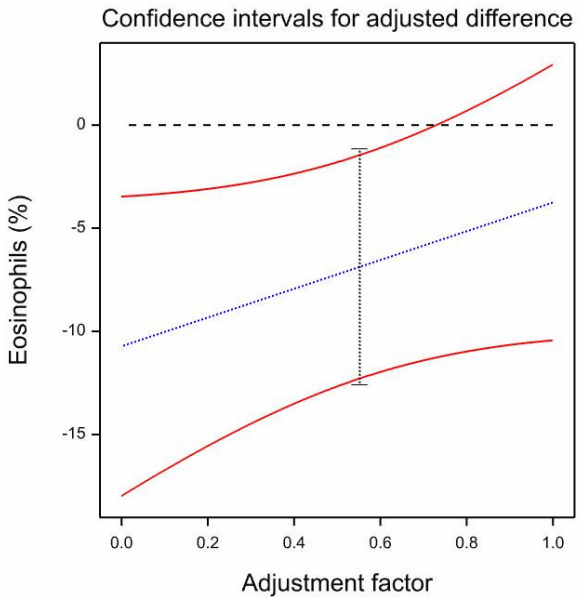
Test statistic t = 1.19 on approximately 29.72 d.f.

Probability = 0.243



Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Baseline eosinophils	1	2233.48	2233.48	35.19	<.001
+ Treatment	1	380.12	380.12	5.99	0.020
Residual	32	2030.87	63.46		
Total	34	4644.47	136.60		



Of Course…

- This is only one example
- But theory shows that by using change score rather than ANCOVA you typically increase the necessary sample size by 50%
- By dichotomising you increase it by at least a further 57%
- 1.5 x 1.57 = 2.36
- You are more than doubling your sample size !

The only people who find dichotomies easy to interpret don't understand them

Finally, a modest proposal…

- Next time you are tempted to do a responder dichotomy
- Do the power calculation both ways
  - ANCOVA
  - Dichotomy
- Then write an essay justifying the extra millions you propose to spend

Motto

Cut costs, not measures