La ciencia acompaña sus afirmaciones
de medidas de incertidumbre

Ventura

Apéndice

Numbers glossary

All examples are based on the following scenario:

In a randomized trial, 200 adults were given either DRUG or placebo for 5 years. Here's what happened:

	EXPOSED DRUG (100 adults)	CONTROL Placebo (100 adults)
Died during study	10 people	30 people

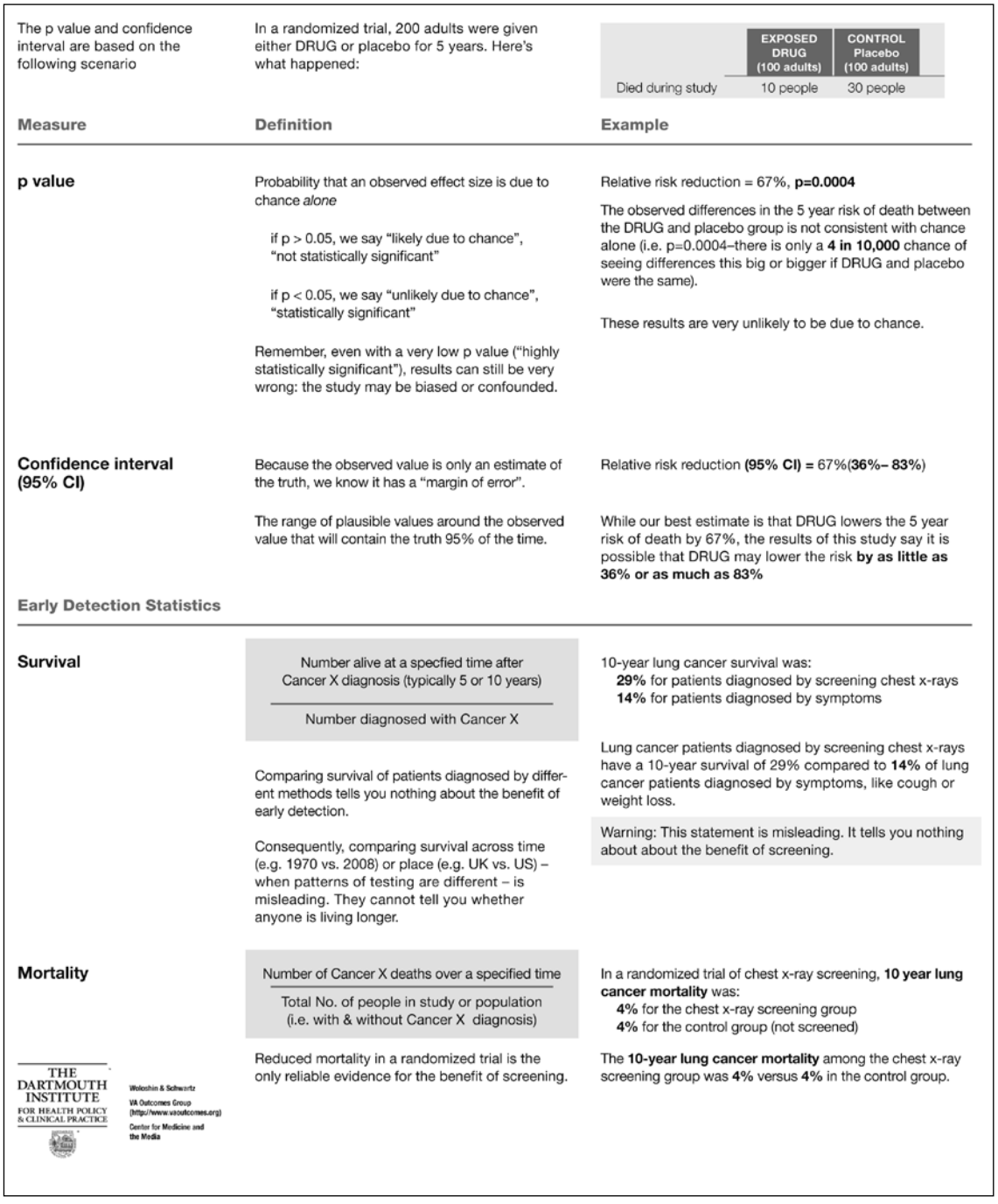
Measure	Definition	Example
Absolute risk Analogy: Price Absolute risk (*control*) is the regular price. Absolute risk (*exposed*) is the *sales* price.	Number who had outcome / Number who could have had outcome	**Absolute risk** (DRUG group) = 10/100=0.10=**10%** **Absolute risk** (Placebo group) = 30/100=0.30=**30%** Over 5 years, **10%** of the DRUG group died compared to **30%** of the placebo group. DRUG lowered the chance of dying compared to placebo: **10%** vs. **30%** died over 5 years.
Absolute risk reduction (ARR) "percentage points lower" Analogy: Savings from a sale. Subtract the sales price from the regular price.	Absolute risk (control) – Absolute risk (exposed)	**Absolute risk reduction** = 30% - 10% = **20% = 20 in 100** **For every 100 people** who take DRUG instead of placebo for 5 years, **20 fewer** people would die. DRUG lowered the chance of dying over 5 years by **20 percentage points** compared to placebo: 10% vs 30%.
Number needed to treat (NNT)	1 / Absolute risk reduction	**Number needed to treat** = 1 / 20% = 1/0.20 = **5** **5 adults** would have to take DRUG for five years to prevent **1 death**.
Relative risk (RR)	Absolute risk (exposed) / Absolute risk (control)	**Relative Risk** = 10% / 30% = 0.1/0.3 = **0.33** The DRUG group had **0.33** times the chance of dying compared to placebo: 10% vs 30% died over 5 years. The DRUG group had **one third** the deaths of the placebo group: 10% vs 30% died over 5 years.
Relative risk reduction (RRR) "% lower" Analogy: "% off" for the sale ("67% off regular price")	1 – Relative Risk	**Relative risk reduction** = 1 - 0.33 = 0.67 or **67%** DRUG reduced the chance of dying by **67%** compared to placebo: 10% vs 30% died over 5 years. DRUG lowered deaths by **two thirds** compared to placebo: 10% vs 30% died over 5 years.

Bottom Line Always report absolute risks for each group (no matter what other numbers are used).

For all risks, you need to be clear about 3 things: exactly what the outcome is (e.g. having a heart attack), over what time period the outcome occurred (e.g. 5 years) and in whom (e.g. adults with diabetes).

Steven Woloshin and Lisa Schwartz.
Center for Medicine and the Media, Dartmouth Institute for Health Policy and Clinical Practice.

Statistics glossary



The p value and confidence interval are based on the following scenario

In a randomized trial, 200 adults were given either DRUG or placebo for 5 years. Here's what happened:

	EXPOSED DRUG (100 adults)	CONTROL Placebo (100 adults)
Died during study	10 people	30 people

Measure	Definition	Example
p value	Probability that an observed effect size is due to chance *alone* if p > 0.05, we say "likely due to chance", "not statistically significant" if p < 0.05, we say "unlikely due to chance", "statistically significant" Remember, even with a very low p value ("highly statistically significant"), results can still be very wrong: the study may be biased or confounded.	Relative risk reduction = 67%, **p=0.0004** The observed differences in the 5 year risk of death between the DRUG and placebo group is not consistent with chance alone (i.e. p=0.0004–there is only a **4 in 10,000** chance of seeing differences this big or bigger if DRUG and placebo were the same). These results are very unlikely to be due to chance.
Confidence interval (95% CI)	Because the observed value is only an estimate of the truth, we know it has a "margin of error". The range of plausible values around the observed value that will contain the truth 95% of the time.	Relative risk reduction **(95% CI) =** 67%(**36%– 83%**) While our best estimate is that DRUG lowers the 5 year risk of death by 67%, the results of this study say it is possible that DRUG may lower the risk **by as little as 36% or as much as 83%**
Early Detection Statistics		
Survival	Number alive at a specfied time after Cancer X diagnosis (typically 5 or 10 years) / Number diagnosed with Cancer X Comparing survival of patients diagnosed by different methods tells you nothing about the benefit of early detection. Consequently, comparing survival across time (e.g. 1970 vs. 2008) or place (e.g. UK vs. US) – when patterns of testing are different – is misleading. They cannot tell you whether anyone is living longer.	10-year lung cancer survival was: **29%** for patients diagnosed by screening chest x-rays **14%** for patients diagnosed by symptoms Lung cancer patients diagnosed by screening chest x-rays have a 10-year survival of 29% compared to **14%** of lung cancer patients diagnosed by symptoms, like cough or weight loss. Warning: This statement is misleading. It tells you nothing about about the benefit of screening.
Mortality	Number of Cancer X deaths over a specified time / Total No. of people in study or population (i.e. with & without Cancer X diagnosis) Reduced mortality in a randomized trial is the only reliable evidence for the benefit of screening.	In a randomized trial of chest x-ray screening, **10 year lung cancer mortality** was: **4%** for the chest x-ray screening group **4%** for the control group (not screened) The **10-year lung cancer mortality** among the chest x-ray screening group was **4%** versus **4%** in the control group.

THE DARTMOUTH INSTITUTE FOR HEALTH POLICY & CLINICAL PRACTICE

Woloshin & Schwartz
VA Outcomes Group (http://www.vaoutcomes.org)
Center for Medicine and the Media

Steven Woloshin and Lisa Schwartz.
Center for Medicine and the Media, Dartmouth Institute for Health Policy and Clinical Practice.

Questions to guide your reporting

What is the finding?

What is the distinct exposure – or treatment – in each group?

If it is a lifestyle exposure (diet or exercise), how does it translate into what you have to eat or do?

What is the outcome under consideration?

If the outcome is a surrogate (cholesterol test), is it strongly linked to patient outcomes (heart attack)?

If the outcome is a composite (combining multiple components such as heart attack, stroke, or death), can you learn about the role of each component?

If the outcome is a score, (Hamilton Rating Scale for Depression) can you learn what constitutes a clinically important difference (that patients can notice) and what proportion in each group experienced it?

How big is the finding?

What is the chance of the outcome (over what time period) in each group?

Just knowing the relative risk ("0.75 times the risk") or the relative risk reduction ("25% fewer") without knowing the absolute risk is insufficient. Remember that a relative risk of 0.75 can represent an infinite number of combinations (0.003% vs. 0.004%, 3% vs. 4%, 30% vs. 40%)

Guidelines for presenting absolute risks

- Present absolute risks for each exposure group along with the time frame.
- Consider expressing absolute risks as percents (10%). This format is understandable even for decimal percents (0.5%).
- If expressing absolute risks as frequencies, DON'T use the "1 in X" format which makes comparisons hard (1 in 35 vs 1 in 56). Instead use "X in ___" like 2 in 1000.
  For the "in ___" part use multiples of 10, choosing the smallest one which makes "X" a whole number.
  Use the same "in ___" for the whole story.
- Provide context for the absolute risk.
  How dangerous is the disease? Compare absolute risk of getting to dying from disease.
  How does this risk compare to others? Compare absolute risk of dying from cancer to dying from heart disease.

What are the downsides of intervention: life threatening harms, bothersome side effects, inconvenience, costs?

When reporting on a beneficial treatment, make sure you look for associated harms. And report the absolute risks for these harms in the same format, for the same time frame, and the same dose.

Special case: Odds ratios overstate effects when outcomes are common (when the absolute risk is >20%).

Always ask: What are the absolute risks in each exposure group?

What does the finding mean?

How does the finding fit with what is already known about the topic?

Look for a systematic review.

Is the finding clinically meaningful or just "statistically significant" (i.e., p<0.05)?

Is the outcome something people directly experience or really care about? Is the effect size big or small?

Avoid the word significant. Consider using "important" for clinically meaningful and "unlikely to be due to chance" for statistical signficance.

Could the finding be wrong?

If an observational study, consider how likely it is that confounding -- differences between the people in the exposure groups — might explain the finding?

How different are the exposure groups in terms of age, sex, income, illness level, behaviors like smoking?

Did the investigators attempt to deal with confounding? How much did adjustment change the finding?

If a negative study (effect size not statistically significant), ask whether the confidence interval includes a clinically meaningful effect?

Special case: 5-year survival and screening

Improved 5-year survival for screened vs unscreened patients tells you nothing about the benefit of screening.

Bottom Line

If you can't get answers, consider skipping the story. Use numbers (and put them in tables) and highlight cautions.

Steven Woloshin and Lisa Schwartz.
Center for Medicine and the Media, Dartmouth Institute for Health Policy and Clinical Practice.

How to highlight study cautions

Setting	Suggested Language
Preliminary research (unpublished scientific meeting presentations)	These preliminary findings may change because the study has not been independently vetted through peer review [and/or] all the data are not in yet.
Inherently weak designs Animal or lab study	It takes many years to learn if the findings of animal [or lab] studies apply to people. Many promising animal [lab] studies fail to pan out in people.
Cross sectional study	Because all information was collected at the same time, you can't know if [exposure] caused [outcome], or visa versa.
Ecological studies (International comparison of dietary fat consumption vs. colon cancer mortality rate)	The study provides weak evidence connecting [exposure] and [outcome]. It shows that populations with more [exposure] have more/less [outcome]. But the study cannot tell if the people [with exposure] are the ones who actually had the [outcome].
Models (decision analysis)	The findings are based on assumptions including hypothetical relationships which may not exist.
No control group	Because everyone [took the drug / had the exposure], it is extremely hard to know if the [drug/exposure] had anything to do with the outcome.
Small study (less than 30 people)	These findings are based on a small study; larger studies are needed to really understand how much the intervention works.
Surrogate outcomes (lab test or x-ray finding)	This study measured [surrogate outcome] - a lab test/ x-ray finding - that patients do not directly experience. Be cautious about acting on these findings since changes in these kinds of measures do not reliably translate into people feeling better or living longer.
Classic designs Randomized trial	
Extrapolation	The findings may not apply to people who differ from those in the study (people with less severe disease or at lower risk for bad outcomes).
New interventions	The study only lasted a short time - [X days, weeks or months]. The balance of benefits and harms may change over a longer time period. Longer-term studies are needed.
New drugs	[Drug] is new: it was approved in [year]. As with all new drugs, we don't know how its safety record will hold up over time. In general, if there are unforeseen, serious side effects, they emerge after the drug is on the market when a large enough number of people have used the drug.
Observational studies (with a control group)	
Trial not possible (harmful exposure)	Because the study was not a true experiment, the findings may be explained by differences in the people who happened to be [exposed] rather than [drug/exposure].
Trial possible (beneficial exposure)	Because the study was not a true experiment, we cannot know whether changing [exposure] will change [outcome]. The findings may be explained by differences in the people who happened to be [exposed] rather than [drug/exposure]. A randomized trial is needed before widespread adoption of [intervention].
All studies	The benefit of [any action/intervention] should be weighed against the [side effects, inconveniences, costs, etc.].

Steven Woloshin and Lisa Schwartz.
Center for Medicine and the Media, Dartmouth Institute for Health Policy and Clinical Practice.