Ethics Based Medicine:
A modest proposal

Publishing guidelines
for clinical evidence

Please check-in to this session

You can use the QR code to the right which should also be available in written form at the entrance to the room.

You can also use this code in your app: **MS3Y3**

Disclosures

No disclosures

Content covered includes cancer screening, mortality, depression & chronic pain

Goals: ***Answer the question: what data do we need and why aren't we getting it?***

Change publication standards for research intended for clinical use.

Objectives: Clinicians and teachers of EBM will be able to

- Distinguish priorities of researchers from needs of clinicians and patients.
- Ask **2 ethical questions** re clinical data and answer with **patient-oriented language**:
  How likely is it that my patient will experience benefit or harm?
  How big is the benefit or harm?
- Look beyond reported relative risk reduction and odds ratios to **absolute risk reduction** and from P-values and mean differences to meaningful **effect sizes**.
- Consider **expected length of treatment, duration of effect and cost**
- Distinguish between **disease-modifying, preventive** and **palliative** treatments and **screening** vs **diagnostic** tests; apply appropriate statistics to each.

Evidence and Ethics

To practice ethically, clinicians must ***know*** and ***communicate*** to patients two things:

- **How likely** is the benefit or harm from an intervention?
- **How big** is the benefit or harm? (includes costs, opportunity costs)

There are three kinds of treatments:
- **Curative**
- **Preventive**
- **Palliative**

and two kinds of tests:
- **Screening**
- **Diagnostic**

As clinicians, we must insist on **publishing standards** regarding the **specific statistics** we need for each of these in order to employ the **four principles of medical ethics: non-maleficence, beneficence, autonomy, justice**

Case 1: Preventive Treatments

55 yo male and his 76 yo mother present for management of hypertension. After optimizing lifestyle, both have BP 145/88. How should we treat them?

Get them to a goal of 140 systolic?
Get their blood pressure to 120?
Leave them alone?
Something else?

How should we counsel them?

- you need to start a medicine to lower yr BP.
- I'm gonna have to start you on a medicine
- if you don't take a medicine you might get a heart attack or stroke – or even die.
- something else?

SPRINT trial 2015: treating SBP to 120 vs 140 in patients with 1 risk factor (but not DM)

Reported Results:

- **25%** reduction in cardiovascular events (including ACS, MI, CHF, stroke, CV death) (RR = 0.75, CI 0.64-0.89)

- **27%** reduction in all cause mortality. (RR = 0.73, CI 0.60 - 0.90)

- **43%** reduction in CV deaths (RR = 0.57, CI 0.38 – 0.85)
- patients over 75 yo had more benefit than those under 75.

Trial stopped early (median 3.25 yrs): it would be "unethical" to continue.

https://www.nejm.org/doi/full/10.1056/nejmoa1511939

Case 1: Preventive Tx

New York Times Nov 9, 2015.
Data on Benefits of Lower Blood Pressure Brings Clarity for Doctors and Patients

Dr. Pfeiffer, cardiologist, says he now feels obligated to lower patients' blood pressure even further, otherwise "I would have lost the opportunity to help another human being."

Drug company payments in 2013: $55,000 (propublica.org)

- Absolute risk reduction (**likelihood of benefit**) was reported as **number needed to treat** over 3.26 years:
- The difference in combined major cardiovascular events was 6.8 vs 5.2%, a difference of **1.6%** so **NNT= 61**.
- All cause mortality fell from 4.5 to 3.3% or **1.2%, NNT= 90**.
- CV deaths dropped from 1.4 to 0.8% or **0.6%, NNT= 172**

BUT:

- **No one takes BP meds for only 3.26 years**
  Chance of benefitting over 10 years? 20 years?
- Benefit reported as "greater" if over 75 (not as "less" if under 75!)
  Δ was big: RRR = **33%** vs **20%**, but ARR = **3%** vs **1%**; **NNT 33** vs **100**
  However, **expected length of treatment** is longer for younger people, so benefits for both are overstated – and same!
- ***What if patients did nothing?*** At the start, 80% either already had heart disease or >15% 10-yr risk. 3.25 yr risk >5%

https://www.nytimes.com/2015/11/10/health/data-on-benefits-of-lower-blood-pressure-brings-clarity-for-doctors-and-patients.html

What about harm?

Harms made to appear small:

Among those without Chronic Kidney Disease "*a decrease in eGFR of ≥30% to less than 60*" occurred in 3.5% in intensive treatment group vs 1.1% in standard therapy group.
"Could be reversible."

Relative increase: 318%
Occurred at about twice the rate of the benefit

Intensive therapy required, on average 3 medications; standard required 2.

Relative increase in meds: 50%

- Not reported in relative terms similar to benefits

  "Decrease in GFR" not reported as "developed CKD Stage 3";
  Cardiac events not reported as "elevated troponins," "abnormal ECG" or "reduction in ejection fraction."

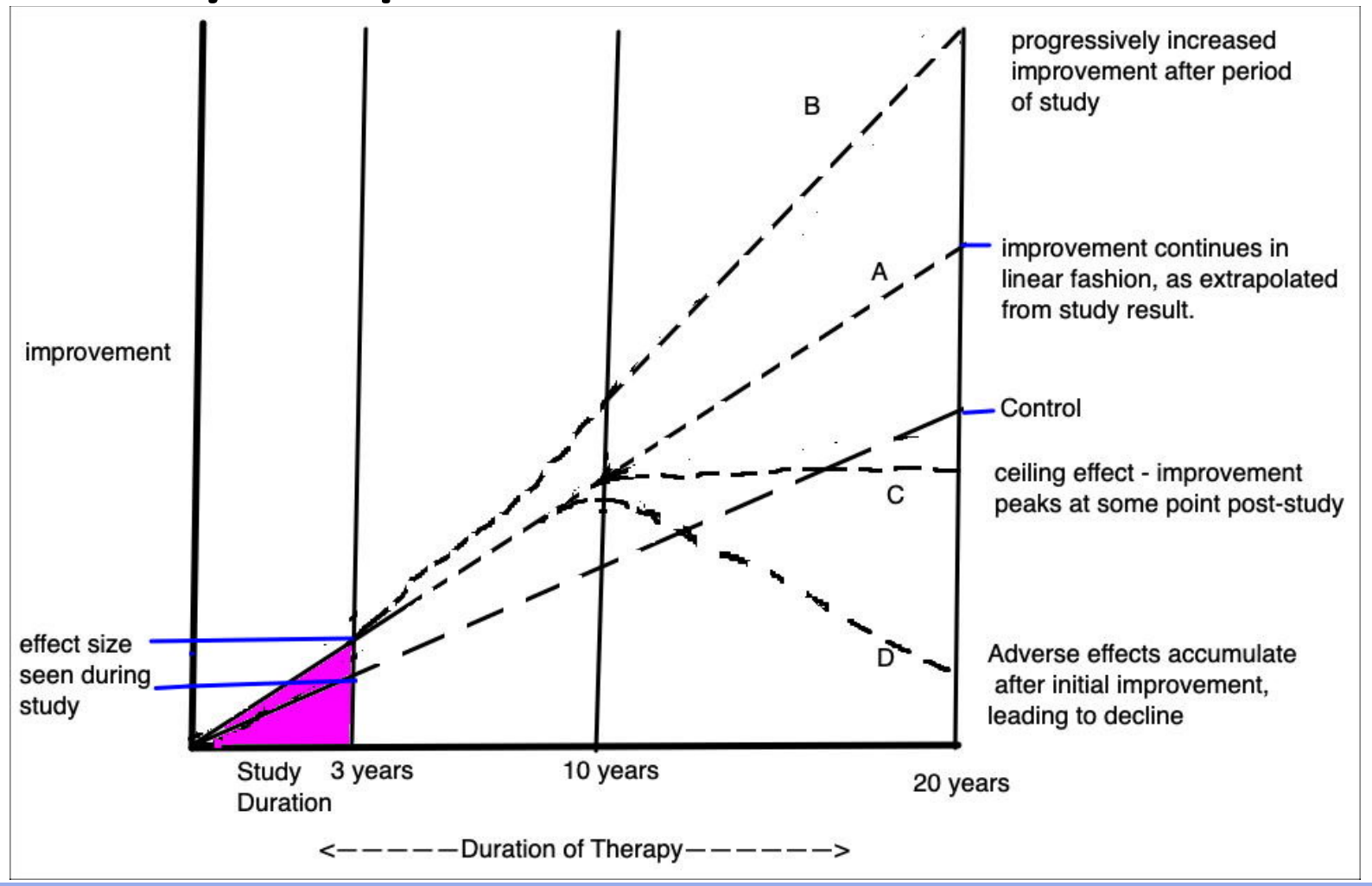
  No attempt to explain away any of the benefits.

- Authors did not comment on the 50% increase in costs to patients and the system.

I spend so much time on this paper because:
1) it is paradigmatic of methods and language
2) it is influential, cited in recent guidelines that increase the number of Americans with HTN to almost half of all adults.

https://www.ajmc.com/view/under-2017-guideline-over-105-million-americans-have-hypertension

Can you predict chance of benefit over time?



Difficulties

1. decisions based on the pink.

2. By the time long studies are completed, tests, treatments, and populations may have dramatically changed.

3. Ethical prohibitions on certain kinds of controlled trials (ie. Not treating a group of people for long periods of time)

Case 2: Disease Modifying Treatment

86 yo female diagnosed with stage 4 pancreatic cancer with multiple liver metastases. Oncologist recommends biopsy followed by dual chemotherapy. Prognosis: "With treatment she could live for another 2 years."

Alternatives offered: none
Mention of hospice: no

Study in question (NEJM, 2013):

Participants: age 27-88 with metastatic pancreatic cancer (≥1 met site).
Study design: comparison of single vs dual drug therapy.

Results:

9% of patients still alive at 2 yrs.

Median survival w dual therapy: **8.5** mo

Median survival w monotherapy: **6.7** mo
(1/3 to 2/3 fewer major side effects)

Survival without treatment: **3-6 months***

*Not reported – I had to look it up!

https://www.nejm.org/doi/full/10.1056/nejmoa1304369

Case 3: screening tests

40 yo female asks if she should start breast cancer screening. A friend's FP didn't start her till age 50, her sister's gyn started her at 45, and she heard on WNPR that she could start at 40

Q: How should we counsel her?

A: It's complicated!

This April, USPSTF issued a *draft statement* on mammography - start at 40 again, repeat every 2 years.

Rationale:
- increased invasive breast cancer in women 40-50 since 2009 guidelines (0.5% per year)
- earlier onset, more aggressive cancers, increased mortality in Black women.

- return to age 40 will prevent 1.3 additional breast cancer deaths per 1000 women screened over a lifetime, and 1.8 deaths per 1000 black women.

Reduced breast cancer death in US women by decade ranges from 12% to 33%; 3 Swedish studies show overall 15% relative risk reduction for women 40-74.

Absolute reduction not given, sources say 5-8 per 1000

www.uspreventiveservicestaskforce.org/uspstf/draft-recommendation/breast-cancer-screening-adults

Case 3: screening tests

In USPSTF 2009 guidelines, rationale for recommending women not be screened routinely before age 50 was
- absolute benefits very small
- harms significantly outweighed

USPSTF draft proposal discusses harms in the following language:

- screening biennially from age 40-74 results in an estimated **14 cases of overdiagnosis** per 1000 women over a lifetime of screening (models vary from 4 to 37)

- screening 40-50 would cause **2 of them.**

No studies show overall mortality benefit to breast cancer screening.
Why?

www.uspreventiveservicestaskforce.org/uspstf/draft-recommendation/breast-cancer-screening-adults

Case 3: screening tests

So, how should the weight of potential benefit and harm be presented to women?

A final consideration:
cost to the system, opportunity costs.
Is more screening the best way to overcome disparities in cancer treatment and survival rates?

According to best estimates:

Chance of getting breast cancer is 1 in 8
Chance of dying from it is about 1 in 40.
Alternatively: 97.5% of women won't die of breast cancer.

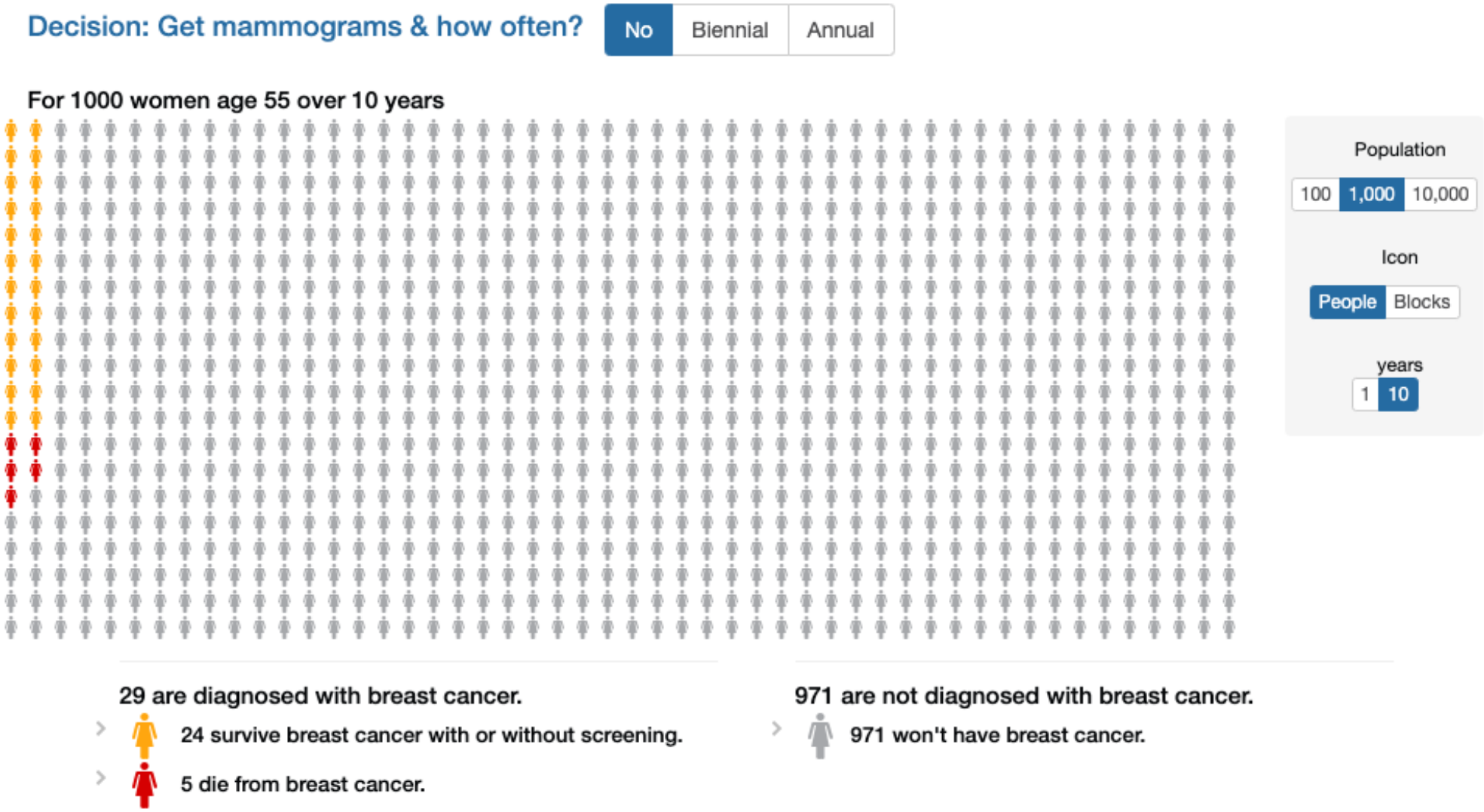
1 in 150 avoid dying of breast cancer by screening at age 50.
1 in 700 women will get this benefit by starting at 40.
99% of women will not get this benefit from screening.

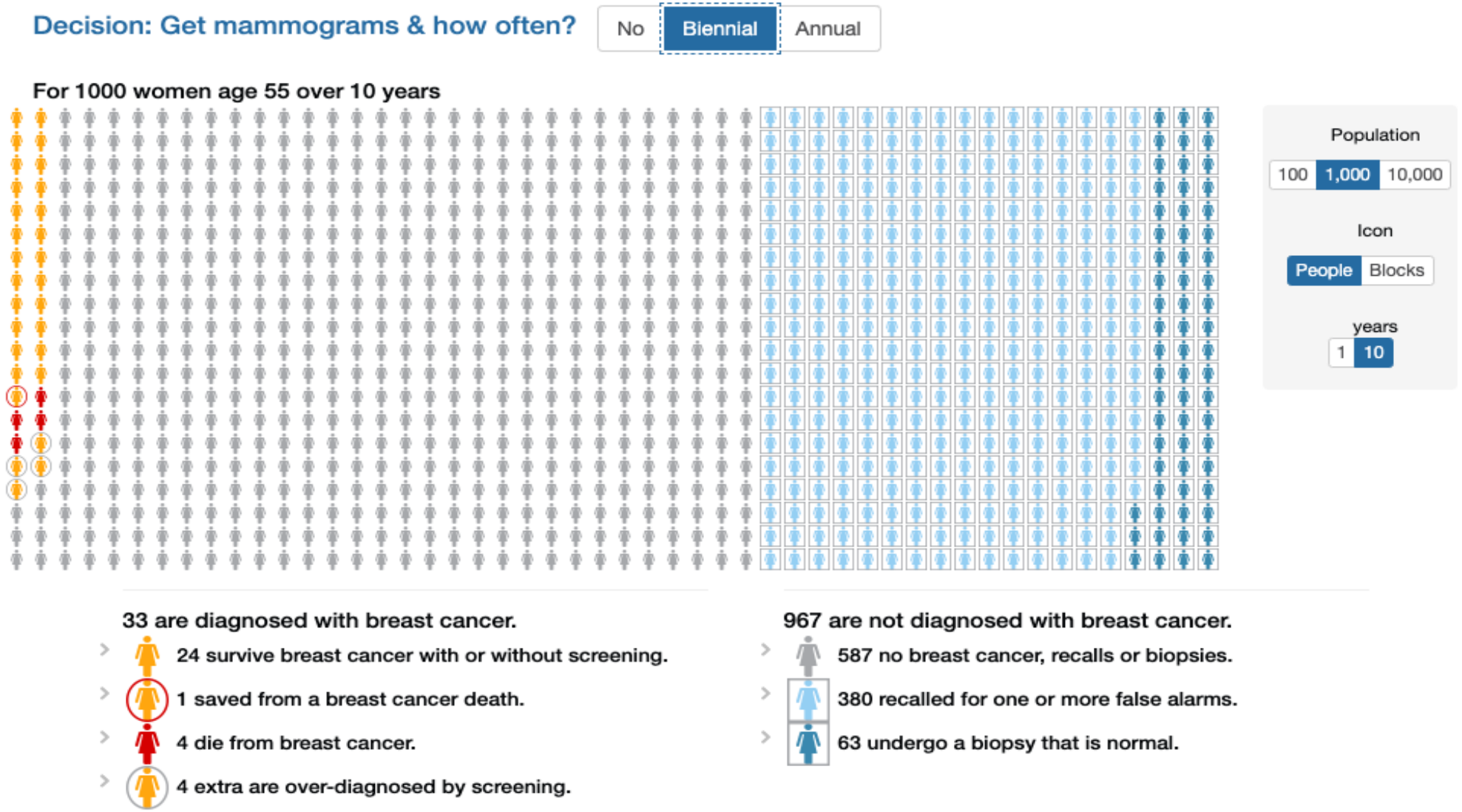
If you are Black, your risk of dying of breast cancer is higher, but your chance of benefitting from screening is less. However, 1 of 400 Black women will get this benefit by starting at 40.

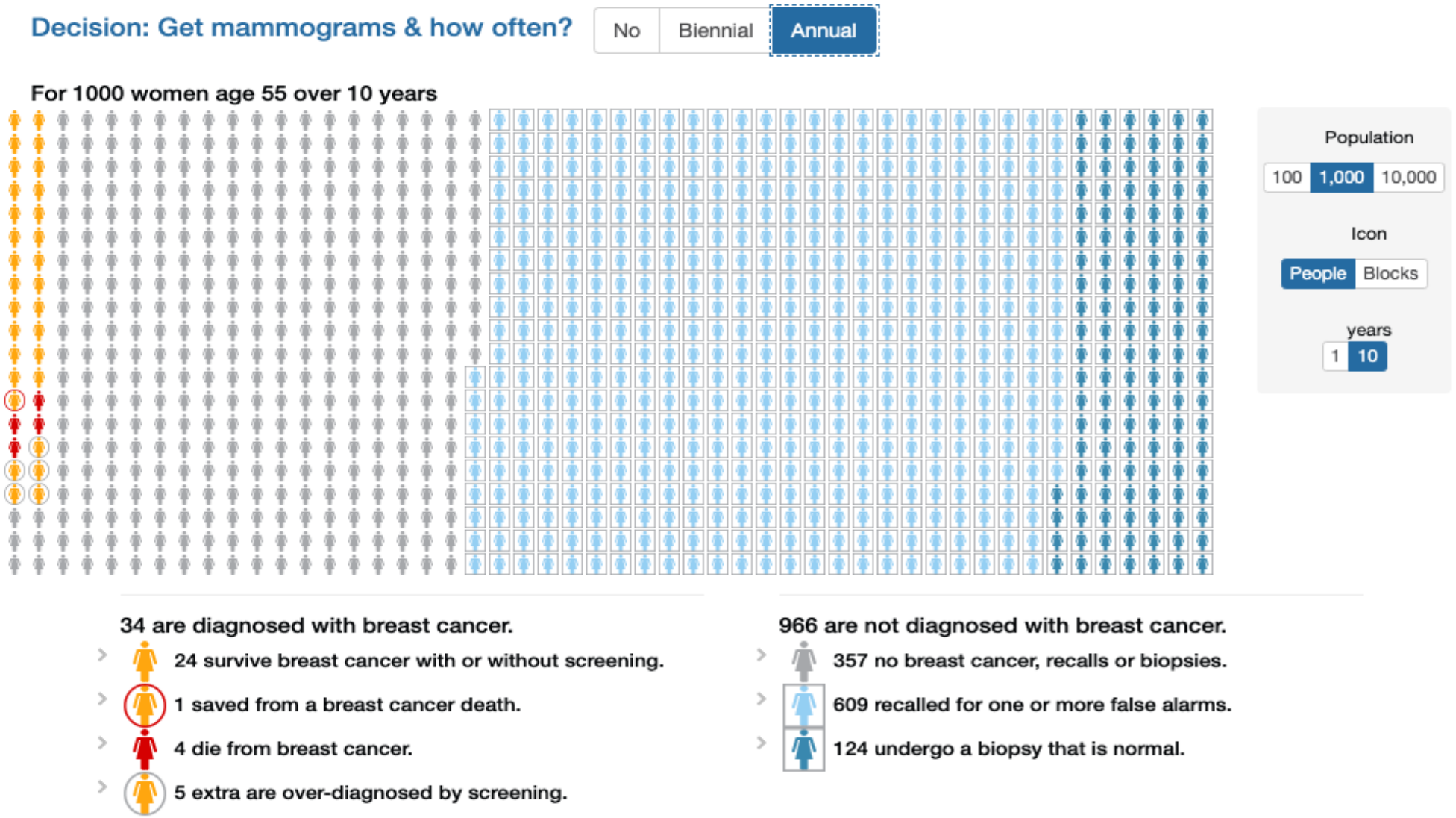
With screening, chance of being diagnosed and treated for a cancer that would not have hurt you is probably twice the chance you'll avoid dying of breast cancer.

No studies have shown that if you get screened for breast cancer you will live longer.

www.uspreventiveservicestaskforce.org/uspstf/draft-recommendation/breast-cancer-screening-adults


Decision: Get mammograms & how often?
No
Biennial
Annual
For 1000 women age 55 over 10 years
Population
100
1,000
10,000
Icon
People
Blocks
years
1
10
29 are diagnosed with breast cancer.
24 survive breast cancer with or without screening.
5 die from breast cancer.
971 are not diagnosed with breast cancer.
971 won't have breast cancer.


Decision: Get mammograms & how often?
No
Biennial
Annual
For 1000 women age 55 over 10 years
Population
100
1,000
10,000
Icon
People
Blocks
years
1
10
33 are diagnosed with breast cancer.
24 survive breast cancer with or without screening.
1 saved from a breast cancer death.
4 die from breast cancer.
4 extra are over-diagnosed by screening.
967 are not diagnosed with breast cancer.
587 no breast cancer, recalls or biopsies.
380 recalled for one or more false alarms.
63 undergo a biopsy that is normal.


Decision: Get mammograms & how often?
No
Biennial
Annual
For 1000 women age 55 over 10 years
Population
100
1,000
10,000
Icon
People
Blocks
years
1
10
34 are diagnosed with breast cancer.
24 survive breast cancer with or without screening.
1 saved from a breast cancer death.
4 die from breast cancer.
5 extra are over-diagnosed by screening.
966 are not diagnosed with breast cancer.
357 no breast cancer, recalls or biopsies.
609 recalled for one or more false alarms.
124 undergo a biopsy that is normal.

Case 4: Palliative Treatment

A 43 yo male comes with a 12 month history of fibromyalgia, but denies depression. You have read articles that say antidepressants can help.

What is the evidence?
Cochrane Review!

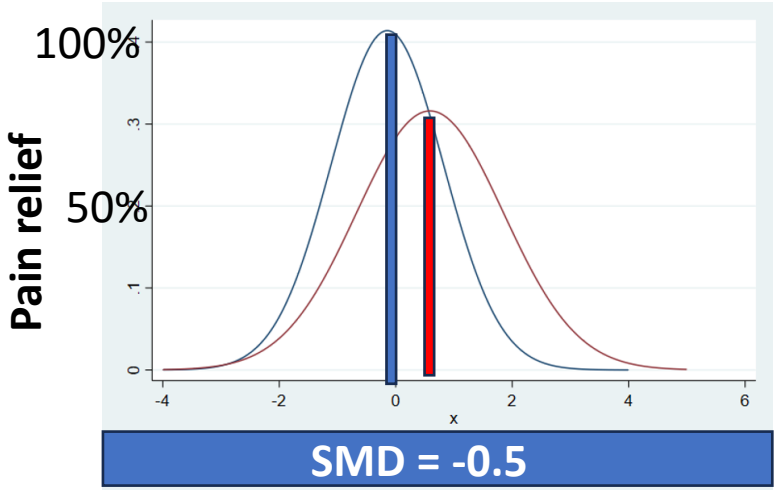
25 antidepressants, 176 trials, 28,000 patients:
only duloxetine was significantly better than placebo.

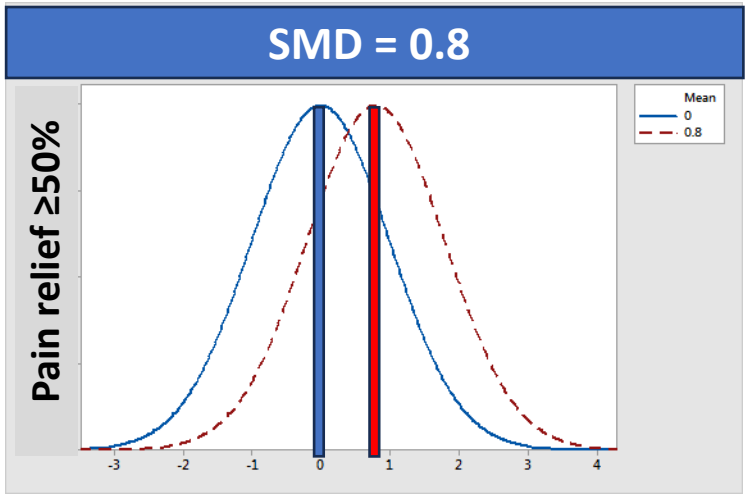
Compared to placebo, "small to moderate effect"

odds ratio 1.9 (CI 1.69-2.17) for "substantial" (≥50%) pain relief

-0.31 for "continuous pain intensity" ***standard mean difference***

What do these stats even mean?





https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD014682.pub2/full

Case 4: Palliative Treatments

Each trial used its own symptom scale, time frame, population, measure of improvement.

So how are we to understand and **use** this kind of pooled data?

The "plain language summary" provides more useful info:

For every 1000 people, **435 will get at least 50%** pain relief vs 287 for placebo.

In other words, **almost half will have their pain cut at least by half.**

(Note: *45 vs 24 per 1000 would give the same Odds Ratio*)

For effect size, mean SMD for pain intensity it's about 12% better than placebo. Is this perceptible?
But what about perceived pain?
How many get 30%, 50%, 75%, total pain relief?
Average/median pain score (out of 10) before/after medicating?

Also, especially in primary care: what about functional status?

Finally, average *duration* of studies: **10 weeks.**
Who treats chronic pain for only 10 weeks?

https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD014682.pub2/full

Case 5: Disease- altering/ Palliative Treatments

A patient reports feeling depressed. You do a PHQ-9 (SIGE-CAPS) and the score is 18. You (eventually) diagnose them with MDD and are considering starting an antidepressant.

Should you do so – and which medication would you pick?

Lancet 2018 systematic review of 21 antidepressants: 522 trials, over 100,000 patients.

Results

All drugs better than placebo in improving depression scores by ≥ 50%

Some work better than others (more likely to help and bigger effect)

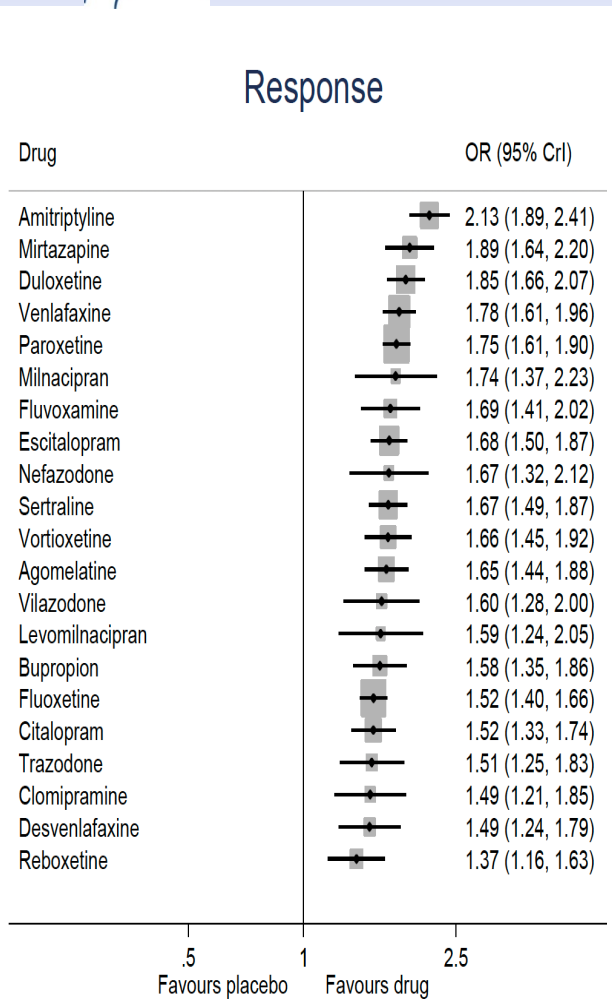
. The best only twice as likely as placebos to help

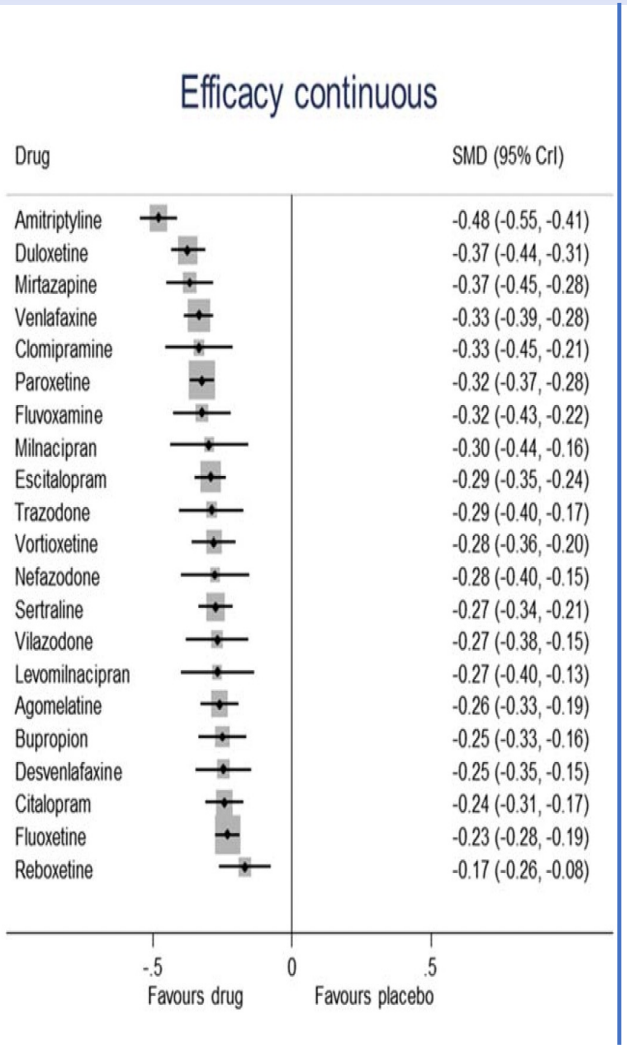
. Average difference in effect size: 0.3 SD

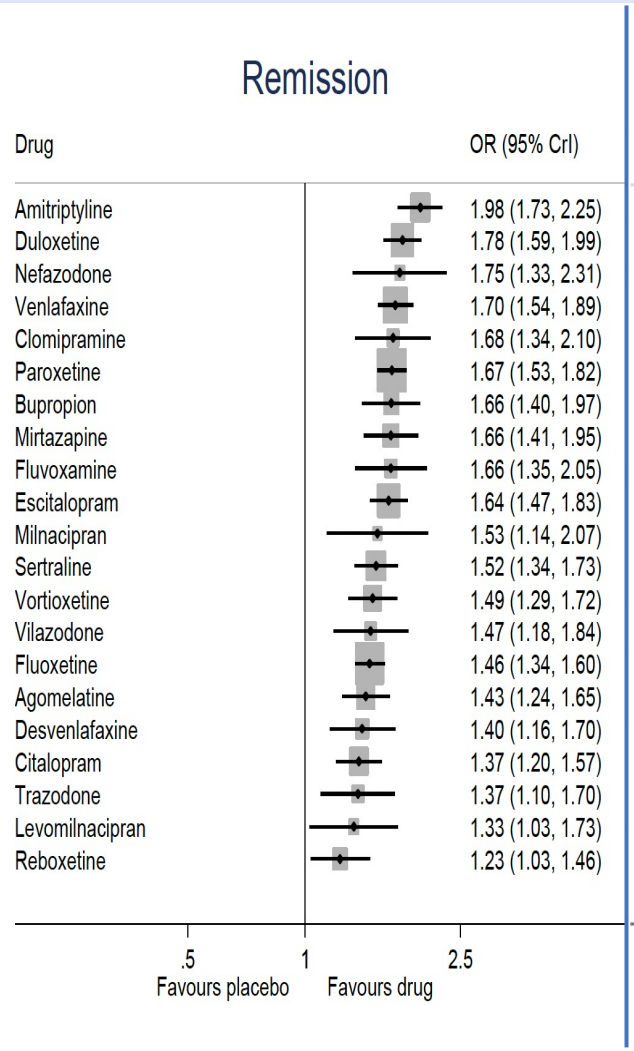
. Some better tolerated than others.

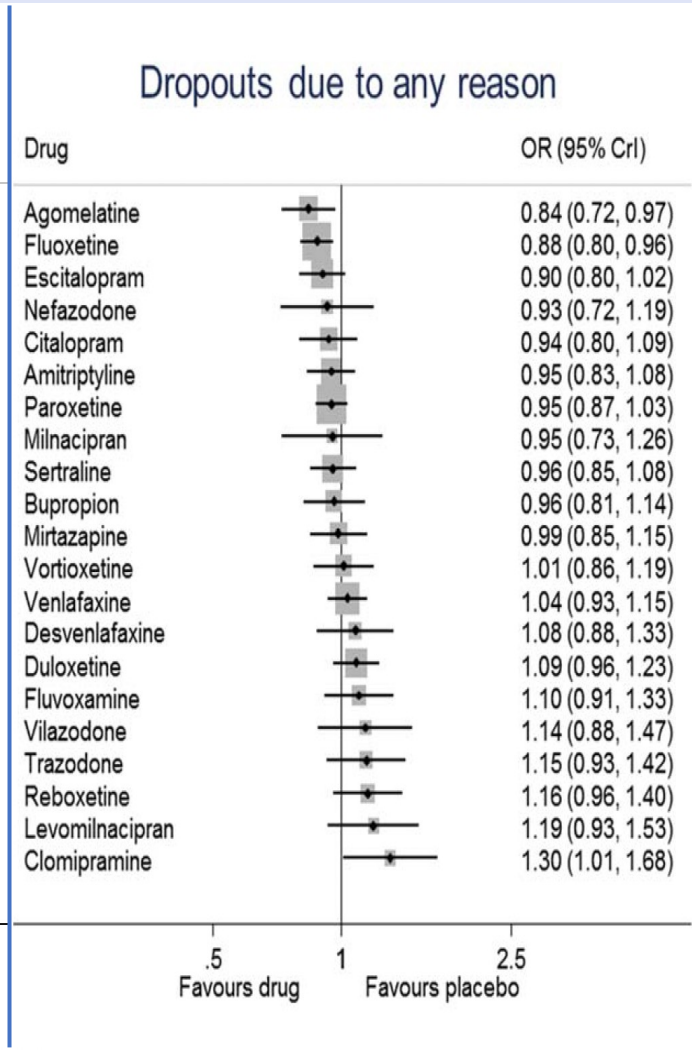
https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)32802-7/fulltext

The 44th Forum for Behavioral Science in Family Medicine









https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)32802-7/fulltext

"The world is made of stories, not atoms." – **Muriel Rukeyser**

Problems: what you'd expect (by now!)

1. Odds Ratios don't tell how many people achieve pre-determined threshold response.

2. SMDs don't tell how much symptom control drugs provide (how good was placebo?).

3. Depression scores are themselves composite outcomes of a number of symptoms that are not unique to depression and not of similar consequence. Can you compare guilt and insomnia? Appetite and suicidal ideation?

4. Average length of trial: 8 weeks.
   Average time for antidepressants to work: 4-5 weeks
   Suggested minimum duration of an initial trial of an antidepressant: 6-12 months.

https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)32802-7/fulltext

Non-Clinical vs Clinical Evidence

1. **Associations**: retrospective studies find *potential* causes.
   Odds ratios: how likely is it that someone who got condition X had exposure Y? This is a necessary but *initial* step; to establish **cause** and calculate **risk reduction** requires **prospective Randomized Control Trials**.
2. **Disease Oriented Evidence (DOE)**: studies that find changes in BP, cholesterol, etc. also represent an *initial* step. Clinical trials need to provide **Patient Oriented Evidence (POE)**: do interventions improve mortality, morbidity, QOL?
3. **Relative Risk** and **P values**. Researchers need **statistical significance** to publish: power to detect a relative difference and $P < 0.05$ to insure differences are real. Clinicians need **clinical significance**: **absolute risk** is the chance a patient will *experience* benefit or harm; **effect size** is how big that benefit or harm will be.

Reporting clinically useful, patient-centered data

General Recommendations:

- Report Absolute Risk Reductions not just Relative Risk Reductions or Odds Ratios. Account for expected length of treatment.

- Report effect sizes (absolute, not just relative) in addition to P values.

- Report harms and benefits using the same terms so they can be compared.

- Report size of benefit/harm compared to no treatment as well as placebo

- Number Needed to Treat (NNT) assumes a provider point of view; use patient-centered language to report likelihood of benefit or harm.

- Consider translating statistics into graphs or quartiles for clinical use

Clinically useful, patient-centered data

Recommendations for preventive treatments (difficult):

- Consider likelihood of benefit for *expected duration of treatment* (e.g. absolute risk reduction per 10 years, 20 years, lifetime). For patient info, give data as both a percentage and "1 in ___ people."

- Where comparison to existing or alternative treatments is the outcome, comparison to no treatment should also be stated (i.e. what if patient chooses to do nothing?)

- For composite outcomes, provide chance for most significant outcomes separately (e.g. deaths and hospitalizations cannot just be combined as a "primary outcome").
- Where disease-specific mortality is given, all cause mortality should also be stated.
- If benefits vary significantly for sub-populations, quantify differences (average, median, maximum effect)
- Where there are significant harms, report these in the same format as benefits (composite and individually) so risks and benefits can be compared directly.
- Who benefits? Note when primary benefit is for patient vs others, ie. Public health initiatives. e.g. immunizations: universal Covid & flu vaccinations mainly protect old/vulnerable; rubella mainly protects pregnant women; HPV mainly protects women

Clinically useful, patient-centered data

Recommendations for disease curing/modifying treatments:

- Chance that a patient will achieve cure or remission (i.e. response rate).
- Average, median, and maximum survival with and without treatment: where appropriate, provide info for significant sub-populations (age, sex, race).
- Frequency/magnitude of adverse effects. When adverse effects are significant (e.g. chemotherapy, surgery), consider reporting survival in years as well as QALYs.

Clinically useful, patient-centered data

Recommendations for palliative treatments

- Response rate: Chances of achieving a predetermined, clinically significant reduction in symptoms from baseline: 30% (minimum), 50%, total remission.
- Effect size: Report average, median, maximum reduction in symptoms as % of initial severity. Difference from placebo alone is not clinically useful.
- Average/median initial and final symptom scores should also be stated as a percentage of total inventory score, or inventory score should be normalized to a 1-10 scale. In reviews, “standard mean difference” alone is not useful.

Clinically useful, patient-centered data

Recommendations for Screening Tests (probably most challenging):

- Report benefits for lifetime as well as shorter screening intervals
- When applicable, report chance of reducing incidence of disease (ie. Primary prevention. Colon & cervical CA screens mostly find pre-cancerous lesions, while breast CA screening finds lesions defined and treated as cancers)
- Report estimated chance of harms: false positives requiring follow-up testing are clinically distinct from overdiagnosis/overtreatment: the latter should be reported in terms allowing direct benefit/harm comparisons
- Consider costs of screening vs not screening (ie. screening results in more cases, but earlier stage diagnoses may cost less to treat)
- Consider opportunity costs: can the same resources be used more effectively?

Clinically useful, patient-centered data

Recommendations for diagnostic tests:

- In addition to **sensitivity and specificity**, report **positive and negative predictive value** in common clinical settings and among specific patient populations using known incidence/prevalence rates.
- Report history and exam findings that increase or decrease pre-test probability
- Estimate harms of both false positive diagnosis followed by unnecessary treatment and false negative diagnosis followed by failure to treat.
- Recommend sequence of tests to most accurately establish a diagnosis

Medical Ethics: Questions

- **Non-maleficence:** are we paying enough attention to harms caused by our clinical guidelines and interventions? Are we reporting harms so they can be directly compared to benefits?
- **Beneficence:** are we measuring results in ways that provide patient-centered estimates of likelihood and magnitude of benefits?
- **Autonomy:** are we providing patients with the information they need to make informed decisions based on their values, preferences and resources?
- **Justice:** Are we using common resources wisely or spending disproportionately large sums to get incrementally better results? Are we getting equal results for all groups? Are socioeconomic causes of disease and unequal access being adequately addressed in our use of evidence?

Coda (thanks to my mentor Thomas Agresta MD, Medical Informatics specialist)

- Ultimately, practicing ethics-based medicine will need to rely on the use of Big Data to monitor long-term safety, efficacy and real world use

- Those of us who rely on evidence must insist on ongoing background monitoring programs; observational studies need funding in addition to the primary research.

- Neutral agencies need to be among those monitoring big data outputs.

- Undoubtedly, we will come to change our understanding of risks and benefits.

- The next step in Evidence/Ethics-Based Medicine is to reduce the role of chance; to find sub-groups of patients who will benefit most from specific interventions.

Session Evaluation Link

- *Session Evaluation Link*

- Please complete an evaluation of this session. There is a direct link in your conference app. You can also use this QR code.