20.1 Design Choices

We are interested in growth rates of preteen kids.
Which provides smaller SE's?

- Lots of kids, few time points for each, or
- Fewer kids, lots of time points on each.

Why?
Does that reasoning extend to more levels? Like kids within schools within cities?
Are group-level factors (uranium effect) measured with the same precision as individual ones (floor level effects)?

Meta-analysis

Analysize other analyses.
General plan:

1. Locate all studies which estimated some treatment.
   1. Publication bias might make "large p-value" studies difficult to locate.
   2. Researchers adjusted for background variables in different ways.
   3. Responses might not be the same.
2. Ask each for permission to use data.
   How could that be a problem?
3. Find a common scale and combine all data together.

Simple case: the $\hat{\theta}_i$ from the ith study has SE_i, so do a weighted average using ___ for weights. Increases sample size, provides broader scope of inference.

We're Never Happy

Why are we never satisfied with sample size n?

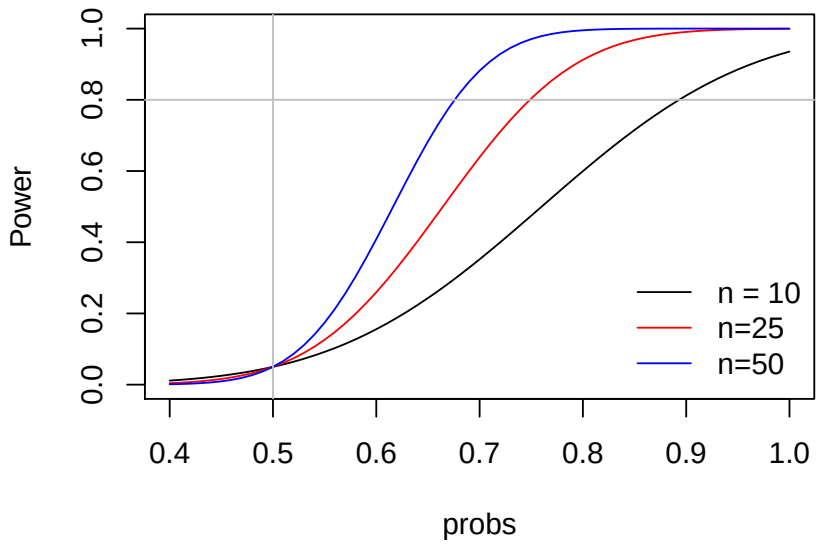
20.2 General Ideas on Power

We want a high probability of detecting a certain size effect, θ.
Restatement: need a SE smaller than half the effect size.
Choices:

- Large effect size and moderate n.
- Moderate effect size and large n.

Gelman: "Better to get conclusive effects on a subgroup than inconclusive effects with broader scope."
Power and sample size computations are inherently hypothetical.

Power plot testing $H_0 : p \leq .5$ vs $H_a : p > .5$



Rules of thumb

Cutoff is 1.96 SE's above p_0 to get $\alpha = .05$, and
0.84 SE's below p_a to get 80% power.
We need 2.8 SE's separation.

$$n \approx p_a(1 - p_a)\left(\frac{2.8}{p_a - p_0})^2\right)^2$$

$\text{SE}(\hat{p}_1 - \hat{p}_2) \leq .5\sqrt{1/n_1 + 1/n_2} \approx 1/\sqrt{n}$ for equal sample sizes.
Given a bound for $SE(\hat{p}_1 - \hat{p}_2)$ of k, $n \geq k^{-2}$.

20.3 Continuous Responses

Gelman says shift of 2.8 SE's is conservative
($\alpha = .05, 1 - \beta = .80$). Due to skew caused by ncp?
Four studies looking at effects of zinc supplement on diarrhea in kids (side effect).
Responses differ:

1. Avg count of episodes per year $\pm$ SE.
2. Avg count of episodes in 100 days (95% CI).
3. % days with diarrhea, Prevalence rate (cases/ total) 95% CI.
4. Person-Days with diarrhea / (total person days) (counts shown).

As do treatments.
Note how they compute effect sizes in days per year and SE's.

Include predictors

Reduces residual variance. Little other change.
Power computation is made based on untestable (in the pilot data) assumptions.
Do not expect "small p-value" even if computations were done properly.