Main text

1. Intro paragraph
   (a) Describe E-value and why extend to regression
   (b) If you have raw data, could choose somewhere to dichotomize and use E-value for SMD or RR
   (c) We provide E-value for cases where only standard regression info is reported
   (d) Approach involves converting regression results to SMD with well-defined effect size (say why critical for E-value): show table with RR and lower and upper RR
2. β-amyloid example
   (a) .
3. Caveats
   (a) What to do with preventative
   (b) Multiple regression
   (c) Be careful with interpretation: depends on Δ
   (d) Interpretation: E-value for a Δ-unit increase in X and any dichotomization of Y

In general, we would recommend using $\Delta = 1$ in order to assess the E-value for the effect size corresponding directly to the regression coefficient, which represents a 1-unit contrast in X. However, if the units of X are very fine-grained (e.g., if X is blood pressure in mmHg), then a 1-unit increase may not be considered clinically meaningful, and a different choice of Δ may be used (e.g., $\Delta = 10$ to represent an increase in blood pressure of 10 mmHg), which is equivalent to rescaling the regression coefficient. It is imperative to report the choice of Δ if it is not taken to be 1, since it directly impacts the size and interpretation of the E-value analog.

β-amyloid example

```
##              point     lower     upper
## RR       0.3783315 0.2887908 0.4956347
## E-values 4.7272287        NA 3.4504987
```

```
##             point     lower    upper
## RR       1.037101 0.7406488 1.452211
## E-values 1.233257 1.0000000       NA
```

```
##             point     lower     upper
## RR       0.505352 0.3366446 0.7586062
## E-values 3.370546        NA 1.9658663
```

```
##              point     lower     upper
## RR       0.6880847 0.4813375 0.9836354
## E-values 2.2649736        NA 1.1466895
```

Appendix - NEW VERSION

Cover somewhere:

*Assumptions inherited from Chinn's conversion from d to log-OR: distrubtion of Y within each outcome group is logistic, but according to Chinn, this is basically the same as normality

Converting univariable regression results to Cohen's d

Lemma 1. *As usual in OLS regression, suppose that $Y = \beta_0 + \beta_1 X + \epsilon$ with $X \amalg \epsilon$ and $E[\epsilon] = 0$. Then $\sigma_Y^2 =$*

Proof.

$$\begin{aligned}
\sigma_Y^2 &= E[\sigma_{Y|X}^2] + Var\left(E[Y|X]\right) \\
&= \sigma_{Y|X}^2 + Var\left(\beta_0 + \beta X\right) && \text{(homoskedasticity)} \\
&= \sigma_{Y|X}^2 + \beta^2 \sigma_X^2 \\
&= \sigma_{Y|X}^2 + r_{YX}^2 \sigma_Y^2 \\
\sigma_{Y|X}^2 &= \left(1 - r_{YX}^2\right) \sigma_Y^2
\end{aligned}$$

Thus, QED.

□

Suppose we estimate the conditional mean of an outcome Y given a continuous exposure of interest, X, via ordinary least squares:

$$E\big[Y \mid X\big] = \beta_0 + \beta_X X$$

where β_0 denotes the intercept and β_X denotes the estimated coefficient of X. Suppose that the effect size of interest is the increase in Y caused by a Δ-unit increase in X.

Assume that (X, Y) are bivariate normal and consider the Cohen's d associated with an increase of Δ units in X:

$$\begin{aligned}
d &= \frac{E[Y \mid X = c + \Delta] - E[Y \mid X = c]}{\sigma_{Y|X}} \\
&= \frac{\Delta \beta}{\sigma_Y \sqrt{1 - \rho_{XY}^2}} \\
&= \frac{\Delta \beta}{\sigma_Y \sqrt{1 - \frac{\beta^2 \sigma_X^2}{\sigma_Y^2}}} \\
&= \frac{\Delta \rho_{XY}}{\sigma_X \sqrt{1 - \rho_{XY}^2}}
\end{aligned}$$

where the denominator comes from a well-known property of the bivariate normal distribution.

An approximate standard error can be derived using the delta method. Let $z_{fis} = XXX$ be the Fisher-transformed correlation, which is approximately normal with variance $\frac{1}{N-3}$. Define the transformation:

$$
\begin{aligned}
g(z_{fis}) = d &= \frac{\Delta \tanh\left(z_{fis}\right)}{\sigma_X \sqrt{1 - \tanh^2\left(z_{fis}\right)}} \\
SE_d &\approx \sqrt{\text{Var}(z_{fis})}\left(g'\left(z_{fis}\right)\right) \\
&= \frac{1}{\sqrt{N-3}} \times \frac{\Delta}{\sigma_X \sqrt{\text{sech}^2(z_{fis})}} \\
&= \frac{\Delta}{\sigma_X \sqrt{(N-3)\left(1 - \rho_{XY}^2\right)}} \\
&= \frac{\Delta}{\sigma_X \sqrt{(N-3)\left(1 - \beta^2 \frac{\sigma_X^2}{\sigma_Y^2}\right)}}
\end{aligned}
$$

E-value for a univariable regression

$$
RR \approx \exp\left(0.91 \times \frac{\Delta \rho_{XY}}{\sigma_X \sqrt{1 - \rho_{XY}^2}}\right)
$$

CI:

$$
\begin{aligned}
RR_{lb} &\approx \exp\left(0.91 \times \frac{\Delta \rho_{XY}}{\sigma_X \sqrt{1 - \rho_{XY}^2}} - 1.78 \times SE_d\right) \\
&= RR_{ub} \approx \exp\left(0.91 \times \frac{\Delta \rho_{XY}}{\sigma_X \sqrt{1 - \rho_{XY}^2}} + 1.78 \times SE_d\right)
\end{aligned}
$$

Converting multivariable regression results to Cohen's d

Extend the regression model in (EQUATION) to include arbitrary measured covariates **Z**:

$$
\widehat{E}[Y \mid X, \mathbf{Z}] = \widehat{\beta}_0 + \widehat{\beta}_X X + \widehat{\boldsymbol{\beta}}_{\mathbf{Z}}' \mathbf{Z}
$$

where $\widehat{\boldsymbol{\beta}}_Z$ denotes a vector of estimated coefficients for **Z**.

$$
\begin{aligned}
R^2_{Y \sim X | Z} &= 1 - \frac{SSE_{full}}{SSE_{red}} \\
&= 1 - \frac{\sigma^2_{Y|X,\mathbf{Z}}}{\sigma^2_{Y|\mathbf{Z}}} \\
\sigma^2_{Y|X,\mathbf{Z}} &= \sigma^2_{Y|\mathbf{Z}}\left(1 - R^2_{Y \sim X|Z}\right)
\end{aligned}
$$

DOES THIS USE MVN ASSUMPTION? FLESH OUT THIS STEP. SEEMS LIKE WE ONLY NEED NORMAL ERRORS.

$$
\begin{aligned}
d &= \frac{E[Y \mid X = c + \Delta, \mathbf{Z}] - E[Y \mid X = c, \mathbf{Z}]}{\sigma_{Y|X,\mathbf{Z}}} \\
&= \frac{\Delta\beta}{\sigma_{Y|\mathbf{Z}} \sqrt{1 - R^2_{Y\sim X|Z}}} \\
&\geq \frac{\Delta\beta}{\sigma_{Y} \sqrt{1 - R^2_{Y\sim X|Z}}}
\end{aligned}
$$

Unlike in the univariable case, a simple relationship between β and $R^2_{Y\sim X|Z}$ is not available with additional distributional assumptions, so both quantities are needed to approximate d.

Note that you can use partial correlation or partial Rˆ2 from regression table.

Standard error: Going to be hard because inference for $R^2_{YX|Z}$ won't be available. For an approximation, can use $Var\left(z^f_{YX|Z}\right) \approx Var\left(z^f_{YX}\right) = \frac{1}{N-3}$

- describe what to do for preventive: just set delta to be negative?
- say that, for univariable case, if SD of X isn't available, can set Delta = SD of X

References