efficient science + meta-analysis

Bayes comes in, and p-values are out:

especially for frequentists!

Accumulation Bias in meta-analysis: the need to consider *time* in error control

Judith ter Schure, Peter Grünwald

INTRO

- Processes of **study accumulation** $T \geq t$ and **meta-analysis timing** $\mathscr{A}^{(t)}$ define *meta* stopping rules.
- These processes should receive attention in recommendations to reduce *research waste*.
- Resulting biases have been recognized before, but were not confronted, and effects on testing were never investigated.

RESULTS

- Not possible to define valid tests based on **p-values** due to nonnormal but unknown sampling distributions; could for example partly look like: ***Gold Rush Accumulation Bias →***
- Possible to define tests based on **likelihood ratios**:

$$\mathbf{LR}_{10}^{(t)}\left(z_1,\ldots,z_t,\mathscr{A}^{(t)},T\geq t\right) := \frac{f_1\left(z_1,\ldots,z_t\right)\cdot \mathbf{P}_1\left(\mathscr{A}^{(t)},T\geq t \mid z_1,\ldots,z_t\right)}{f_0\left(z_1,\ldots,z_t\right)\cdot \mathbf{P}_0\left(\mathscr{A}^{(t)},T\geq t \mid z_1,\ldots,z_t\right)}$$

DISCUSSION

Table 3. Possible 2001 state of a database of study series per topic, visualizing what study series are taken into account in the two approaches to error control: conditional on time (blue and grey) and surviving over time (orange).

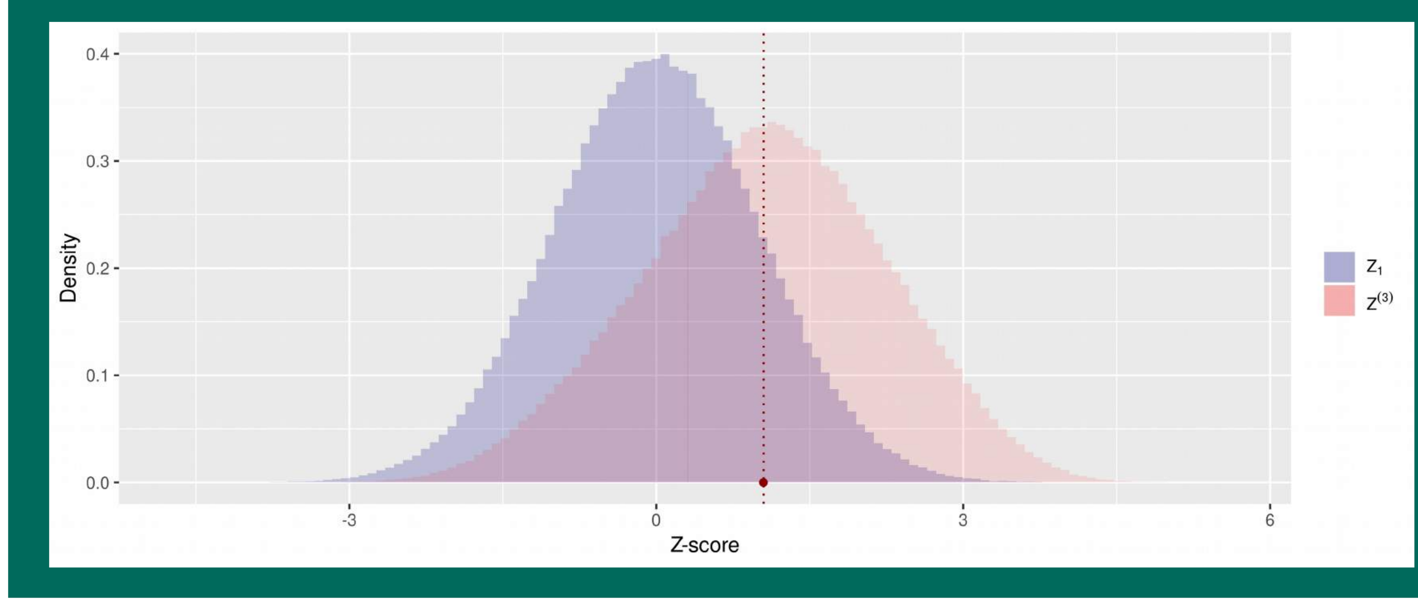
Study series size (t)	Topics: 1	2	3	4	5	6	7	8	9	10	…	9 998	9 999	10 000
1	$z_{1,1}$	$z_{1,2}$	$z_{1,3}$	$z_{1,4}$	$z_{1,5}$	$z_{1,6}$	$z_{1,7}$	$z_{1,8}$	$z_{1,9}$	$z_{1,10}$	…	$z_{1,9998}$	$z_{1,9999}$	$z_{1,10000}$
2	$z_{2,1}$	$z_{2,2}$	$z_{2,3}$	$z_{2,4}$	$z_{2,5}$		$z_{2,7}$	$z_{2,8}$		$z_{2,10}$		$z_{2,9998}$		$z_{2,10000}$
3	$z_{3,1}$	$z_{3,2}$	$z_{3,3}$		$z_{3,5}$		$z_{3,7}$			$z_{3,10}$		$z_{3,9998}$		$z_{3,10000}$
4		$z_{4,2}$	$z_{4,3}$		$z_{4,5}$		$z_{4,7}$					$z_{4,9998}$		$z_{4,10000}$
5		$z_{5,2}$			$z_{5,5}$							$z_{5,9998}$		
6		$z_{6,2}$			$z_{6,5}$							$z_{6,9998}$		
…												…		
136												$z_{136,9998}$		

Two approaches:

Error control conditioned on *t*: *(conventional meta-analysis)*
→ needs prior odds
→ composite H0/H1: future work

Error control 'surviving' over *t*: *(living systematic reviews)*
→ 'universal bound'/Robbins (1970)
→ composite H0/H1: Safe Tests (Grünwald et al., 2019)

Gold Rush Accumulation Bias: *more studies after significance*
→ the larger the series, the larger the fraction of significant studies



- ***Analysis time probabilities***

$$A(t \mid z_1,\ldots,z_t) := \mathbf{P}\left[\mathscr{A}^{(t)} \,\middle|\, T\geq t, z_1,\ldots,z_t\right] \cdot S(t-1 \mid z_1,\ldots,z_{t-1}),$$

where we define

$$S(t-1\mid z_1,\ldots,z_{t-1}) := \mathbf{P}[T\geq t \mid z_1,\ldots,z_{t-1}] = \prod_{i=0}^{t-1}(1-\lambda(i\mid z_1,\ldots,z_i))$$

$$\lambda(t\mid z_1,\ldots,z_t) := \mathbf{P}[T=t \mid T\geq t, z_1,\ldots,z_t]$$

- **Conventional meta-analysis**

$$f_0\left(z^{(t)} \,\middle|\, \mathscr{A}^{(t)}, T\geq t\right) = \frac{f_0(z^{(t)})\cdot \mathbf{P}_0\left[\mathscr{A}^{(t)}, T\geq t \,\middle|\, z^{(t)}\right]}{\mathbf{P}_0[\mathscr{A}^{(t)}, T\geq t]} = \frac{f_0(z^{(t)})\cdot \bar{A}_0\left(t \,\middle|\, z^{(t)}\right)}{\bar{A}_0(t)},$$

where we define:

$$\bar{A}_0\left(t\,\middle|\, z^{(t)}\right) := \mathbf{E}_0\left[A(t\mid Z_1,\ldots,Z_t) \,\middle|\, Z^{(t)}=z^{(t)}\right]$$
$$\bar{A}_0(t) := \mathbf{E}_0[A(t\mid Z_1,\ldots,Z_t)],$$

- **Living systematic reviews**

$$f_0\left(z^{(1)},\ldots,z^{(t)}, T=t\right) = f_0\left(z^{(1)},\ldots,z^{(t)}\right)\cdot \mathbf{P}_0\left[T=t \,\middle|\, z^{(1)},\ldots,z^{(t)}\right],$$

$$\mathbf{P}_0\left[T=t \,\middle|\, z^{(1)},\ldots,z^{(t)}\right] := \mathbf{E}_0\left[S(t-1 \mid Z_1,\ldots,Z_{t-1}) \,\middle|\, Z^{(1)}=z^{(1)},\ldots\right] - \mathbf{E}_0\left[S(t\mid Z_1,\ldots,Z_t)\,\middle|\, Z^{(1)}=z^{(1)},\ldots\right],$$

- **Posterior odds error control conditioned on time *t***

$$\frac{1-\mathbf{P}_1\left[\mathscr{E}^{(t)}_{\text{TYPE-II}} \,\middle|\, \mathscr{A}^{(t)}, T\geq t\right]}{\mathbf{P}_0\left[\mathscr{E}^{(t)}_{\text{TYPE-I}} \,\middle|\, \mathscr{A}^{(t)}, T\geq t\right]} = \frac{\mathbf{P}_1\left[O_{\text{post}}\left(Z_1,\ldots,Z_t \mid \mathscr{A}^{(t)}, T\geq t\right)\geq \gamma\cdot\frac{\pi}{1-\pi}\right]}{\mathbf{P}_0\left[O_{\text{post}}\left(Z_1,\ldots,Z_t \mid \mathscr{A}^{(t)}, T\geq t\right)\geq \gamma\cdot\frac{\pi}{1-\pi}\right]} = \frac{\mathbf{P}_1\left[O_{\text{post}}(Z_1,\ldots,Z_t)\geq\gamma\cdot\frac{\pi}{1-\pi}\right]}{\mathbf{P}_0\left[O_{\text{post}}(Z_1,\ldots,Z_t)\geq\gamma\cdot\frac{\pi}{1-\pi}\right]} = \frac{\mathbf{P}_1[\mathscr{R}]}{\mathbf{P}_0[\mathscr{R}]} \geq \frac{\mathbf{P}_1[\mathscr{R}]}{\mathbf{P}_1[\mathscr{R}]\cdot\frac{1}{\gamma}} = \gamma,$$

- **Likelihood Ratio error control 'surviving' over time *t***

$$\mathbf{P}_0\left[\text{there exists } t\leq T \text{ with } \mathscr{E}^{(t)}_{\text{TYPE-I}} \text{ and } \mathscr{A}^{(t)}\right] = \mathbf{P}_0\left[\exists t\leq T:\ \mathscr{E}^{(t)}_{\text{TYPE-I}}; \mathscr{A}^{(t)}\right] = \mathbf{P}_0\left[\exists t\leq T:\ \mathbf{LR}_{10}^{(t)}(Z_1,\ldots,Z_t)\geq\frac{1}{\alpha}; \mathscr{A}^{(t)}\right] \leq \mathbf{P}_0\left[\exists t>0:\ \mathbf{LR}_{10}^{(t)}(Z_1,\ldots,Z_t)\geq\frac{1}{\alpha}\right]\leq\alpha$$

- ***Gold Rush* Accumulation Bias**

$$\omega^{(t)}_{\text{S}} = \omega_{\text{S}} := \mathbf{P}\left[T\geq t+1 \,\middle|\, T\geq t, Z_t\geq z_{\frac{\alpha}{2}}\right] = 1$$
$$\omega^{(t)}_{\text{X}} = \omega_{\text{X}} := \mathbf{P}\left[T\geq t+1 \,\middle|\, T\geq t, Z_t\leq -z_{\frac{\alpha}{2}}\right] = 0$$
$$\omega^{(t)}_{\text{NS}} = \omega_{\text{NS}} := \mathbf{P}\left[T\geq t+1 \,\middle|\, T\geq t, |Z_t| < z_{\frac{\alpha}{2}}\right] = 0.02.$$

$\lambda(0)=0$ and for all $i\geq 1$, $\lambda(i)$ is defined as follows:

$$\lambda(i\mid z_i,\alpha) = 1-\left(\omega^{(i)}_{\text{S}}\cdot \mathbb{1}_{z_i\geq z_{\frac{\alpha}{2}}} + \omega^{(i)}_{\text{NS}}\cdot\mathbb{1}_{|z_i|<z_{\frac{\alpha}{2}}}\right)$$
$$\bar{\lambda}_0(i\mid\alpha) := \mathbf{E}_0[\lambda(i\mid Z_i;\alpha)] = 1-\left(\omega^{(i)}_{\text{S}}\cdot\frac{\alpha}{2} + \omega^{(i)}_{\text{NS}}\cdot(1-\alpha)\right).$$

CWI
Centrum Wiskunde & Informatica

Take a picture to **download** the **full paper**