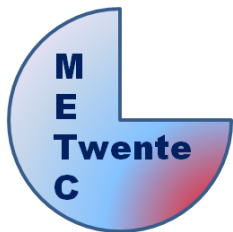


# Voortijdig stoppen van studies

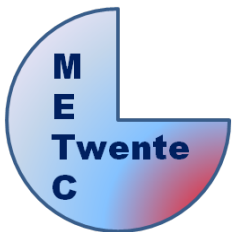
Job van der Palen

METC Twente



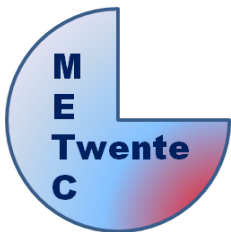
# Waarom voortijdig stoppen?

- Veiligheid
- Effectiviteit
- Futiliteit
- Andere redenen



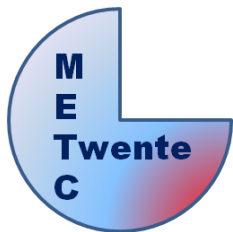
# Veiligheid (1)

- Data and Safety Monitoring Board (DSMB)
  - Onafhankelijk!
  - Samenstelling: vaak twee artsen, één statisticus/epidemioloog
  - Samenstelling en charter goedgekeurd door METc
- Studies met (sterk) verhoogd risico
  - met uitzondering van fase I geneesmiddelen trials
- Onderzoeker rapporteert periodiek aan DSMB
  - Frequentie hangt af van risico, duur en omvang studie: op maat
    - Bv éénmaal halverwege
    - Na elke 200 patiënten
    - Na elke patiënt



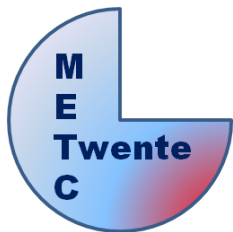
# Veiligheid (2)

- DSMB evalueert en geeft advies aan hoofdonderzoeker over voortzetting of aanpassing studie
- Bij aanpassing studie of stoppen: via amendement of melding beëindiging studie melden aan METc
- Als onderzoeker advies DSMB niet opvolgt: melden aan METc
- Praktisch punt: Het is moeilijk mensen te vinden voor DSMB's
  - Vooral epidemiologen....



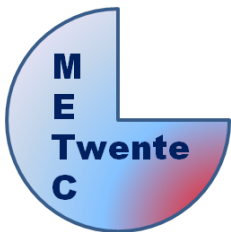
# Veiligheid (3)

- SAEs en SUSARs in elk investigator initiated onderzoek verplicht bij oordelende METc melden via ToetsingOnline
- Indien veel SAE's worden verwacht (bv overlijdens bij kankerstudies) kunnen SAE's ook via line listings gemeld worden
- METc kan studie on hold zetten of stoppen



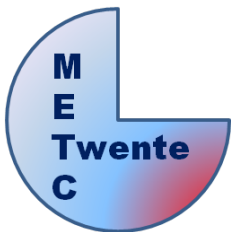
# Waarom voortijdig stoppen?

- Veiligheid
- **Effectiviteit**
- Futiliteit
- Andere redenen



# Effectiviteit

- Interim analyse door onderzoeker
- Onverwacht groot effect van interventie
- Onderzoeker wil studie stoppen
  - “Doorgaan niet ethisch”
  - “Onthouden van effectieve behandeling van controlegroep”
- Mag onder voorwaarden
  - Interim analyse voor effectiviteit in protocol opgenomen
  - Stopregels helder



# Probleem van interim analyse

- Hebben behandeling A en B hetzelfde effect?
- $H_0$ : effect A = effect B (geen verschil)
- Bij statistische toets met significantieniveau  $\alpha$  (meestal 5%).
  - Kans van 5% om ten onrechte  $H_0$  te verwerpen
  - Foute conclusie: Effect A  $\neq$  Effect B
- Interim analyses veranderen de totale kans om minstens één keer de  $H_0$  ten onrechte te verwerpen (als alle nulhypothesen waar zijn)
  - Overall significantieniveau verandert!



# Bij $x$ onafhankelijke toetsen, elk met $\alpha = 0,05$

- Kans om  $\geq 1$  keer ten onrechte een  $H_0$  te verwerpen: *overall*  $\alpha = 1 - (0,95)^n$

- Aantal toetsen                      overall  $\alpha$

3	0,143
---	-------

10	0,401
----	-------

100	0,994
-----	-------

- Interim analyse en eind analyse zijn niet onafhankelijk van elkaar
- Maar: overall  $\alpha$  moet 0,05 blijven
- 3 veelgebruikte methoden

R	Interim analyse nummer	O'Brien-Fleming $\alpha$	Haybittle-Peto $\alpha$	Pocock $\alpha$
2	1	0.0054	0.002	0.0294
	2	0.0492	0.0500	0.0294
3	1	0.0006	0.0010	0.0221
	2	0.0151	0.0010	0.0221
	3	0.0471	0.0500	0.0221

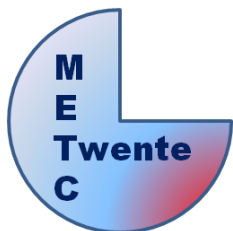
- Pocock:
  - Grootste kans op voortijdig stoppen; aan einde grote “straf” in p-waarde
  - (Veel) meer mensen nodig
- Haybittle-Peto & O'Brien-Fleming:
  - Grootste kans op significantie aan eind; erg moeilijk voortijdig te stoppen
  - Niet (veel) meer mensen nodig (nagenoeg ‘gratis’)

# Interim analyse in protocol opnemen!

- Aangeven wanneer (bv na  $\frac{1}{3}$ ,  $\frac{2}{3}$  aantal patiënten en einde studie)
- Berekening aantal patiënten moet uitgaan van laatste  $\alpha$

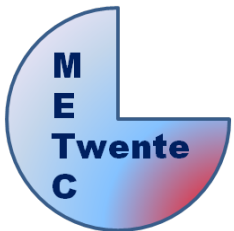
Als je aan wilt tonen dat therapie A 5% succes heeft en B 10%

- Eén analyse aan het eind: n=870
- Twee interim analyses en eindanalyse
  - Pocock: N=1086
  - O'Brien-Fleming: N=884
  - Haybittle-Peto: N=870



# Fout voorbeeld interim analyse

- Behandeling met AZT bij AIDS

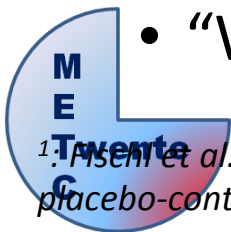


# Analyses en aantallen (1)

- Effect van Azidothymidine (AZT) bij behandeling AIDS <sup>1</sup>

	AZT	Placebo	RR	p-waarde
Sterfgevallen	1/145 (0.7%)	19/137 (13.9%)	0.07	<0.001

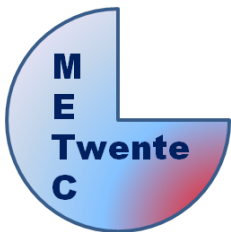
- Conclusie: AZT werkt, data zijn overtuigend
- Toch werkt AZT niet zo goed. Aanwijzingen destijds:
  - RR extreem laag
  - Trial is te vroeg gestopt. Niet a-priori in protocol opgenomen
- Geplande FU was 24 weken
- “When the study was terminated, 27 subjects had completed 24 weeks”



<sup>1</sup>: Fischl et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *N Engl J Med* 1987;317:185-91

# Waarom voortijdig stoppen?

- Veiligheid
- Effectiviteit
- **Futiliteit**
- Andere redenen



# Futiliteit

- Stoppen wegens gebrek aan effectiviteit
- Op basis van interim analyse(s)
- Vooraf in protocol opnemen: welk (klein) verschil vinden we niet meer relevant?
- Als grootte van het effect onbekend is
  - Voorlopige powerberekening in protocol
  - Op basis van interim analyse definitieve berekening doen en beslissing nemen

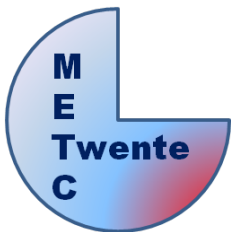
# Voorbeeld

- **Verwachting: behandeling A is beter dan B**
  - Behandeling A: 25% remissie; behandeling B: 15% remissie
  - Minimaal interessant verschil: 5% (arbitrair bepaald)
- **Aantal nodig: 2 groepen van 250**
- **Interimanalyse na bv 2\*50 patiënten**
  - Behandeling A: 21% remissie; behandeling B: 17% remissie
  - Aantal nodig: 2 groepen van 1500
- **Studie stoppen. Verschil in effectiviteit < 5% en aantal (waarschijnlijk) onhaalbaar.**



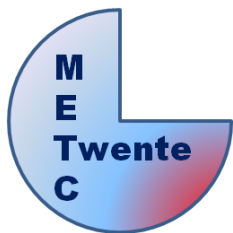
# Waarom voortijdig stoppen?

- Veiligheid
- Effectiviteit
- Futiliteit
- Andere redenen



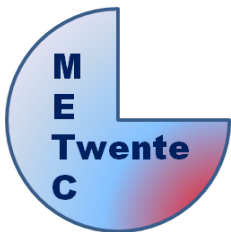
# Anders (1): Onderzoekers ingehaald door andere studies/werkelijkheid

- Nieuwe publicaties over zelfde middel. Doorgaan onethisch?
- Verandering in vergoedingen (bv stoppen-met-roken studies)
- E-cigarettes zijn dodelijk voor stoppen-met-roken studies gebleken



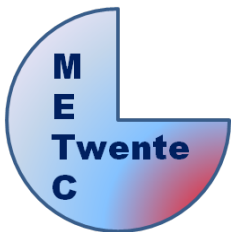
# Anders (2): Studie duurt te lang

- Personele redenen
  - Vertrek personeel, bv klaar met opleiding
    - METc: let hier op! Co-assistenten project? Max 20 weken, inclusief METc, analyseren en schrijven!
- Geld is op/financiering stopt



# Anders (3): Problemen met inclusiesnelheid

- Is inclusiesnelheid realistisch?
  - ..... Meestal niet ☹️
  - Vaak: te strenge in- en exclusie criteria (moeten we als METc op letten!)
    - Bv Depressie als exclusie criterium
      - aangetoond met een diagnose?
      - behandeld hiervoor?
      - zelfgerapporteerd?
  - Angst (of luiheid) om multi-center onderzoek te doen
  - Concurrerende studies: moeten we als METc op letten!
  - In het hoofdcentrum gaat het vaak goed; in andere centra (veel) minder



# Anders (4):

- Voorbeelden uit de zaal!
  - Waar moeten we als METc nog meer op letten?

