**Interim Analysis**

clinical trials are longitudinal and takes lots of time to complete

impossible to enroll all patients at the same time

interim analysis at primary endpoint data prior to all patients being enrolled or completing treatment to reevaluate or for adaptive design

* early efficacy = convincing evidence of benefit of the new product
* safety concerns = evidence of harm of the new product
* futility = change of conditional power of significant beneficial effect by the end of the study is small given the observed data
* re-estimate final sample size required to yield adequate power to obtain significant result

**Group Sequential Design**

data is analyzed at regular intervals

= # of total planned analysis

after every patients are enrolled and followed for 30 days, perform an interim analysis on all patients followed cumulatively

if there is significant treatment difference at any point, stop the trial early for superior/inferior efficacy of new drug

Issues with Multiple Testing

inflated Type I error if use at each interim

interim analysis shouldn’t be performed using the family-wise error rate

compare to preset efficacy significant levels at each interim

data at each interim analysis contains data from previous interims, so aren’t independent

Pocock Tests

constant Z critical values across all stages to maintain overall

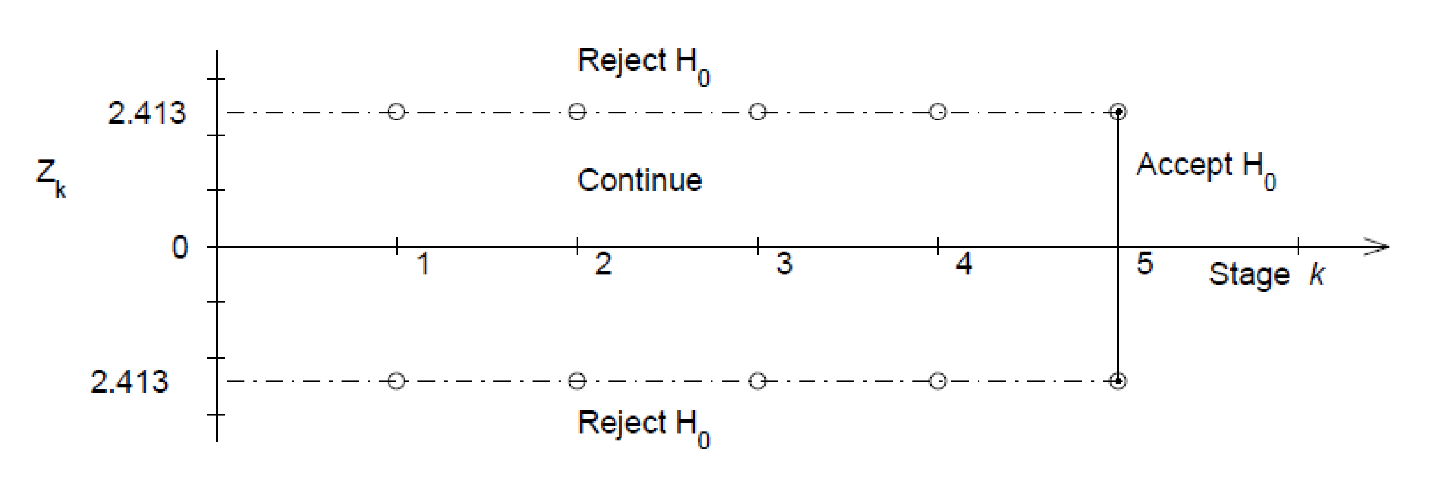
compare two proportions using chi-squared test

critical value depends on number of interim analysis, but is the same for each interim look

after group , if , stop the trial and reject H0

if otherwise, continue to group

after group , if , reject H0, otherwise accept H0



|  |  |
| --- | --- |
| Pocock Z critical values | |
|  |  |
| 1 | 1.960 |
| 2 | 2.178 |
| 3 | 2.289 |
| 4 | 2.361 |
| 5 | 2.413 |
| 6 | 2.453 |
| 7 | 2.485 |
| 8 | 2.512 |
| 9 | 2.535 |
| 10 | 2.555 |
| 15 | 2.626 |
| 20 | 2.672 |

O’Brien-Fleming Approach

sequential testing procedure where critical values decrease over the stages

chi-squared critical values depend on total number of interim analysis and the stage of the interim analysis

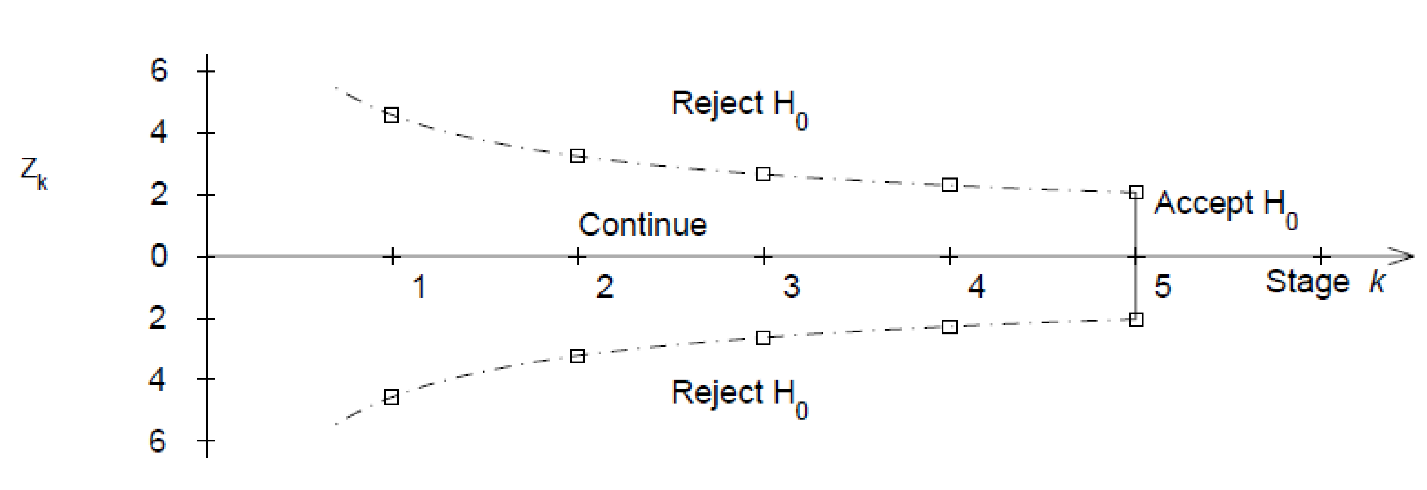
after group , if , stop the trial and reject H0

if , continue to group

after group , if , reject H0, otherwise accept H0

more difficult to declare superiority at the earlier looks

more conservative than Pocock at earlier stages



|  |  |
| --- | --- |
| O’Brien Fleming Z critical values | |
|  |  |
| 1 | 1.960 |
| 2 | 1.977 |
| 3 | 2.004 |
| 4 | 2.024 |
| 5 | 2.040 |
| 6 | 2.053 |
| 7 | 2.063 |
| 8 | 2.072 |
| 9 | 2.080 |
| 10 | 2.087 |
| 15 | 2.110 |
| 20 | 2.126 |

**Alpha-Spending**

= proportion of sample size accrued

beginning of the study 0% of information is accrued

no α is spent

end of the study 0% of information is accrued

amount of α used up after the analysis

= alpha-spending function, probability of Type I error willing to be spent up to time

alpha-spending is not the same as significance level

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
|  |  | **Z** | **Significance Level** |
| 0.2 | 0.000001 | 4.8769 | 0.000001 |
| 0.4 | 0.00079 | 3.3571 | 0.00078 |
| 0.6 | 0.00762 | 2.6803 | 0.00736 |
| 0.8 | 0.02442 | 2.2898 | 0.02204 |
| 1 | 0.05 | 2.0310 | 0.04226 |

If the first interim analysis occurs after 20% of information is gathered, reject treatment equality if p-value < 0.000001

If the second interim analysis occurs after 40% of information is gathered, reject treatment equality if p-value < 0.00078

If the third interim analysis occurs after 60% of information is gathered, reject treatment equality if p-value < 0.0074

If the fourth interim analysis occurs after 80% of information is gathered, reject treatment equality if p-value < 0.022