検証的試験において、Cox回帰などの統計モデルを主解析におく 妥当性について

議論をする上で想定する場

- 新治療vs標準治療の比較試験
- 検定を主な目的とする
- がん領域などのアウトカムがtime-to-event

- 試験計画時に主解析の方法を決める状況
 - (層別)ログランク検定
 - Cox回帰

妥当性に関する論点

- アルファエラーが名義水準以下に担保されるか
- (それが成り立っているとして)
検出力が高いのは？

要するに

- そもそもCox回帰を使っていいか？
- 使って何が嬉しいのか？
- ログランク検定でダメな場面は？

(層別)ログランク検定の メリット・デメリット

- デザインベースの解析
 - ランダム化に基づく解析なので、デザインと整合性がある
 - アルファエラーが担保される
 - 動的割付ではどうか？
- 割付因子を考慮した層別ログランク検定
 - (全てではないかもしれないが) 共変量を考慮した解析ができる
 - 検出力もあがる
 - でも連続量の共変量は難しい (but がん領域では強い予後因子はカテゴリーカル:ステージ)

Cox回帰のメリット

- 層別ログランク検定よりも多くの共変量を考慮できる
- 連続量の共変量も考慮できる
- 結果として検出力が層別ログランク検定よりも上がることもある？
- 非劣性試験では主解析でCox回帰に基づく信頼区間を出している

Cox回帰のデメリット

- モデルの仮定がある
 - 比例ハザード性
 - 対数線形性
- 仮定が崩れたときにどうなるのかは不明？
 - アルファエラー、検出力etc...
- 比例ハザード性が成り立たない場面における性能
 - ワクチンなどの遅れて効果が出るもの
 - バイオマーカーと治療との質的交互作用がある場面

まとめ

- Cox回帰を主解析として使っていいのか？
- 使ってはいけない、という明確な理由はなさそう
- ただし使うなら結果の頑健性をちゃんと考えなければならない(感度解析が必要)
 - 感度解析の具体的な方法までは議論できなかった