

Impact d'une erreur sur le choix des hypothèses p_0 et p_1 en terme de risque de première et deuxième espèce dans un plan de Simon

Charlotte Baey Marie-Cécile LeDeley

Institut Gustave Roussy

Congrès EpiClin - 26 mai 2010

Plan

- 1 Contexte
- 2 Problématique - Plan de Simon
- 3 Exemple
- 4 Généralisation
- 5 Conclusion

Contexte

- Les essais de phase II sont une des étapes clés du développement clinique \Rightarrow ils doivent fournir suffisamment d'information concernant la décision de poursuivre ou non en phase III

Contexte

- Les essais de phase II sont une des étapes clés du développement clinique \Rightarrow ils doivent fournir suffisamment d'information concernant la décision de poursuivre ou non en phase III
- En cancérologie, le plus souvent, ce sont des essais à un bras \Rightarrow cette décision repose alors sur une comparaison implicite du critère d'efficacité à des données historiques

Contexte

- Le taux de faux positifs généralement accepté lors d'un essai de phase II (= risque de première espèce) se situe autour de 10%

1. Gan et al. *Journal of Clinical Oncology*, Avril 2010

Contexte

- Le taux de faux positifs généralement accepté lors d'un essai de phase II (= risque de première espèce) se situe autour de 10%
- Ce taux contraste fortement avec le fort taux d'attrition des traitements testés en phase III, autour de 60%¹

1. Gan et al. *Journal of Clinical Oncology*, Avril 2010

Contexte

- Le taux de faux positifs généralement accepté lors d'un essai de phase II (= risque de première espèce) se situe autour de 10%
 - Ce taux contraste fortement avec le fort taux d'attrition des traitements testés en phase III, autour de 60%¹
- ⇒ Cette différence peut être en partie expliquée par une mauvaise spécification des taux de réponse p_0 et p_1

1. Gan et al. *Journal of Clinical Oncology*, Avril 2010

Plan de Simon

Plan séquentiel en deux étapes défini pour :

- Un seuil p_0 en-dessous duquel le traitement est jugé inefficace, et qui permet de définir l'hypothèse nulle $H_0 : p \leq p_0$
- Un risque de première espèce α associé à l'hypothèse H_0
- Un seuil p_1 d'efficacité, associé à une certaine puissance $1 - \beta$

A partir des taux p_0 et p_1 on définit une règle de décision unique à 4 paramètres, respectant les risques α et β fixés.

Plan de Simon

- Les taux p_0 et p_1 sont fixés *a priori*, à partir de données historiques, d'avis d'experts, ...

Plan de Simon

- Les taux p_0 et p_1 sont fixés *a priori*, à partir de données historiques, d'avis d'experts, ...
- Il peut exister une forte incertitude sur ces taux de réponses : biais de sélection des patients, pas de prise en compte de la variabilité du taux historique, modification des standards de soins, ...

Plan de Simon

- Les taux p_0 et p_1 sont fixés *a priori*, à partir de données historiques, d'avis d'experts, ...
 - Il peut exister une forte incertitude sur ces taux de réponses : biais de sélection des patients, pas de prise en compte de la variabilité du taux historique, modification des standards de soins, ...
- ⇒ quelles peuvent être les conséquences d'une mauvaise spécification du taux de réponse p_0 sur le risque de première espèce α ? Et de la même façon, pour le taux p_1 et le risque de deuxième espèce β ?

Exemple

Plan H défini par :

$$p_{0,h} = 0.20, \quad \alpha_h = 0.10, \quad p_{1,h} = 0.40, \quad \text{et} \quad \beta_h = 0.10$$

- Plan de Simon optimal : $n_{1,h} = 17$, $k_{1,h} = 3$, $N_h = 37$ et $K_h = 10$
- Les risques de première et deuxième espèces calculés sont $\alpha = 0.0948$ et $\beta = 0.0967$

Exemple

Plan H défini par :

$$p_{0,h} = 0.20, \quad \alpha_h = 0.10, \quad p_{1,h} = 0.40, \quad \text{et} \quad \beta_h = 0.10$$

- Plan de Simon optimal : $n_{1,h} = 17$, $k_{1,h} = 3$, $N_h = 37$ et $K_h = 10$
- Les risques de première et deuxième espèces calculés sont $\alpha = 0.0948$ et $\beta = 0.0967$
- Supposons qu'en réalité, un taux de réponse inférieur ou égal à 0.25 aurait du être considéré comme inefficace.

Exemple

Plan H défini par :

$$p_{0,h} = 0.20, \quad \alpha_h = 0.10, \quad p_{1,h} = 0.40, \quad \text{et} \quad \beta_h = 0.10$$

- Plan de Simon optimal : $n_{1,h} = 17$, $k_{1,h} = 3$, $N_h = 37$ et $K_h = 10$
- Les risques de première et deuxième espèces calculés sont $\alpha = 0.0948$ et $\beta = 0.0967$
- Supposons qu'en réalité, un taux de réponse inférieur ou égal à 0.25 aurait du être considéré comme inefficace.
- Le risque de première espèce α associé à la règle de décision ($n_{1,h} = 17$, $k_{1,h} = 3$, $N_h = 37$, $K_h = 10$) doit être recalculé pour $p_0 = 0.25$

Exemple

Calcul des risques α et β :

$$\alpha = P(s_1 > K \mid n_1, p_0) + \sum_{i=k_1+1}^{\min(n_1, K)} P(s_1 = i \mid n_1, p_0) P(s_2 > K-i \mid n_2, p_0)$$

$$\beta = P(s_1 \leq k_1 \mid n_1, p_1) + \sum_{i=k_1+1}^{\min(n_1, K)} P(s_1 = i \mid n_1, p_1) P(s_2 \leq K-i \mid n_2, p_1)$$

- Pour $p_0 = 0.20$, $\alpha = 0.0948$ (plan initial)
- Pour $p_0 = 0.25$, $\alpha = 0.284$

⇒ On est très loin du $\alpha = 0.10$ fixé !

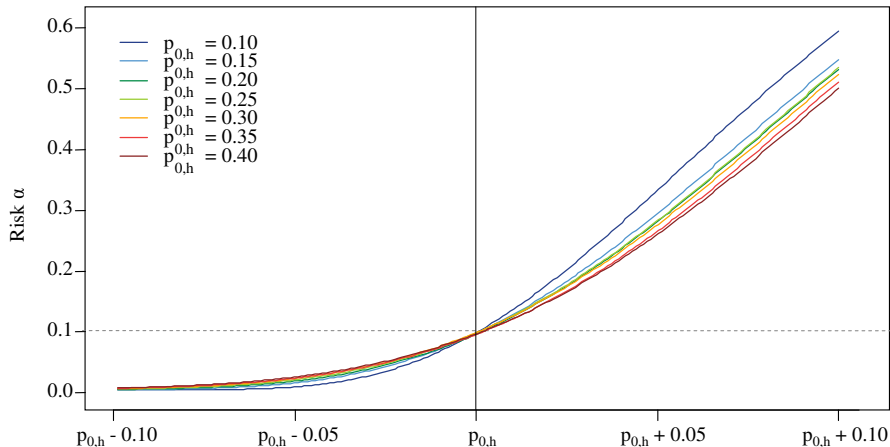
Généralisation

Sept scénarii ont été considérés, en faisant varier les taux p_0 et p_1 , mais en conservant une taille d'effet $p_1 - p_0$ constante ($= 0.20$)

- L'impact sur le risque de première espèce α est calculé en fonction de l'erreur commise sur p_0
- L'impact sur le risque de deuxième espèce β est calculé en fonction de l'erreur commise sur p_1

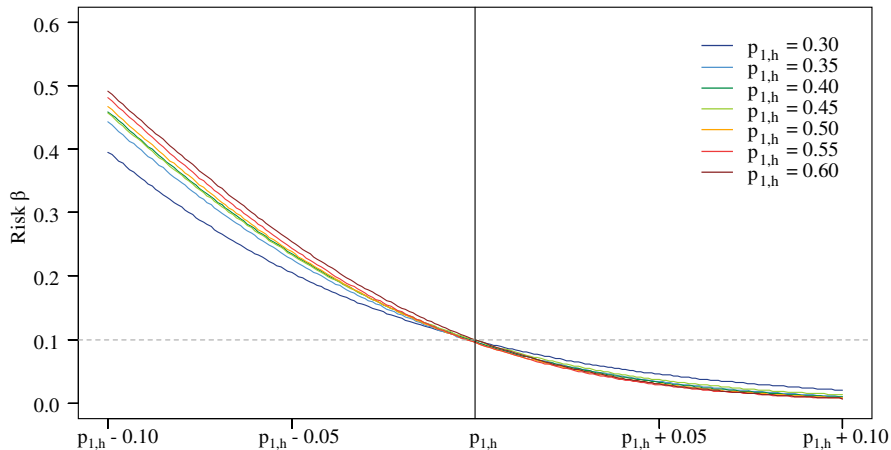
Généralisation

Impact sur le risque α :



Généralisation

Impact sur le risque β :



Généralisation

- Une erreur de 5% sur p_0 conduit à une augmentation du risque de première espèce de l'ordre de 20%
- Cette inflation du risque de première espèce conserve la même amplitude lorsqu'on fait varier :
 - ▶ taille d'effet
 - ▶ risques initiaux
 - ▶ type de design (optimal ou minimax)

Conclusion

- Ce type de design à un bras ne permet pas de contrôler efficacement les risques α et β en cas de mauvaise spécification des taux p_0 et/ou p_1

Conclusion

- Ce type de design à un bras ne permet pas de contrôler efficacement les risques α et β en cas de mauvaise spécification des taux p_0 et/ou p_1
- Les essais de phase II fréquentistes non comparatifs sont conduits à partir d'une seule règle de décision, en oubliant l'incertitude liée au choix des paramètres

Conclusion

- Ce type de design à un bras ne permet pas de contrôler efficacement les risques α et β en cas de mauvaise spécification des taux p_0 et/ou p_1
- Les essais de phase II fréquentistes non comparatifs sont conduits à partir d'une seule règle de décision, en oubliant l'incertitude liée au choix des paramètres
- Perspectives ? :

Conclusion

- Ce type de design à un bras ne permet pas de contrôler efficacement les risques α et β en cas de mauvaise spécification des taux p_0 et/ou p_1
- Les essais de phase II fréquentistes non comparatifs sont conduits à partir d'une seule règle de décision, en oubliant l'incertitude liée au choix des paramètres
- Perspectives ? :
 - ▶ essais randomisés avec bras contrôle
 - ▶ designs bayésiens