

# GROUP SEQUENTIAL METHODS BASED ON SUPREMUM LOGRANK STATISTICS UNDER PROPORTIONAL AND NON PROPORTIONAL HAZARDS

JEAN MARIE BOHER

Unité de Biostatistique et de Méthodologie  
Institut Paoli-Calmettes

Equipe QUANTIM, UMR 1252

VENDREDI 09 JANVIER 2025

---

## CONTEXTE GÉNÉRAL

- L'immunothérapie (IT) est devenue incontournable dans le traitement de nombreux cancers
  - Les principaux critères d'évaluation de l'efficacité de ces traitements restent la survie globale la survie sans progression ou décès.
  - De manière générale la planification, l'analyse d'essais en oncologie reposent sur l'hypothèse de proportionnalité des risques instantanés.
  - Les méthodes classiques (test du log-rang, modèle de Cox) sont optimales sous de telles hypothèses.
-

---

## CAS PARTICULIER DES ESSAIS CLINIQUES EN IMMUNO-ONCOLOGIE

- Le non-respect des hypothèses de proportionnalité a souvent été observé dans des essais évaluant des traitements par IT efficaces.
  - Cela s'accompagne généralement d'une perte de chances de démontrer de réelles différences des méthodes d'analyse classiques.
  - La définition de méthodes d'analyse capables de détecter des différences tardives ou autres est un enjeu important.
-

---

## PROPORTIONALITÉ DES RISQUES INSTANTANÉES SOUVENT REJÉTÉE

- Gefitinib en association avec carboplatin et paclitaxel ou taxol, *IPASS*; *Mok et al. 2009*)
  - Revlimid + Dexamethasone versus Standard Thalidomide dans les myelomes non éligibles à une greffe, *FIRST*; *Benboubker et al. 2014*)
  - Pembrolizumab vs chimiothérapie standard dans les carcinomes urotheliaux avancés *KEYNOTE-045*; *Bellmunt et al. 2017*)
  - Pembrolizumab versus CT standard dans le cancer du poumon à petites cellules localement avancé ou métastatique PDL1 + , *KEYNOTE-042*; *Mok et al. 2019*)
-

---

## MÉTHODES ALTERNATIVES

### ON NOTE

$$(X_i, \Delta_i, G_i), i = 1, \dots, N$$

- $X_i$  délai avant censure ou évènement d'intérêt,
- $\Delta_i$  le statut,  $\Delta_i = 1$  si évènement, 0 sinon
- $G_i$  indicatrice du groupe alloué par Rx.

$G_i = 1$  si groupe expérimental, 0 sinon

---

---

## PRINCIPALES MÉTHODES ALTERNATIVES PROPOSÉES POUR PRÉSERVER LA PUISSANCE DES ESSAIS

### TESTS DE LOG-RANG PONDÉRÉS

$$T = S/s.d., \quad S = \sum_{j=1}^D w(t_j) \left( d_{1j} - n_{1i} \frac{d_j}{n_j} \right)$$

où

- $t_1 < \dots < t_d$  désigne la séquence ordonnée des temps d'évènements
  - $w(t)$  une fonction de poids
  - $d_{1j}, n_{1j}$  le total des évènements et des patients à risque au temps  $t_j$  dans le groupe  $G = 1$ ,
  - $d_j, n_j$  le total des évènements et des patients à risque au temps  $t_j$  dans la population totale,
-

---

— **Fleming-Harrington weighted logrank tests** *Fleming and Harrington, Comm in Stat 1981, Hasegawa SIM 2014*

$$FH(0, 0) : w(t) = 1 \text{ (logrank test)}$$

$$FH(0, 1) : w(t) = 1 - \hat{S}(t-) \text{ (différences tardives)}$$

$$FH(1, 0) : w(t) = \hat{S}(t-) \text{ (différences précoces)}$$

— **Piecewise weighted logrank tests** *Xu et al., SIM 2017, 2018*

$$w(t) = 1 \text{ si } t > t_0, 0 \text{ sinon}$$

---

# EXEMPLE

## □ Keynote-045 Trial (NCT02256436)

### Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma

J. Bellmunt, R. de Wit, D.J. Vaughn, Y. Fradet, J.-L. Lee, L. Fong, N.J. Vogelzang, M.A. Climent, D.P. Petrylak, T.K. Choueiri, A. Necchi, W. Gerritsen, H. Gurney, D.I. Quinn, S. Culine, C.N. Sternberg, Y. Mai, C.H. Poehlein, R.F. Perini, and D.F. Bajorin, for the KEYNOTE-045 Investigators\*

#### BACKGROUND

Patients with advanced urothelial carcinoma that progresses after platinum-based chemotherapy have a poor prognosis and limited treatment options.

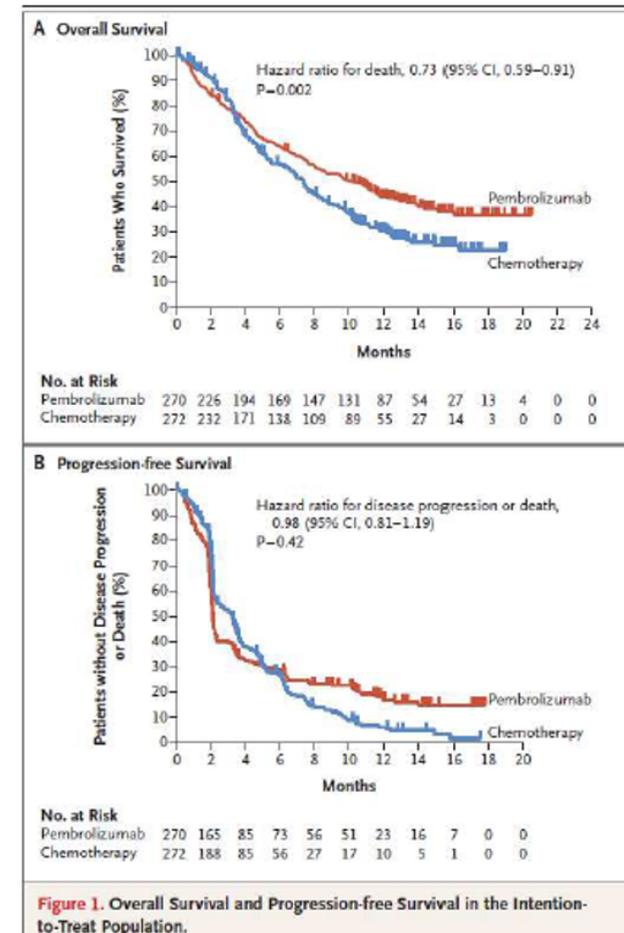
#### METHODS

In this open-label, international, phase 3 trial, we randomly assigned 542 patients with advanced urothelial cancer that recurred or progressed after platinum-based chemotherapy to receive pembrolizumab (a highly selective, humanized monoclonal IgG4κ isotype antibody against programmed death 1 [PD-1]) at a dose of 200 mg every 3 weeks or the investigator's choice of chemotherapy with paclitaxel, docetaxel, or vinflunine. The coprimary end points were overall survival and progression-free survival, which were assessed among all patients and among patients who had a tumor PD-1 ligand (PD-L1) combined positive score (the percentage of PD-L1-expressing tumor and infiltrating immune cells relative to the total number of tumor cells) of 10% or more.

#### CONCLUSIONS

Pembrolizumab was associated with significantly longer overall survival (by approximately 3 months) and with a lower rate of treatment-related adverse events than chemotherapy as second-line therapy for platinum-refractory advanced urothelial carcinoma. (Funded by Merck; KEYNOTE-045 ClinicalTrials.gov number, NCT02256436.)

N ENGL J MED 376;11 NEJM.ORG MARCH 16, 2017

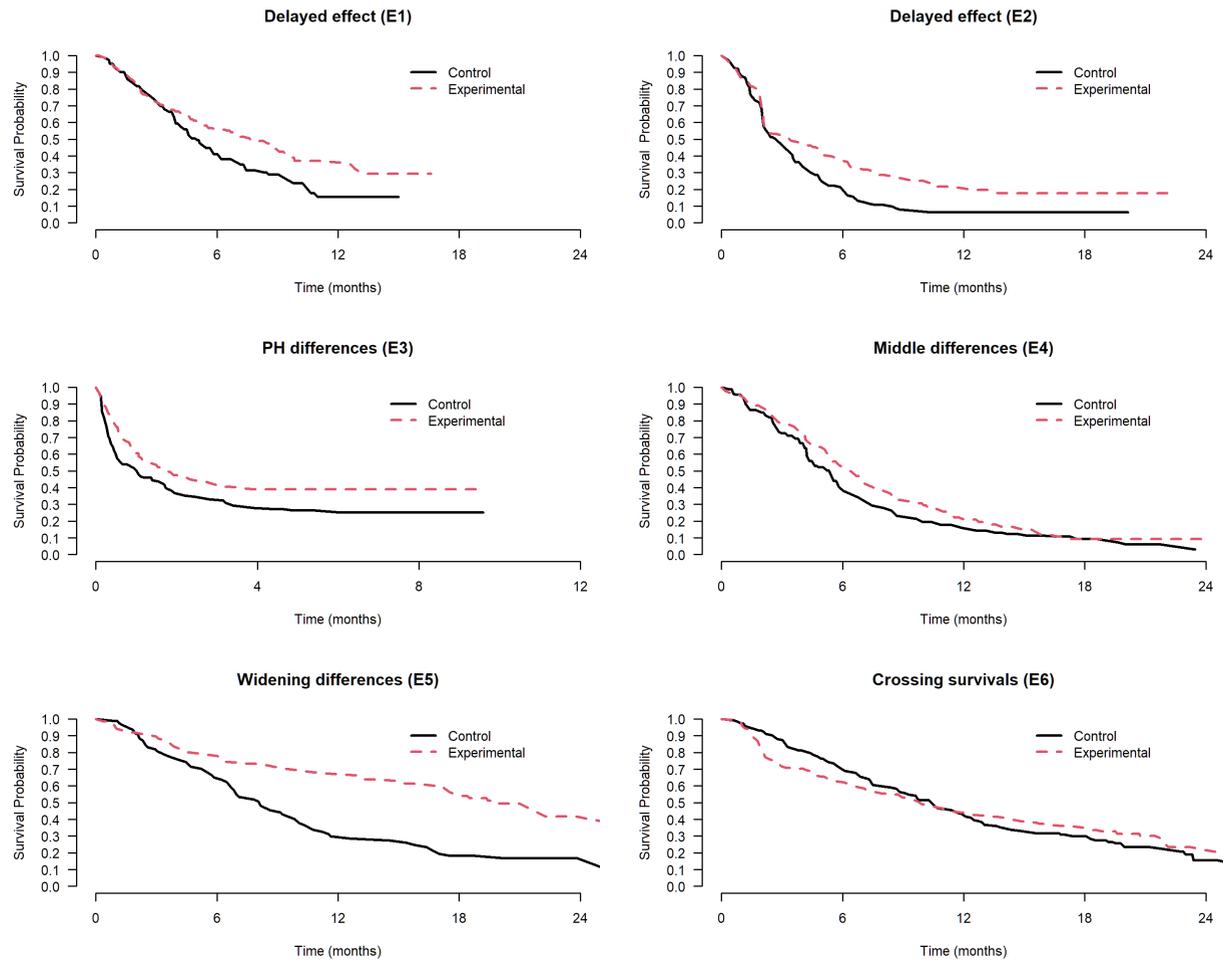


Keynote – 45  
 Survie sans progression

Study Data	Log-rank tests			RMST tests	
	$W(t) = 1$	$W(t) = \hat{S}(t-)$	$W(t) = 1 - \hat{S}(t-)$	Diff	Ratio
<b>Keynote-045</b> <i>All patients (N=542)</i>	0.917	0.006	< .001	0.077	0.071



# EXEMPLES DE DIFFÉRENCES OBSERVÉES



---

— **Augmented logrank test**

$$\max(FH(0, 0), RMST)$$

*Royston and Parmar, BMC Med Res Meth 2016*

— **Maxcombo test**

$$\max(FH(0, 0), FH(1, 0), FH(0, 1), FH(1, 1))$$

*Karrison et al., Stata Journal 2016; Shen et al., Stat in Biopharm Res 2023; Mukhopadhyay et al., JAMA Oncology 2023*

---

---

PROPOSITION

TESTS BASÉS SUR LA VALEUR EXTRÊME D'UN PROCESSUS

$$T_j(t) = \text{Sup}_t S_j(t) / s.d., \quad j = 1, 2$$

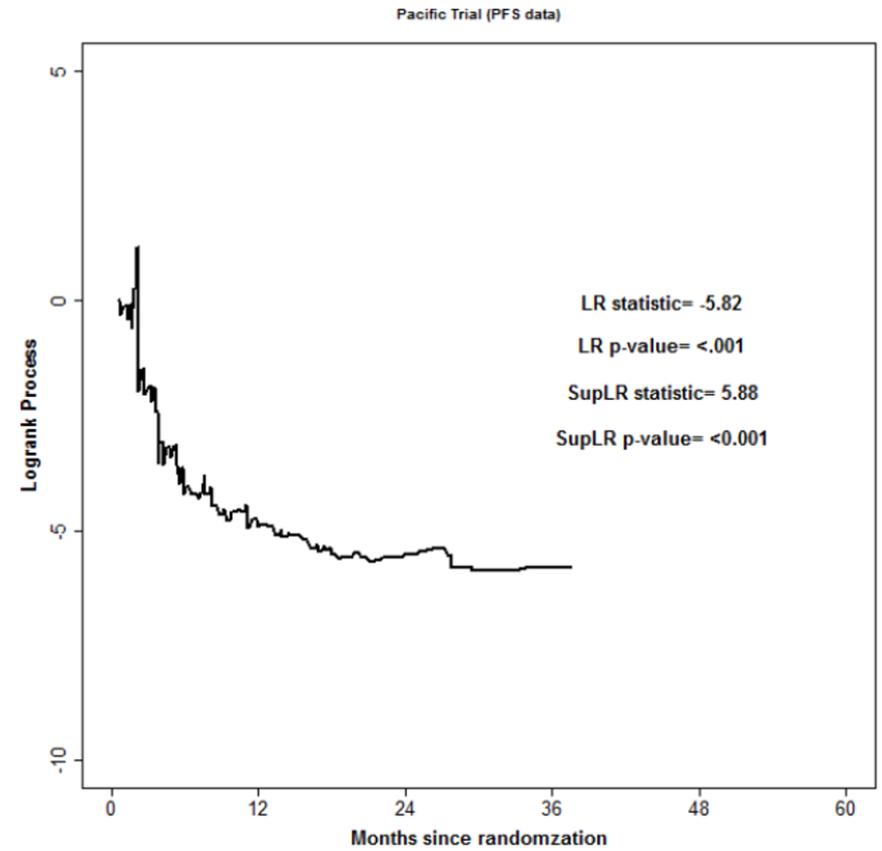
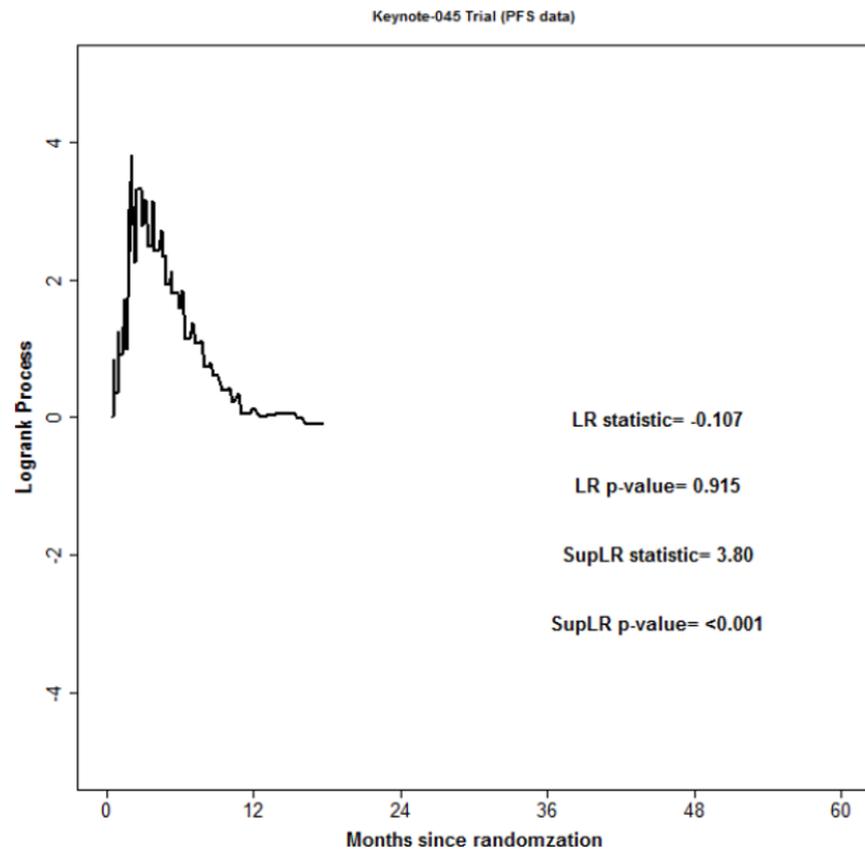
où

$$S_1(t) = \sum_{j:t_j \leq t} \left( d_{1j} - n_{1i} \frac{d_j}{n_j} \right),$$

$$S_2(t) = \sum_{j:t \leq t_j} \left( d_{1j} - n_{1i} \frac{d_j}{n_j} \right).$$

---

## EXEMPLE KEYNOTE-045



---

Si l'on suppose une randomisation équilibrée (1 :1), la distribution des statistiques sous l'hypothèse  $H_0$  est estimée en réalisant différents tirages aléatoires d'un processus équivalent (de même loi).

$$\tilde{S}_j(t) = \sum_{i=1, \dots, n} Y_{ji}(t) \Delta_i (G_i - 0.5) U_i$$

—  $Y_{1i}(t) = I(X_i \leq t)$

—  $Y_{2i}(t) = I(t \leq X_i)$

—  $\{U_i : i = 1, \dots, n\}$  échantillon d'une loi normale  $N(0, 1)$ .

*Boher et al, Stat in Medecine 2023*

---

L'algorithme peut être décomposé en 3 étapes clés

ETAPE 1 : Réaliser le tirage de  $ns$  échantillons de taille  $n$ , de loi  $N(0, 1)$

$$\{\mathbf{U}_{i\ell}, \mathbf{i} = \mathbf{1}, \dots, \mathbf{n}\}, \ell = \mathbf{1}, \dots, ns$$

.

ETAPE 2 : Dériver les  $ns$  trajectoires correspondantes

$$\tilde{S}_{j\ell}(t) = \sum_i Y_{ji}(t) \Delta_i (G_i - 0.5) U_{i\ell}, \ell = 1, \dots, ns$$

ETAPE 3 : Calculer les  $ns$  statistiques

$$\tilde{T}_{j\ell} = \tilde{\sigma}_j^{-1} \text{Sup}_t \tilde{S}_{j\ell}(t)$$

où  $\hat{\sigma}_j^2 = \text{variance empirique observée sur l'échantillon } \{\tilde{T}_{j\ell} : \ell = 1, \dots, ns\}$ .

## VERSION SEQUENTIELLE DES TESTS

Des tests séquentiels sont souvent pratiqués dans les essais randomisés.

Soit  $(Z_1, Z_2, \dots, Z_K)$  la séquence des tests répétés,  $(\alpha_1, \dots, \alpha_K)$ ,

$$\alpha_1 \prec \alpha_2 \prec \dots \prec \alpha_K$$

les risques cumulés d'erreurs de type I, les valeurs critiques  $b_1, \dots, b_K$  des différents tests sont sélectionnés pour satisfaire les équations suivantes

$$P_{H_0}[Z_1 > b_1] = \alpha_1$$

et

$$P_{H_0} \left[ \bigcap_{k'=1}^{k-1} Z_{k'} \leq b_{k'}, Z_k \succ b_k \right] = \alpha_k - \alpha_{(k-1)},$$

$k = 2, \dots, K$ .

## NOTONS

$$\{(\mathbf{X}_i(\mathbf{d}), \Delta_i(\mathbf{d}), \mathbf{G}_i), i = 1, \dots, n\}$$

les données observées à la date d'analyse prévue  $d$ .

Les valeurs critiques des différents tests peuvent être déterminées en simulant un grand nombre de trajectoires

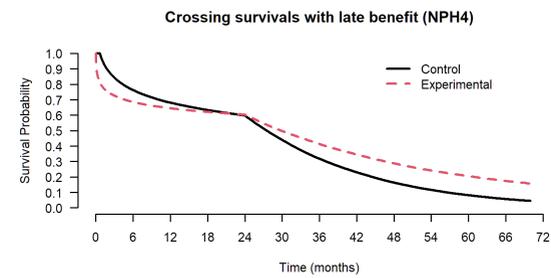
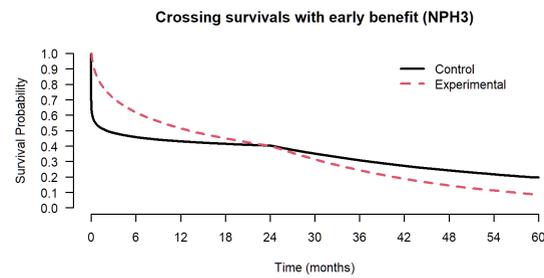
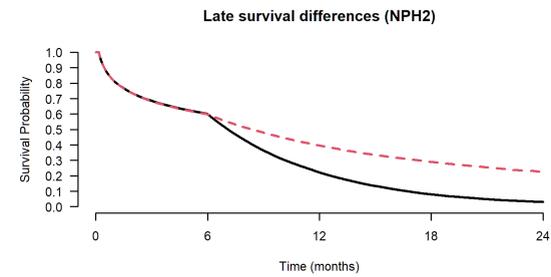
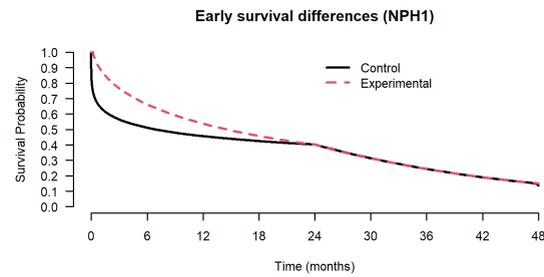
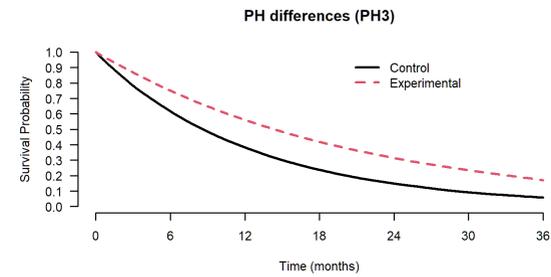
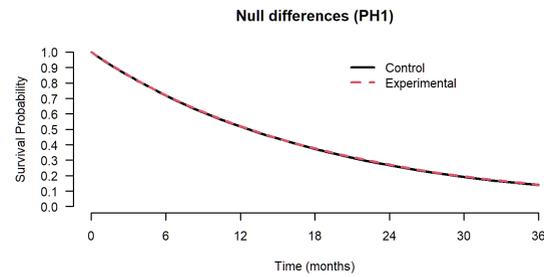
$$\tilde{S}_j(d, t) = \sum_i Y_{ji}(d, t) \Delta_i(d) (G_i - 0.5) U_{il},$$

$$\text{— } Y_{1i}(d, t) = I(t \leq X_i(d)), Y_{2i}(d, t) = I(t \leq X_i(d))$$

—  $l = 1, \dots, ns$ ,  $d = d_1, \dots, d_K$  dates des différents analyses (intermédiaires ou finale)

# PERFORMANCE DES DIFFÉRENTS TESTS

## SIMULATIONS



---

## RÉSULTATS

- Maintien de l'erreur de type I
  - Toutes les statistiques ont des performances très proches des performances du test du log-rang sous l'hypothèse de proportionnalité,
  - La statistique  $S_1$  est plus sensible à des écarts précoces
  - Le test MAXCOMBO plus sensible pour la détection de différences tardives
  - La combinaison des statistiques  $S_1$  et  $S_2$  plus sensible lorsque les courbes de survie se croisent
-

## PERSPECTIVES

Extension à l'analyse des risques compétitifs

Extension possible à l'analyse de critères multiples (OS et PFS par exemple)

Plus de détails :

*Boher JM, Filleron T, Bunouf P et al. New late-emphasis and combination tests based on infimum and supremum versions of logrank statistics with applications in oncology trials. Stat Med 2023.*

*Boher JM, Filleron T, Bunouf P et al. Group sequential methods based on supremum logrank proportional and nonproportional hazards. Stat Med in Med Res 2024.*

**MERCI DE VOTRE ATTENTION**

---