

JEM SFES

atelier 1 : estimation des retombées à long terme des traitements

**Extrapolation des modèles de survie, comment s'y
prendre?**

**Florence Joulain
Sanofi, département G-HEOR**

Paris, 22 Jan 2015

Agenda

- Contexte et Objectif

- I. Analyse de survie

- II. Pour extrapoler, quelle distribution choisir?

Contexte

- En données de survie le bénéfice du traitement est généralement exprimé en terme de Hazard Ratio ou en différence de médiane
 - Dans un essai clinique l'estimation des durées moyennes de survie se révèle souvent impossible car il faudrait suivre tous les patients jusqu'à la survenue de l'événement considéré
 - Le caractère asymétrique des données incite à préférer la médiane
- L'évaluation économique utilise comme indicateurs les années de vie et les années de vie ajustée par sa qualité (ou QALYs Quality-adjusted life-years)
- L'horizon temporel est la «vie entière»
- La moyenne de survie doit être estimée plutôt que la médiane
 - Prends en compte l'ensemble des données
- → Besoin d'extrapoler les données au-delà de l'horizon temporelle de l'étude
- Les modèles paramétriques doivent être utilisés

Objectif

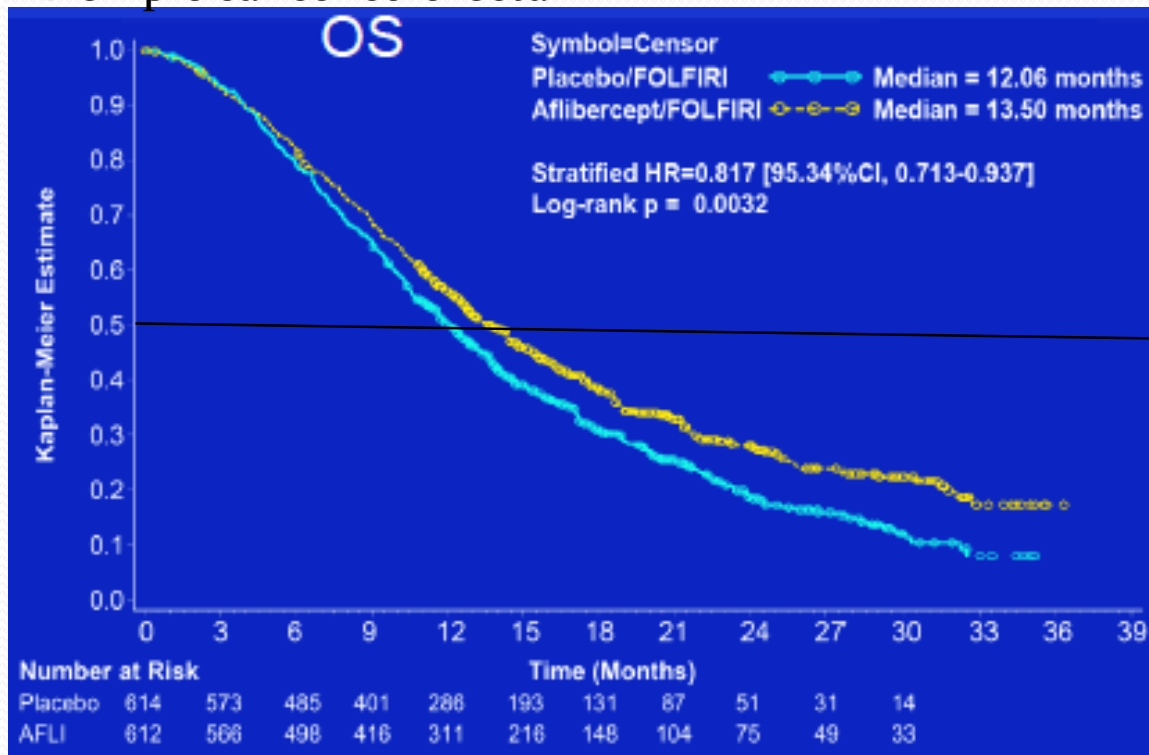
- Présenter les méthodes pour définir la distribution qui représentera au mieux les données expérimentales de Kaplan-Meier afin d'obtenir des extrapolations fiables
- Approche rigoureuse

Analyse de survie

- Analyser des données sous la forme de temps entre une origine bien définie (randomisation) jusqu'à l'occurrence d'un évènement particulier ou endpoint (décès ou progression de la maladie)
- Les temps de survie peuvent être censurés
 - Certains patients n'atteignent pas l'endpoint avant la fin de l'étude, ou s'ils sont perdu de vue
- S'il existe des données censurées, la moyenne de l'échantillon est alors un estimateur biaisé de la vraie moyenne
- La moyenne de survie peut être estimée en projetant la courbe jusqu'à ce qu'elle atteigne zéro et en calculant l'aire sous la courbe

Courbe de Kaplan-Meier (KM)

Exemple cancer colorectal



Median OS

- reflète le moment où $P(\text{survie})=50\%$
- observé pendant l'étude: pas d'hypothèse sur la survie à long terme au delà de la période de suivi de l'étude
- approche pragmatique
- mais peut ne pas refléter le bénéfice complet du produit

Van Cutsem, JCO 2012

$P(\text{survie})$ à la fin de suivi de l'étude:

17% dans le bras aflibercept vs 8% dans le bras placebo

Modèle économique: besoin d'estimer la moyenne de survie

- Les modèles économiques modernes raisonnent aujourd'hui en termes de temps moyen de survie ajusté ou non sur des coefficients de qualité de vie obtenus par transformation des scores des échelles psychométriques.
- Payers intéressés par une décision basée sur le bénéfice thérapeutique global et son impact attendu sur les dépenses de santé
- Estimation de la moyenne de survie est demandée par de nombreux payers pour estimer le ratio coût-efficacité et informer la décision de remboursement
 - NICE au UK, PBAC en Australie, ...

Moyenne de survie restreinte

- Utilise toutes les données disponibles pendant l'étude pour estimer l'aire sous la courbe de Kaplan-Meier jusqu'à la dernière observation
- En présence de censure, sous-estime la vraie moyenne
- Approche à utiliser quand les données sont relativement complète avec suivi de tous les patients jusqu'à survenue de l'évènement
- Rarement le cas car cela nécessiterait une durée d'étude très longue et reculerait de façon non acceptable le moment d'obtention de la réponse à la question posée

Extrapoler: quelle distribution choisir?

- Rechercher quelle loi de survie théorique est à même de représenter au mieux les données expérimentales de Kaplan - Meier pour obtenir des extrapolations fiables.
- Plusieurs méthodes disponibles
 - Distribution Exponentiel, Weibull, Gompertz*, log-logistic or log normal
 - Modèles plus complexes et flexibles
- Differences d'estimation potentiellement large
 - Particulièrement quand la quantité de données à extrapoler est importante
- Important de justifier l'approche choisie, pour démontrer que l'extrapolation a été faite de manière appropriée
- Pour une confiance dans les résultats associés à l'analyse économique

* Gompertz pas disponible sous SAS mais sous STATA

Méthodes

- Comparaisons graphiques
 - Graphe Log-cumulative hazard
 - Courbe KM vs distributions
- Méthodes statistiques
 - Critères AIC, BIC
- Validité Clinique et utilisation de données extérieures

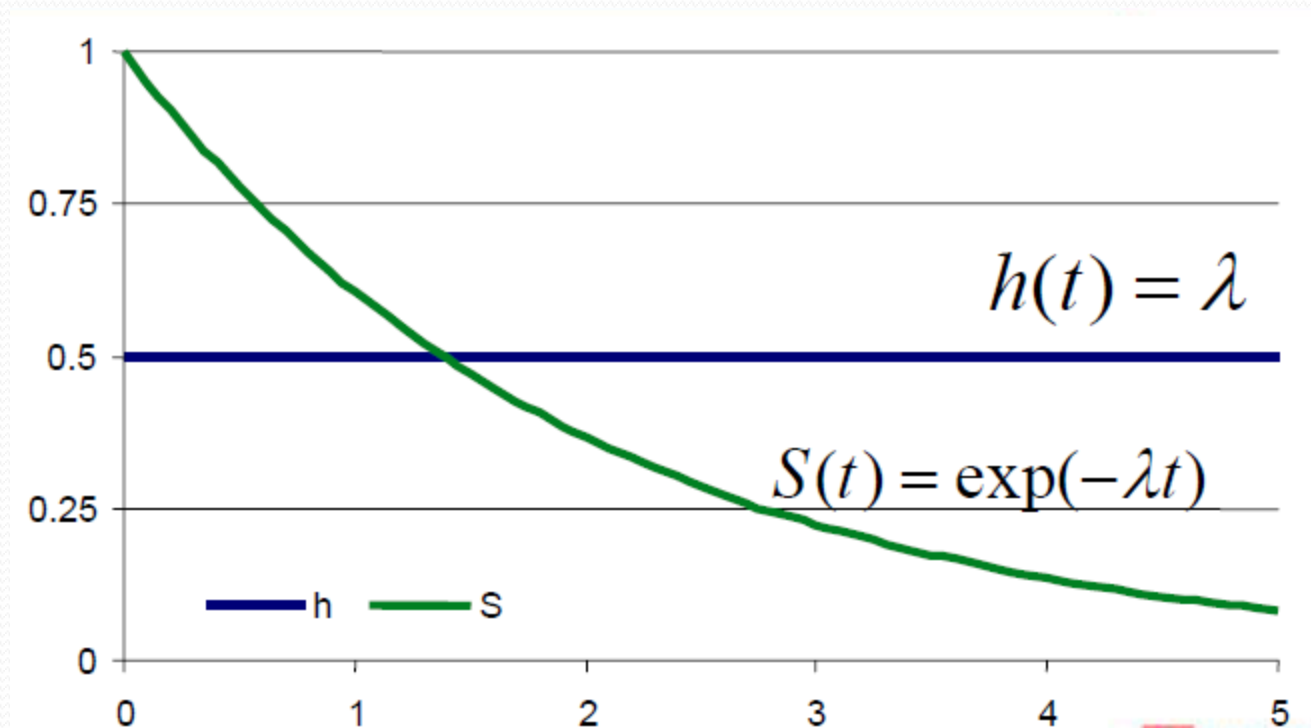
Graphe Log-cumulative hazard

- Important de considérer l'évolution au cours du temps du risque instantané (hazard rate) pour déterminer le modèle paramétrique adéquate
- Modèle exponentiel si risque instantané relativement constant
- Weibull et Gompertz models si risque monotone
- Loglogistic et log normal peuvent incorporer des risques non-monotones
 - problème typique: risque instantané diminue avec le temps à partir d'un certain point(long tail)

Graphe Log-cumulative hazard (suite)

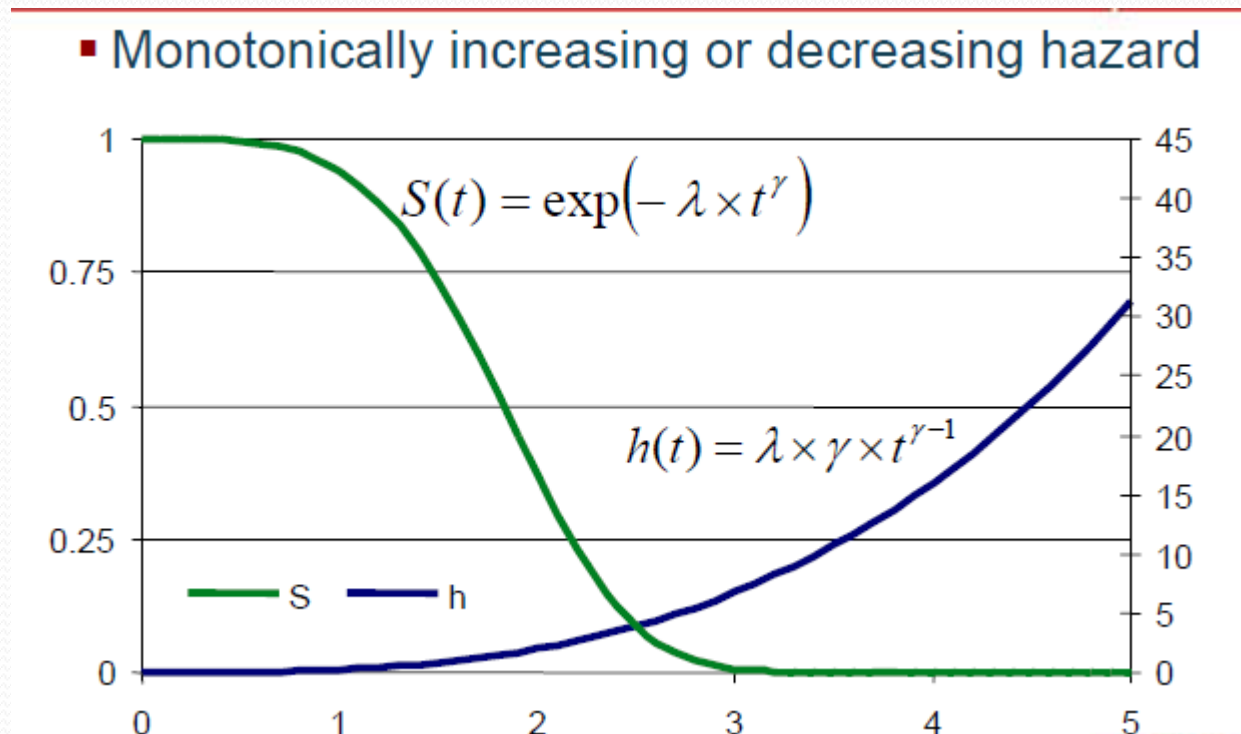
- Evaluation de l'hypothèse des risques proportionnels
 - Nécessaire pour certains modèles
- Permet de déterminer s'il existe des changements significatifs dans l'évolution du risque instantané
 - Modèle piecewise si différent modèle paramétrique sont nécessaire sur différente période
 - Besoin de validation clinique

Distribution exponentielle: risque constant



Détection: si $\log(S)$ vs. t relativement linéaire

Distribution de Weibull: monotone



Détection: si $\log(-\log(S))$ vs. $\log(t)$ relativement linéaire

Critères AIC, BIC

- Critères AIC (Akaike's Information criteria) et BIC (Bayesian Information Criteria) et méthodes graphiques pour évaluer les qualités des différentes distribution pour représenter les données (goodness-of-fit)
- Le modèle préféré est celui avec la plus petite valeur AIC et BIC
 - Explique le modèle avec le moins de paramètre

Exemple cancer colorectal

TABLE 3. AIC/BIC Criteria by Distribution - Treatment Arm Separately

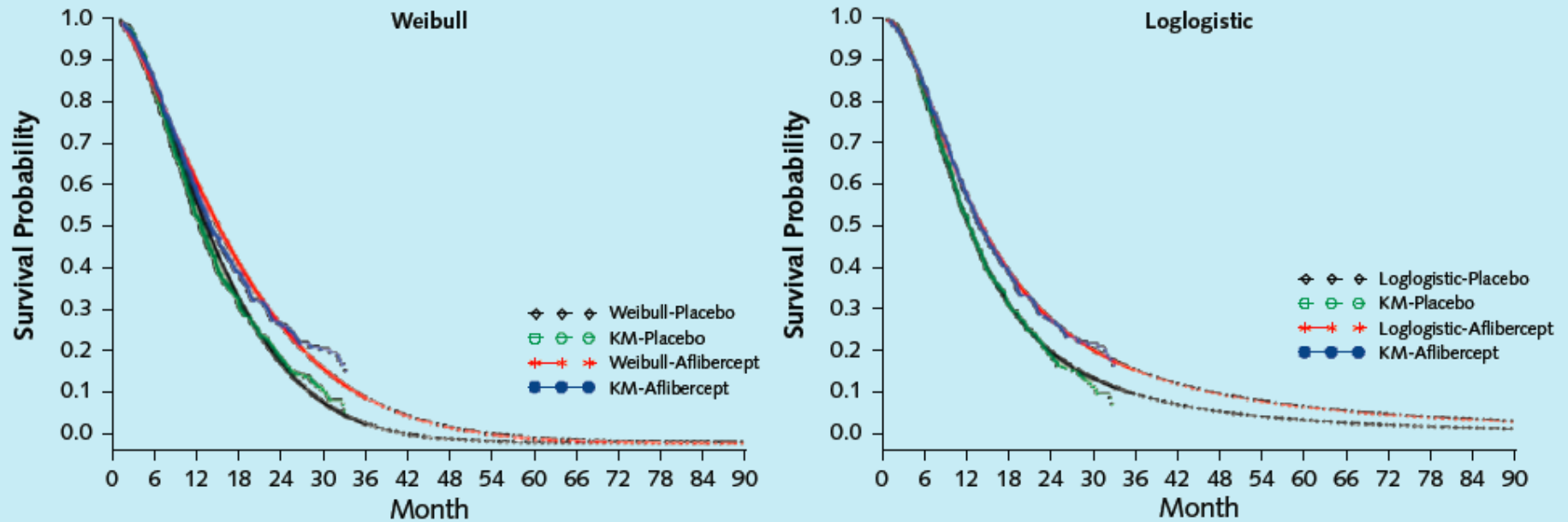
Distribution	Log Likelihood (LL)	n	Df	AIC	BIC
Placebo arm					
Exponential	-773.867	614	1	1549.73	1554.15
Weibull	-732.669	614	2	1469.34	1478.18
Lognormal	-732.542	614	2	1469.08	1477.92
Loglogistic	-726.595	614	2	1457.19	1466.03
Gompertz	-753.299	614	2	1510.60	1519.44
Aflibercept arm					
Exponential	-766.702	612	1	1535.40	1539.82
Weibull	-746.528	612	2	1497.06	1505.89
Lognormal	-743.572	612	2	1491.14	1499.98
Loglogistic	-737.851	612	2	1479.70	1488.53
Gompertz	-760.235	612	2	1524.47	1533.30

Df, degree of freedom; $AIC=2 \cdot df - 2 \cdot LL$; $BIC=-2 \cdot LL + df \cdot \ln(N)$.

- Plus petite valeur AIC et BIC pour la distribution loglogistic dans les 2 bras
→ suggère une meilleure adéquation aux données pour loglogistic

Comparaisons graphiques

FIGURE 7. Graphical Comparison Between KM and Weibull or Loglogistic Fitting Treatment Arms Separately

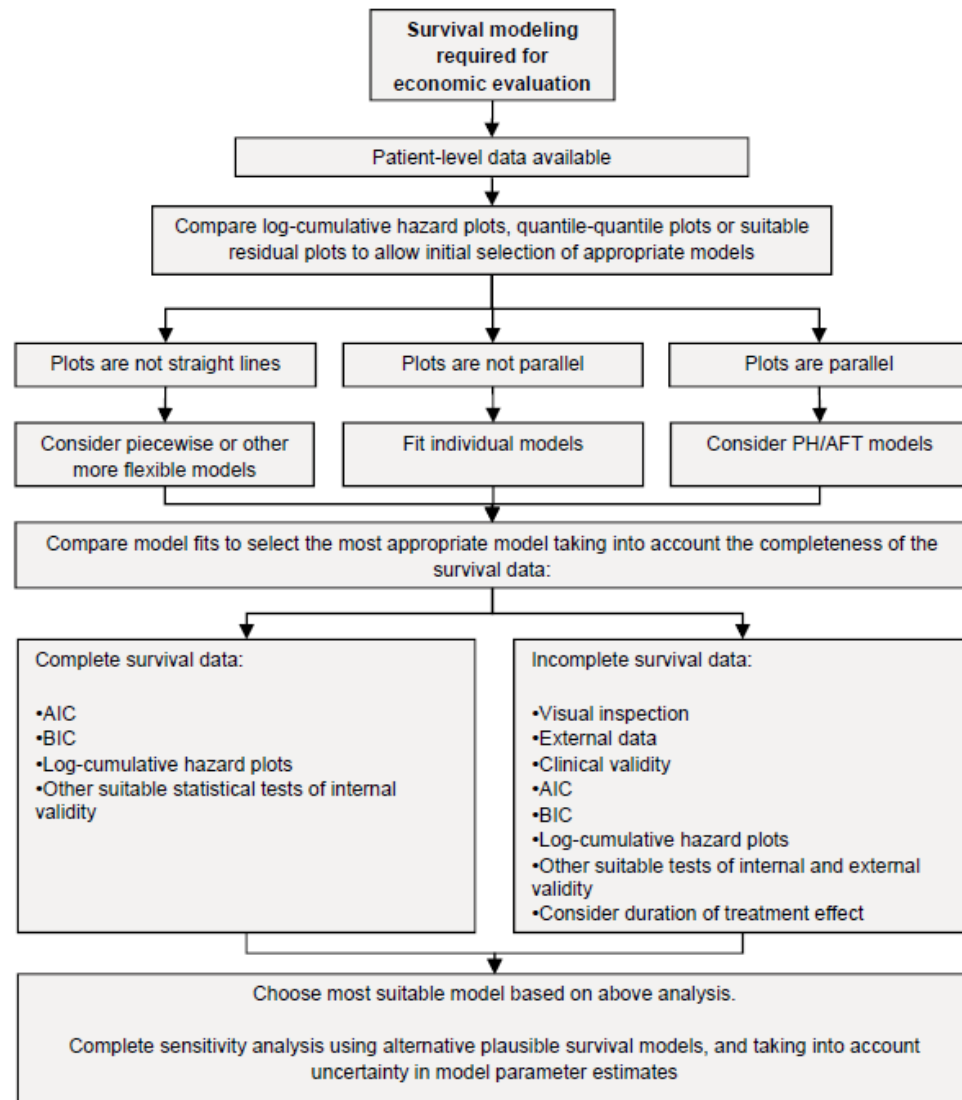


- Confirmation graphique de la meilleur adéquation avec la distribution loglogistique

Modèle Proportional Hazards (PH)

- Implique une survie paramétrique de base pour estimer un groupe control
- HR ensuite appliqué pour estimer le groupe intervention
- Souvent utilisé quand plusieurs comparateurs sont inclus dans l'évaluation
 - Données individuelles non disponible pour tous les comparateurs
- Attention ne peut pas être utilisé avec toutes les distributions (log-logistic, log-normal sont des modèles AFT (Accelerated Failure time) et non PH

Algorithm de sélection du modèle de survie



NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA.

Incertitude

- **PSA (Probabilistic Sensitivity Analysis)**
- Matrice de variance-covariance des distributions
- Décomposition de Cholesky

Besoin d'analyses de sensibilité

- Tester l'impact de différents modèles paramétriques
- Appliquer des durées d'effet traitement différent

Tester différentes distributions

- Example cancer colorectal

Table 4. Estimated Mean OS (treatment arms fitted separately)

	Mean OS (months)		
	Placebo/ FOLFIRI	Aflibercept/ FOLFIRI	Difference (95% CI)
ITT			
Best fit: log-logistic	18.9	25.5	6.6 (2.4; 11.6)
Log-logistic with cutoff applied (15 years)	18.1	22.8	4.7 (2.1; 6.1)
Weibull	14.9	17.9	3.0 (1.2; 4.2)
Log-logistic with cutoff applied (5 years)	16.5	19.5	3.0 (1.2; 4.9)
Log-logistic with cutoff applied (10 years)	17.6	21.9	4.2 (1.7; 6.7)

Conclusion

- Essentiel d'estimer avec précision le bénéfice associé à une nouvelle intervention
- Difficulté liée aux données de survie, avec existence de censures
- Des techniques d'extrapolation sont nécessaires pour obtenir l'estimation du bénéfice de survie complet
- Sans extrapolation le bénéfice de survie sera restreint à celui observé directement pendant l'étude clinique, représentant probablement une sous-estimation du gain de survie réelle et des résultats de coût-efficacité inexacte
- Besoin d'une approche systématique
 - Tester différentes distributions
 - Comparaisons graphiques et critères statistiques (AIC, BIC)

References

- Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974, **19** (6): 716–723.
- NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS - EXTRAPOLATION WITH PATIENT-LEVEL DATA. Available at :
http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis_finalv2.pdf
- Schwarz G. Estimating the dimension of a model. The annals of statistics, 1978, **6** (2), p. 461-464
- Collett, D. Modelling survival data in medical research (2nd ed.), Boca Raton: Chapman & Hall/CRC, 2003
- Joulain F., Mean overall survival gain with aflibercept plus FOLFIRI vs placebo plus FOLFIRI in patients with previously treated metastatic colorectal cancer; British Journal of Cancer 2013