



Disponible en ligne sur

ScienceDirect
www.sciencedirect.com

Elsevier Masson France

EM|consulte
www.em-consulte.com



COMMUNICATION BRÈVE

Le concept de l'analyse de survie : vérifier l'applicabilité



Key concepts of survival analysis: Checking appropriateness

P. Mouracade

Department of urology, Strasbourg university hospital, university of Strasbourg, 1, place de l'Hôpital, 67000 Strasbourg, France

Reçu le 22 février 2017 ; accepté le 30 mars 2017

Disponible sur Internet le 5 mai 2017

MOTS CLÉS

Analyse de survie ;
Oncologie ;
Éducation ;
Méthodologie

Résumé L'analyse des données de survie en cancérologie est nécessaire pour évaluer l'efficacité des traitements. L'analyse de survie, bien que considérée comme la pierre angulaire des travaux sur l'oncologie, est le point faible pour l'interprétation des résultats lorsqu'elle est utilisée sans données robustes, sans avoir vérifié les conditions d'application et sans définir la terminologie à utiliser ; conduisant à des fausses conclusions. Un grand nombre de travaux scientifiques en oncologie ne respecte pas les hypothèses nécessaires à l'utilisation de ces tests, par conséquent les conclusions de ces articles ne peuvent pas être assumées par la méthodologie utilisée.

© 2017 Elsevier Masson SAS. Tous droits réservés.

KEYWORDS

Survival analysis;
Oncology;
Assumption;
Education;
Methodology

Summary Analysis of cancer survival data and related outcomes is necessary to assess cancer treatment. The survival analysis, although considered a cornerstone of oncology papers, is the weak point for the interpretation of results when used without robust data, without verifying the conditions of usage, and without defining the terminology we are using; leading to missing the real message. Unfortunately, a lot of scientific papers dealing with oncology violates the assumptions necessary for the usage of these tests, consequently, the conclusions in these papers cannot be assumed by the methodology used.

© 2017 Elsevier Masson SAS. All rights reserved.

Adresse e-mail : pmouracade@hotmail.com

<http://dx.doi.org/10.1016/j.purol.2017.03.012>

1166-7087/© 2017 Elsevier Masson SAS. Tous droits réservés.

Introduction

L'analyse des données de survie en cancérologie est nécessaire pour évaluer l'efficacité des traitements. L'analyse de survie, bien que considérée comme la pierre angulaire des travaux sur l'oncologie, est le point faible pour l'interprétation des résultats lorsqu'elle est utilisée sans données robustes, sans avoir vérifié les conditions d'application et sans définir la terminologie à utiliser ; conduisant à des fausses conclusions.

Un grand nombre de travaux scientifiques en oncologie ne respecte pas les hypothèses nécessaires à l'utilisation de ces tests, par conséquent, les conclusions de ces articles ne peuvent pas être assumées par la méthodologie utilisée.

L'estimateur de la fonction de survie le plus utilisé lorsqu'aucune hypothèse ne veut être faite sur la distribution des temps de survie est l'estimateur de Kaplan-Meier (test non paramétrique). Après avoir estimé les courbes de survie sur deux groupes d'individus ou plus, on est souvent amené à vouloir les comparer. Sous certaines conditions, le test de *log-rank* a le plus fort pouvoir dans la classe des tests de rangs linéaire. Le test de *log-rank* permet de prendre en compte toute l'information sur l'ensemble du suivi sans la nécessité de faire des hypothèses sur la distribution des temps de survie (test non paramétrique). Par souci de simplicité, ce test est présenté par la comparaison de deux groupes mais il est généralisable à un nombre indéfini de groupes de comparaison. Parmi les modèles multivariés d'analyse des durées de survie, le plus utilisé et le plus populaire est le modèle de Cox. Il permet de calculer une courbe de survie avec un ajustement sur l'influence de plusieurs variables.

Ces tests nécessitent le respect de certaines conditions d'utilisation.

L'estimation de Kaplan Meier

L'estimation de Kaplan Meier n'est valide que sous l'hypothèse de censure non informative : le mécanisme de censure est indépendant de l'évènement observé. Autrement dit, la survie des patients censurés ne doit pas différer de la survie des patients non censurés. Considérons un exemple de cancer de rein métastatique, ou les perdus de vue se sont pas rendus en consultation parce qu'ils étaient décédés à leur domicile. Le mécanisme de censure est dépendant de l'évènement observé (le décès). Ignorer la censure informative implique une estimation biaisée de la survie [1].

L'estimation de Kaplan Meier n'est valide que si les patients inclus dans la cohorte au début de l'étude et plus tard ont la même fonction de survie. Autrement dit, absence d'effet période.

L'estimation de Kaplan Meier peut être biaisée quand l'évènement auquel on s'intéresse survient à un temps imprécis sans exactitude. Ce qui est le cas, par exemple, de la récurrence biologique dans le cancer de la prostate, pour laquelle on saura qu'elle est survenue entre 2 dates d'examen du dosage du PSA (censure par intervalle) mais avec une imprécision sur la date exacte de l'évènement.

Lorsque les effectifs des individus à risque diminuent, la précision des estimateurs se dégrade. Il est donc

recommandé de surveiller l'évolution de ces effectifs avec l'augmentation des temps de survie afin de s'assurer qu'un nombre raisonnable d'observations sont utilisées dans la construction des estimateurs de la survie. Il est recommandé de rajouter un tableau avec les nombre de patients exposés au risque à chaque temps (t) permettant d'apprécier la précision de l'information portée par la courbe de survie.

L'estimation de Kaplan Meier requiert que le nombre de censures ne soit pas trop important et déséquilibré entre les groupes. La survenue de censures réduit la précision de l'estimation. En présence d'un taux de censures important, l'intervalle de confiance de la courbe de survie devient large rendant l'interprétation clinique de la fonction de survie compliquée.

Dans l'estimation de Kaplan-Meier, la probabilité cumulée définit la probabilité au début et tout au long d'un intervalle de temps. Les intervalles de temps sont déterminés par les temps de survie non censurés distincts. Cela signifie que plus la taille de l'échantillon est petite, plus les intervalles seront longs, ce qui soulève la question de savoir si l'hypothèse d'une probabilité de survie constante dans chaque intervalle est appropriée [2].

Le test de *log-rank*

Le test de *log-rank* n'est valide que sous l'hypothèse de censure non informative : les perdus de vue doivent être comparable aux non perdus de vue [1].

Le test de *log-rank* n'est valide que si les patients inclus dans la cohorte au début de l'étude et plus tard ont la même fonction de survie. Autrement dit, absence d'effet période.

Lorsque les courbes de survie se coupent alors la puissance du test peut être affectée : il arrive parfois que deux courbes se croisent, alors les risques s'inversent pour chaque groupe à un moment donné. Le test de *log-rank* perd alors en puissance. L'intersection des courbes remet en cause l'hypothèse de risque proportionnel et donc l'optimalité du test de *log-rank*. Il faut que le rapport des risques entre les deux groupes reste constant tout au long du temps, ce qui rend le test de *log-rank* cohérent avec le modèle de Cox [3].

Lorsque le nombre de patients restant à risque devient limité (moins de 20 % de l'effectif à $t=0$), la précision des estimateurs se dégrade. Dans ce cas, l'étendue temporelle des courbes de survie pour l'interprétation des résultats doit s'arrêter au temps correspondant.

Le test de *log-rank* peut être biaisé quand l'évènement auquel on s'intéresse survient à un temps imprécis sans exactitude : censure par intervalle.

Le modèle de Cox

C'est un modèle à risques proportionnels : le rapport des fonctions de risque instantané pour 2 sujets ne dépend que de leurs caractéristiques et ne dépend pas du temps [4]. C'est-à-dire que les fonctions de risque instantané des deux individus a et b sont proportionnelles et que leur rapport de proportionnalité ne dépend pas du temps (t). Il est présumé que si le risque de décès dans un groupe X a un moment (t) est deux fois plus important que dans le groupe Y, le rapport de risque restera le même à n'importe quel autre moment

(t'). La vérification de l'hypothèse de risques proportionnels peut se faire visuellement par la vérification du parallélisme des courbes de survie en échelles log ($-\log S[t]$) versus (t) des deux groupes. Si l'hypothèse des risques proportionnels n'est pas valide en pratique, il se peut que l'effet d'une covariable ne soit pas visible dans le modèle de Cox ou soit atténué.

Le modèle de Cox suppose que la censure est non informative. C'est-à-dire que la censure est indépendante du risque d'acquiescer un événement.

Informations indispensables pour une analyse critique d'une étude de survie

L'analyse critique des études de survie nécessitent la connaissance d'un certain nombre de données. La terminologie est importante afin de comparer les différentes études [5,6].

La date d'origine qui est la date d'entrée du patient dans l'étude : au diagnostic du cancer, la date d'opération, ou la date de greffe. . .

La date des dernières nouvelles qui est la date la plus récente à laquelle on a pu avoir des informations sur le statut du sujet. Cette date peut correspondre à la date de la dernière consultation pour les sujets encore vivants ou à la date de décès pour les sujets décédés, elle peut aussi correspondre à la date du dernier examen scannographique, ou examen biologique.

La date de point qui est la date à laquelle on décide de faire le bilan de l'étude. Cette date est la même pour tous les sujets. La plupart des études rétrospectives ne fixent pas une date de point.

L'évènement étudié peut prendre plusieurs formes : mort, récurrence, progression, ou rechute. Une définition exacte de l'évènement étudié est indispensable. Par exemple, la définition de la récurrence n'est souvent pas standardisée entre les études comparant la néphrectomie partielle et les traitements ablatifs. Dans les études sur la néphrectomie, la récurrence pouvait prendre en compte les récurrences locales et métastatiques alors que dans les séries de traitement ablatif, la récurrence sous entend récurrence locale.

Le critère de jugement utilisé doit être défini clairement, la survie globale est un critère objectif facile à recueillir mais qui nécessite un suivi particulier afin de ne pas avoir trop de perdus de vue. La survie spécifique est bien connue dans le domaine de la recherche clinique et consiste à ne compter comme événements que les décès attribués à la pathologie considérée et eux seuls (les autres sont considérés comme des censures). Cette méthode nécessite de disposer de causes de décès suffisamment fiables. La survie sans progression est un critère composite mesurant directement le rapport bénéfice-risque (regroupement de plusieurs événements délétères-décès, récurrence et progression mesurant un seul effet négatif). Les regroupements d'évènements dans la survie sans progression diffèrent entre les études. Une définition exacte des événements inclus doit être notée dans la section matériel et méthode.

Erreurs dans la présentation des résultats d'analyse de survie

Il y a quelques erreurs courantes réalisées par les chercheurs lors de l'application des outils d'analyse de survie pour leurs recherches [7] :

- la période de suivi de la population n'est pas toujours mentionnée. Il est recommandé de mentionner le suivi chez les survivants ;
- l'analyse de la survie se base parfois sur l'évaluation brut des fréquences d'évènements (décès) sans prendre en compte le temps dans l'analyse de la survie, ce qui produit des résultats biaisés ;
- beaucoup d'études ne donnent pas d'information sur le nombre d'évènements recensés (par exemple, décès, récurrence). Aucune distinction n'est faite pour savoir si un patient a subi un événement ou a été censuré, d'autant plus que le nombre de patients exposés au risque ne sont pas toujours mentionnés ;
- comme avec toutes estimations statistiques, la précision de l'estimation d'une courbe de survie est représentée par un intervalle de confiance (IC à 95 %). L'intervalle de confiance prend la forme de deux courbes correspondant à la limite supérieure et à la limite inférieure. Cet intervalle a une probabilité de 95 % d'inclure la vraie courbe de survie. L'intervalle de confiance est rarement établi dans la littérature.

Support financier

Aucun.

Déclaration de liens d'intérêts

L'auteur déclare ne pas avoir de liens d'intérêts.

Références

- [1] Kramar A, Mathoulin-Pélissier S. Méthodes biostatistiques appliquées à la recherche clinique en oncologie. Édition John Libbey; 2011.
- [2] Mouracade P. Do the data violate Kaplan-Meier assumptions? Eur Urol 2017, <http://dx.doi.org/10.1016/j.eururo.2017.01.047>, pii: S0302-2838(17)30085-4.
- [3] Kleinbaum DG, Klein M. Survival analysis: a self-learning text. Springer; 2013.
- [4] Bewick V, Cheek L. Ball Statistics review 12: survival analysis. J Crit Care 2004;8(5):389-94.
- [5] Miller Jr RG. Survival analysis. Ed. 2 Hoboken, NJ: Wiley-Interscience; 2011.
- [6] Jekel JF, Katz DL, Elmore JG, Wild D. Epidemiology, biostatistics and preventive medicine. Ed. 3 Philadelphia, PA: Elsevier Saunders; 2007.
- [7] Singh R, Mukhopadhyay K. Survival analysis in clinical trials: basics and must know areas. Perspect Clin Res 2011;2(4): 145-8.