

Survival analysis: Part I — analysis of time-to-event

Junyong In¹ and Dong Kyu Lee²

Department of Anesthesiology and Pain Medicine, ¹Dongguk University Ilsan Hospital, Goyang, ²Guro Hospital, Korea University School of Medicine, Seoul, Korea

Length of time is a variable often encountered during data analysis. Survival analysis provides simple, intuitive results concerning time-to-event for events of interest, which are not confined to death. This review introduces methods of analyzing time-to-event. The Kaplan-Meier survival analysis, log-rank test, and Cox proportional hazards regression modeling method are described with examples of hypothetical data.

Keywords: Censored data; Cox regression; Hazard ratio; Kaplan-Meier method; Log-rank test; Medical statistics; Power analysis; Proportional hazards; Sample size; Survival analysis.

서론

실험 또는 임상연구는 연구 대상에게 연구목표인 처치(treatment)를 시행하고 일정 시간이 경과한 후 결과를 모아 분석하는 과정을 거친다. 대부분의 통계분석방법에서는 시간에 대한 개념이 누락되어 있으며, 연구 계획에 설정된 시간이 종료된 후의 결과만 분석한다. 하지만, 관찰 시간이 수 시간 또는 수분 이내로 짧더라도, 서로 다른 시점에서 나타난 결과가 다른 의미를 가진다면, 시간을 포함하는 통계 분석방법이 유리하다[1]. 시간이 포함된 통계 분석방법은 포아송 회귀분석(Poisson regression)과 생존분석(survival analysis)이 대표적이다[2].

*Korean Journal of Anesthesiology*에도 결과에 생존분석을 이용한 논문들이 있다. 전신마취 하 수술 후 압박궤양(pressure ul-

cer)의 발생여부에 따른 90일 생존율 비교[3], 유방암 환자의 수술에 사용된 마취제의 종류에 따라 수술 후 5년간 재발률 비교에 생존분석에 적용되었다[4]. 또 다른 연구에서는 기관삽관이 어려운 상황에서 두 종류의 video laryngoscopes의 기관삽관 성공률과 삽관 시간의 비교에서 생존분석이 적용되었다[5]. 첫 번째와 두 번째는 쉽게 생존분석이 가능한 자료임을 짐작하지만, 세 번째 논문은 기관삽관 시간 비교에서 생존분석을 먼저 떠올리기는 쉽지 않다. 이처럼 “생존”이라는 단어 때문에 환자의 사망이나 혹은 특정 질병의 치료 성공 혹은 실패에 대한 자료에만 사용될 것이라는 오해 때문에 생존분석은 널리 사용되지 못하고 있다. 세 번째 논문과 같이, 생존분석은 사망 자료뿐만 아니라 관찰하는 시간 내에 ‘관심이 있는 사건(event) 이 발생’하는 모든 자료에 적용할 수 있다.¹⁾ 생존분석이 가능한 예는 다음과 같다.

1) 허리통증을 치료하기 위해 경막외 주입(epidural injection)을 시행하면, 호전된 증상이 얼마나 오랫동안 유지되는가? (허리통증 재발까지의 기간 및 유무)

¹⁾두 반응만 나타나는 범주형 변수(categorical variable)인 결과에서 연구자가 관찰하려는 것을 ‘사건’이라고 표현한다. 예를 들자면, 참여자의 사망 혹은 생존 중에서 “사망”, 허리통증의 재발 혹은 완화 중 에서 “완화” 등이다. 노출(exposure), 중재(intervention), 처치(treatment)와 같이 가설의 시작이 되는 원인들을 생존분석에서는 최초의 사건(initial event)라고 하며, 결과에 해당하는 사건(subsequent event)의 발생여부와 이때까지의 시간이 생존시간(survival time)이 자료에 포함된다. 혼란을 방지하기 위해 이 글에서는 최초의 사건(initial event)을 처치(treatment)로, 결과에 해당하는 사건(subsequent event)을 사건(event)로 설명하도록 한다.

Corresponding author: Dong Kyu Lee, M.D., Ph.D.
Department of Anesthesiology and Pain Medicine, Guro Hospital,
Korea University School of Medicine, 148, Gurodong-ro, Guro-gu,
Seoul 08308, Korea
Tel: 82-2-2626-3237, Fax: 82-2-2626-1438
Email: entopic@naver.com
ORCID: <https://orcid.org/0000-0002-4068-2363>

Received: March 12, 2018.

Revised: May 7, 2018.

Accepted: May 8, 2018.

Korean J Anesthesiol June 71(3): 182-191

<https://doi.org/10.4097/kja.d.18.00067>

2) 마약류 진통제를 포함한 정맥내 자가진통조절장치를 사용할 때 발생하는 구역, 구토를 예방하기 위하여 투여한 항구역제의 효과 지속시간이 얼마나 되는가? (구역/구토 예방 기간 및 발현 유무)

3) 전신마취후 발생할 수 있는 인지기능 장애는 언제 나타나는가? (인지장애 발생까지의 기간 및 발생 유무)

위의 세 가지 예는 모두 처치(경막외 주입, 항구역제 투여, 전신마취 시행)에 뒤따르는 사건(허리통증의 재발, 구역/구토의 발현, 인지기능 장애의 발생)이 관찰기간 내에 발생하거나 사건 발생없이 관찰기간이 종료될 수 있으며, 참여자에게는 사건 발생 여부에 대한 상태변수와 관찰기간이라는 시간변수가 각각 부여된다.

시간경과와 관계없이 사건의 발생여부만 분석한다면 chi-squared test나 Fisher's exact test 등으로도 충분하다. 사건이 발생할 때까지의 시간(time)을 분석하더라도 모든 참여자에게서 사건이 발생했다면, Student's *t*-test나 Mann-Whitney's *U* test 등의 방법이 가능하다. 하지만 모든 참여자에게 사건이 발생하는 것은 아니며, 연구 중 참여자가 탈락되는 경우도 있다(상태 호전 또는 악화로 인한 참여포기, 참여자의 거주이전, 다른 원인에 의한 사망 등). 게다가, 탈락 외에도 한정된 연구기간, 결과가 각기 다른 시기에 발생한다는 점 등 관찰기간도 동일하지 않다. 이런 이유로 자료는 불완전할 수밖에 없는데, 이런 자료는 분석에서 제외하거나, 별도의 방법으로 분석해야 한다. 생존분석은 사건이 발생할 때까지의 시간(time-to-event)을 일차 결과(primary outcome)로 두면서 이런 불완전한 자료를 분석하기 위해 필요하다. 이와 같이 생존분석은 환자의 사망에 대한 자료만 분석하는 방법이 아닌, 어떤 관심의 대상이 되는 사건이 관찰기간내에 발생하는 여부에 따라 분석하는 방법으로 여러 가지 상황에서 적용할 수 있다[1].

다른 통계방법과 마찬가지로, 생존분석 방법에는 모수적 방법(parametric methods)과 비모수적 방법(non-parametric methods)이 있다. 마취통증의학 분야에는 모수적 생존분석의 가정들을 만족시키지 못하는 자료가 많기 때문에 대체로 비모수적 생존분석이 사용된다. 모수적 생존분석 방법은 결과 변수가 정규분포, 이분형 분포(binomial distribution), 또는 포아송 분포(Poisson distribution)와 같이 기존에 알고 있는 분포를 따른다는 가정하에 회귀분석을 통하여 생존함수를 구축하는 방법이다. 흔하게 사용되는 모수적 생존분석 방법(parametric survival method)에는 Weibull, exponential, log-logistic, lognormal, generalized gamma model 등이 있다. 주로 사용되는 비모수적 생존분석 방법에는 Kaplan-Meier 곡선(Kaplan-Meier curve)를 이용하는 방법과 Cox 비례위험회귀모형(Cox proportional hazard regression model)이 있으며, 본 글에서는 비모수적 분석방법만 살펴본다.

중도절단 자료(Censored Data)

생존분석을 위한 가장 이상적인 자료는 처치가 행해지는 시점이 명확하고, 모든 참여자가 사건을 경험할 때까지 관찰하는 경우이다. 그러나, 모든 참여자에서 사건이 발생할 때까지 연구하는 것은 현실적으로 어려움이 많기 때문에 사건 발생 없이 관찰기간이 종료되기도 하고, 혹은 관찰기간이 종료되기 전에 중도 탈락되는 경우도 있다. 이와 같이 처치 이후 관찰이 종료되는 시점사에서 사건 발생 여부를 확인할 수 없는 자료를 중도절단 자료(censored data)라 하며,²⁾ 다른 통계분석 방법에서는 결측값(missing value)으로 처리되지만, 생존분석에서는 매우 중요한 자료로 포함된다[6,7].

예를 들면, 허리통증 치료에 신약의 효과를 알아보기 위해 ‘신약의 경막외 주입’이라는 처치를 시행하고 ‘허리통증의 재발’이라는 사건을 연구한다고 하자. 연구기간 도중 교통사고와 같은 다른 원인으로 허리통증이 재발하거나, 다른 질병으로 사망하거나, 혹은 환자가 병원을 옮겨서 허리통증의 재발여부를 알 수 없다면 중도절단 자료로 처리되어야 한다. 중도절단 자료는 관찰이 이루어진 시점까지의 허리통증의 재발여부는 확인할 수 있지만, 관찰이 중단된 시점부터는 재발여부를 알 수 없다. 좀 더 구체적인 예를 보자(Fig. 1). 이 가상의 연구는 6주간 진행되었으며, 모두 6명의 환자가 참여하였다. 1번 환자는 처음부터 연구에 참여하였다. 4주간의 관찰기간이 경과된 뒤 교통사고로 허리통증이 재발하여 관찰이 종료되고 중도절단 자료로 처리되었다. 이 자료는 “4주간의 관찰기간” 동안 신약의 효과를 입증한다. 2번 환자는 연구

²⁾ 중도절단 자료를 명확하게 구분하는 것은 시행되는 생존분석이 비뚤림(bias)이 없도록 하는 중요한 과정이다. 생존분석을 위한 자료의 수집은 특정시점에서 시작되어 다른 특정시점에서 종료되므로, 이들과 관련되는 특성에 따라 중도절단 자료를 구분할 수 있다. 흔하게 발생하는 우측 중도절단(right censored)은 연구기간이 종료되는 시점에서 피험자가 생존하거나(end-of-study censoring), 사망하지 않은 상태로 다른 원인에 의해 관찰이 종료되는 경우이다(loss-to-follow-up censoring). 우측중도절단 자료들은 절단된 시점 이후에도 생존하여 자료 전체의 생존시간을 증가시킬 수 있지만, 중도절단되어 생존시간의 비뚤림을 유발할 수 있는 원인이 된다. 좌측중도절단(left censored)은 연구에 포함되기 이전의 명확하지 않은 시점부터 사건 발생의 원인이 되는 최초의 사건에 노출된 경우이다. 예를 들면, 수술 전 혈장 크레아티닌이 증가된 환자를 대상으로 한다면 수술전 검사에서 혈장 크레아티닌이 증가된 것은 확인 가능하나, 언제부터 증가되어 있었는지에 대해서는 알 수 없다. 이런 경우 좌측중도절단이 발생한다. 환자의 상태를 확인하는 간격에 따라 우측 혹은 좌측중도절단이 발생할 수 있으며, 이런 모든 경우를 구간중도절단(interval censored)이라 한다. 예를 들어, 수술 후 3일 간격으로 혈장 크레아티닌을 검사하여 상승하는 경우를 사건으로 정의한다면, 검사를 시행하는 구간 내에서 어느 시점에 상승하였는지는 정확하게 알 수 없다. 이 경우 우측중도절단에 해당되며, 3일간의 관찰 간격 내에 생존시간의 절단이 발생하므로 구간중도절단에 해당된다. 임상연구는 관찰 대상이 사건이 발생할 때까지 기억없이 지속할 수 없으므로, 연구의 종료시점을 정하고 수행한다. 연구 종료시점이 되면 그 시점까지 생존한 모든 대상은 우측중도절단되며 이런 경우를 Type I censoring이라고 한다.

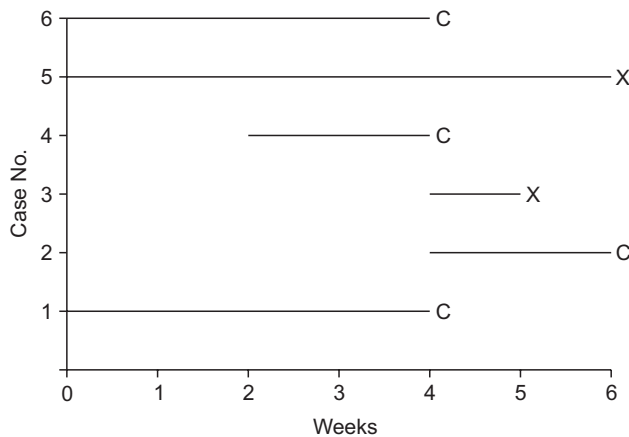


Fig. 1. Illustration of survival analysis data from an imaginary experiment with a 6-week study period. Cases 3 and 5 experienced the event and are coded as “event cases”, denoted “X”. Their follow-up periods were 6 and 1 weeks, respectively. The others are all “censored cases”, denoted “C.” Patient 1 participated in the study from the start, but the follow-up was terminated due to unexpected death from an irrelevant cause. Patient 2 participated in the study from 4 to 6 weeks, but then refused to participate further. Case 4 was enrolled in the study in the 2nd week, lost contact in the 4th week, and was considered lost to follow-up. Case 6 was excluded in the 4th week because of uncovered random allocation.

시작 후 4주차에 참여하였으며 6주차까지도 사건이 발생하지 않았다. 즉, 사건의 발생 없이 연구가 종료된 경우로 역시 중도절단 자료이며, 2주간의 관찰기간이 분석에 포함된다. 3번 환자는 4주차에 참여하여 5주차에 사건이 발생하면서 관찰이 종료된 완전한 자료로 1주간의 관찰기간 및 사건 발생이 분석대상이다. 4번 환자는 2주차에 연구에 참여하였으나 4주차 이후 연락이 되지 않아 (loss of follow-up) 중도절단 되었고, 역시 2주간 관찰기간이 의미를 가진다. 5번 환자는 연구 시작부터 참여하였으며, 연구 종료시점인 6주차에 사건이 발생하여 완전한 자료를 제공하였다. 6번 환자는 연구 초기에 포함되었으나 4주차에 무작위배정이 노출되어 (uncovered random allocations) 탈락된 경우로 중도절단 자료이며, 관찰이 지속된 4주간의 기간은 분석에 포함된다. 결론적으로, 처치와 관련된 사건이 발생한다면 완료된 자료이고(X), 사건의 발생 여부를 확인할 수 없거나 다른 원인에 의해 사건이 발생한다면 중도절단 자료이다[8].

Survival Function, Hazard Function

생존함수(survival function)는 사건, 중도절단과 함께 생존분석의 대표적 용어이며, 생존분석을 이해하기 위한 필수 개념이지만 친숙하지 않아 쉽게 다가가기 어렵다.

생존함수는 특정시점을 포함하여 특정 시점까지 사건이 발생하지 않을 확률을 함수로 나타낸 것이며 보통 $S(t)$ 로 표현한다. 즉,

사건이 ‘허리통증의 재발’이라면 특정 시점까지 ‘허리통증이 없을 확률’을 예측해 준다(Equation 1).³⁾ 생존함수에서, 시간이 0이면 1.0 (즉, 100% 생존)의 확률을 보여주며, 50%의 생존확률을 보이는 시점이 생존시간 중앙값(median survival time)이다.

$$S(t) = P(T > t) = 1 - F(t) \dots \dots \dots \text{equation 1}$$

T : Random variable denoting the time of the event.

t : Any specified time point.

$P(T > t)$: Probability of not experience the event up to and including time t .

$F(t)$: Cumulative distribution function

전체 연구기간 동안 발생한 사건수를 전체 관찰수로 나눈 비율은 전체 사건 발생률(incidence rate)이다. 예를 들어, 사건이 사망이라면 사망률(mortality)이다. 하지만, 연구기간 내내 사건 발생률이 일정한 것은 아니므로, 특정 시점(t)에서의 사건 발생률이 필요한 경우가 있다. 먼저, 특정 시점(t)에서 다음 시점($t + \alpha$ 라고 둔다) 사이에서의 사건 발생률을 계산해 보자. 특정 시점(t)에서 α 의 시간이 경과하는 동안 발생한 사건수를 특정시점(t)에서의 전체 관찰수로 나누면 된다. 이때, α 를 0으로 수렴하면, 즉, 두 시점 사이의 기간이 0에 무한히 가까워진다면, 특정시점(t)에서의 사건발생률이 나오는데, 이것이 위험(hazard)이다.⁴⁾ 그리고, 임의의 시점에서의 사건발생률을 구하는 함수를 위험함수(hazard function)라고 하며, $h(t)$ 로 표현한다[9].

Survival Analysis by Kaplan-Meier Curves (Estimates)

Kaplan-Meier method는 관찰기간이 짧은 경우부터 긴 순서로 나열하여 사건이 발생한 시점에서의 사건 발생률을 계산하는 방법이다. Fig. 1의 자료를 Kaplan-Meier method로 분석하려면 먼저 모든 참여자의 관찰 시작 지점을 0으로 두어야 한다(Fig. 2). 모든 자료를 나열한 뒤, 사건이 발생할 때마다 각 사건 발생 시간을 기준으로 사건 발생률을 계산하여, 생존곡선(survival curve)을 추정한다. Kaplan-Meier method는 관찰기간과 사건 발생 여부의 두 가지에만 초점을 맞추어 분석하는 방법이기 때문에 사건 발생에 영향을 주는 다른 변수(허리통증의 재발의 예에서 보면 성별, 나

³⁾ 만약 중도절단 자료가 전혀 없다면 생존함수($S(t)$)는 “특정시간(t)까지 사건을 경험하지 않은 환자수/전체 연구 참여 환자수”로 간단하게 표현할 수 있다. 다만 현실에서는 중도절단 자료가 하나도 없을 가능성은 희박하기 때문에 equation 1을 사용하여 생존분석이 이루어진다.

⁴⁾ Hazard는 일부 통계학 교과서에서는 hazard rate와 같은 개념으로 다루고 있다. 용어에서 유발되는 혼란의 여지를 줄이기 위해 이 글에서는 hazard rate에 해당하는 경우는 hazard로 표기하였다.

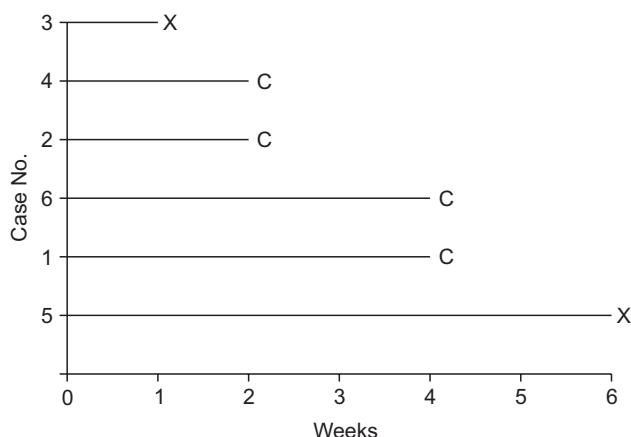


Fig. 2. Data arranged for Kaplan-Meier analysis. Note that for all participants the observations are rearranged to start at zero. Events are denoted "X," and censored data are denoted "C."

이 등)에 대해서는 분석할 수 없다[10].

생존곡선을 나타낸 그래프에 대한 해석은 Fig. 3을 참고한다. 생존곡선에 대한 표준은 없지만, 보통은 중도절단 자료를 점 또는 기호로, 사건의 발생으로 생존율이 감소하는 경우 계단형태의 표시를 사용한다. 또한, 관찰기간이 긴 쪽으로 갈수록 연구에 포함된 환자의 수는 급격하게 감소하여 임상적으로 의미 없는 사건 발생률을 나타낼 수 있으므로, 각각의 시각에서 환자수를 표시하는 것을 권장한다. 학술지에 따라 생존곡선 작성기준을 제시하기도 한다[11].

Kaplan-Meier 생존분석은 생존곡선들 간의 유의한 차이 검정, 중앙 혹은 평균 생존기간(median or mean survival time)의 차이 검정에 사용된다. 따라서, 연구의 주된 목적이 생존기간, 즉 사건 발생까지의 시간인 경우에 적절하다.

Multiple comparison method for Kaplan-Meier survival analysis

둘 이상의 Kaplan-Meier 생존곡선들을 비교하는 방법은 Mantel-Haenszel method, Log-rank method (Mantel-Cox), Gehan's generalized Wilcoxon method (Breslow test), Likelihood ratio 등이 있지만, 흔히 Log-rank method가 사용된다.⁵⁾ Log-rank method는 '비교 대상인 생존곡선에 전반적으로 차이가 없다(no overall difference)'는 귀무가설을 설정하고, 실제 관측된 사건 발생수(observed count)와 사건의 기대 발생 수(expected count)를 카이제곱 검정법(chi-squared method)으로 비교한다. Log-rank method는 비교하는 군들의 사건 발생률이 상당히 차이 나는 경우 검정력이 우수하다고 알려져 있다. 세 군 이상 비교하는 경우, Bonferroni test나 Holm-Sidak method와 같은 사후검정(post-hoc

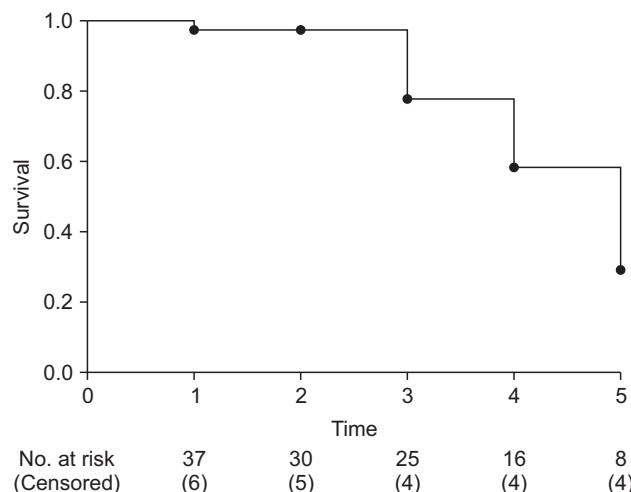


Fig. 3. Example of a Kaplan-Meier survival curve. The axes represent time (x-axis) and survival rate (y-axis). At every time point of event occurrence, the survival curve steps down to the next lower level. Black dots represent points of censored data. At the bottom of the graph, the number of risk and censored cases are presented. It is recommended that specify these numbers with adequate time interval to clarify the clinical meanings.

test)을 통해 군 간 짝비교(pairwise multiple comparisons)도 가능하다[12].

Independent t-test를 적용하기 위해 정규성 검정을 통과해야 하는 것처럼, Log-rank method를 이용하기 위해서는 몇 가지 전제조건이 필요하다. 첫째, 위험비(hazard ratio)가 연구기간 내 내 일정해야 한다(비례위험가정, proportional hazards assumption).⁶⁾ 여기에서 위험비가 일정하다는 것의 예를 보겠다. 전신마취에 사용된 약제가 흡입마취제이든, 정맥마취제이든 수술후 오심/구토의 발생률은 수술 직후 가장 높으며, 시간이 경과함에 따라 감소한다. 즉, 어떤 마취제를 사용한 군이든 시간이 경과함에 따라 수술후 오심/구토의 위험률은 달라지지만(감소하지만), 흡입마취제가 사용된 군에서의 수술후 오심/구토의 위험률은 정맥마취제 사용군에 비해 만일 2배 높다면, 모든 시점에서 항상 2배 이어야 한다는 의미이다. 그리고, 2배라는 것은 1보다 크므로 정맥마취제군에 비하여 흡입마취제군의 수술후 오심/구토의 위험이 크다는 의미이다.⁷⁾ 이렇듯, 위험비는 사건 발생에 대한 위험에

⁵⁾이들은 특정 시점에서 예상되는 사건 발생 수와 실제 관찰된 사건 발생 수의 차이를 이용하여 가중치를 부여한 뒤, 비교하는 방법들이다. Log-rank method는 모든 시점의 가중치를 1로 하여 순위에 기반한 비교결과를 보여준다.

⁶⁾Log-rank method나 Cox proportional hazard regression model의 기본 가정(assumption)이며, 이후 설명되어 있는 내용을 참고한다.

⁷⁾Hazard와 risk 모두 한글로는 "위험"이라고 번역된다. 사전적 의미의 hazard는 사건 발생의 원인이 될 수 있는 것(something could be dangerous)이며, risk는 사건 발생의 가능성(possibility)을 뜻한다.

대한 정보를 한눈에 알 수 있도록 제공한다[13] (Equation 2).

$$\text{Hazard ratio} = \frac{h_2(t)}{h_1(t)} \dots \dots \dots \text{equation 2}$$

Where, $h_1(t)$ and $h_2(t)$: hazard functions of groups 1 and 2

t : any specified time point during study period

여기서 소개된 비례위험가정과 위험비는 다음에 소개되는 Cox 비례회귀모형(Cox proportional regression model)에서도 중요하다[14]. 비례위험가정이 위반된다면 가중치에 1이 아닌, 관찰시점에서 위험에 노출된 환자수를 가중치로 사용하는 Gehan's generalized Wilcoxon test (Breslow test) 또는 다른 방법들을 적용하여야 한다[15].

An example of Kaplan-Meier survival analysis

다음의 생존분석 결과는 수술환자를 대상으로 두 가지 항구토제에 대하여 수술 후 24시간 이내에 구역/구토의 발현 여부에 대한 가상의 연구 자료다(Table 1). 해석의 편의를 위해 나이와 체중, 수술 중 사용한 두 가지 마약류 진통제에 대해서만 분석을 시행하였다. 분석 방법 역시 이해를 돕기 위해 가능한 몇 가지 분석 방법을 중복하여 제시하였다.

Table 2에는 연구에 포함된 환자의 일반적 특성이다. 수술 후 24시간 관찰하는 연구라면 증상 발현 전에 중도에 탈락하는 경우는 흔하지 않지만, 예제를 위해 중도에 탈락하는 환자 수가 꽤 포함되어 있다. 전체 104명의 환자가 포함되었으며, 항구토제 A를 투여 받은 환자는 51명 (A군), 항구토제 B는 53명이다(B군). A군에서는 수술 후 24시간 이내 구역/구토 증상이 25명(49.0%)이, B군에서는 38명(71.7%)이 나타났으며, Chi-squared test 분석결과 유의한 차이를 보였다($\chi^2(1) = 5.597$, P value = 0.018, Cohen's Φ [phi] = 0.23). B군에 대하여 A군의 상대위험은 0.7(95%신뢰구간: 0.5–0.9)이다.

중복된 분석이지만, 두 군의 수술 후 24시간 이내에 발생한 구역/구토의 발생률은 A군에서 49.0%(95%신뢰구간: 62.7–35.3%), B군에서 71.7% (95%신뢰구간: 83.8–59.6)였으며, 역시 유의한 차이를 보였다(z-test, P value = 0.018, Cohen's $h = 0.47$). 이 결과는 24시간 이내에 구역/구토의 발현 없이 중도절단된 환자도 모두 포함된 결과이다. 중도절단되어 연구자가 관찰할 수 없게 된 환자들에서도 연구기간(24시간) 이내에 구역/구토가 발현될 수 있기 때문에 위의 결과가 정확하다고는 할 수 없다. 만약 이 자료에서 수술 후 24시간이 되기 전에 수술 후 구역/구토의 발현 없이 관찰이 종료된 경우를 모두 배제하면 연구에 포함된 환자수는 73명이었고, A군에서의 발생률은 89.3% (95%신뢰구간: 100.0–77.8%), B군에서는 84.4% (95%신뢰구간: 95.0–73.9%)가 된다. 배제된 환자는 모두 수술 후 구역/구토 없이 관찰기간 24시간을 채우지 못한 경우였기 때문에 발생률이 증가하였다. 이 경우 연구자가 의도하지는 않았지만 증상 발현이 없는 환자 수가 감소하였기 때문에 실제 증상 발생률은 높게 추정된다.

이 자료를 Kaplan-Meier 생존분석을 이용한 결과는 다음과 같다(위의 결과와 중복되는 결과인 수술 후 구역/구토 증상의 발현 수 등은 제외하였다). A군의 생존시간(증상 발현이 없는 시간)의 중앙값은 13.0 시간(25%–75%: 8.0–24.0 시간), B군의 생존시간 중앙값은 6.0 시간(25%–75%: 3.0–17.0 시간)이었고, log-rank method에 의한 두 군의 차이는 유의하였다($\chi^2(1) = 6.802$, P value = 0.009). 두 군의 Kaplan-Meier 곡선들을 보면, B군의 누적생존률은 A군에 비하여 시간이 지날수록 급격하게 감소하였다(Fig. 4).

Cox 비례위험회귀모형(Cox Proportional Hazard Regression Model)

Cox 비례위험회귀모형은 회귀분석 방법을 응용하여 중도절단 자료를 처리하는 방법이다. 이 방법은 특정 분포를 가정하지 않

Table 1. Sample of Imaginary Data from an Experiment Testing Two Antiemetics for Use in Controlling Postoperative Nausea and Vomiting

No	Antiemetic	Age (yr)	Weight (kg)	Opioid used	Onset time (h)	PONV	Event	Inclusion
1	A	48	78.5	Remifentanyl	4	Not reported	Censored	Excluded
2	B	48	86.7	Remifentanyl	17	Reported	Event	Included
3	A	54	88.3	Fentanyl	21	Not reported	Censored	Excluded
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
25	B	77	53	Remifentanyl	24	Not reported	Censored	Included
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
104	B	58	60	Fentanyl	1	Reported	Event	Included
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

In this imaginary experiment, two antiemetics (A or B) were administered, and each patient was randomly allocated to a group receiving one of them. Intraoperative opioid used was recorded in the "opioid used" column. Postoperative nausea and vomiting (PONV) and its onset time were also recorded. For statistical tests such as Student's t -test, patients 1 and 3 patients would normally be excluded (as indicated in the column labeled "Inclusion"). For survival analysis, all the patients would be included either as censored cases (numbers 1, 3, 25) or as ordinary event cases (numbers 2, 104).

Table 2. Demographic Data of Patients in the Imaginary Experiment Described in Table 1

Variables	Group A	Group B
Number of patients	51	53
Age (yr)	56.1 (13.7)	61.2 (14.9)
Body weight (kg)	65.7 (13.8)	57.7 (11.6)
Observation time (h)	11.3 (6.7)	8.9 (7.7)
Fentanyl used	19 (37.3%)	22 (41.5%)

All data are presented as mean (SD) or number (%).

으며, Kaplan-Meier 생존분석과 달리 사건의 발생에 영향을 줄 수 있는 다른 변수들의 분석도 가능하므로, semiparametric methods 로 여겨지기도 한다[16,17].

Cox 비례위험회귀모형은 생존함수가 지수함수(exponential function)을 따른다는 것과 두 군의 위험비가 연구기간 동안 일정하게 유지된다는 두 가지 중요한 가정에서 출발한다. 즉, 위험비는 하나의 상수($HR = \lambda$)가 되며, 특정 시점에서의 생존함수는 위험비에 대한 지수함수로 표시된다 ($s(t) = \exp^{-\lambda t}$). 두 생존곡선을 비교하기 위한 귀무가설은 '비교되는 군 간의 위험비는 1이다'가 된다.

Cox 비례위험회귀모형은 각 사건이 발생한 시점에서 사건 발생에 영향을 줄 수 있는 변수에 대해, 사건이 발생한 경우와 그렇지 않은 경우의 위험의 수준을 비교하여 결과를 도출한다[18]. 이 비교의 결과가 위험비이며, Cox 비례위험회귀모형에서 회귀계수로 추정된다. 어떤 변수의 위험비가 1보다 크면서 통계적 유의성도 유지되는 경우, 그 변수에 의해 사건이 발생할 확률이 증가한다는 의미이다.

만약, 연구기간 동안 군 간 위험비가 일정하게 유지되지 않는다면 다른 분석 방법을 사용해야 한다. 왜냐하면, 시간의 변화에 따라 위험비가 달라진다는 의미는 시간의 변화에 따라 회귀계수가 하나의 추정값이 아니라 시간에 따라 다른 추정값들이 존재한다는 뜻이기 때문에 Cox 비례위험회귀모형으로는 분석이 불가능하다. 물론 위험비에 대한 가정과 함께 회귀분석이 가지는 기본적인 가정들(연속변수, 상호작용 등)도 Cox 비례위험회귀모형의 분석에서 고려되어야 한다[19,20].⁸⁾

Cox 비례위험회귀모형의 회귀분석 사례(An example of regression analysis with the Cox proportional hazards regression model)

이전 Kaplan-Meier 생존분석을 하였던 자료를 Cox 비례위험회귀모형으로 다시 분석해 보았다. 이 방법은 회귀분석 방법과 유사하기 때문에 Table 1에 나와있는 공변량을 사용하여 각 공변량의 위험비를 추정하고 유의성을 검증한다.

Table 3은 Cox 비례위험회귀모형 분석 후 추정된 각 공변량들

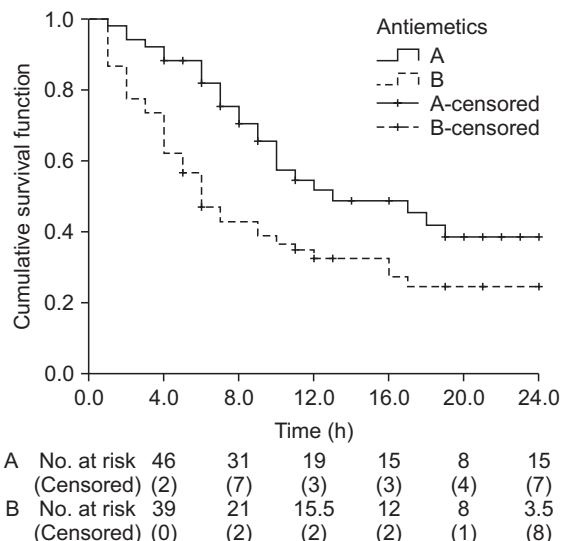


Fig. 4. Kaplan-Meier curves for groups A and B. The median survival time (postoperative nausea and vomiting-free time) for group A was 13.0 h (Q1, Q3: 8.0, 24.0 h), and for group B 6.0 h (Q1, Q3: 3.0, 17.0 h). The median survival time was significantly longer in group A than in group B (log-rank test, $\chi^2(1) = 6.802$, $P = 0.009$). Black crosses indicate censored data.

의 위험비와 유의수준이다. 이 결과는 네 개의 공변량을 모두 모형에 넣어서 분석한 결과로 항구토제만 유의한 위험비가 추정되었다. 이 결과를 바탕으로 우도비(likelihood-ratio)에 바탕을 둔 전진선택법(forward conditional method)을 사용하여 Cox 비례위험회귀모형을 추정하면 항구토제와 수술중 사용한 마약류 진통제가 방정식에 포함되었다. 추정된 항구토제의 위험비는 2.0 (95%CI: 1.2–3.3, P value = 0.009)로 수술 후 24시간 동안 항구토제 B는 항구토제 A에 비하여 수술 후 구역/구토 발생의 위험이 약 2배이며, 수술 중 fentanyl 사용은 remifentanyl 사용에 비해 수술 후 구역/구토 발생의 위험이 약 4.7배 많았다 (위험비 4.7, 95%CI: 2.8–8.1, P value < 0.001). 시간이 지남에 따라 B군의 누적생존곡선이 A군에 비하여 현저하게 낮아지는 것(수술 후 구역/구토가 많이 발생하는 것)을 확인할 수 있다(Fig. 5).

위의 글에서 잠깐 소개되었듯이, Cox 비례위험회귀모형을 추정된 이후에는 비례위험 가정을 위반하지 않았는지 확인하는 것이 반드시 필요하다. 본 예제에서는 Cox 비례위험회귀모형을 가장 간단하게 진단하는 log minus log plot (LML plot)을 소개한다. LML plot은 생존함수에 log-log 변환(log-log transformation)을 하여 하나의 범주형 변수의 각 값에 따른 그래프를 그린 것이다. 몇

⁸⁾ Cox 비례위험회귀모형의 회귀진단 방법은 추정된 모델과 실제자료를 비교해보는 Martingale residuals과 시간의 도표 확인, 이상점(outlier)을 확인할 수 있는 deviance residual 분석, 모형 적합에 영향을 줄 수 있는 지렛대값(leverage value)을 확인하는 delta-beta value 분석, 위험비의 시간에 따른 변화여부를 확인하는 Schoenfeld residual을 분석하는 방법 등이 있다.

몇 통계 프로그램들은 Cox 비례위험회귀모형 추정 후 선택적으로 이 그래프를 출력해주며, 이를 통해 간단하게 연구기간 동안 위험비가 일정하게 유지되어야 한다는 기본 가정을 확인할 수 있다. 비례위험가정을 위반하였는지에 대한 확인 방법은 두 가지 생존분석에 필수적이며, 시간의 로그값을 x축으로, 생존함수에 대한 음의 로그값에 다시 한번 로그를 취한 값을 y축으로 하는 그림을 통하여 확인할 수 있다. LML plot에서 비교하려는 군들의 그래프가 교차하거나 만나지 않고 연구기간 동안 평행을 유지한다면, 비례위험가정을 위반하지 않는 것으로 판단한다[21].

Fig. 6은 예제의 LML plot이다. 항구토제에 따라 두 군으로 표시된 곡선들은 서로 만나거나 교차하지 않고 평행한 곡선을 보이므로 항구토제의 수술 후 구역/구토 발생에 대한 두 군의 위험비는 시간의 경과에도 불구하고 달라지지 않는다고 판단한다. 만약 이 가정이 충족되지 않으면 Cox 비례위험회귀모형은 정확하지 않으므로, 시간에 따른 영향이 분석 가능한 time-dependent Cox regression analysis (non-proportional hazard model)등을 사용하는 것이 바람직하다.

생존분석에 필요한 표본수(Sample Size for Survival Analysis)

생존분석은 관심의 대상이 되는 사건 발생에 대해 시간을 고려하여 분석하는 방법으로, log-rank test나 Cox 비례위험회귀모형에서는 위험비를 이용하여 표본수를 계산한다[22]. 다른 통계분석 방법에서 검정력(statistical power)을 확인하거나 필요한 표본수를 계산하기 위하여 정규분포(normal distribution)와 각 통계 방법에 해당하는 효과크기(effect size, 예를 들어 Cohen's h 등)가 필요하듯이, 생존분석에서 표본수를 계산하기 위해서는 해당되는 생존분석 방법의 기본이 되는 생존모델(survival model)과 효과크기(effect size)에 대한 정의가 필요하다. Log-rank test나 Kaplan-Meier survival analysis는 비례위험가정(proportional hazards assumption)에 기초한 생존함수를 추정하며 생존함수와 위험비는 지수함수 모형(exponential model)의 가정하에 일정한

관계를 가진다.⁹⁾ 비례위험가정을 기본으로 하는 생존분석은 관찰기간동안 일정하게 유지되는 위험비가 효과크기(effect size)로 사용된다[23].

생존분석이 계획된 연구는 가설 설정 단계에서 특정 사건 발생률의 기댓값을 정하고,¹⁰⁾ 두 군의 위험비를 설정하는 것에서 표본수 계산을 시작한다. 예를 들어 대조군에 비하여 실험군에서 발생률이 20% 감소한다고 하면, 위험비는 0.8이 된다(대립가설). 비례위험가정에 의하면 모든 관찰기간동안 위험비는 일정하게 유지되며, 귀무가설은 “두 군의 사건발생에 대한 위험이 차이가 없다” 즉, “위험비 = 1”이 된다[24].

위험비에 대한 대립가설과 귀무가설이 설정되면, 이 가정을 이용하여 적절한 유의수준(α)과 검정력(β)을 유지하기 위한 사건 발생 수를 추정할 수 있다 (Equation 3).

$$\text{Total number of events} = \frac{\left(\frac{z_{\alpha/2} + z_{\beta}}{2} \right)^2}{\pi_1 \pi_2 (\log HR)^2} \dots \dots \dots \text{equation 3}$$

$z_{\alpha/2}$, z_{β} : Standard normal percentiles according to preset significance level and power

π_1 , π_2 : the proportions to be allocated into groups 1 and 2

HR: predetermined hazard ratio

연구에 참여하는 모든 대상이 사건을 경험하는 것은 아니기 때문에, 사건의 총 발생수와 사건 발생의 확률(probability of events)을 이용하여 표본수 계산을 해야 한다(Equation 4).

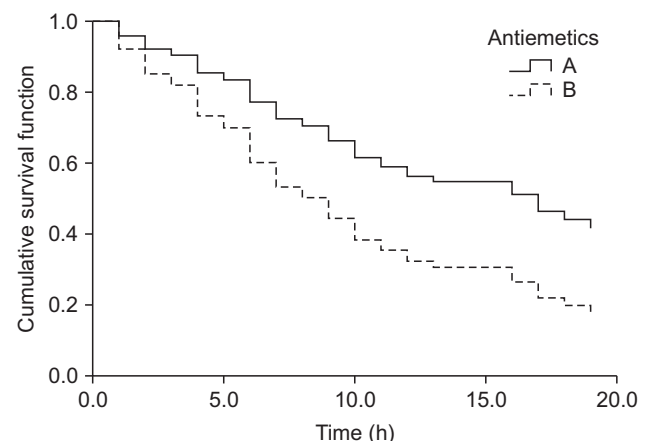


Fig. 5. Cumulative survival curves of groups A and B estimated using the Cox proportional hazards regression model. The survival curve of group B decreases considerably compared to that of group A as time passes. The hazard ratio for antiemetics was estimated as 2.0 (95% CI: 1.2–3.3, $P = 0.009$). Group B showed a 2-fold greater hazard of postoperative nausea and vomiting (PONV) compared to group A. Intraoperative fentanyl increased the PONV risk about 4.7-fold (95% CI: 2.8–8.1, $P < 0.001$) compared to non-use.

Table 3. Estimated Hazard Ratio of Each Covariate Included in Cox Proportional Hazards Regression Modeling

Variables	Hazard ratio	P value
Antiemetic	1.95 (95% CI: 1.14–3.33)	0.014*
Age	0.99 (95% CI: 0.98–1.01)	0.494
Weight	0.99 (95% CI: 0.97–1.02)	0.612
Opioid used	4.87 (95% CI: 2.82–8.39)	< 0.001*

“Antiemetic” and “Opioid used” showed statistically significant hazard ratios. To enhance comprehension, all hazard ratios are presented with 2 decimal points. According to this result, stepwise regression should be carried out.

$$\text{Sample size required } (n) = \frac{\text{Total number of events}}{\text{Probability of an event}} \dots\dots\dots \text{equation 4}$$

본모에 해당하는 사건발생의 확률은 간단하게 나타내면 “1 - 생존확률”이다. 이 확률값은 두 군의 생존함수와 환자수의 비율로 추정이 가능하다(Equation 5).

$$\text{Probability of an event} = 1 - (\pi_1 S_1(t) + \pi_2 S_2(t)) \dots\dots\dots \text{equation 5}$$

π_1, π_2 : the proportions to be allocated into groups 1 and 2

$S_1(t), S_2(t)$: Survival function of groups 1 and 2

생존확률은 앞서 설명된 생존함수에 의해 추정이 가능하다. 그러나 생존분석을 이용하는 연구의 통계학적 목표는 생존함수를 추정하는 것이므로, 연구 설계단계에서 생존함수를 이용하여 사

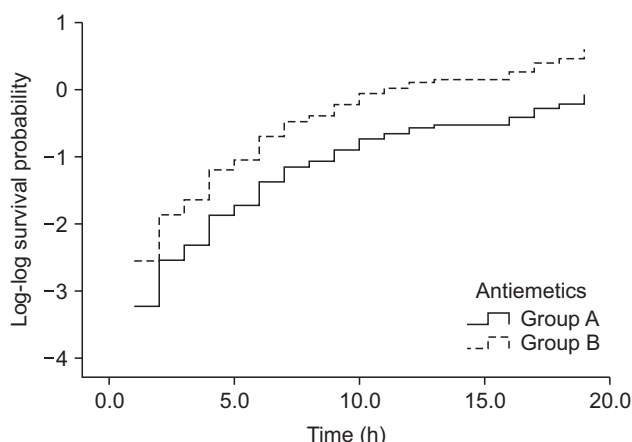


Fig. 6. Log-log function curves for the antiemetic comparison experiment. The Y-axis represents the log-log transformed survival function estimated using the Cox proportional hazards regression model. The basic assumption of Cox proportional hazards regression modeling, a constant hazard ratio over the time (the proportional hazards assumption), can be checked using this plot for only one covariate, the antiemetic. The two curves do not cross over in the early period of observation and remain parallel after that. This plot suggests that this model does not violate the proportional hazards assumption (Created with: IBM® SPSS® Statistics ver. 23 [IBM Corp., USA]).

⁹⁾ 생존함수(survival function, $s(t)$)과 위험함수(hazard function, $h(t)$)은 서로 일정한 관계가 있다. 위험함수는 특정 시간(t)까지 사건이 발생하지 않은 한 명의 환자에게 다음 시간 간격(time segment)동안 사건이 발생할 가능성(instantaneous potential)을 나타낸다. 연구기간 동안 위험함수가 변하지 않는다면 ($h(t) = \lambda$), 생존함수와 위험함수의 관계는 다음과 같다.

$$s(t) = e^{-\lambda t}$$

생존함수는 상수인 λ 와 지수 관계이므로 이와 같은 생존분석 모델을 exponential model이라 한다.

¹⁰⁾ 기댓값은 pilot test나 기존에 알려진 발생률을 사용한다.

건의 확률을 추정하는 것은 불가능하다. 따라서 기존의 알려져 있는 발생률을 이용하여 생존확률을 추정한다. 다만 일반적으로 알려져 있는 발생률은 시간의 개념을 고려하지 않은 경우가 많으므로, 참고하는 연구의 내용을 잘 살펴보아야 한다.¹¹⁾ 시간의 개념이 포함된 발생률이 보고된 선행연구가 없는 경우에는 pilot test를 계획하여 발생률을 추정하여야 한다. 시간의 개념이 포함된 발생률은 인일(person-day)이나 인년(person-year)처럼 규칙적인 시간 간격(time segment)에 대한 사건 발생 수 이므로 그 시간 간격에서의 위험의 수준을 나타낸다. 위험의 수준을 이용하면 비례위험 가정을 기초로 하는 생존함수값의 추정이 가능하며, 가설 수립과정에서 설정한 위험비에 따라 비교하려는 군의 생존함수값도 추정이 가능하다.

허리통증의 재발여부에 대한 예를 다시 고려해보자. 기존 약물 A에 대하여 신약 B의 효과가 허리통증의 재발을 30% 감소시킨다고 가정하자. 이 가설을 검증하기 위한 80% 검정력, 5% 유의수준에 대한 표본수를 구해보자. 표본수 산정을 위해 pilot test를 시행하였고, 기존 약물 A를 경막외 주입한 5명 환자 중 2명에서 통증이 재발하였으며, 5명의 환자를 관찰한 평균 시간이 4주라고 가정하자. 이 경우 위험의 수준(λ)은 $2/(5 \times 4\text{주}) = 0.1/\text{person-week}$ 이다. 생존함수와 위험함수의 관계를 이용하여 약물 A에 대한 4주간의 생존함수값은 $S_A(4) = \exp(-0.1 \times 4) = 0.018$ 이다. 신약 B는 재발을 30% 감소시키므로 위험비(hazard ratio)는 0.7이고, 신약 B에 대한 4주간의 생존함수값은 $S_B(4) = \exp((-0.1 \times 0.7) \times 4) = 0.061$ 이다. 두 군의 표본수를 같게 설정한다면 $\pi_1 = \pi_2 = 0.5$ 이며, 표본수 계산의 분모 값인 사건 발생의 확률은 $1 - (\pi_1 S_1(t) + \pi_2 S_2(t)) = 1 - (0.5 \times 0.018 + 0.5 \times 0.061) = 0.960$ 이다. 표본수 계산의 분자 값인 사건의 총 발생수는 Equation 3에서 구할 수 있다. $z_{\alpha/2}, z_\beta$ 는 표준정규분포에서 해당되는 확률값으로 보통 유의수준 0.05, 검정력 80%일 때 각각 1.96, 0.842가 된다. π_1, π_2 는 각각 0.5, 0.5이며, 위험비는 0.7로 설정하였으므로, 필요한 총 사건수는 $(1.96 + 0.842)^2 / (0.5 \times 0.5 \times (\log 0.7)^2) = 246.9$, 즉 247건이다. 이 값과 사건의 발생률을 Equation 4에 대입하면 $247/0.960 = 257.3$, 즉 258명이 계산된다. 이 값에 일반적으로 고려하는 탈락률 10%를 적용하면 $258/(1 - 0.1) = 286.7$, 즉, 총 287명이 필요한 표본수가 되며, 두 군의 비율을 0.5로 같게 설정하였기 때문에 각 군당 144명이 된다.¹²⁾

¹¹⁾ 시간의 개념을 고려한 발생률은 흔히 알고 있는 발생률과 분모에서 차이가 생긴다. 예를 들면, 5명의 환자를 관찰하였고 이들 중 1명에서 사건이 발생하였다면 발생률은 $1/5 = 0.2$ 이다. 이와 같은 계산에는 시간의 개념이 포함되어 있지 않다. 만약, 이 5명의 관찰 시간이 모두 2일동안 이었다면 시간의 개념이 포함된 발생률은 $1/(5 \times 2\text{ days}) = 0.1/\text{person-day}$ 가 되며 1명의 환자에게 1일동안 사건이 발생할 확률은 0.1이 된다. 마취통증의학 분야 연구들은 인일(person-day)이나 인년(person-year)으로 추정된 발생률을 보고하는 경우는 드물지만, 시간을 변수로 사용하는 연구에서는 관찰시간을 기초로 하는 발생률을 사용한다.

Conclusions

생존분석은 기존의 자료에 시간의 개념을 더하여 최종적인 결과 비교와 함께 관찰기간의 차이를 알 수 있도록 하는 통계방법이다. 마취통증의학과에서 시행되는 많은 연구들은 짧은 길든 간에 시간 변화에 따른 결과를 살펴보는 경우가 많으며, 생존분석을 적용하면 더욱 효과적인 자료 비교와 가설을 뒷받침할 수 있는 강력한 근거를 제시할 수 있다. 이 글은 생존분석을 사용한 연구논문

의 해석에 초점을 맞추었지만, 향후 생존분석이 적극 활용되어 시간 변수까지 포함된 가설 및 연구가 설계 및 진행됨으로써 더 많고 흥미로운 결과가 제시되기를 기대해 본다.

ORCID

Junyong In, <https://orcid.org/0000-0001-7403-4287>

Dong Kyu Lee, <https://orcid.org/0000-0002-4068-2363>

¹²⁾이 표본수는 연구에 포함된 대상자 대부분이 연구기간에 비해 짧게 중도 절단 되지 않는다는 가정하에서 계산된 것이다. 하지만 연구가 진행되면서 모든 환자가 적절한 관찰기간을 다 채울 수는 없기 때문에, 연구 중간에 분석을 다시 시행하여(interim analysis) 적절한 조정을 해야 한다. 연구 진행 도중에 검정력 분석을 다시 시행하는 경우는 처음 설정한 검정력이 조금씩 감소하게 되므로, 이를 보상하기 위해 계산된 표본수 보다 조금 더 많이 설정하여야 한다. 이런 설정을 위해서 Pocock 등은 보수적인 승수(multipliers)를 제시하였으며, 검정력을 80%로 설정하였을 때, 2번의 중간 분석이 계획된 경우에는 1.11, 3번이 계획된 경우에는 1.17을 사용하도록 권장한다. 위의 예에서 만약 2번의 중간 분석이 계획되었다면, $247 \times 1.11 = 274.2$ 즉, 총 275명이 처음 연구 설계 단계에서 필요한 사건수가 된다.

References

1. Kleinbaum DG, Klein M. Survival Analysis. A Self-Learnin Text. 2nd ed. New York, Springer Science+Business Median, Inc. 2005, pp 4-5.
2. Hancock MJ, Maher CG, Costa Lda C, Williams CM. A guide to survival analysis for manual therapy clinicians and researchers. Man Ther 2014; 19: 511-6.
3. Kim JM, Lee H, Ha T, Na S. Perioperative factors associated with pressure ulcer development after major surgery. Korean J Anesthesiol 2018; 71: 48-56.
4. Lee JH, Kang SH, Kim Y, Kim HA, Kim BS. Effects of propofol-based total intravenous anesthesia on recurrence and overall survival in patients after modified radical mastectomy: a retrospective study. Korean J Anesthesiol 2016; 69: 126-32.
5. Shim JH, Jeon WJ, Cho SY, Choe GH. Comparison of the GlideScope and the McGrath method using vascular forceps and a tube exchanger in cases of simulated difficult airway intubation. Korean J Anesthesiol 2016; 69: 133-7.
6. Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. Annu Rev Public Health 1997; 18: 83-104.
7. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. Br J Cancer 2003; 89: 232-8.
8. Selvin S. Survival analysis for epidemiologic and medical research: a practical guide. Cambridge, Cambridge university press. 2008, pp 1-26.
9. Zwiener I, Blettner M, Hommel G. Survival analysis: part 15 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2011; 108: 163-9.
10. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. Int J Ayurveda Res 2010; 1: 274-8.
11. Backman S, Baker A, Beattie S, Brasher P, Bryson G, Cheng D, et al. 2011 Canadian Journal of Anesthesia Guide for Authors. Can J Anesth 2011; 58: 668-96.
12. Mathew A, Pandey M, Murthy NS. Survival analysis: caveats and pitfalls. Eur J Surg Oncol 1999; 25: 321-9.
13. Zwiener I, Blettner M, Hommel G. Survival analysis: part 15 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2011; 108: 163-9.
14. Bewick V, Cheek L, Ball J. Statistics review 12: survival analysis. Crit Care 2004; 8: 389-94.
15. Macin D, Cheung YB, Parmar MK. Survival analysis – A practical approach. 2nd ed. West Sussex, John Wiley & Sons, Ltd. 2006, pp 226-9.
16. Ishak KJ, Kreif N, Benedict A, Muszbek N. Overview of parametric survival analysis for health-economic applications. Pharmacoeconomics 2013; 31: 663-75.
17. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis--an introduction to concepts and methods. Br J Cancer 2003; 89: 431-6.
18. George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol 2014; 21: 686-94.
19. Ng'andu NH. An empirical comparison of statistical tests for assessing the proportional hazards assumption of Cox's model. Stat Med 1997; 16: 611-26.

20. Abairra V, Muriel A, Emparanza JI, Pijoan JI, Royuela A, Plana MN, et al. Reporting quality of survival analyses in medical journals still needs improvement. A minimal requirements proposal. *J Clin Epidemiol* 2013; 66: 1340-6.
21. Macin D, Cheung YB, Parmar MK. *Survival analysis – A practical approach*. 2nd ed. West Sussex, John Wiley & Sons, Ltd. 2006, pp 223-5.
22. Schmidt R, Kwiecien R, Faldum A, Berthold F, Hero B, Ligges S. Sample size calculation for the one-sample log-rank test. *Stat Med* 2015; 34: 1031-40.
23. Schmoor C, Sauerbrei W, Schumacher M. Sample size considerations for the evaluation of prognostic factors in survival analysis. *Stat Med* 2000; 19: 441-52.
24. Julious SA. *Sample sizes for clinical trials*. Boca Raton, CRC Press. 2009, pp 63-9.