

상대생존율

Relative Survival Rate

배 중 면
Jong-Myon Bae

제주대학교 의과대학 예방의학교실
Department of Preventive Medicine, Cheju National University College of Medicine

생존율 산출의 의의와 고려사항

환자의 진단, 치료, 예후 등을 주로 연구하는 의학 연구뿐만 아니라, 질병 예방과 관리대책을 개발하는 보건 연구에 있어서, 연구 대상자의 생존율은 각종 의료행위 및 보건정책 사업의 결과를 평가하는데 매우 중요하게 사용하는 통계 값이다 [1]. 그 예로 2003년 현재 국민 사망원인 1위인 악성 암에 있어서 암 관리 측면에서 사망률은 각종 진단검진 사업과 암 진단 및 치료 효능에 대한 평가지표로 활용할 수 있는 것이다 (Table 1).

생존율을 산출하려면 관찰을 시작한 기준, 관찰을 마치는 기준, 그리고 관찰 단위의 3가지 내용을 정해야 한다 [2]. 예를 들어 새로이 개발된 항암제의 치료 효과를 평가하기 위하여 임상시험을 수행한다면, 대상자가 무작위 배정을 받은 날과 사망한 날을 관찰 시작과 관찰 종료로 각각 정하고, 그 동안의 기간을 추적 관찰기간으로 삼아 생존율을 산출한 뒤, 비교치료군의 생존율과 비교하여 그 효능의 정도를 평가하는 것이다.

그러나, 사망률을 산출하는 과정에서 현

실적으로 불완전 관찰례 (censored case)가 발생하는 것을 피할 수가 없다. 관찰기간 동안 대상자가 다른 원인으로 사망하거나, 행방불명으로 생사여부를 더 이상 확인할 수 없게 되거나, 연구를 종료할 때까지 생존할 경우에 해당 대상자의 생존기간은 완전히 관찰되었다고 할 수 없기 때문이다. 예를 들어 2000년도에 특정 암으로 첫 진단을 받은 10명의 환자를 1년간 추적관찰을 했을 때 4명이 살아있는 반면, 3명은 해당 암으로 사망하고, 교통사고, 자살, 폐혈중으로 각각 1명씩 사망했다고 가상해보자. 이 경우 1년 생존율은 40%라고 할 수 있지만, 해당 암이 아닌 다른 원인으로 사망한 3명을 고려한다면 40%라는 것이 완전한 의미의 1년 생존율은 아닌 것이다.

따라서, 연구 대상 질병 이외의 사망이 없다고 가정할 가운데 해당 질병을 가진 대상에서 생존할 확률인 순 생존율 (Net survival)을 산출할 필요성이 있으며, 이를 위하여 원인별 생존확률 (cause-specific survival rate; 이하 CSR)과 상대생존율 (relative survival rate; 이하 RSR)의 2가지 지표로 제시하고 있다 [3].

원인별 생존확률 (CSR)

사망을 관찰 종료로 정의하여 생존율을 산출하는 연구로서 임상 시험을 수행할 경우라면, 대부분 연구 대상자 모두에 대하여 완전한 추적이 가능하다. 즉, 대상자 개개인 별로 사망원인을 명확히 알 수 있기에, 해당 질병이 아닌 이유로 사망한 경우는 불완전 관찰례로 처리한 뒤, Kaplan-Meier 법이나 Life table 법으로 생존율을 산출한다 [2]. 이처럼 사망의 원인을 알 수 있는 경우 원인별 생존확률을 구하여서 특정 질병의 순 생존율을 제시할 수 있는 것이다.

그러나, 지역주민에 있어 통계청의 사망자 데이터베이스를 근거로 사망 여부를 추적하는 역학연구를 수행할 경우라면, 앞서 임상시험과 달리 사망자 개개인 별로 사망원인을 명확히 알 수 없는 한계에 접한다. 물론 통계청의 사망자료에는 사망의 원인이 기재하도록 되어 있으나, 해당 정보에 대한 타당성을 추가적으로 확보하지 않은 상태에서 연구에 직접 사용하는 것은 문제가 될 수 있다. 예를 들면 어떤 암환자로 항암제 투여를 받다가 폐혈증이 생겨 사망한다고 가정할 때, 기재되는 사망원인이 해당 암으로 기록될 수도 있겠지만, 폐혈중으로도 기재될 수 있어서 원인별 생존확률을 구하려면 지역주민에서 사망자 모두를 추적 확인해야만 가능하다.

상대생존율 (RSR)

앞서의 통계청 사망자료를 이용한 역학 연구와 같이 개별적으로 사망 원인을 명

Table 1. Components of cancer control and outcome measures

Components	Outcome measures
Primary prevention	incidence, mortality
Screening	mortality, %localized, survival
Early diagnosis	survival, %localized, mortality
Treatment	survival, mortality
Rehabilitation	quality of life
Palliative care	quality of life

Source: Ref [1]

2004년도 대한 예방의학회 하계 워크숍 발표 내용을 특강형식으로 정리하였음.
책임저자: 배중면 (제주도 제주시 이내동, 전화: 064-754-3856, 팩스: 064-725-2593, E-mail: jmbae@cheju.ac.kr)

확이 알지 못하는 상황에서, 연구 대상자에 대한 순 생존율을 제시하기 위하여 개발된 지표로 상대생존율이 있다.

상대생존율은 연구 대상자로부터 얻어진 관찰생존율 (observed survival rate)을 기대생존율 (expected survival rate)로 나누어 얻어낸다 [3]. 여기서 기대생존율이란 연구 대상자가 속해 있는 집단에서 해당 질병이 없는 가운데 생존할 확률을 의미한다. 즉, 해당 질병이 없다는 가정하에 생존할 확률을 얻기 위하여 관찰생존율을 기대생존율로 나누어 구한다는 개념이다.

이러한 상대생존율을 산출하는 공식에 따르면 상대생존율의 측정 단위는 원칙적으로는 비 (ratio)에 속한다. 그러나, 상대생존율의 의미가 해당 질병이 없다는 가정하에서는 생존할 확률인 점에서 측정단위는 분율 (proportion)에 해당될 수 있으며, 통상 생존확률을 생존율로 언급하면서 survival rate로 기술하고 있어 연구자들 사이에 혼선이 있는 것이 사실이다. 본 논고에서는 생존율 (survival rate)로 일관하게

사용하지만, 측정단위는 분명 비율(rate)가 아님을 분명히 한다.

Table 2는 가상의 연구자료인데, 구간 상대 생존율 (c율)은 구간 관찰 생존율 (a율)을 구간 기대 생존율 (b율)로 나누어 구한다. 여기서 기대 생존율은 통상 통계청에서 제시하는 성별, 연령별, 연도별 생명표 자료를 이용한다. 이럴 경우 개개인 별로 사망원인을 추적하지 않아도 된다는 장점이 있다. 그렇지만 전국민을 대상으로 만들어진 생명표 자료는 연구 대상질병에 의한 사망을 포함하여 만든다는 점에서 완전한 의미의 기대 생존율 자료가 될 수 없다고 볼 수 있다. 이에 Ederer 등 [4]은 생명표 자료를 이용한 상대생존율은 순 생존율의 의미를 손상하지 않는다고 발표한 이래, 특정 집단에서의 기대 생존율로 해당 집단의 성별, 연령별, 연도별 생명표 자료를 적용하고 있다.

그러나, 암 같은 만성질환인 경우 5년 이상의 장기간 생존율을 얻을 필요가 있을 경우, 해당 구간별 생존율을 누적한 누적

생존율 (cumulative survival rate)을 산출하고 있다. 이 경우에 있어서도 누적 상대생존율 (cumulative relative survival rate)은 누적 관찰 생존율에 누적 기대 생존율을 나누어 산출하면 된다. 가상의 연구자료인 Table 2에서, 구간 관찰 생존율 (a율)을 누적하여 구한 누적 관찰 생존율 (d율)을 누적 기대 생존율 (e율)로 나누어 f율이 누적 상대 생존율이 된다. 여기서 누적 기대 생존율을 어떻게 적용하는가가 누적 상대 생존율 산출의 핵심이 된다. 개념적으로는 각 대상자마다 추적 연도별로 성별, 연령별 기대 생존확률을 일일이 붙인 후, 이를 누적하여 누적 기대 생존율을 구하게 된다. 이러한 산출 과정에서 누적하는 방법에 따라 Ederer I, Ederer II, Hakulinen 법 등이 개발되어 있으며, 현재 대부분의 상대생존율 산출에는 Hakulinen법을 사용하고 있다 [4-6].

이러한 산출과정을 쉽게 해 주는 소프트웨어가 현재 몇 가지 개발되어 있지만, 무료로 쉽게 사용할 수 있는 것 2가지를 소개하고자 한다. 첫째는 핀란드 암등록본부가 만든 SURV3이란 간단한 프로그램이다 [7]. 해당 홈페이지에서 쉽게 다운로드할 수 있으며, 간단한 매뉴얼이 제공되고 있다. 압축된 파일내에 있는 POPMORT.data란 파일을 열어 연구대상의 추적연도별 성별, 연령별 사망률과 기대 생존율을 입력하고, MEL7594.dat란 파일은 연구대상자의 추적결과를 변수형식에 맞게 입력하면 바로 누적 상대생존율과 이의 95% 신뢰구간을 얻어낼 수 있다. 둘째는, SAS 통계프로그램에서 분석이 가능한 macro 명령문을 사용하는 것이다. Paul Dickman이 만든 것으로 해당 홈페이지에서 다운로드 받을 수 있다 [8]. 특히 split.sas는 개별 대상자에게 해당 연도별로 기대 생존확률을 붙여주는 명령문이 담긴 것으로 분석때 같은 폴더에 위치해 놓으면 된다. 분석을 수행하려면 연구대상자의 추적 결과 자료와 기대 생존 확률 자료를 준비하면 된다.

Table 2. A hypothetical example for calculating the relative survival rate

Age	Interval Observed Survival Rate (a)	Interval Expected Survival Rate (b)	Interval Relative Survival Rate (c)	Cumulative Observed Survival Rate (d)	Cumulative Expected Survival Rate (e)	Cumulative Relative Survival Rate (f)
40-44	0.824	0.982	0.839	0.824	0.982	0.839
45-49	0.679	0.815	0.833	0.559	0.800	0.699

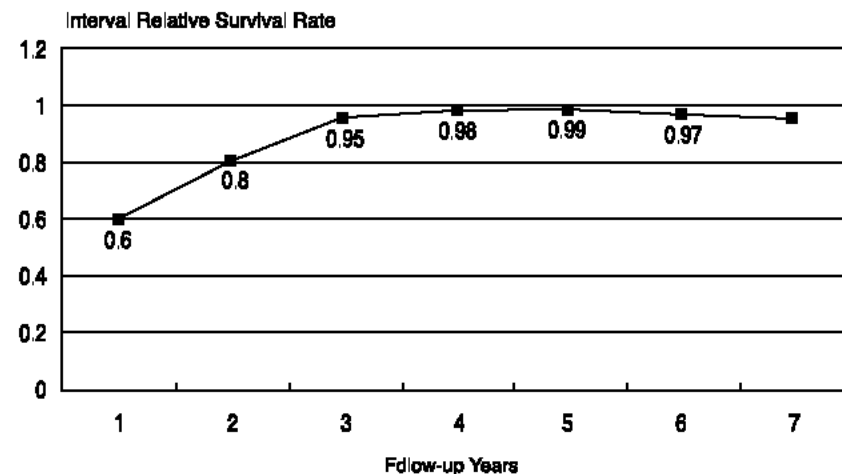


Figure1. An example of statistical cure.

¹ the proportion of patients who would have survived t years or more following diagnosis in the hypothetical situation where the disease of interest were the only possible cause of death

상대생존율의 활용

상대생존율은 순 생존율을 제시하는 통계지표라는 측면뿐만 아니라, 역학 연구에서 다른 측면에서도 활용이 가능하다. 여기서는 2가지 활용에 대하여 간단히 설명한다.

첫째, 통계적 완치 (statistical cure)라는 개념을 제시할 수 있다 [9]. 통상 악성 암인 경우 5년 이상 생존할 경우를 임상적으로 완치되었다고 판단하고 있는 것과 달리, 해당 암환자의 생존율이 일반 국민과 같다면 통계적으로 완치했다고 할 수 있다는 것이다. 만약 구간별 관찰 생존율을 해당 구간별 기대 생존율로 나누어 구한 구간별 상대 생존율이 1에 수렴한다면 이는 해당 질병의 생존율과 일반 국민의 생존율이 같다는 의미가 되는 것이다. 다시 말해 해당 질병에 의한 사망률이 일반국민보다 더 높지 않다는 의미로 해석될 수 있으며, 암 질환에서는 통계적으로 완치라는 의미로 해석이 가능한 것이다. Figure 1에서 보듯이 5년 차에서 구간 상대 생존율이 0.99로 1에 가까워 지면서, 5년까지 생존한 대상에게는 통계적으로 완치라는 의미를 부여할 수 있다는 것이다.

둘째, 암과 같이 예후가 상대적으로 나쁜 만성 질환에 있어 추적자료의 질 평가 지표로 사용할 수 있다 [10]. 즉 누적 상대 생존율이 1보다 크게 나온다면 이는 일반 인구의 생존보다 더 높은 생존확률을 가진다는 의미이다. 이럴 경우, 암 같은 질환에

서는 정상적으로 나타날 수 없는 현상이므로, 실제 사망했음에도 불구하고 추적이 제대로 이루어지지 못하여서 통계산출 과정에서 아직도 생존하고 있는 것으로 반영되면서 벌어진 것으로 해석하는 것이 보다 타당하다. 따라서, 암 통계연구에서 추적과정과 수집 정보의 질 평가를 위한 지표로 활용할 수 있다는 것이다. 반면, 질병발생이 건강한 생활패턴을 유도하여 질병발생자가 오히려 더 생존할 경우 health patient effect-에는 이러한 내용이 활용될 수 없음을 염두에 두어야 한다. 예를 들어 흡연자로 뇌졸중이 발생한 후 생명을 건진 다음부터 금연을 하고, 절주, 운동, 식이 조절을 적극적으로 하게 될 경우 일반 국민보다 더 생존할 확률이 높아진다면 누적 상대 생존율은 1보다 크게 나올 수 있기 때문이다.

맺음말

통계 결과의 대표성을 확보하기 위하여 지역주민 혹은 일반 국민을 대상으로 생존율을 산출하고자 하는 경우, 상대 생존율은 해당 구성원의 순 생존율을 제시한다는 측면에서 매우 유용하게 사용할 수 있다 [11,12]. 특히 지역 암등록자료와 통계청 사망자료를 이용하여 특정 암의 조기 검진, 진단, 치료, 예후 등을 평가하기 위한 생존율 산출을 해야 할 경우, 누적 상대 생존율 산출을 적극 활용해야 할 것이다.

참고문헌

1. Armstrong BK. The role of the cancer registry in cancer control. *Cancer Causes Control* 1992; 3: 569-579
2. Le CT. *Applied survival analysis*. John Wiley & Sons, 1997
3. Esteve J, Benhamou E, Raymond L. Descriptive Epidemiology. IARC 1994: 229-236
4. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Natl Cancer Inst Monograph* 1961; 6: 101-121
5. Ederer F, Heise H. Instruments to IBM 650 programmers in processing survival computations. Methodological note No. 10, End Results Evaluation Section, National Cancer Institute, Bethesda MD, 1959
6. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982; 38: 933-942
7. SURV3. <http://www.cancerregistry.fi/surv3/>
8. Dickman P. <http://www.pauldickman.com>
9. Cutler SJ, Axtell LM. Partitioning of a patient population with respect to different mortality patterns. *J Am Stat Assoc* 1963; 58: 701-712
10. Kerr GP, Kunkler IH, Langlands AO, Rodger A. (In)curability of breast cancer: a 30-year report of a series of 3933 patients. *Breast* 1998; 7: 190-194
11. Bae JM, Won YJ, Jung KW, Suh KA, Yun YH, Shin MH, Ahn YO, Lee DH, Shin HR, Oh DK, Park JG. 134 KCCR-affiliated hospitals. Survival of Korean cancer patients diagnosed in 1995. *Cancer Res Treat* 2002; 34(5): 319-325
12. Lee KH, Kwon Y, Lee TY. Survival rate and factors affecting survival among patients of lung cancer lived in Daejeon city. *Korea J Epidemiol* 2003; 25(2): 62-75 (Korean)