creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

연세대학교
YONSEI UNIVERSITY

Hypo-exponential 분포 하에서 로그-순위 검정을 이용한 생존함수 비교에 대한 표본수와 검정력 연구

연세대학교 대학원
의학전산통계학협동과정
의학통계학전공
고 희 중

Hypo-exponential 분포 하에서 로그-순위 검정을 이용한 생존함수 비교에 대한 표본수와 검정력 연구

지도교수 남 정 모

이 논문을 박사 학위논문으로 제출함

2016년 12월 일

연세대학교 대학원

의학전산통계학협동과정

의학통계학전공

고 희 중

연세대학교
YONSEI UNIVERSITY

고희중의 박사 학위논문을 인준함

심사위원 남 정 모 인

심사위원 정 인 경 인

심사위원 박 소 희 인

심사위원 송 기 준 인

심사위원 김 윤 남 인

연세대학교 대학원

2016년 12월 일

감사의 말씀

먼저 본 논문을 완성하는 데에 있어 연구계획에서부터 결론에 이르기까지 아낌없이 많은 지도와 도움을 주신 남정모 교수님께 진심으로 깊이 감사드립니다. 그리고 바쁘신 중에도 세심한 지도와 배려를 아끼지 않으신 정인경 교수님, 박소희 교수님, 송기준 교수님, 김윤남 박사님께 감사드립니다.

대학을 졸업하고 의학통계학에 뜻을 두었던 순간부터 지금까지 연세대학교 의학통계학과는 제게 큰 의미가 되었습니다. 그 동안 학교에서 배우고 겪은 수 없이 많은 일들이 지금의 저를 있게 해준 바탕입니다. 의학통계학과를 통해 만났던 많은 후배님들 한명 한명을 떠올리며 감사드리고, 그 인연을 계속 이어나갈 수 있기를 기원합니다.

학교에서 배운 지식을 마음껏 펼칠 수 있는 토대를 마련해 주고, 때론 그 보다 더 큰 지혜를 발휘할 수 있도록 저를 믿어준 드림씨아이에스의 동료들께도 감사의 마음을 전합니다.

기나긴 시간 동안 어려움을 함께 참고 격려해 준 사랑하는 아내에게 감사하며, 세상에서 가장 아름답고 소중한 예람이에게 앞으로도 영원히 부끄럽지 않은 아빠가 되고자 최선을 다하겠다는 다짐을 끝으로 모든 분들께 다시 한 번 머리 숙여 감사드립니다.

2016년 12월

고희중 올림

연세대학교
YONSEI UNIVERSITY

차 례

그림 차례 ····· iii
표 차례 ····· iv
국문 요약 ····· v

제 1 장 서론 ····· 1
1.1 연구 배경 및 목적 ····· 1
1.2 논문의 구성 ····· 4

제 2 장 이론적 배경 ····· 5
2.1 로그-순위 검정 통계량의 분포 ····· 6
2.2 대립가설 하에서 로그-순위 검정 통계량의 분포 (비 중심 평균 모수) ····· 7
2.3 전체 생존기간의 분포 ····· 9

제 3 장 제안 방법 ····· 12
3.1 Hypo-exponential 분포를 갖는 로그-순위 검정 통계량의 분포 ····· 12
3.2 Hypo-exponential 분포를 갖는 로그-순위 검정의 표본 크기 ····· 15

제 4 장 모의 실험 ····· 16
4.1 모의 실험의 조건 ····· 16
4.1.1 기본 조건 ····· 16
4.1.2 세부 조건 1 ····· 16
4.1.3 세부 조건 2 ····· 17
4.2 모의 실험의 결과 ····· 17
4.2.1 유의수준 5% & 검정력 80% ····· 17
4.2.2 유의수준 5% & 검정력 90% ····· 13

연세대학교
YONSEI UNIVERSITY

4.2.3 유의수준 1% & 검정력 80% ····· 28
4.2.4 유의수준 1% & 검정력 90% ····· 33

제 5 장 고찰 및 결론 ····· 38
5.1 고찰 ····· 38
5.2 결론 ····· 40

참고문헌 ····· 41

부록 ····· 44

ABSTRACT ····· 56

연세대학교
YONSEI UNIVERSITY

그림 차례

그림 1. 무진행 생존기간이 각 9개월, 3개월, 진행 후 생존기간이 3개월인 hypo-exponential 분포를 갖는 두 군의 전체 생존기간에 대한 위험함수. ····· 11
그림 2. 전체 생존 기간에 대한 구간의 분할. ····· 12
그림 3. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.05, \beta = 0.2$). ····· 45
그림 4. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.05, \beta = 0.1$). ····· 46
그림 5. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.01, \beta = 0.2$). ····· 47
그림 6. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.01, \beta = 0.1$). ····· 48
그림 7. 무진행 생존기간 시험군 9개월, 대조군 3개월에 대한 위험도와 위험함수비. ····· 49
그림 8. 무진행 생존기간 시험군 9개월, 대조군 6개월에 대한 위험도와 위험함수비. ····· 50

연세대학교
YONSEI UNIVERSITY

표 차례

표 1. 본 논문에서 사용되는 기호와 그에 대한 설명 ····· 5
표 2. Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.05, \beta = 0.2$) ····· 21
표 3. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.05, \beta = 0.2$) ····· 22
표 4 Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.05, \beta = 0.1$) ····· 26
표 5. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.05, \beta = 0.1$) ····· 27
표 6. Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.01, \beta = 0.2$) ····· 31
표 7. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.01, \beta = 0.2$) ····· 32
표 8. Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.01, \beta = 0.1$) ····· 36
표 9. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.01, \beta = 0.1$) ····· 37

연세대학교
YONSEI UNIVERSITY

국 문 요 약

Hypo-exponential 분포 하에서 로그-순위 검정을 이용한 생존함수 비교에 대한 표본수와 검정력 연구

본 논문에서는 전체 생존기간의 비교를 목적으로 하는 항암제 임상시험에서 시험군과 대조군의 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정했을 때, 로그-순위 검정을 이용한 생존함수 비교에 대한 표본수와 검정력에 대하여 연구하였다. 이를 위하여 로그-순위 검정에 대한 검정 통계량을 결정하는 비 중심 평균 모수를 구함에 있어, 검정 통계량의 기댓값을 이용하는 방법(Zhang 방법)과 비 중심 평균 모수를 구성하는 각 함수의 정확한 형태를 이미 알고 있다는 사실을 이용하여 직접 적분을 적용(Schoenfeld 방법)하는 두 가지 방법을 제안하였다.

Zhang 방법에서는 시험군과 대조군의 위험비의 차이가 크면 검정력이 계획한 수준에 비해 낮게 나타났으나, 진행 후 생존기간이 9개월 이상이면 대조군의 무진행 생존기간에 관계없이 계획한 수준의 검정력을 만족하는 것으로 나타났으며, Schoenfeld 방법에서는 진행 후 생존기간과 대조군의 무진행 생존기간에 관계없이 모든 경우에 대하여 계획한 수준의 검정력을 만족하는 것으로 나타났으나, 시험군과 대조군의 위험비의 차이가 크거나 진행 후 생존기간이 3개월 또는 6개월로 짧은 경우에는 계획한 수준의 검정력을 초과하는 결과도 나타났다. 결과적으로 임상시험에 필요한 환자 수는 전체 생존기간의 분포가 hypo-exponential 분포를 따른다고 가정하면, 지수분포를 따른다고 가정하였을 때에 비해 감소하는 결과를 얻을 수 있었다.

핵심되는 말 : hypo-exponential, 로그-순위 검정, 비 중심 평균 모수, 표본수, 검정력, 전체 생존기간, 무진행 생존기간, 진행 후 생존기간

제 1 장 서론

1.1 연구 배경 및 목적

식품의약품안전처에서 발표한 2015년 임상시험 계획 승인 현황 분석에 따르면, 국내에서 승인된 전체 임상시험 계획 건수는 675건으로 2014년의 652건과 비교하여 3.5% 증가하였다 이 중 1상 임상시험은 199건, 2상 임상시험은 115건, 3상 임상시험은 223건이고 기타 임상시험이 4건, 연구자 주도 임상시험은 134건으로 분석되었다. 식품의약품안전처는 특히 2014년과 비교하여 국내 제약사의 2상과 3상 임상시험이 각각 25건에서 42건으로 68%, 42건에서 53건으로 26% 증가한 것은 의약품 개발을 위한 연구개발 투자가 활성화되고, 1상 임상시험에서 제품화 가능성이 확인되어 다음 단계로 진입하는 품목들이 많아졌기 때문으로 해석했다. 또한 다국적 제약사는 그 동안 3상 임상시험에 집중하고 있었으나, 2015년에는 1상과 2상 임상시험이 각각 51건과 73건으로 증가하였으며, 이는 국내 임상시험기관의 인력과 시설 등 우수한 인프라와 국내 임상시험 수행에 대한 신뢰도가 높아져 제품 개발의 성패를 좌우하는 초기 단계 임상시험을 국내에서 많이 실시하고 있기 때문이라고 밝혔다. 2015년에 승인된 임상시험 계획 현황을 효능군 별로 살펴보면 항암제가 255건으로 가장 많았고, 심혈관계 76건, 중추신경계 46건, 소화기계 43건, 내분비계 42건 등의 순으로 나타났다. 그 동안 항암제 임상시험은 지속적으로 증가 추세를 보이면서 2014년 210건에서 2015년에는 255건으로 늘어나 전체 임상시험의 38%를 차지하는 것으로 나타났는데, 이는 면역세포를 활성화하여 암세포를 죽이는 면역항체항암제등에 대한 개발이 활발히 진행되는 것으로 해석되고 있다. 특히 다국적 제약사에서 진행하는 임상시험 중 항암제 관련 임상시험은 절반 이상인 58%를 차지하는 것으로 나타났다(MFDS, 2016).

우리나라에서 암은 1984년 이후 지금까지 전 국민 사망률 1위를 차지하고 있으며, 특히 고령화 추세와 함께 다른 만성질환에 대한 치료법의 향상으로 인하여 상대적으로 암 환자의 발생률과 사망률은 빠르게 증가하고 있어, 새로운 약제개발과 임상시험

을 통한 신약 연구개발은 제약 산업계 입장에서도 매우 중요하다고 할 수 있다. 또한 신약개발과정은 개발에 필요한 비용과 장기간의 연구기간, 그리고 낮은 성공률로 인해 매우 고위험의 투자가 필요한 산업으로 인식되고 있어, 보다 효과적이고 효율적인 약제 개발전략을 통해 성공적인 약물을 빠르게 임상현장과 시장에 진입시킬 수 있도록 하는 체계적인 전략과 다양한 연구 방안이 필요하다. 따라서 이러한 문제점을 해결하기 위해서는 보다 효과적이고 효율적인 임상시험 기술의 개발도 필요하며, 임상시험의 효율성을 증가시킴으로써 새로운 암 치료제 또는 치료법에 투입 되는 산업체의 연구개발 및 임상적용과정에 대한 비용부담을 감소시켜 궁극적으로 환자에게 보다 저렴한 비용으로 보다 효과적인 치료를 제공하고, 국가 사회적으로도 그 부담을 감소시키는 노력이 필요하다(Asan medical center, 2013).

많은 항암제 비교 임상시험에서 시험군의 치료 효과를 확인하기 위하여 전체 생존기간(overall survival, OS)을 일차 유효성 평가 변수로 설정하고, 무진행 생존기간(progression free survival, PFS)을 이차 유효성 평가 변수로 설정한다. 전체 생존기간은 임상시험에 등재된 대상자의 치료 시작 후부터 사망에 이르는 기간을 말하고, 무진행 생존기간은 치료를 시작한 후 질병의 진행(progression of disease, PD)에 이르기까지의 생존기간을 말한다. 몇몇 항암제 비교 임상시험에서 무진행 생존기간에 대한 유의성 검정에서는 통계학적으로 유의한 차이를 보이는 반면, 전체 생존기간에 대한 유의성 검정에서는 유의한 차이를 보이지 못하는 경우가 나타났다(Escudier et al., 2009; Sparano et al., 2010). 전이성 유방암(metastatic breast cancer)에서 ixabepilone/capecitabine 병용군과 capecitabine 단독군의 전체 생존기간을 비교하는 Sparano 등(2010)의 연구 결과, 두 군의 무진행 생존기간의 비교는 유의확률(p-value) 0.0005로 두 군 간에 통계학적으로 유의한 차이가 있었지만, 전체 생존기간의 유의확률은 0.1162로 유의하지 않았다. 콩팥세포암종(renal cell carcinoma)에서 sorafenib 약물의 효과를 보기 위한 sorafenib군과 placebo군의 비교 임상시험인 Escudier 등(2009)의 연구 결과에서도 두 군의 무진행 생존기간의 비교는 유의확률 0.001보다 작아 매우 유의하였지만, 전체 생존기간의 유의확률은 0.146으로 유의하지 않았다. 시간-사건 자료(time-to-event data)에 대한 분포의 차이는 대부분 일정한

위험비를 갖는 비례 위험 모형으로 표현되고Cox, 1972), 로그-순위 검정(log-rank test)은 비례 위험 대립 가설에 대해 근사적으로 가장 검정력(power)이 높은 비 모수적 검정으로 알려져 있다(Fleming and Harrington, 1991). 시간-사건 자료의 표본 크기 계산에 있어, Schoenfeld and Richter(1982)는 지수 분포 가정하에서 임상시험에 필요한 표본 크기 계산을 위한 계산표를 개발하였고, Lachin and Foulkes(1986)는 지수 분포와 최대 우도(maximum likelihood) 방법을 이용하여 일정하지 않은 환자 등록(non-uniform patient entry), 추적 실패(losses to follow-up), 비 순응(non-compliance) 등 실제 상황에서의 표본 크기와 검정력에 대해 논하였고, Lakatos(1988)는 일반적인 대립가설에 대한 로그-순위 검정의 검정력을 계산하기 위한 공식을 유도하였으며, Lakatos and Lan(1992)은 로그-순위 검정의 표본 크기 계산을 위한 여러 방법을 비교하였다. Broglio and Berry (2009)는 두 군의 무진행 생존기간과 전체 생존기간의 비교에 관한 모의실험을 하였는데, 진행 후 생존기간(post-progression survival, PPS)을 질병이 진행한 시점부터 사망할 때까지의 시간으로, 전체 생존기간은 무진행 생존기간과 진행 후 생존기간의 합으로 정의하였고, 몇 가지 가정 아래서 두 군의 무진행 생존기간에 대한 유의성 검정의 검정력은 높았으나, 진행 후 생존기간이 길어질수록 전체 생존기간을 비교하는 로그-순위 검정의 검정력은 낮아짐을 모의실험을 통해 확인하였다. 박소영(2013)은 두 군의 전체 생존기간의 비교에서 차이가 없는 이유에 대해, 무진행 생존기간과 진행 후 생존기간이 각각 지수분포(exponential distribution)를 따르면 전체 생존기간은 더 이상 지수분포가 아닌 hypo-exponential 분포를 따르게 됨으로써, 일반적으로 많이 사용하는 로그-순위 검정의 비례위험 관계를 가정할 수 없으므로 검정력이 낮아질 것으로 생각하였고, 로그-순위 검정외에 Gehan, Peto-Peto, Fleming-Harrington (p=0.5, q=0.5)과 Renyi type test의 검정력을 비교하여 전반적으로 로그-순위 검정통계량이 높은 검정력을 보이고 있음을 확인하였다.

본 논문에서는 전체 생존기간의 비교를 목적으로 하는 항암제 임상시험에서 시험군과 대조군의 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정했을 때,

Schoenfeld(1981)와 Zhang and Quan(2009)에 의해 정리된 로그-순위 검정에 대한 검정 통계량을 결정하는 비 중심 평균 모수(non centrality mean parameter, ϕ)를 구하는 방법을 제시하고, 이를 통해 임상시험에 필요한 표본 크기를 산출하여, 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때 필요한 군당 환자 수 대비 어느 정도의 효용성을 제공할 수 있는지 알아보고자 한다.

1.2 논문의 구성

본 논문에서는 전체 생존기간의 비교를 목적으로 하는 항암제 임상시험에서 시험군과 대조군의 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정했을 때, Schoenfeld(1981)와 Zhang and Quan(2009)에 의해 정리된 로그-순위 검정에 대한 검정 통계량을 결정하는 비 중심 평균 모수를 구하는 방법을 제시하고자 한다. 2장에서는 로그-순위 검정 통계량의 분포와 대립가설 하에서의 로그-순위 검정 통계량을 결정하는 비 중심 평균 모수를 구하는 방법과 함께, hypo-exponential 분포를 따르는 전체 생존기간의 분포를 정리하고, 3장에서는 hypo-exponential 분포를 갖는 로그-순위 검정에 대한 검정 통계량을 결정하는 비 중심 평균 모수를 구함에 있어 검정 통계량의 기댓값을 이용하는 방법(Zhang 방법)과 비 중심 평균 모수를 구성하는 각 함수(function)의 정확한 형태(form)을 이미 알고 있다는 사실을 이용하여 직접 적분을 적용(Schoenfeld 방법)하는 두 가지 방법으로 비 중심 평균 모수를 추정하여 표본 크기를 산출하는 방법을 제안할 것이다. 4장에서는 모의 실험(simulation)을 통해 제안된 방법에 대하여 다양한 조건의 무진행 생존기간과 진행 후 생존기간, 비 중심 평균 모수를 계산하는 기본 단위인 구간, 전체 연구기간에 따른 검정력과 표본 크기를 알아볼 것이다. 마지막으로 5장에서는 본 논문의 의미에 대한 고찰과 함께 결론을 제시할 것이다.

제 2 장 이론적 배경

본 논문에서 사용되는 기호는 다음과 같다.

표 1. 본 논문에서 사용되는 기호와 그에 대한 설명

기 호	설 명
k	시험군(A)과 대조군(B)
t	사건이 발생하는 시점
π	환자가 시험군에 배정될 확률
$h_k(t)$	시험군과 대조군의 위험함수
$n_k(t)$	시점 t에서 시험군과 대조군에 대해 위험 대상이 되는 환자 수
$d_k(t)$	시점 t에서 시험군과 대조군의 사건 수
T	로그-순위 검정 통계량
γ	두 군의 로그 위험비 ($\log\frac{h_A(t)}{h_B(t)}$)
ϕ	비 중심 평균 모수
λ_{1k}	시험군과 대조군의 무진행 생존기간의 위험도
λ_2	시험군과 대조군의 진행 후 생존기간의 위험도
$f_k(t)$	시간-사건 자료의 밀도함수
$F_k(t)$	시간-사건 자료의 누적밀도함수
$S_k(t)$	시간-사건 자료의 생존함수
$H_k(t)$	중도절단 시간의 누적분포함수
$C_k(t_i)$	i번째 구간의 위험함수

2.1 로그-순위 검정 통계량의 분포

Zhang and Quan(2009)에 의하면, 시험군과 대조군의 위험함수(hazards functions)를 $h_k(t)$라고 하고, 연구에 참여하는 n명의 환자가 시험군에 배정될 확률이 π일 때, 사건이 발생하는 각 시점 t에서 시험군과 대조군에 대해 위험 대상이 되는 환자 수(number of patient at risk)는 $n_k(t)$, 각 군의 사건 수는 $d_k(t)$라고 하면, 귀무가설 $H_0 : h_A(t) = h_B(t)$를 검정하기 위한 로그-순위 검정 통계량(T)은 다음과 같고, 귀무가설 아래에서 근사적으로 평균이 0, 분산이 1인 표준정규분포를 따르며, 미리 정해진 제1종 오류(type I error) α에 대해, $|T| \geq z_{\alpha/2}$이면 귀무가설을 기각할 수 있다.

$$T = \frac{\Sigma_t \left\{ d_A(t) - n_A(t) \frac{d(t)}{n(t)} \right\}}{\sqrt{\Sigma_t \left\{ \frac{n_A(t) n_B(t) d(t) \{n(t) - d(t)\}}{n^2(t)\{n(t)-1\}} \right\}}} \cong N(0,1) \qquad (1)$$

또한, Schoenfeld(1981)는 두 군의 로그 위험비가 $\log \frac{h_A(t)}{h_B(t)} = O(\frac{1}{\sqrt{n}}) = \gamma$를 만족하는 대립가설 $H_a : h_A(t) \neq h_B(t)$하에서 n이 무한이 증가할 때, 수식(1)에 정의된 로그-순위 검정 통계량 T의 분포가 비 중심 평균 모수 ϕ를 갖고, 분산이 1인 근사적 정규 분포를 따름을 밝혔다.

$$\phi = \frac{\sqrt{n} \int_0^\infty \log \left\{ \frac{h_A(t)}{h_B(t)} \right\} \pi(t)\{1-\pi(t)\} V(t) dt}{\sqrt{\int_0^\infty \pi(t)\{1-\pi(t)\} V(t) dt}} \qquad (2)$$

여기서,

$$V(t) = (1-\pi) f_B(t)\{1 - H_B(t)\} + \pi f_A(t)\{1 - H_A(t)\}$$

$$\pi(t) = \frac{\pi S_A(t)\{1 - H_A(t)\}}{(1-\pi) S_B(t)\{1 - H_B(t)\} + \pi S_A(t)\{1 - H_A(t)\}}$$

이고, 만일 두 군의 중도절단 프로세스(censoring process)가 같으면, 즉 $H_A(t) = H_B(t)$인 비례 위험 대립가설은,

$$H_a : h_A(t) = h_B(t) e^{\gamma} \tag{3}$$

이고, 수식 (2)의 비 중심 평균 모수 ϕ는 $\gamma\sqrt{\pi(1-\pi)D}$로 간단히 표현할 수 있으며, 여기서 D는 두 군에서의 총 기대 사건 수가 된다. 따라서 제1종 오류 α와 검정력 $1-\beta$를 갖는 총 기대 사건 수는 다음과 같다.

$$D = \frac{(z_{\alpha/2} + z_{\beta})^2}{\gamma^2 \pi (1-\pi)}$$

2.2 대립가설 하에서 로그-순위 검정 통계량의 분포 (비 중심 평균 모수)

Zhang and Quan(2009)은 수식 (1)의 로그-순위 검정 통계량 T의 분자를 S라고 할 때, t시점까지 이용가능한 모든 정보 $H(t)$를 이용한 근사적 조건부 분포(approximate conditional distribution) $d_k(t)|H(t) \sim Bin\{n_k(t), h_k(t)\Delta t\}$와 두 군이 동일한 중도절단 프로세스를 가지고 치료효과(treatment effect)가 0에 가까우면(즉, $\gamma \approx 0$), $\frac{n_B(t) n_A(t)}{\{n_A(t) + n_B(t)\}^2} \approx \pi(1-\pi)$임을 이용하여,

$$E[S(t)|H(t)] \approx \pi(1-\pi)\{E[d_A(t)|H(t)](1-e^{-\gamma})+E[d_B(t)|H(t)](e^{\gamma}-1)\}$$

이고, 따라서

$$E(S) \approx \pi(1-\pi)\left[\sum_t (1-e^{-\gamma})\{n_A(t)h_A(t)\Delta t\}+\sum_t (e^{\gamma}-1)\{n_B(t)h_B(t)\Delta t\}\right]$$

이며, 분모는 구간 $[t,t+\Delta t)$에 최대 하나의 사건이 나타날 수 있어 $d(t)$는 0 또는 1이 되어, $\frac{d(t)\{n(t)-d(t)\}}{n(t)-1}=d(t)$임을 이용하여,

$$E\left[\sum_t \frac{n_A(t)n_B(t)d(t)\{n(t)-d(t)\}}{n^2(t)\{n(t)-1\}}\right] \approx \pi(1-\pi)\sum_t \{n_A(t)h_A(t)\Delta t+n_B(t)h_B(t)\Delta t\}$$

이므로, 비 중심 평균 모수를 아래 수식(4)와 같이 정리하였다.

$$\phi=\frac{\sqrt{\pi(1-\pi)}\left[\sum_t (1-e^{-\gamma})\{n_A(t)h_A(t)\Delta t\}+\sum_t (e^{\gamma}-1)\{n_B(t)h_B(t)\Delta t\}\right]}{\sqrt{\sum_t \{n_A(t)h_A(t)\Delta t+n_B(t)h_B(t)\Delta t\}}} \quad (4)$$

Schoenfeld(1981)에 의하면, 검정 통계량은 분산이 1이고, 비 중심 평균 모수가 다음과 같은 근사적 정규 분포를 따르고,

$$\phi=\frac{\sqrt{n}\int_0^{\infty} g(t)\log\left\{\frac{h_A(t)}{h_B(t)}\right\}\pi(t)\{1-\pi(t)\}V(t)dt}{\sqrt{\int_0^{\infty}\{g(t)\}^2\pi(t)\{1-\pi(t)\}V(t)dt}} \quad (5)$$

$$V(t) = P_B f_B(t)\{1 - H_B(t)\} + P_A f_A(t)\{1 - H_A(t)\}$$

$$\pi(t) = \frac{P_A\{1 - F_A(t)\}\{1 - H_A(t)\}}{P_B\{1 - F_B(t)\}\{1 - H_B(t)\} + P_A\{1 - F_A(t)\}\{1 - H_A(t)\}}$$

여기서, 생존 곡선의 비교에 가장 많이 사용되는 로그-순위 검정에서는 가중치(weight)를 나타내는 $g(t)$가 1이 되어, 수식 (2)의 비 중심 평균 모수와 같음을 정리하였다.

2.3 전체 생존기간의 분포

박소영(2013)은 시험군과 대조군의 무진행 생존기간과 진행 후 생존기간의 분포를 각각 위험도(hazard rate)가 λ_{1k}, λ_2인 지수분포를 따르며 서로 독립이라고 가정하고, t시점에서의 무진행 생존기간의 위험함수와 진행 후 생존기간의 위험함수를 각각 다음과 같이 정리하였다.

$$h_{PFS,k}(t) = \lambda_{1k}, \ h_{PPS,k}(t) = \lambda_2$$

여기서 전체 생존기간을 두 지수분포의 합으로 정의하면, 전체 생존기간은 더 이상 지수분포가 아닌 hypo-exponential 분포 $HYPO(\lambda_{1k}, \lambda_2)$를 따르며, 확률밀도함수 $f_k(t)$와 누적밀도함수 $F_k(t)$는 다음과 같이 정의할 수 있다.

$$f_k(t) = \frac{\lambda_{1k}\lambda_2}{\lambda_2 - \lambda_{1k}}(e^{-\lambda_{1k}t} - e^{-\lambda_2 t})$$

$$F_k(t) = 1 - \frac{\lambda_2}{\lambda_2 - \lambda_{1k}}e^{-\lambda_{1k}t} + \frac{\lambda_{1k}}{\lambda_2 - \lambda_{1k}}e^{-\lambda_2 t}$$

여기서, 위험함수 $h_k(t) = \dfrac{-\dfrac{d}{dt}\{S_k(t)\}}{S_k(t)}$ 이므로,

$$h_k(t) = \frac{\lambda_{1k}\lambda_2(e^{-\lambda_{1k}t} - e^{-\lambda_2 t})}{\lambda_2 e^{-\lambda_{1k}t} - \lambda_{1k}e^{-\lambda_2 t}}$$

가 된다. 따라서 시험군과 대조군에 대한 전체 생존기간의 위험함수와 두 군의 위험비는 다음과 같다.

$$h_{OS,k}(t) = \frac{\lambda_{1k}\lambda_2(e^{-\lambda_{1k}t} - e^{-\lambda_2 t})}{\lambda_2 e^{-\lambda_{1k}t} - \lambda_{1k}e^{-\lambda_2 t}}$$

즉, 두 군의 무진행 생존기간의 위험도 λ_{1k}와 진행 후 생존기간의 위험도 λ_2가 각각 지수분포를 따르면 전체 생존기간의 분포는 더 이상 지수분포가 아니고, 두 군의 위험함수는 시간에 따라 비례하지 않음을 알 수 있다(그림 1).

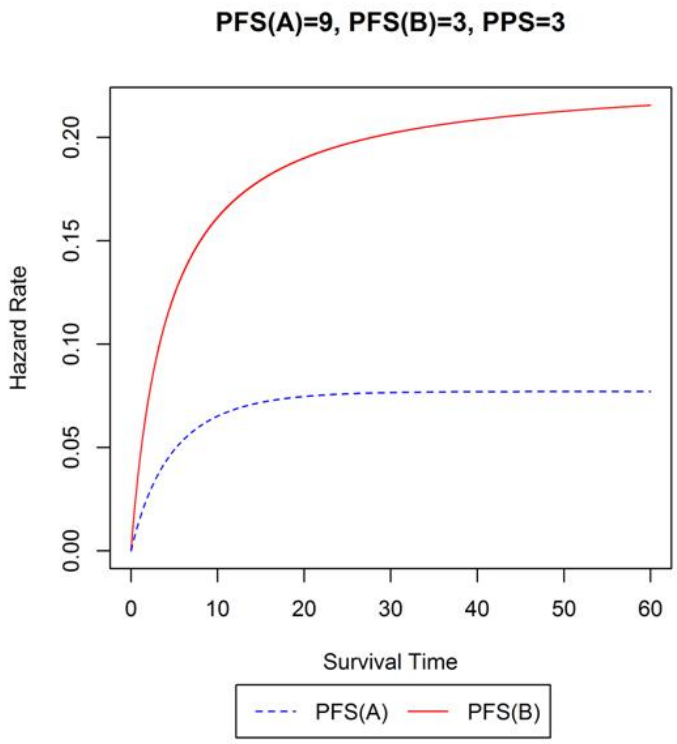

그림 1. 무진행 생존기간이 각 9개월, 3개월, 진행 후 생존기간이 3개월인 hypo-exponential 분포를 갖는 두 군의 전체 생존기간에 대한 위험함수.

여기서 무진행 생존기간의 위험도 λ_{1k}와 진행 후 생존기간의 위험도 λ_2가 같을 때에는(즉, $\lambda_{1k} = \lambda_2$), 전체 생존기간의 분포는 hypo-exponential 분포가 아닌 감마 분포(Gamma distribution)의 성질을 이용하여 정리할 수 있다(Hogg and Craig, 1978).

제 3 장 제안 방법

3.1 Hypo-exponential 분포를 갖는 로그-순위 검정 통계량의 분포

전체 생존기간이 hypo-exponential 분포를 가질 때, 두 군 간의 위험함수는 더 이상 시간에 따라 비례하지 않고 가설(3)의 비례 위험 대립가설이 만족하지 않는다.

$$H_a : \frac{h_A(t)}{h_B(t)} = e^{\gamma_i} \tag{6}$$

두 군 간의 위험함수가 비례하지 않는 대립가설 하에서 비 중심 평균 모수를 구하기 위하여, 전체 연구기간을 총 I개의 구간으로 구분하고, 각 구간에서의 위험도는 해당 구간의 중간 시점($\frac{t_{i-1}+t_i}{2}$)을 이용하여 계산한다(그림 2).

그림 2. 전체 생존 기간에 대한 구간의 분할.

여기서, 수식(1)의 로그-순위 검정 통계량 T는 $\gamma = O(\frac{1}{\sqrt{n}})$이고 두 군의 중도절단 프로세스가 같을 때, 수식 (6)의 대립가설 아래에서 새로운 비 중심 평균 모수를 갖

는 근사적 정규 분포를 따르고, 이를 Zhang and Quan(2009)의 반복 조건부 기댓값 증명(iterative conditional expectation argument)에 따라 정리하면 수식 (4)로부터 다음과 같이 근사시킬 수 있다.

$$\phi = \frac{\sqrt{\pi(1-\pi)}\left[\sum_i (1-e^{-\gamma_i})\left\{n_A(t_{i-1})h_A\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1})\right\} + \sum_i (e^{\gamma_i}-1)\left\{n_B(t_{i-1})h_B\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1})\right\}\right]}{\sqrt{\sum_i \left\{n_A(t_{i-1})h_A\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1}) + n_B(t_{i-1})h_B\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1})\right\}}}$$

여기서, 각 시험군과 대조군에 대해 각 구간 $[i, i+1)$에서의 위험 대상이 되는 환자 수 $n_k(t_i)$는 이전 구간 $[i-1, i)$에서의 위험 대상이 되는 환자 수 $n_k(t_{i-1})$과 이전 구간 $[i-1, i)$에서의 위험함수 $h_k\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1})$을 이용하여 다음과 같이 계산할 수 있다.

$$n_k(t_i) = n_k(t_{i-1})\left\{1 - h_k\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1})\right\}$$

$$= n_k \prod_i \left\{1 - h_k\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1})\right\}$$

여기서, 각 구간의 위험함수 h_k는 해당 구간의 중간 시점 $\frac{t_{i-1}+t_i}{2}$의 위험도를 직접 이용하고, 연구 시작 시점의 위험 대상이 되는 환자 수 $n_k(t_0) = n_k$로 표현한다. 이때, 각 구간의 위험함수 $C_k(t_i)$를 다음과 같이 정의하면,

$$C_k(t_i) = \begin{cases} 0 & , i = 0 \\ h_k\left(\frac{t_{i-1}+t_i}{2}\right)(t_i - t_{i-1}) & , i = 1, \cdots, I \end{cases}$$

이고, 각 구간 $[i, i+1)$에서의 위험 대상이 되는 환자 수는

$$n_k(t_i) = n_k \prod_i \{1 - C_k(t_i)\}$$

로 표현할 수 있다. 따라서 수식 (6)의 대립가설에 대한 로그-순위 검정 통계량의 비중심 평균 모수는

$$\phi = \frac{\sqrt{\pi(1-\pi)}\left[\sum_i (1-e^{-\gamma_i})\left\{n_A \prod_{j=1}^{i}\{1-C_A(t_{j-1})\}C_A(t_i)\right\} + \sum_i (e^{\gamma_i}-1)\left\{n_B \prod_{j=1}^{i}\{1-C_B(t_{j-1})\}C_B(t_i)\right\}\right]}{\sqrt{\sum_i \left\{n_A \prod_{j=1}^{i}\{1-C_A(t_{j-1})\}C_A(t_i) + n_B \prod_{j=1}^{i}\{1-C_B(t_{j-1})\}C_B(t_i)\right\}}}$$

이고, 시험군과 대조군의 비율(π)이 동일하다고 하면, $n_A = n_B = n_k$이므로,

$$\phi = \frac{\sqrt{n_k}\sqrt{\pi(1-\pi)}\left[\sum_i (1-e^{-\gamma_i})\left\{\prod_{j=1}^{i}\{1-C_A(t_{j-1})\}C_A(t_i)\right\} + \sum_i (e^{\gamma_i}-1)\left\{\prod_{j=1}^{i}\{1-C_B(t_{j-1})\}C_B(t_i)\right\}\right]}{\sqrt{\sum_i \left\{\prod_{j=1}^{i}\{1-C_A(t_{j-1})\}C_A(t_i) + \prod_{j=1}^{i}\{1-C_B(t_{j-1})\}C_B(t_i)\right\}}} \quad (7)$$

가 된다.

수식(2)의 비 중심 평균 모수를 적분을 이용하여 직접 얻기 위해서는 2.3절에서 제시한 위험함수 $h_k(t)$, 밀도함수 $f_k(t)$, 중도절단 시간의 분포 함수 $H_k(t)$, 누적밀도함수 $F_k(t)$를 수식 (5)에 직접 이용한다. 여기서 연구 종료 이외의 사유로 인한 중도절단은 없고 시험군과 대조군의 비율은 서로 동일하다($P_A = P_B = \frac{1}{2}$)고 가정한다.

3.2 Hypo-exponential 분포를 갖는 로그-순위 검정의 표본 크기

Hypo-exponential 분포를 갖는 로그-순위 검정의 표본 크기는 비 중심 평균 모수 ϕ를 이용하여, 수식 (4)의 검정 통계량의 기댓값을 이용한 방법과, 수식 (5)의 직접 적분을 이용한 방법으로 계산할 수 있다. 수식 (4)의 검정 통계량의 기댓값을 이용하면 시험군과 대조군 각각에 대한 표본 크기 n_k는

$$n_k = \frac{(z_{\alpha/2}+z_\beta)^2}{\left[\dfrac{\sqrt{\pi(1-\pi)}\left[\sum_i (1-e^{-\gamma_i})\left\{\prod_{j=1}^{i}\{1-C_A(t_{j-1})\}C_A(t_i)\right\}+\sum_i (e^{\gamma_i}-1)\left\{\prod_{j=1}^{i}\{1-C_B(t_{j-1})\}C_B(t_i)\right\}\right]}{\sqrt{\sum_i \left\{\prod_{j=1}^{i}\{1-C_A(t_{j-1})\}C_A(t_i)+\prod_{j=1}^{i}\{1-C_B(t_{j-1})\}C_B(t_i)\right\}}}\right]^2}$$

이고, 수식 (5)의 직접 적분을 이용하면 시험군과 대조군에 대한 총 표본 크기 n은

$$n = \frac{(z_{\alpha/2}+z_\beta)^2}{\left[\dfrac{\dfrac{\sqrt{n}}{\sqrt{2}}\int_0^\infty \log\left\{\dfrac{h_A(t)}{h_B(t)}\right\}\pi(t)\{1-\pi(t)\}2V(t)dt}{\sqrt{\int_0^\infty \pi(t)\{1-\pi(t)\}2V(t)dt}}\right]^2}$$

가 된다.

제 4 장 모의 실험

4.1 모의 실험의 조건

4.1.1 기본 조건

모의 실험을 위하여 시험군의 무진행 생존기간에 대한 중앙생존기간(median survival time)을 9개월로 하고, 대조군의 무진행 생존기간에 대한 중앙생존기간은 3개월부터 8개월까지 6가지로 구분하였다. 두 군의 진행 후 생존기간에 대한 중앙생존기간은 3, 6, 9, 12개월의 4가지로 구분하고, 환자의 등록기간은 총 12개월 동안 매월 초에 균등하게 등록하는 경우를 가정하였다. 모든 모의 실험에서 검정력을 계산하기 위해 매 10,000번의 반복을 수행하였다. 모든 조건에서 유의수준(α)은 각각 5%와 1%의 2가지로, 검정력은 80%와 90%의 2가지로 적용하였다. 결과는 모든 조건에서 Zhang 방법과 Schoenfeld 방법으로 얻은 각각의 군당 환자 수(n_k)와 검정력을 정리하여 제시하였다.

4.1.2 세부 조건 1

시험군과 대조군의 각 중앙생존기간 조건에 대해 모든 사건이 일어날 수 있을 만큼 충분한 정도의 연구기간(PFS와 PPS의 합에 대한 10배)을 기본으로 설정하였다. 결과로는 만일 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때 필요한 군당 환자 수와 검정력을 제시하고, 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정하였을 때 Zhang 방법과 Schoenfeld 방법으로 얻은 각각의 군당 환자 수와 검정력을 정리하여 제시하였다.

4.1.3 세부 조건 2

시험군과 대조군의 각 중앙생존기간 조건에 대해 전체 연구기간을 실제 연구 설계에 적합한 수준(PFS와 PPS의 합에 대한 약 3배)으로 설정하고, 보다 정확한 결과를 확인하기 위해 각 구간의 크기를 1개월($\Delta t = 1$), 1/2개월($\Delta t = 1/2$), 1/3개월($\Delta t = 1/3$)의 3가지 경우로 구분하여 설정하였다. 결과로는 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정하였을 때 Zhang 방법과 Schoenfeld 방법으로 얻은 각각의 군당 환자 수와 검정력을 정리하여 제시하였다.

4.2 모의 실험의 결과

4.2.1 유의수준 5% & 검정력 80%

표 2는 시험군의 무진행 생존기간에 대한 중앙생존기간이 9개월이고, 대조군의 무진행 생존기간에 대한 중앙생존기간이 3개월부터 8개월까지이며, 진행 후 생존기간에 대한 중앙생존기간이 각각 3개월, 6개월, 9개월, 12개월인 경우에서 얻은 군당 환자 수와 검정력에 대한 결과를 정리한 것이다. Zhang 방법에 의한 검정력은 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 80%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 6개월부터 3개월로 짧아질수록 77.68%, 77.33% 76.35%, 76.42%까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 7개월부터 5개월로 짧아질수록 78.25%, 77.44%, 74.57%까지 떨어지고, 4개월과 3개월인 경우에는 68.40%, 63.64%까지도 낮게 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 80%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의

차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 7개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 80%를 만족하는 것으로 나타났다. Schoendeld 방법에 의한 검정력은 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 80% 이상을 만족하였다. 하지만 Schoendeld 방법에서는 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 83.41%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 3개월인 경우 84.86%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 4개월과 3개월인 경우 83.99%, 86.93%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다. 각 모의실험 조건에 대해, 만일 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때 필요한 군당 환자 수와 비교하면, Zhang 방법에서는 지수분포에서의 군당 환자 수 대비 약 50~60% 수준으로 크게 감소하였고, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 3개월인 경우 43.9% 수준으로, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 4개월과 3개월인 경우 48.4%와 38.9% 수준까지도 떨어졌다. 그러나 무진행 생존기간과 진행 후 생존기간이 동시에 낮은 경우에는 검정력 또한 감소한 것을 고려하여 표본 크기의 감소를 주의 깊게 판단할 필요가 있다. Schoenfeld 방법에서도 마찬가지로 약 50~60% 수준의 표본 크기 감소 경향을 보였고, 진행 후 생존기간이 3개월인 경우에는 61%~66.7% 정도까지의 표본 크기 감소가 나타났다. 그러나 진행 후 생존기간이 3개월인 경우에는 검정력이 주어진 80%를 초과한 것을 고려하면 전반적으로 50%~60% 수준의 표본 크기 감소가 나타난 것으로 판단할 수 있다.

표 3은 표 2와 동일한 조건(두 군의 무진행 생존기간, 진행 후 생존기간, $\alpha = 0.05, \beta = 0.2$)을 유지하면서, 전체 연구기간은 실제 연구 설계에 적합한 수준(PFS와 PPS의 합에 대한 약 3배)으로 설정하여, 대조군의 무진행 생존기간이 3개월, 6개월, 9개월, 12개월일 때, 전체 연구기간은 각각 36개월, 48개월, 60개월, 60개월로

가정하였다. 또한 비 중심 평균 모수를 계산하는 기본 단위인 구간을 보다 잘게 구분하여 결과의 정확도(precision)를 높이기 위해 각 1개월, 1/2개월, 1/3개월의 3가지 경우로 구분하여 살펴보았다. Zhang 방법에 의한 검정력은 기본 단위 구간의 크기가 1개월이고 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 80%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 4개월과 3개월인 경우 76.84%, 76.33% 까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 5개월부터 3개월로 짧아질수록 75.26%, 72.03%, 65.95%까지 떨어지는 것으로 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 80%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의 차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 6개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 80%를 만족하는 것으로 나타났다(그림 3). Schoendeld 방법에 의한 검정력은 기본 단위 구간의 크기와 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 80% 이상을 만족하였다. 하지만 Schoendeld 방법에서는 기본 단위 구간의 크기가 1개월일 때, 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 83.26%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 3개월인 경우 84.70%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 6개월과 3개월인 경우 83.90%, 86.87%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다(그림 3). Zhang 방법에서는 기본 단위 구간의 크기가 1개월인 경우에 비해 1/2개월로 세분화하면, 대조군의 무진행 생존기간이 3개월부터 6개월까지는 진행 후 생존기간이 길이에 따라 1명이나 2명의 표본 크기 증가가 있었고, 대조군의 무진행 생존기간이 7개월이나 8개월인 경우에는 진행 후 생존기간이 길이에 따라 1명에서 최대 16명까지 표본 크기가 증가하였다. 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해

대조군의 무진행 생존기간이 4개월부터 7개월까지에서 진행 후 생존기간이 9개월 이상인 경우, 1명까지의 표본 크기 증가가 나타났고, 대조군의 무진행 생존기간이 8개월일 때에는 진행 후 생존기간이 길어질수록 1명부터 최대 5명까지의 표본 크기 증가가 나타났다. Schoendeld 방법에서 대조군의 무진행 생존기간이 3개월부터 6개월까지는 기본 단위 구간의 크기에 관계없이 동일한 표본 크기를 보였고, 대조군의 무진행 생존기간이 7개월인 경우에는 진행 후 생존기간이 9개월과 12개월이면서 기본 단위 구간의 크기가 1개월인 경우에 비해 1/2개월로 세분화하면, 각 1명씩 표본 크기가 증가하였으나, 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해 추가 증가는 없었다. 대조군의 무진행 생존기간이 8개월인 경우에는 기본 단위 구간의 크기를 1개월인 경우에 비해 1/2개월로 세분화하면, 진행 후 생존기간이 3개월부터 9개월까지는 각 1명씩, 진행 후 생존기간이 12개월이면 2명까지의 표본 크기 증가가 있었고, 기본 단위 구간의 크기를 1/2개월에서 1/3개월로 세분화하면, 대조군의 무진행 생존기간이 8개월이면서 진행 후 생존기간이 9개월인 경우에만 추가로 1명의 표본 크기 증가가 있었다. 즉, 모든 경우에 있어 표본 크기의 증가는 크지 않았고, 검정력도 비슷한 수준을 유지하였다.

연세대학교
YONSEI UNIVERSITY

표 2. Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.05, \beta = 0.2$)

Follow up time	Median (month)			Exponential		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power
120	9	3	3	36	80.21	14	63.64	24	86.93
150			6	64	80.30	31	76.42	37	83.99
180			9	99	80.01	55	78.91	58	81.71
210			12	143	80.21	87	79.34	88	80.01
120	9	4	3	57	79.77	25	68.40	37	84.86
150			6	99	80.01	47	76.35	55	82.86
180			9	152	80.04	81	78.04	86	81.07
210			12	218	80.36	127	79.38	130	80.36
120	9	5	3	99	80.01	50	74.57	62	83.41
150			6	167	80.05	80	77.33	90	82.24
180			9	254	80.27	132	79.15	138	80.79
210			12	360	80.50	202	79.52	207	80.72
120	9	6	3	193	79.93	106	77.44	119	81.23
150			6	320	80.11	156	77.68	168	80.91
180			9	479	80.23	245	78.92	253	80.31
210			12	675	80.60	369	79.83	376	79.38
120	9	7	3	476	79.98	278	78.25	291	80.39
150			6	773	80.11	386	78.95	398	80.04
180			9	1,146	80.34	580	79.62	590	80.22
210			12	1,601	80.71	856	79.63	868	80.69
120	9	8	3	2,079	79.99	1,263	80.21	1,274	80.42
150			6	3,316	80.13	1,684	80.06	1,696	80.34
180			9	4,862	80.39	2,442	79.61	2,461	80.47
210			12	6,741	80.79	3,537	79.15	3,573	79.73

연세대학교
YONSEI UNIVERSITY

표 3. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.05, \beta = 0.2$)

Follow up time	Median (month)			$\Delta t = 1$				$\Delta t = 1/2$				$\Delta t = 1/3$			
				Zhang		Schoenfeld		Zhang		Schoenfeld		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power
36	9	3	3	15	65.95	24	86.87	15	65.95	24	86.87	15	65.95	24	86.87
48			6	31	76.33	37	83.90	31	76.33	37	83.90	31	76.33	37	83.90
60			9	55	79.04	58	81.91	55	79.04	58	81.91	55	79.04	58	81.91
60			12	84	79.52	85	79.79	85	79.79	85	79.79	85	79.79	85	79.79
36	9	4	3	27	72.03	37	84.70	27	72.03	37	84.70	27	72.03	37	84.70
48			6	48	76.84	55	82.26	48	76.84	55	82.26	48	76.84	55	82.26
60			9	81	78.44	85	80.88	81	78.44	85	80.88	82	80.04	86	81.15
60			12	123	79.02	126	80.40	124	79.68	126	80.40	124	79.68	126	80.40
36	9	5	3	52	75.26	63	83.26	53	76.34	63	83.26	53	76.34	63	83.26
48			6	82	78.25	91	81.91	83	78.55	91	81.91	83	78.55	91	81.91
60			9	132	79.16	138	80.55	132	79.16	138	80.55	133	79.47	138	80.55
60			12	197	79.65	201	79.82	198	79.99	201	79.82	198	79.99	201	79.82
36	9	6	3	112	78.60	122	81.70	112	78.60	122	81.70	112	78.60	122	81.70
48			6	161	78.92	171	81.14	162	79.43	171	81.14	162	79.43	171	81.14
60			9	246	79.36	254	80.61	247	79.28	254	80.61	248	79.45	254	80.61
60			12	360	79.55	367	80.40	362	79.92	367	80.40	363	80.20	367	80.40
36	9	7	3	293	78.87	304	80.41	294	79.55	304	80.41	294	79.55	304	80.41
48			6	400	79.78	411	80.05	401	79.03	411	80.05	401	79.03	411	80.05
60			9	585	79.24	595	80.14	587	79.88	596	80.40	588	79.51	596	80.40
60			12	841	79.81	852	79.73	846	80.06	853	79.83	847	80.12	853	79.83
36	9	8	3	1,337	79.57	1,353	81.59	1,341	80.39	1,354	81.06	1,342	80.33	1,354	81.06
48			6	1,756	80.78	1,773	80.03	1,760	80.74	1,774	80.13	1,762	80.20	1,774	80.13
60			9	2,480	79.82	2,502	80.54	2,489	79.89	2,503	80.50	2,492	79.52	2,504	80.08
60			12	3,507	79.97	3,540	79.50	3,523	79.62	3,542	79.94	3,528	79.34	3,542	79.94

4.2.2 유의수준 5% & 검정력 90%

표 4에서 Zhang 방법에 의한 검정력은 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 90%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 6개월부터 3개월로 짧아질수록 88.66%, 87.85% 87.74%, 86.87%까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 6개월, 5개월일 때 87.52%, 84.92%까지 떨어지고, 4개월과 3개월인 경우에는 81.32%, 74.71%까지도 낮게 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 90%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의 차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 7개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 90%를 만족하는 것으로 나타났다. Schoenedld 방법에 의한 검정력은 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 90% 이상을 만족하였다. 하지만 Schoenedld 방법에서는 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 92.71%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 6개월과 3개월인 경우 92.61%, 93.17%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 6개월과 3개월인 경우 91.80%, 94.97%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다. 각 모의실험 조건에 대해, 만일 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때 필요한 군당 환자 수와 비교하면, Zhang 방법에서는 지수분포에서의 군당 환자 수 대비 약 50~60% 수준으로 크게 감소하였고, 진행 후 생존기간이 6개월인 경우 대조군의 무진행 생존기간이 6개월부터 3개월로 짧아질수록 48.8%, 47.8%, 47.4%, 48.2%까지 감소하고, 진행 후 생존기간이 3개월인 경우 대조군의 무진행 생존기간이 5개월부터 3개월로 짧아질수록 49.6%, 44.2%, 37.5% 수준까지도 떨어졌다. 그

러나 무진행 생존기간과 진행 후 생존기간이 동시에 낮은 경우에는 검정력 또한 감소한 것을 고려하여 표본 크기의 감소를 주의 깊게 판단할 필요가 있다. Schoenfeld 방법에서도 마찬가지로 약 50~60% 수준의 표본 크기 감소 경향을 보였고, 진행 후 생존기간이 3개월인 경우에는 61%~66.7% 정도까지의 표본 크기 감소가 나타났다. 그러나 진행 후 생존기간이 3개월인 경우에는 검정력이 주어진 90%를 초과한 것을 고려하면 전반적으로 50%~60% 수준의 표본 크기 감소가 나타난 것으로 판단할 수 있다.

표 5에서 Zhang 방법에 의한 검정력은 기본 단위 구간의 크기가 1개월이고 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 90%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 5개월부터 3개월인 경우 87.98%, 87.26%, 86.69%까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 6개월부터 3개월로 짧아질수록 87.81%, 86.71%, 83.68%, 80.24%까지 떨어지는 것으로 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 90%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의 차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 7개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 90%를 만족하는 것으로 나타났다(그림 4). Schoendeld 방법에 의한 검정력은 기본 단위 구간의 크기와 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 90% 이상을 만족하였다. 하지만 Schoendeld 방법에서는 기본 단위 구간의 크기가 1개월일 때, 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 92.78%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 6개월과 3개월인 경우 92.47%, 93.41%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 3개월인 경우 94.94%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다(그림 4). Zhang 방법에서는 기본 단위 구간의 크기가 1개월인 경우에 비해 1/2개월로 세분화하면, 대

조군의 무진행 생존기간이 3개월부터 7개월까지는 진행 후 생존기간이 길이에 따라 1명이나 3명까지의 표본 크기 증가가 있었고, 대조군의 무진행 생존기간이 8개월인 경우에는 진행 후 생존기간이 길이에 따라 5명에서 최대 21명까지 표본 크기가 증가하였다. 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해 대조군의 무진행 생존기간이 3개월부터 7개월까지에서 진행 후 생존기간에 따라, 표본 크기 증가가 없거나 최대 2명까지의 증가가 나타났고, 대조군의 무진행 생존기간이 8개월일 때에는 진행 후 생존기간이 길어질수록 1명부터 최대 7명까지의 표본 크기 증가가 나타났다. Schoendeld 방법에서 기본 단위 구간의 크기가 1개월인 경우에 비해 1/2개월로 세분화하면, 대조군의 무진행 생존기간이 3개월부터 6개월까지는 대조군의 무진행 생존기간이 5개월이면서 진행 후 생존기간이 6개월인 경우에만 1명의 표본 크기 증가가 있었을 뿐, 그 외에는 기본 단위 구간의 크기에 관계없이 동일한 표본 크기를 보였고, 대조군의 무진행 생존기간이 7개월인 경우에는 진행 후 생존기간이 12개월일 때 1명의 표본 크기가 증가하였며, 대조군의 무진행 생존기간이 8개월인 경우에는 진행 후 생존기간이 길어질수록 1명에서 2명까지의 표본 크기 증가가 나타났다. 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해 추가 증가는 없었다. 즉, 모든 경우에 있어 표본 크기의 증가는 크지 않았고, 검정력도 비슷한 수준을 유지하였다.

연세대학교
YONSEI UNIVERSITY

표 4. Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.05, \beta = 0.1$)

Follow up time	Median (month)			Exponential		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power
120	9	3	3	48	89.86	18	74.71	32	94.97
150			6	85	89.87	41	86.87	49	91.80
180			9	133	90.02	74	89.62	78	90.99
210			12	192	90.17	116	89.82	118	89.86
120	9	4	3	77	89.95	34	81.32	49	93.17
150			6	133	90.02	63	87.74	74	92.61
180			9	204	90.04	109	89.55	115	90.58
210			12	293	90.33	170	89.95	173	90.59
120	9	5	3	133	90.02	66	84.92	83	92.71
150			6	224	90.03	107	87.85	120	91.12
180			9	340	90.15	177	88.99	184	90.09
210			12	482	90.34	271	89.40	277	90.04
120	9	6	3	259	89.97	142	87.52	159	91.37
150			6	428	90.03	209	88.66	224	90.74
180			9	642	90.18	328	88.80	339	91.18
210			12	904	90.43	494	89.40	503	89.96
120	9	7	3	638	90.00	372	89.19	389	91.01
150			6	1,035	90.07	516	89.31	533	90.37
180			9	1,534	90.23	776	89.56	790	89.95
210			12	2,144	90.51	1,145	89.79	1,161	89.86
120	9	8	3	2,784	90.00	1,691	89.98	1,705	90.61
150			6	4,439	90.09	2,254	89.77	2,270	89.93
180			9	6,509	90.28	3,268	89.91	3,294	89.57
210			12	9,025	90.57	4,735	89.57	4,783	90.56

연세대학교
YONSEI UNIVERSITY

표 5. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.05, \beta = 0.1$)

Follow up time	Median (month)			$\Delta t = 1$				$\Delta t = 1/2$				$\Delta t = 1/3$			
				Zhang		Schoenfeld		Zhang		Schoenfeld		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power
36	9	3	3	20	80.24	32	94.94	20	80.24	32	94.94	20	80.24	32	94.94
48			6	41	86.69	49	91.74	42	87.72	49	91.74	42	87.72	49	91.74
60			9	73	89.15	77	90.36	73	89.15	77	90.36	74	89.74	77	90.36
60			12	112	88.99	114	89.46	113	89.77	114	89.46	113	89.77	114	89.46
36	9	4	3	36	83.68	50	93.41	36	83.68	50	93.41	36	83.68	50	93.41
48			6	64	87.26	74	92.47	64	87.26	74	92.47	65	88.18	74	92.47
60			9	108	89.15	114	90.37	109	89.58	114	90.37	109	89.58	114	90.37
60			12	164	89.77	168	90.00	165	89.70	168	90.00	166	90.12	168	90.00
36	9	5	3	70	86.71	84	92.78	70	86.71	84	92.78	70	86.71	84	92.78
48			6	110	87.98	121	90.93	110	87.98	122	91.12	110	87.98	122	91.12
60			9	176	89.30	184	90.12	177	89.31	184	90.12	177	89.31	184	90.12
60			12	263	89.76	269	90.32	265	89.75	269	90.32	265	89.75	269	90.32
36	9	6	3	149	87.81	164	91.80	150	88.64	164	91.80	150	88.64	164	91.80
48			6	216	89.30	229	90.67	216	89.30	229	90.67	217	89.13	229	90.67
60			9	329	89.37	339	90.90	331	89.64	339	90.90	331	89.64	339	90.90
60			12	482	89.35	491	90.07	485	89.51	491	90.07	486	90.53	491	90.07
36	9	7	3	392	89.50	407	90.34	393	89.54	407	90.34	393	89.54	407	90.34
48			6	535	89.35	550	90.90	537	89.57	550	90.90	537	89.57	550	90.90
60			9	783	90.06	797	90.41	786	89.60	797	90.41	787	90.54	797	90.41
60			12	1,126	89.65	1,141	90.09	1,132	90.14	1,142	90.25	1,134	90.10	1,142	90.25
36	9	8	3	1,790	89.99	1,811	90.56	1,795	90.24	1,812	90.10	1,796	89.85	1,812	90.10
48			6	2,351	90.11	2,373	90.16	2,357	90.33	2,374	90.27	2,358	88.86	2,374	90.27
60			9	3,320	90.15	3,350	90.29	3,332	90.48	3,351	89.91	3,335	90.05	3,351	89.91
60			12	4,695	89.71	4,739	89.63	4,716	89.59	4,741	89.96	4,723	89.44	4,741	89.96

4.2.3 유의수준 1% & 검정력 80%

표 6에서 Zhang 방법에 의한 검정력은 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 80%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 6개월부터 3개월로 짧아질수록 77.21%, 76.13% 75.79%, 74.13%까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 6개월, 5개월일 때 76.03%, 73.68%까지 떨어지고, 4개월과 3개월인 경우에는 66.72%, 57.72%까지도 낮게 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 80%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의 차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 7개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 80%를 만족하는 것으로 나타났다. Schoenedld 방법에 의한 검정력은 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 80% 이상을 만족하였다. 하지만 Schoenedld 방법에서는 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 83.88%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 6개월과 3개월인 경우 83.53%, 86.17%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 6개월과 3개월인 경우 83.82%, 86.85%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다. 각 모의실험 조건에 대해, 만일 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때 필요한 군당 환자 수와 비교하면, Zhang 방법에서는 지수분포에서의 군당 환자 수 대비 약 50~60% 수준으로 크게 감소하였고, 진행 후 생존기간이 6개월인 경우 대조군의 무진행 생존기간이 6개월부터 3개월로 짧아질수록 48.8%, 48.0%, 47.6%, 47.9%까지 감소하고, 진행 후 생존기간이 3개월인 경우 대조군의 무진행 생존기간이 4개월과 3개월일 때 44.7%, 37.7% 수준까지도 떨어졌다. 그러나 무진행 생존

기간과 진행 후 생존기간이 동시에 낮은 경우에는 검정력 또한 감소한 것을 고려하여 표본 크기의 감소를 주의 깊게 판단할 필요가 있다. Schoenfeld 방법에서도 마찬가지로 약 50~60% 수준의 표본 크기 감소 경향을 보였고, 진행 후 생존기간이 3개월인 경우에는 61%~66.0% 정도까지의 표본 크기 감소가 나타났다. 그러나 진행 후 생존기간이 3개월인 경우에는 검정력이 주어진 80%를 초과한 것을 고려하면 전반적으로 50%~60% 수준의 표본 크기 감소가 나타난 것으로 판단할 수 있다.

표 7에서 Zhang 방법에 의한 검정력은 기본 단위 구간의 크기가 1개월이고 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 최소 77.62%까지 떨어졌으나 대부분 80%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 5개월부터 3개월인 경우 77.22%, 75.91%, 75.16%까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 5개월부터 3개월로 짧아질수록 75.26%, 69.49%, 63.30%까지 떨어지는 것으로 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 80%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의 차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 7개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 80%를 만족하는 것으로 나타났다(그림 5). Schoendeld 방법에 의한 검정력은 기본 단위 구간의 크기와 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 80% 이상을 만족하였다. 하지만 Schoendeld 방법에서는 기본 단위 구간의 크기가 1개월일 때, 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 83.54%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 6개월과 3개월인 경우 83.15%, 85.95%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 6개월과 3개월인 경우 83.65%, 86.77%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다(그림 5). Zhang 방법에서는 기본 단위 구간의 크기가 1개월인

경우에 비해 1/2개월로 세분화하면, 대조군의 무진행 생존기간이 3개월부터 7개월까지는 진행 후 생존기간이 길이에 따라 1명이나 최대 6명까지의 표본 크기 증가가 있었고, 대조군의 무진행 생존기간이 8개월인 경우에는 진행 후 생존기간이 길이에 따라 5명에서 최대 23명까지 표본 크기가 증가하였다. 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해 대조군의 무진행 생존기간이 5개월부터 7개월까지에서 진행 후 생존기간에 따라, 표본 크기 증가가 없거나 최대 2명까지의 증가가 나타났고, 대조군의 무진행 생존기간이 8개월일 때에는 진행 후 생존기간이 길어질수록 2명부터 최대 7명까지의 표본 크기 증가가 나타났다. Schoende1d 방법에서 기본 단위 구간의 크기가 1개월인 경우에 비해 1/2개월로 세분화하면, 대조군의 무진행 생존기간이 3개월일 때 진행 후 생존기간이 3개월인 경우, 대조군의 무진행 생존기간이 5개월일 때 진행 후 생존기간이 3개월과 9개월인 경우, 대조군의 무진행 생존기간이 7개월일 때 진행 후 생존기간이 6개월과 12개월인 경우에만 각 1명의 표본 크기 증가가 있었을 뿐, 그 외에는 기본 단위 구간의 크기에 관계없이 동일한 표본 크기를 보였고, 대조군의 무진행 생존기간이 8개월인 경우에는 진행 후 생존기간이 길어질수록 1명에서 2명까지의 표본 크기 증가가 나타났다. 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해 추가 증가는 없었다. 즉, 모든 경우에 있어 표본 크기의 증가는 크지 않았고, 검정력도 비슷한 수준을 유지하였다.

연세대학교
YONSEI UNIVERSITY

표 6. Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.01, \beta = 0.2$)

Follow up time	Median (month)			Exponential		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power
120	9	3	3	53	79.93	20	57.72	35	86.85
150			6	94	79.87	45	74.13	55	83.82
180			9	147	80.00	82	78.53	86	81.28
210			12	213	80.36	129	79.11	131	79.63
120	9	4	3	85	79.95	38	66.72	55	86.17
150			6	147	80.00	70	75.79	82	83.53
180			9	226	80.08	121	78.61	128	80.83
210			12	324	80.43	188	78.95	192	80.88
120	9	5	3	147	80.00	74	73.68	92	83.88
150			6	248	80.01	119	76.13	134	82.70
180			9	377	80.24	196	79.02	205	80.38
210			12	536	80.66	301	79.25	307	80.18
120	9	6	3	287	79.93	158	76.03	177	81.73
150			6	475	80.05	232	77.21	249	81.97
180			9	713	80.32	365	78.83	376	79.88
210			12	1,004	80.72	549	79.13	559	80.72
120	9	7	3	708	79.98	413	79.22	433	81.17
150			6	1,150	80.14	574	78.69	592	80.74
180			9	1,705	80.41	862	79.26	878	80.09
210			12	2,382	80.86	1,273	79.91	1,291	80.25
120	9	8	3	3,094	80.01	1,879	80.31	1,895	80.24
150			6	4,833	79.17	2,506	80.32	2,523	80.76
180			9	7,234	80.48	3,633	79.37	3,661	79.73
210			12	10,030	80.97	5,263	79.18	5,316	80.35

연세대학교
YONSEI UNIVERSITY

표 7. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.01, \beta = 0.2$)

Follow up time	Median (month)			$\Delta t = 1$				$\Delta t = 1/2$				$\Delta t = 1/3$			
				Zhang		Schoenfeld		Zhang		Schoenfeld		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power
36	9	3	3	22	63.30	35	86.77	22	63.30	36	88.28	22	63.30	36	88.28
48			6	46	75.16	55	83.65	46	75.16	55	83.65	46	75.16	55	83.65
60			9	81	77.62	86	81.22	82	78.97	86	81.22	82	78.97	86	81.22
60			12	125	78.55	127	79.50	126	79.54	127	79.50	126	79.54	127	79.50
36	9	4	3	40	69.49	55	85.95	40	69.49	55	85.95	40	69.49	55	85.95
48			6	71	75.91	82	83.15	72	76.26	82	83.15	72	76.26	82	83.15
60			9	120	77.96	127	80.97	121	78.77	127	80.97	121	78.77	127	80.97
60			12	183	79.68	187	80.25	184	78.99	187	80.25	184	78.99	187	80.25
36	9	5	3	78	75.26	93	83.54	78	75.26	94	84.11	78	75.26	94	84.11
48			6	122	77.22	135	81.70	123	77.59	135	81.70	123	77.59	135	81.70
60			9	196	78.78	204	80.06	197	79.34	205	80.25	197	79.34	205	80.25
60			12	292	79.34	299	80.14	294	79.46	299	80.14	295	80.17	299	80.14
36	9	6	3	166	78.02	182	82.04	166	78.02	182	82.04	166	78.02	182	82.04
48			6	240	78.78	254	81.17	240	78.78	254	81.17	241	77.93	254	81.17
60			9	366	79.44	377	80.39	368	79.98	377	80.39	368	79.98	377	80.39
60			12	536	80.01	546	80.40	539	79.82	546	80.40	540	79.33	546	80.40
36	9	7	3	436	79.26	452	81.13	437	79.20	452	81.13	437	79.20	452	81.13
48			6	595	79.37	611	80.61	597	79.82	612	80.67	597	79.82	612	80.67
60			9	870	79.94	886	80.26	874	80.02	886	80.26	875	79.09	886	80.26
60			12	1,252	79.76	1,268	79.60	1,258	80.04	1,269	79.73	1,260	80.00	1,269	79.73
36	9	8	3	1,990	80.03	2,013	80.04	1,995	80.52	2,014	80.12	1,997	80.36	2,014	80.12
48			6	2,613	79.63	2,638	80.33	2,619	79.63	2,639	80.40	2,621	79.42	2,639	80.40
60			9	3,691	79.04	3,723	80.24	3,703	80.13	3,725	79.52	3,707	79.25	3,725	79.52
60			12	5,219	79.25	5,268	80.44	5,242	79.75	5,270	80.00	5,249	79.23	5,270	80.00

4.2.4 유의수준 1% & 검정력 90%

표 8에서 Zhang 방법에 의한 검정력은 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 90%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 5개월부터 3개월로 짧아질수록 87.62%, 87.09% 86.51%까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 6개월, 5개월일 때 87.34%, 84.70%까지 떨어지고, 4개월과 3개월인 경우에는 79.28%, 72.33%까지도 낮게 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 90%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의 차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 7개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 90%를 만족하는 것으로 나타났다. Schoendeld 방법에 의한 검정력은 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 90% 이상을 만족하였다. 하지만 Schoendeld 방법에서는 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 92.35%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 6개월과 3개월인 경우 92.08%, 94.19%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 6개월과 3개월인 경우 92.08%, 95.50%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다. 각 모의실험 조건에 대해, 만일 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때 필요한 군당 환자 수와 비교하면, Zhang 방법에서는 지수분포에서의 군당 환자 수 대비 약 50~60% 수준으로 크게 감소하였고, 진행 후 생존기간이 6개월인 경우 대조군의 무진행 생존기간이 7개월부터 3개월로 짧아질수록 49.9%, 48.8%, 47.9%, 47.3%, 47.9%까지 감소하고, 진행 후 생존기간이 3개월인 경우 대조군의 무진행 생존기간이 4개월과 3개월일 때 44.0%, 38.2% 수준까지도 떨어졌다. 그러나 무진행 생존기간과

진행 후 생존기간이 동시에 낮은 경우에는 검정력 또한 감소한 것을 고려하여 표본 크기의 감소를 주의 깊게 판단할 필요가 있다. Schoenfeld 방법에서도 마찬가지로 약 50~60% 수준의 표본 크기 감소 경향을 보였고, 진행 후 생존기간이 3개월인 경우에는 61%~66.2% 정도까지의 표본 크기 감소가 나타났다. 그러나 진행 후 생존기간이 3개월인 경우에는 검정력이 주어진 90%를 초과한 것을 고려하면 전반적으로 50%~60% 수준의 표본 크기 감소가 나타난 것으로 판단할 수 있다.

표 9에서 Zhang 방법에 의한 검정력은 기본 단위 구간의 크기가 1개월이고 진행 후 생존기간이 9개월과 12개월로 충분히 큰 경우, 대조군의 무진행 생존기간에 대한 중앙생존기간에 관계없이 대부분 90%에 가깝게 유지되었다. 그러나 진행 후 생존기간이 6개월일 때는, 대조군의 무진행 생존기간이 5개월부터 3개월인 경우 87.70%, 87.06%, 86.65%까지 떨어졌고, 진행 후 생존기간이 3개월일 때는, 대조군의 무진행 생존기간이 5개월부터 3개월로 짧아질수록 85.92%, 81.91%, 76.23%까지 떨어지는 것으로 나타났다. 즉, 진행 후 생존기간이 짧을수록 검정력은 90%를 만족하지 못했고, 특히 시험군과 대조군의 위험비의 차이가 클수록 이러한 현상은 두드러지게 나타났다. 하지만 대조군의 무진행 생존기간이 7개월 이상으로 시험군의 무진행 생존기간에 가까워지면, 진행 후 생존기간의 길이와 관계없이 검정력은 대부분 90%를 만족하는 것으로 나타났다(그림 6). Schoendeld 방법에 의한 검정력은 기본 단위 구간의 크기와 대조군의 무진행 생존기간이나 진행 후 생존기간에 관계없이 모두 90% 이상을 만족하였다. 하지만 Schoendeld 방법에서는 기본 단위 구간의 크기가 1개월일 때, 대조군의 무진행 생존기간이 5개월이고 진행 후 생존기간이 3개월인 경우 92.39%, 대조군의 무진행 생존기간이 4개월이고 진행 후 생존기간이 6개월과 3개월인 경우 92.22%, 93.97%, 대조군의 무진행 생존기간이 3개월이고 진행 후 생존기간이 3개월인 경우 95.45%로 주어진 검정력에 비해 크게 나타났다. 즉, 시험군과 대조군의 위험비의 차이가 크고, 진행 후 생존기간이 짧을수록 주어진 검정력을 초과하는 결과를 나타냈다. 하지만 대조군의 무진행 생존기간이 6개월 이상이거나, 진행 후 생존기간이 9개월 이상일 때에는 적정한 수준의 검정력을 유지하는 것으로 나타났다(그림 6). Zhang 방법에서는 기본 단위 구간의 크기가 1개월인 경우에 비해 1/2개월로 세분화하면, 대조군의 무진행

생존기간이 3개월부터 7개월까지는 진행 후 생존기간이 길이에 따라 1명이나 최대 7명까지의 표본 크기 증가가 있었고, 대조군의 무진행 생존기간이 8개월인 경우에는 진행 후 생존기간이 길이에 따라 6명에서 최대 30명까지 표본 크기가 증가하였다. 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해 대조군의 무진행 생존기간이 4개월부터 7개월까지에서 진행 후 생존기간에 따라, 표본 크기 증가가 없거나 최대 3명까지의 증가가 나타났고, 대조군의 무진행 생존기간이 8개월일 때에는 진행 후 생존기간이 길어질수록 2명부터 최대 10명까지의 표본 크기 증가가 나타났다. Schoende1d 방법에서 기본 단위 구간의 크기가 1개월인 경우에 비해 1/2개월로 세분화하면, 대조군의 무진행 생존기간이 3개월일 때 진행 후 생존기간이 6개월인 경우, 대조군의 무진행 생존기간이 4개월일 때 진행 후 생존기간이 9개월인 경우, 대조군의 무진행 생존기간이 5개월일 때 진행 후 생존기간이 6개월과 12개월인 경우, 대조군의 무진행 생존기간이 6개월일 때 진행 후 생존기간이 3개월인 경우, 대조군의 무진행 생존기간이 7개월일 때 진행 후 생존기간이 9개월인 경우에만 각 1명의 표본 크기 증가가 있었을 뿐, 그 외에는 기본 단위 구간의 크기에 관계없이 동일한 표본 크기를 보였고, 대조군의 무진행 생존기간이 8개월인 경우에는 진행 후 생존기간이 길어질수록 1명에서 3명까지의 표본 크기 증가가 나타났다. 기본 단위 구간의 크기를 1/3개월로 세분화한 경우에는 1/2개월에 비해 대조군의 무진행 생존기간이 3개월일 때 진행 후 생존기간이 12개월인 경우, 대조군의 무진행 생존기간이 6개월일 때 진행 후 생존기간이 12개월인 경우, 대조군의 무진행 생존기간이 7개월일 때 진행 후 생존기간이 12개월인 경우, 대조군의 무진행 생존기간이 8개월일 때 진행 후 생존기간이 9개월인 경우에만 각 1명의 표본 크기 증가가 있었을 뿐, 그 외에는 표본 크기의 추가 증가는 없었다. 즉, 모든 경우에 있어 표본 크기의 증가는 크지 않았고, 검정력도 비슷한 수준을 유지하였다.

연세대학교
YONSEI UNIVERSITY

표 8. Hypo-exponential 분포를 따르는 로그-순위 검정의 표본수와 검정력 ($\alpha = 0.01, \beta = 0.1$)

Follow up time	Median (month)			Exponential		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power
120	9	3	3	68	89.99	26	72.33	45	95.50
150			6	121	90.12	58	86.51	69	92.08
180			9	188	90.03	104	88.60	110	90.44
210			12	272	90.27	164	89.75	167	89.90
120	9	4	3	109	90.03	48	79.28	70	94.19
150			6	188	90.03	89	87.09	104	92.08
180			9	289	90.11	154	88.89	162	91.16
210			12	413	90.27	240	89.67	245	90.17
120	9	5	3	188	90.03	94	84.70	117	92.35
150			6	317	90.06	152	87.62	170	91.62
180			9	481	90.18	250	89.05	261	90.91
210			12	683	90.44	383	89.75	391	90.24
120	9	6	3	367	90.02	201	87.34	225	91.39
150			6	606	90.05	296	88.79	317	91.49
180			9	909	90.22	465	90.05	479	90.32
210			12	1,280	90.51	699	89.28	712	90.29
120	9	7	3	903	90.00	526	88.61	551	90.63
150			6	1,465	90.08	731	89.20	754	90.67
180			9	2,172	90.28	1,099	89.36	1,118	90.45
210			12	3,035	90.60	1,621	89.34	1,645	90.29
120	9	8	3	3,942	90.00	2,394	89.90	2,415	90.16
150			6	6,286	90.11	3,192	90.02	3,214	90.14
180			9	9,217	90.33	4,628	89.69	4,664	89.90
210			12	12,779	90.68	6,705	89.74	6,773	90.51

연세대학교
YONSEI UNIVERSITY

표 9. Hypo-exponential 분포를 따르는 로그-순위 검정의 단위 구간별 표본수와 검정력 ($\alpha = 0.01, \beta = 0.1$)

Follow up time	Median (month)			$\Delta t = 1$				$\Delta t = 1/2$				$\Delta t = 1/3$			
				Zhang		Schoenfeld		Zhang		Schoenfeld		Zhang		Schoenfeld	
	PFS(A)	PFS(B)	PPS	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power	n_k	power
36	9	3	3	28	76.23	45	95.45	28	76.23	45	95.45	28	76.23	45	95.45
48			6	58	86.65	69	91.99	59	86.96	70	92.61	59	86.96	70	92.61
60			9	103	88.79	109	90.85	104	88.86	109	90.85	104	88.86	109	90.85
60			12	159	89.27	161	89.66	160	89.07	161	89.66	160	89.07	162	90.23
36	9	4	3	51	81.91	70	93.97	51	81.91	70	93.97	51	81.91	70	93.97
48			6	91	87.06	105	92.22	91	87.06	105	92.22	91	87.06	105	92.22
60			9	153	89.21	161	90.56	154	89.39	162	90.96	154	89.39	162	90.96
60			12	232	89.50	238	90.03	234	89.59	238	90.03	235	89.98	238	90.03
36	9	5	3	99	85.92	119	92.39	99	85.92	119	92.39	99	85.92	119	92.39
48			6	156	87.70	172	91.33	156	87.70	172	91.33	156	87.70	172	91.33
60			9	249	89.23	260	90.37	251	89.23	261	91.01	251	89.23	261	91.01
60			12	372	89.21	380	90.19	374	89.38	381	89.98	375	89.65	381	89.98
36	9	6	3	211	88.73	231	91.48	212	88.27	232	91.13	212	88.27	232	91.13
48			6	305	88.71	324	91.10	306	89.10	324	91.10	306	89.10	324	91.10
60			9	466	89.61	480	89.82	468	89.20	480	89.82	469	90.14	480	89.82
60			12	683	89.45	695	90.07	687	90.10	695	90.07	688	89.91	696	90.02
36	9	7	3	555	89.54	576	90.73	556	89.63	576	90.73	556	89.63	576	90.73
48			6	758	89.49	779	90.58	760	89.26	779	90.58	761	90.06	779	90.58
60			9	1,109	89.98	1,128	90.58	1,113	89.95	1,129	90.36	1,114	89.69	1,129	90.36
60			12	1,595	89.88	1,616	90.24	1,602	89.94	1,616	90.24	1,605	90.06	1,617	90.10
36	9	8	3	2,535	90.19	2,564	90.62	2,541	89.66	2,566	90.17	2,544	89.83	2,566	90.17
48			6	3,329	90.42	3,360	89.65	3,337	90.26	3,362	89.74	3,339	90.19	3,362	89.74
60			9	4,702	89.95	4,744	89.99	4,718	89.82	4,745	90.12	4,723	89.42	4,746	89.79
60			12	6,648	89.32	6,711	90.47	6,678	90.09	6,714	90.18	6,688	90.19	6,714	90.18

제 5 장 고찰 및 결론

5.1 고찰

4장에서 살펴 본 모의 실험 결과, Zhang 방법에서는 진행 후 생존기간이 9개월 이상인 경우, 대조군의 무진행 생존기간에 관계없이 계획한 수준의 검정력을 만족하는 것으로 나타났다. 그러나 진행 후 생존기간이 3개월 또는 6개월로 짧고, 대조군의 무진행 생존기간이 시험군의 무진행 생존기간에 비해 짧으면, 즉 시험군과 대조군의 위험비의 차이가 클수록 검정력이 계획한 수준에 비해 낮아져, 검정력이 80%인 경우 최대 57.72%까지 낮게 나타났고, 검정력이 90%인 경우 최대 72.33%까지 낮게 나타났다. Schoenfeld 방법에서는 진행 후 생존기간과 대조군의 무진행 생존기간에 관계없이 모든 경우에 대하여 계획한 수준의 검정력을 만족하는 것으로 나타났다. 그러나 시험군과 대조군의 위험비의 차이가 크거나, 진행 후 생존기간이 3개월 또는 6개월로 짧은 경우, 계획한 수준의 검정력을 초과해, 검정력이 80%인 경우 조건에 따라 최대 88.28%까지 높게 나타났고, 검정력이 90%인 경우 조건에 따라 최대 95.50%까지 높게 나타났다. 비 중심 평균 모수를 계산하는 기본 단위인 구간을 보다 잘게 구분하여 결과의 정확도를 높이기 위해 각 1개월, 1/2개월, 1/3개월의 3가지 경우로 구분하였을 때 두 방법 모두에서 기본 단위 구간은 검정력에 크게 영향을 미치지 않는 것으로 확인되었고, 전체 연구기간을 시험군과 대조군의 각 중앙생존기간 조건에 대해 모든 사건이 일어날 수 있을 만큼 충분한 정도의 연구기간(PFS와 PPS의 합에 대한 10배)으로 설정하였을 때와 실제 연구 설계에 적합한 수준(PFS와 PPS의 합에 대한 약 3배)으로 설정하였을 때에도 두 방법 모두에서 검정력에 크게 영향을 미치지 않는 것으로 확인되었다. Schoenfeld 방법에서의 적분은 단위 구간에 대하여 해당 구간의 중간 시점 $\frac{t_{i-1}+t_i}{2}$의 값을 적용하였으며, 이는 적분 결과를 얻는데 있어 약간의 오차(error of trapezoidal rule)를 발생시킨다고 생각할 수 있다. 이 오차의 크기를 확인하기 위하

여 비 중심 평균 모수에 대한 전체 함수를 포함하는 직사각형에 대해 100,000개의 난수를 발생시키고, 전체 난수 중 비 중심 평균 모수가 갖는 함수보다 작은 값을 갖는 경우의 비율을 계산해 면적을 추정하는 몬테카를로 적분(Monte Carlo integration)의 한 방법인 hit and miss 방법(Jones, Maillardet and Robinson, 2009)을 이용하여 계산한 결과, 모든 조건의 경우에 대해 표본 크기에 영향을 미치는 정도의 차이는 발견되지 않았고, 이러한 결과는 본 논문에서 제안한 단위 구간의 중간 시점 $\frac{t_{i-1}+t_i}{2}$값을 이용한 방법이 매우 정확하다는 사실을 뒷받침한다고 할 수 있다. 시험군과 대조군의 위험도와 위험함수비의 관계를 살펴보면(그림 7, 그림 8), 진행 후 생존기간이 3개월부터 6개월, 9개월, 12개월로 길어질수록 시험군과 대조군의 위험도의 차이가 적어 전체 생존기간을 비교에 대한 검정력이 낮아진다는 Broglio and Berry (2009)의 연구 결과를 확인할 수 있다. 결과적으로 각 모의 실험의 조건에 관계없이 임상시험에 필요한 환자 수는 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정하면, 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때에 비해 감소하였으며, 이는 전체 생존기간의 비교를 목적으로 하는 항암제 임상시험에서 비용 감소와 함께 연구 기간의 단축 효과를 제공할 수 있다.

본 논문에서는 검정 통계량을 결정하는 비 중심 평균 모수를 구함에 있어 중도절단 시간의 분포 함수 $H_k(t)$에 대한 고려를 충분히 반영하지 못했다. 중도절단 사건은 시험군과 대조군에서 서로 다르게 나타나기 때문에, 시간이 경과함에 따라 각 군의 위험함수 $h_k(t)$에 영향을 미치고, 때문에 각 시점에서의 위험함수 $h_k(t)$는 중도절단 시간의 분포 함수 $H_k(t)$를 반영해 다시 계산되어야 하며, 이는 본 논문의 제한점으로 남게 된다. 또한 본 논문에서는 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정하였으나, 무진행 생존기간과 진행 후 생존기간의 분포가 와이블 분포(Weibull distribution)를 따르는 등과 같이 지수 분포에 대한 가정이 깨지는 경우도 생각해 볼 수 있고, 이는 본 논문의 개념을 확장하여 추가 연구가 진행되어야 할 것

으로 생각된다.

5.2 결론

본 논문에서는 전체 생존기간의 비교를 목적으로 하는 항암제 임상시험에서 시험군과 대조군의 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정했을 때, Schoenfeld(1981)와 Zhang and Quan(2009)에 의해 정리된 로그-순위 검정에 대한 검정 통계량을 결정하는 비 중심 평균 모수를 구함에 있어 검정 통계량의 기댓값을 이용하는 방법(Zhang 방법)과 비 중심 평균 모수를 구성하는 각 함수의 정확한 형태를 이미 알고 있다는 사실을 이용하여 직접 적분을 적용(Schoenfeld 방법)하는 두 가지 방법으로 비 중심 평균 모수를 추정하여 표본 크기를 산출하는 방법을 제안하였다. Zhang 방법에서는 시험군과 대조군의 위험비의 차이가 크면 검정력이 계획한 수준에 비해 낮게 나타났으나, 진행 후 생존기간이 9개월 이상이면 대조군의 무진행 생존기간에 관계없이 계획한 수준의 검정력을 만족하는 것으로 나타났다. Schoenfeld 방법에서는 진행 후 생존기간과 대조군의 무진행 생존기간에 관계없이 모든 경우에 대하여 계획한 수준의 검정력을 만족하는 것으로 나타났으나, 시험군과 대조군의 위험비의 차이가 크거나, 진행 후 생존기간이 3개월 또는 6개월로 짧은 경우에는 계획한 수준의 검정력을 초과하는 경우도 나타났다. 결과적으로 각 모의 실험의 조건에 관계없이 임상시험에 필요한 환자 수는 무진행 생존기간과 진행 후 생존기간의 분포가 각각 지수분포를 따르고 따라서 전체 생존기간의 분포는 hypo-exponential 분포를 따른다고 가정하면, 전체 생존기간의 분포가 지수분포를 따른다고 가정하였을 때에 비하여 감소하였으며, 이는 전체 생존기간의 비교를 목적으로 하는 항암제 임상시험에서 비용 감소와 함께 연구 기간의 단축 효과를 제공할 수 있다.

참고문헌

Asan medical center. 국가 암 임상시험 제도발전 방향에 대한 연구. KoNECT, 2013.

Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. J. Natl. Cancer Inst. 2009;101:1642-9.

Cox D. Regression models and life-tables. J. R. Stat. Soc. 1972;34:187-220.

Escudier B, Eisen T, Stadler MW, Szczylik C, Oudard S, Staehler M, Negrier S, Chevreau C, Desai AA, Rolland F, Demkow T, Hutson TE, Gore M, Anderson S, Hofilena G, Shan M, Pena C, Lathia C, Bukowski RM. Sorafenib for treatment of renal cell carcinoma: Final efficacy and safety results of the phase III treatment approaches in renal cancer global evaluation trial. J. Cli. Oncol. 2009;27(20):3312-8.

Fleming TR, Harrington DP. Counting processes and survival analysis. New York: Wiley, 1991.

Hogg RV, Craig AT. Introduction to mathematical statistics, 4th edition. New York: Macmillan, 1978.

Jones O, Maillardet R, Robinson A. Introduction to scientific programming and simulation using R. Boca Raton: CRC Press, 2009.

Lachin JM, Foulkes MA. Evaluation of sample size and power for analyses of

survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. Biometrics 1986;42(3):507-19.

Lakatos E. Based on the log-rank statistic in complex clinical trials. Biometrics 1988;44(1):229-41.

Lakatos E, Lan KKG. A comparison of sample size methods for the logrank statistic. Stat. Med. 1992;11:179-91.

Ministry of food and drug safety. 15년 임상시험 계획 승인 현황 분석. 식품의약품 안전처 임상제도과, 2016.

Schoenfeld D. The asymptotic properties of nonparametric tests for comparing survival distributions. Biometrika 1981;68(1):316-9.

Schoenfeld DA, Richter JR. Nomograms for calculating the number of patients needed for a clinical trial with survival as an endpoint. Biometrics 1982;38(1):163-170.

Sparano JA, Vrdoljak E, Rixe O, Xu B, Manikhas A, Medina C, Costa SCVD, Ro J, Rubio G, Rondinon M, Manga GP, Peck R, Poulart V, Conte P. Randomized phase III trial of ixabepilone plus capecitabine versus capecitabine in patients with metastatic breast cancer previously treated with an anthracycline and a taxane. J. Cli. Oncol. 2010;28(20):3256-63.

Zhang D, Quan H. Power and sample size calculation for log-rank test with a time lag in treatment effect. Stat. Med. 2009;28:864-79.

연세대학교
YONSEI UNIVERSITY

박소영. 무진행 생존기간과 전체 생존기간의 관계 [석사학위논문]. 서울: 연세대학교; 2013.

박재빈. 생존분석 이론과 실제. 서울: 신광출판사, 2006.

송혜향, 정갑도, 이원철. 생존분석. 파주: 청문각, 1996.

연세대학교
YONSEI UNIVERSITY

부록

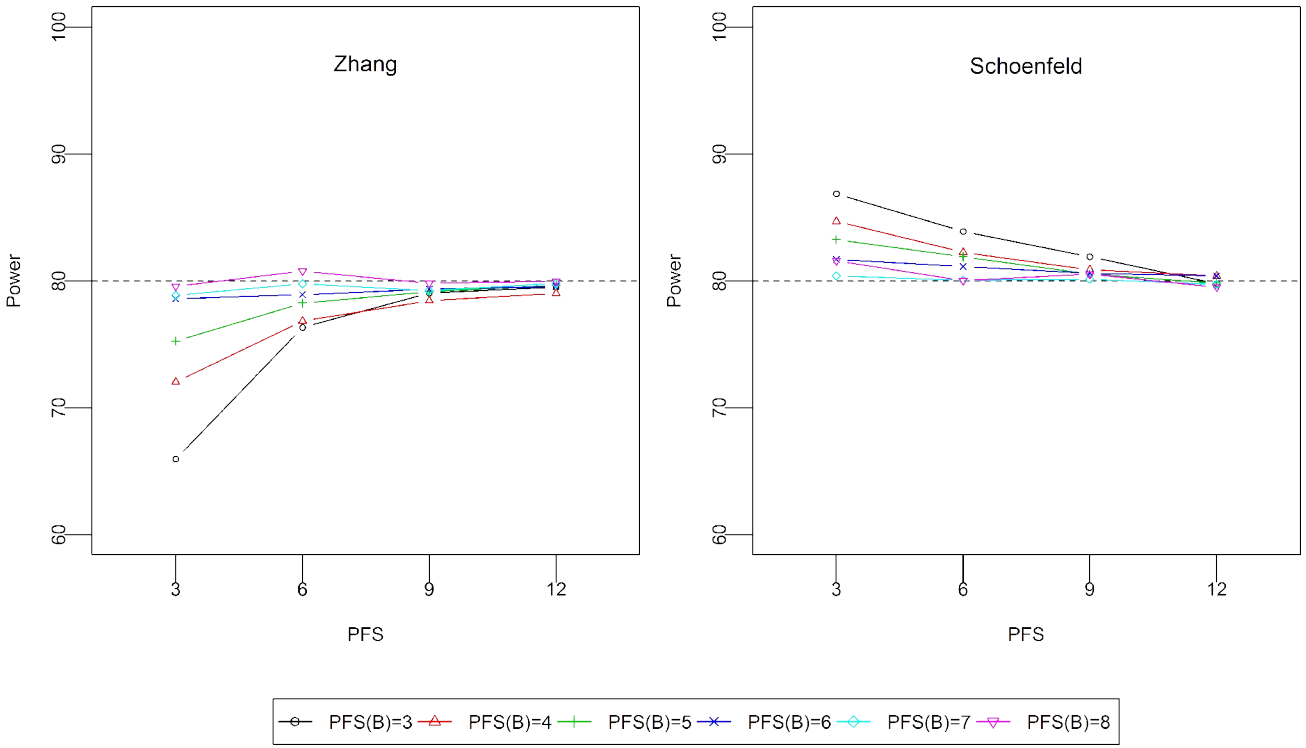

그림 3. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.05, \beta = 0.2$).

연세대학교
YONSEI UNIVERSITY

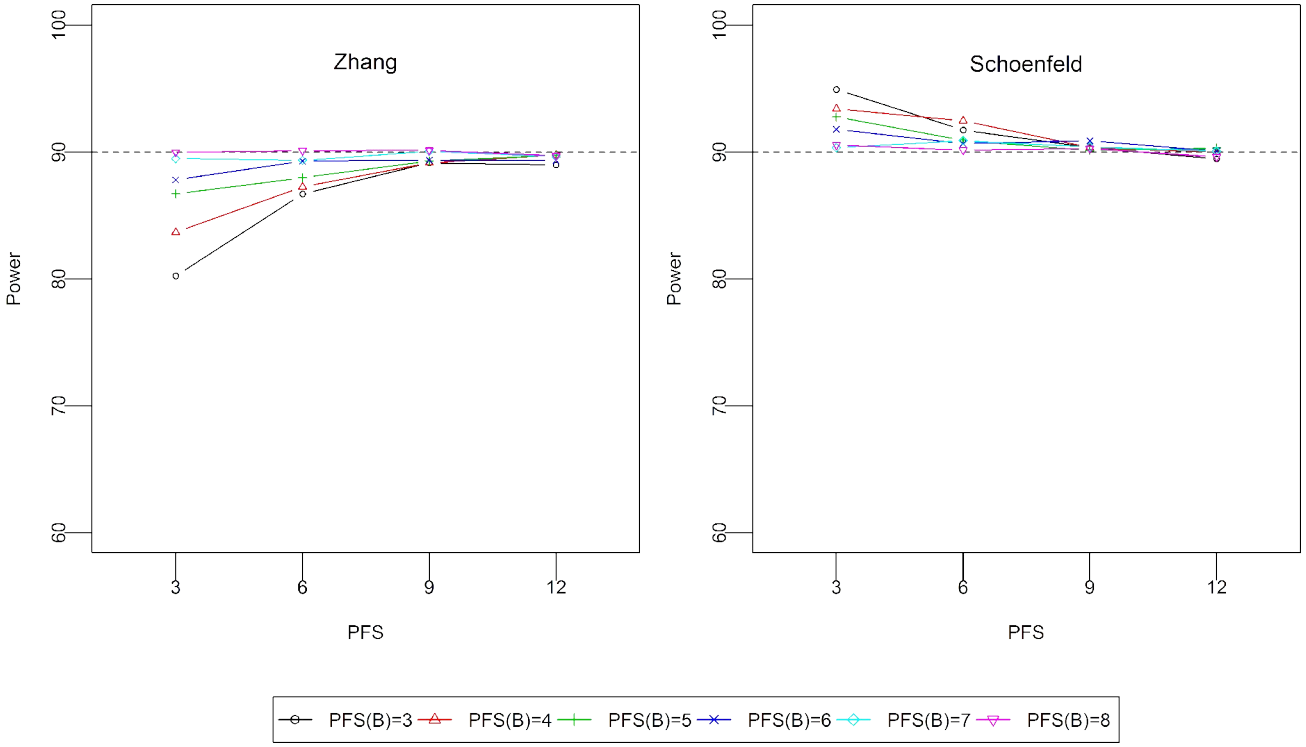

그림 4. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.05, \beta = 0.1$).

연세대학교
YONSEI UNIVERSITY

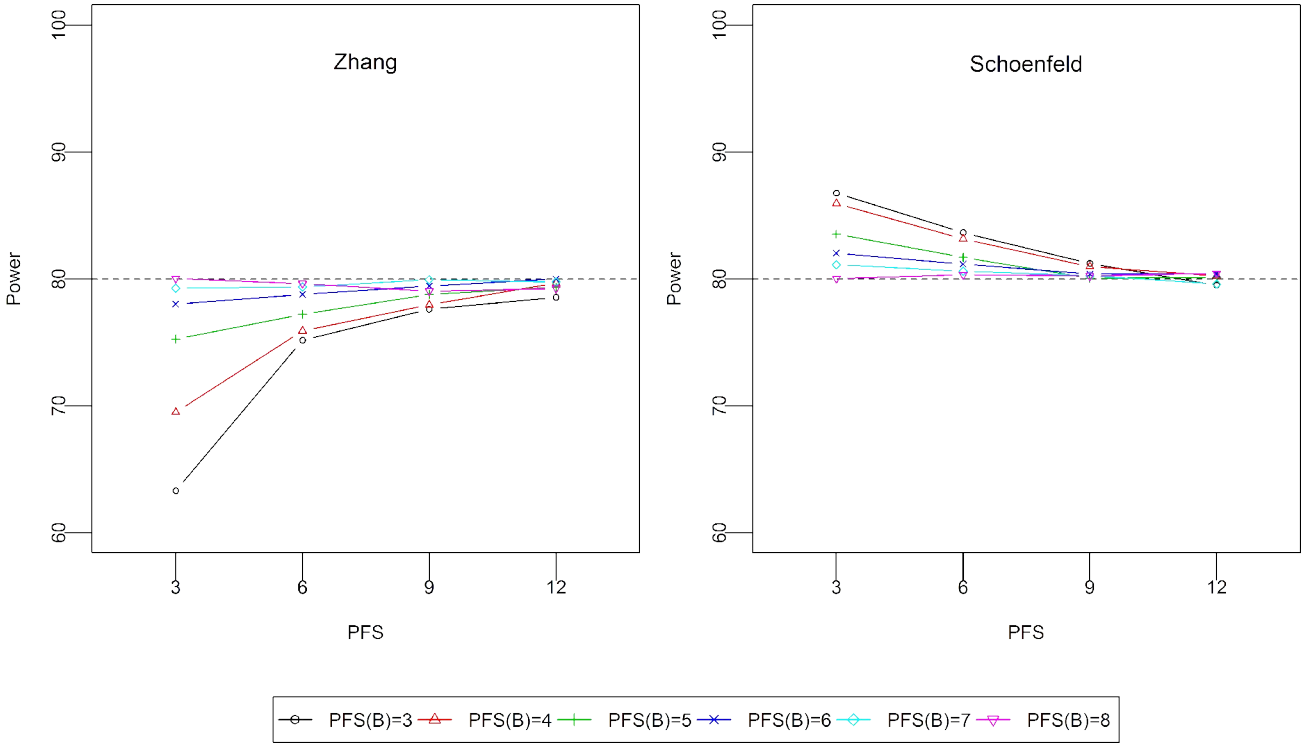

그림 5. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.01, \beta = 0.2$).

연세대학교
YONSEI UNIVERSITY

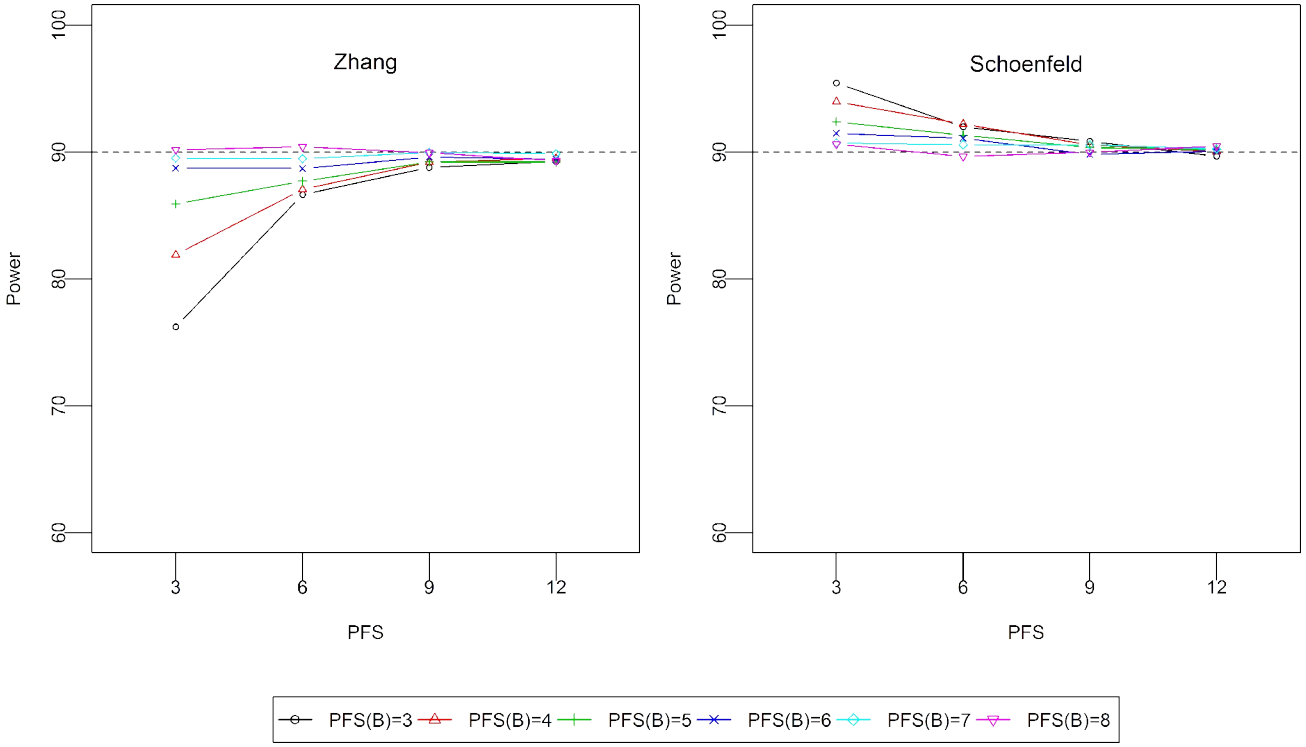

그림 6. 무진행 생존기간과 진행 후 생존기간에 따른 검정력 ($\alpha = 0.01, \beta = 0.1$).

연세대학교
YONSEI UNIVERSITY

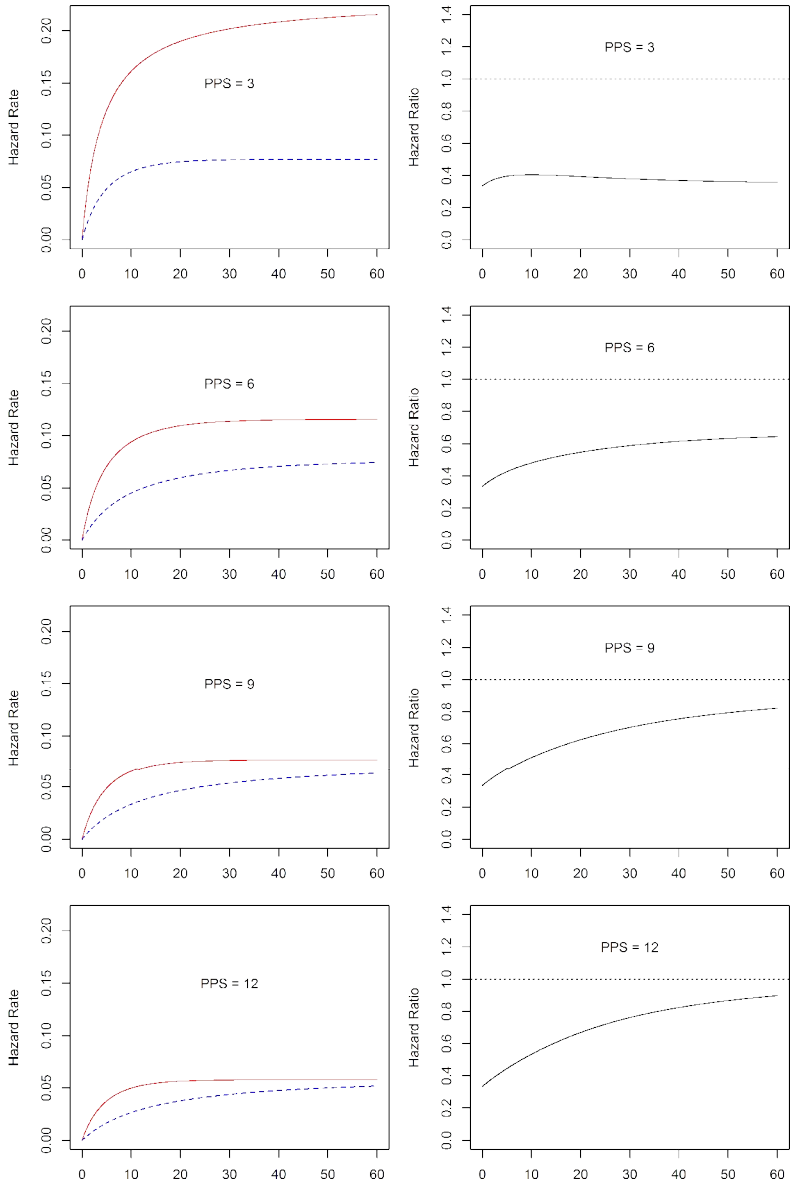

그림 7. 무진행 생존기간 시험군 9개월, 대조군 3개월에 대한 위험도와 위험함수비.

연세대학교
YONSEI UNIVERSITY

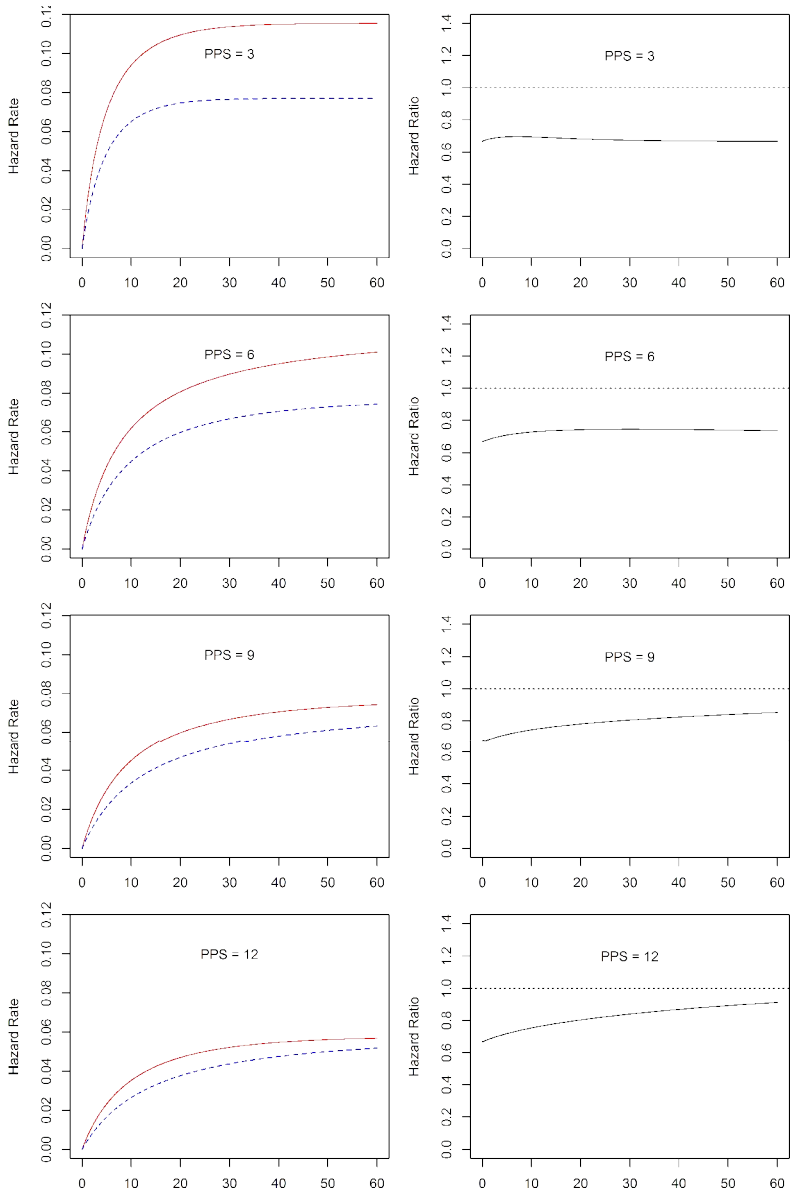

그림 8. 무진행 생존기간 시험군 9개월, 대조군 6개월에 대한 위험도와 위험함수비.

모의 실험 R 프로그램 (Zhang)

```
# Loading survival package
library(survival)

# Zhang method
simul <- function(median_t, median_c, median_pps, alpha, beta, iteration, length, accral, k){
  # Set seed number
  set.seed(20161024)

  # Define parameters
  pfs_t <- log(2)/median_t
  pfs_c <- log(2)/median_c
  pps   <- log(2)/median_pps

  ht <- c();nQt <- 1;nut <- c();det <- c();nut_12 <- c()
  hc <- c();nQc <- 1;nuc <- c();dec <- c();nuc_12 <- c()
  HR <- c()
  nt <- c()

  # Define time conditions
  a <- accral
  l <- (length+accral)
  t <- c(1/k)*c(1:(k*l-1))-(1/(2*k))

  # Calculating hazard functions
  for(i in 1:(k*l-1)){
    # When PFS_t=PPS
    if(median_t==median_pps){
      ht[i] <- (((pfs_t^2)*t[i]) / (1 + pfs_t*t[i]))*(1/k)
      } else {
      ht[i] <-
(pfs_t*pps*(exp(-pfs_t*t[i])-exp(-pps*t[i])))/(pps*exp(-pfs_t*t[i])-pfs_t*exp(-pps*t[i]))*(1/k)
    }

    # When PFS_c=PPS
    if(median_c==median_pps){
      hc[i] <- (((pfs_c^2)*t[i]) / (1 + pfs_c*t[i]))*(1/k)
      } else {
      hc[i] <-
(pfs_c*pps*(exp(-pfs_c*t[i])-exp(-pps*t[i])))/(pps*exp(-pfs_c*t[i])-pfs_c*exp(-pps*t[i]))*(1/k)
    }

    # Calculating hazard ratio
    HR[i] <- ht[i]/hc[i]

    # Calculating cumulative Phi
    nut[i] <- nQt * ht[i]*(1    -1/HR[i])
```

```
    nuc[i] <- nQc * hc[i]*(HR[i]-1      )

    det[i] <- nQt * ht[i]
    dec[i] <- nQc * hc[i]

    nQt <- nQt*(1-ht[i])
    nQc <- nQc*(1-hc[i])
  }

  # Calculating sample size
  for(j in 1:accral){
    nt[j] <-  ((qnorm(1-alpha/2)+qnorm(1-beta))^2*4) /
((sum(nut[1:(l*k-j*k)])+sum(nuc[1:(l*k-j*k)])) /
sqrt((sum(det[1:(l*k-j*k)])+sum(dec[1:(l*k-j*k)]))) )^2
  }

  if (accral==0){
    nt <- ceiling (sum(nt))
    } else {
    nt <- ceiling (sum(nt)/accral)
  }

  # Define p-vlaue vectors
  hypo_pvalue <- c()
  os_pvalue   <- c()

  # Calculating Power
  for(i in 1:iteration){
    # Generating random numbers
    pfs_t <- rexp(  nt, log(2)/median_t  )
    pfs_c <- rexp(  nt, log(2)/median_c  )
    pps   <- rexp(2*nt, log(2)/median_pps)

    # Generating distributions
    hypo_os <- c(pfs_t, pfs_c) + pps
    exp_os  <- c(rexp(nt, log(2)/(median_t+median_pps)), rexp(nt,
log(2)/(median_c+median_pps)))
    group   <- rep(c(1,0), each=nt)

    # Generating study duration
    if (accral==0){
      la      <- length
      } else {
      amtemp <- c()
      am <- c()
      for(j in 1:accral){
        m      <- (nt - nt%%accral) / accral + (j <= nt%%accral)
        amtemp <- rep(c(accral-j),m)
        am     <-  c(am,amtemp)
      }
      la      <- c(length + c(am),length + c(am))
```

```
    }

    # Generating event
    stat_hypo <- hypo_os <= la
    stat_os   <- exp_os  <= la

    # p-value
    hypo_pvalue[i] <- summary(coxph(Surv(hypo_os, stat_hypo) ~ group))$sctest[3]
    os_pvalue[i]   <- summary(coxph(Surv(exp_os , stat_os  ) ~ group))$sctest[3]
  }

  # Save output
  result <- c(nt, sum(hypo_pvalue < alpha)/iteration, sum(os_pvalue < alpha)/iteration)

  return(result)
}
```

모의 실험 R 프로그램 (Schoenfeld)

```
# Loading survival package
library(survival)

# Schoenfeld method
integral <- function(median_t, median_c, median_pps, alpha, beta, iteration, length, accral,
k){
  # Set seed number
  set.seed(20161024)

  # Define parameters
  pfs_t <- log(2)/median_t
  pfs_c <- log(2)/median_c
  pps   <- log(2)/median_pps

  ht <- c();ft <- c();Lft <- c();nu <- c();
  hc <- c();fc <- c();Lfc <- c();de <- c();
  n  <- c()

  # Define time conditions
  a <- accral
  l <- (length+accral)
  t <- c(1/k)*c(1:(k*l-1))-(1/(2*k))

  # Calculating hazard functions
  for(i in 1:(k*l-1)){
    # When PFS_t=PPS
    if(median_t==median_pps){
      ht[i]  <- ((pfs_t^2)*t[i]) / (1 + pfs_t*t[i])
      ft[i]  <- (pfs_t^2)*t[i]*exp(-pfs_t*t[i])
      Lft[i] <- (1 - exp(-pfs_t*t[i]) - (pfs_t)*t[i]*exp(-pfs_t*t[i]))
      } else {
      ht[i]  <- ( pfs_t*pps*(exp(-pfs_t*t[i]) - exp(-pps*t[i])) ) / ( pps*exp(-pfs_t*t[i]) -
pfs_t*exp(-pps*t[i]) )
      ft[i]  <- ((pfs_t*pps) / (pps-pfs_t)) * (exp(-pfs_t*t[i]) - exp(-pps*t[i]))
      Lft[i] <- (1 - (((pps) / (pps-pfs_t)) * (exp(-pfs_t*t[i]))) + (((pfs_t) / (pps-pfs_t)) *
(exp(-pps*t[i]))))
    }

    # When PFS_c=PPS
    if(median_c==median_pps){
      hc[i]  <- ((pfs_c^2)*t[i]) / (1 + pfs_c*t[i])
      fc[i]  <- (pfs_c^2)*t[i]*exp(-pfs_c*t[i])
      Lfc[i] <- (1 - exp(-pfs_c*t[i]) - (pfs_c)*t[i]*exp(-pfs_c*t[i]))
      } else {
      hc[i]  <- ( pfs_c*pps*(exp(-pfs_c*t[i]) - exp(-pps*t[i])) ) / ( pps*exp(-pfs_c*t[i]) -
pfs_c*exp(-pps*t[i]) )
```

```
      fc[i]  <- ((pfs_c*pps) / (pps-pfs_c)) * (exp(-pfs_c*t[i]) - exp(-pps*t[i]))
      Lfc[i] <- (1 - (((pps) / (pps-pfs_c)) * (exp(-pfs_c*t[i]))) + (((pfs_c) / (pps-pfs_c)) *
(exp(-pps*t[i]))))
    }

    # Calculating Phi
    nu[i]  <- (log(ht[i]/hc[i]) * ((1-Lft[i])/((1-Lfc[i])+(1-Lft[i]))) * (1 -
(1-Lft[i])/((1-Lfc[i])+(1-Lft[i]))) * 1/2 * (fc[i] + ft[i])) * (1/k)
    de[i]  <- (                  ((1-Lft[i])/((1-Lfc[i])+(1-Lft[i]))) * (1 -
(1-Lft[i])/((1-Lfc[i])+(1-Lft[i]))) * 1/2 * (fc[i] + ft[i])) * (1/k)

  }


  # Calculating sample size
  for(j in 1:accral){
    n[j] <-  ((qnorm(1-alpha/2)+qnorm(1-beta))^2) /
(sum(nu[1:(l*k-j*k)])/sqrt(sum(de[1:(l*k-j*k)])))^2
  }

  if (accral==0){
    nt <- ceiling (sum(n)/2)
    } else {
    nt <- ceiling (sum(n)/accral / 2)
  }

  # Define p-vlaue vectors
  hypo_pvalue <- c();
  os_pvalue   <- c()

  # Calculating Power
  for(i in 1:iteration){
    # Generating random numbers
    pfs_t <- rexp(  nt, log(2)/median_t  )
    pfs_c <- rexp(  nt, log(2)/median_c  )
    pps   <- rexp(2*nt, log(2)/median_pps)

    # Generating distributions
    hypo_os <- c(pfs_t, pfs_c) + pps
    exp_os  <- c(rexp(nt, log(2)/(median_t+median_pps)), rexp(nt,
log(2)/(median_c+median_pps)))
    group   <- rep(c(1,0), each=nt)

    # Generating study duration
    if (accral==0){
      la      <- length
      } else {
      amtemp <- c()
      am <- c()
      for(j in 1:accral){
        m      <- (nt - nt%%accral) / accral + (j <= nt%%accral)
```

```
        amtemp <- rep(c(accral-j),m)
        am     <-  c(am,amtemp)
      }
      la      <- c(length + c(am),length + c(am))
    }

    # Generating event
    stat_hypo <- hypo_os <= la
    stat_os   <- exp_os  <= la

    # p-value
    hypo_pvalue[i] <- summary(coxph(Surv(hypo_os, stat_hypo) ~ group))$sctest[3]
    os_pvalue[i]   <- summary(coxph(Surv(exp_os , stat_os  ) ~ group))$sctest[3]
  }

  # Save output
  result <- c(nt, sum(hypo_pvalue < alpha)/iteration, sum(os_pvalue < alpha)/iteration, point1,
point2)

  return(result)
}
```

연세대학교
YONSEI UNIVERSITY

ABSTRACT

Sample size and power for comparison of survival functions under hypo-exponential distribution based on log-rank test

Ko, Hee Jung
Dept. of Biostatistics and Computing
The Graduate School
Yonsei University

Most clinical trials of anti-cancer agent, comparison of overall survival rate between treatment and control arms is used for primary end point and progression free survival for secondary end point. Overall survival is partitioned into two parts and expressed it as the sum of progression free survival and post progression survival. If the progression free survival and post progression survival have exponential distribution, then the overall survival has hypo-exponential distribution. In simulation results of Broglio and Berry's study(2009), the statistical power of the log-rank test for the comparison of two overall survival function decreased as the post progression survival increased when two progression free survival is different. Park SH(2013) assumed that log-rank test which is widely used to compare survival times of two groups is no longer satisfied with proportional hazards, so statistical power of log-rank test would not reached at expected level. She

연세대학교
YONSEI UNIVERSITY

compared statistical power for log-rank test which is used to compare of two survival functions with Gehan, Peto-Peto, Fleming-Harrington (p=0.5,q=0.5) and Renyi type test statistics which have different weight depend on times and log-rank test show the most powerful statistical power among other tests in the simulation.

In this thesis, we assume that each progression free survival and post progression survival have exponential distribution therefore the overall survival has hypo-exponential distribution for comparing overall survival between treatment and control arms in the anti-cancer clinical trials. We proposed two methods to calculate non centrality mean parameter of log-rank test statistic using Schoenfeld(1981) and Zhang and Quan(2009). We organized log-rank test statistic, overall survival has hypo-exponential distribution and the methods to calculate non centrality mean parameter of log-rank test statistic. And we proposed a method of using expectation of test statistic and a method of using integral with an exact known form of each functions to calculate non centrality mean parameter. A simulation was designed to assess the performance of proposed methods with various conditions, such as progression free survival, post progression survival, interval unit, total study duration and power.

In Zhang method, the power satisfies the pre specified level when post progression survival is longer than 9 months regardless of the period of progression free survival of control arm, however the power is under estimated in certain situations when the difference of hazards ratio is bigger. In Schoenfeld method, all the power satisfies the expected level regardless any conditions, however the over power appears when the bigger difference of hazards ratio and the shorter post progression survival. At last, we find out small sample size need in clinical trials when the overall survival has

연세대학교
YONSEI UNIVERSITY

hypo-exponential distribution against exponential distribution.

Key words : hypo-exponential, log-rank test, non centrality mean parameter, sample size, power, overall survival, progression free survival, post progression survival