



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](#)

폐암 환자에서 뇌졸중 발생 여부에

따른 사망률 비교:

병원 OMOP-CDM 데이터를

이용한 IPTW 기반 랜드마크 분석

연세대학교 보건대학원

보건정보통계학과 보건정보통계전공

고 창 준

폐암 환자에서 뇌졸중 발생 여부에
따른 사망률 비교:

병원 OMOP-CDM 데이터를
이용한 IPTW 기반 랜드마크 분석

지도 박 소 희 교수

이 논문을 보건학 석사 학위 논문으로 제출함

2025년 6월

연세대학교 보건대학원
보건정보통계학과 보건정보통계전공
고 창 준

고창준의 보건학 석사학위 논문을 인준함

심사위원 _____ 박 소 희 _____ 인

심사위원 _____ 하 민 진 _____ 인

심사위원 _____ 주 재 홍 _____ 인

연세대학교 보건대학원
보건정보통계학과 보건정보통계전공

2025년 6월

감사의 글

2023년 3월에 시작한 대학원 석사과정이 어느덧 마무리를 향해 가고 있습니다. 2년 6개월이라는 시간이 이렇게 빠르게 흘러갈 줄은 몰랐고, 입학 당시 20대였던 제가 이제는 30대가 되었습니다. 직접 수집한 의료 데이터를 분석해 보고 싶다는 열망으로 보건정보통계학과에 입학했던 제가, 이제 졸업논문의 마지막을 준비하고 있다는 사실이 감회가 깊습니다. 학교생활 동안 많은 분들의 도움을 받았기에 이 글을 쓸 수 있었다고 생각하며, 이 자리를 빌려 진심으로 감사의 말씀을 드립니다.

가장 먼저, 통계에 대한 기본 지식 없이 입학하여 막막했던 저에게 기초적인 통계 이론과 체계적인 수업으로 방향을 제시해 주시고, 연구 전반과 논문 작성 과정에서도 끝까지 세심하게 지도해 주신 박소희 교수님께 진심으로 감사드립니다.

SAS, R, Python 등 다양한 통계 프로그램 실습을 통해 실제 데이터를 다루는 능력을 기를 수 있도록 이끌어 주시고, 연구 과정에서의 여러 문제에도 실질적인 조언을 해주신 한민진 교수님께도 진심으로 감사드립니다.

보건정보통계학과의 든든한 버팀목이 되어 주시며, 깊이 있는 강의와 따뜻한 격려로 학업의 방향을 밝혀 주신 남정모 교수님께도 진심으로 감사드립니다.

고도화된 통계분석 실습과 논문 작성 과정에서 아낌없는 피드백을 주신 주재홍 박사님, 그리고 수업 공지와 학업 전반에 큰 도움을 주신 조교 선생님들께도 진심으로 감사드립니다.

새롭고 다양한 데이터 학습과 폭넓은 연구 경험의 기회를 제공해 주시며,
학업과 업무를 병행하는 과정에서도 이해와 배려로 큰 편의를 주신 YouLab의
수장 유승찬 교수님께 진심으로 감사드립니다.

연구실에 처음 적응하는 과정에서 세심한 도움을 주시고, 다양한 데이터
연구와 관련해서도 아낌없는 조언을 해주신 신서정 교수님께 진심으로
감사드립니다.

YouLab의 모든 구성원 여러분께서 보내주신 도움과 따뜻한 격려 덕분에 매
순간 행복하고 뜻깊은 연구실 생활을 이어갈 수 있었습니다. 진심으로
감사드립니다.

힘들 때나 기쁠 때나 언제나 위로와 힘이 되어 준 사랑하는
보건정보통계학과 동기들이 있었기에 포기하지 않고 이 자리까지 올 수
있었습니다. 진심으로 감사드립니다.

함께하는 순간마다 행복과 즐거움을 안겨준 관악팻의 문송아, 신지수,
류효승 선생님께도 진심으로 감사드립니다.

마지막으로, 일과 학업을 병행하며 바쁜 삶을 살아가는 저를 누구보다
열심히 응원해 주시고 사랑으로 지켜봐 주신 아버지 고영수와 어머니 연선숙,
그리고 최근 수술로 기력이 쇠하셨음에도 늘 손주를 응원해 주시는 할머니
서길자께 마음 깊이 감사의 마음을 전합니다. 부모님과 할머니가 계시지
않았다면 결코 여기까지 올 수 없었을 것입니다. 부끄러움이 많아 평소에는
표현하지 못했지만, 이 자리를 빌려 말씀드립니다. 정말 많이 사랑합니다.

감사합니다.

2025년 6월 고창준 배상

차 례

국문 요약	vii
-------------	-----

I . 서론

1. 연구 배경	1
2. 연구 목적	6

II . 연구 방법

1. 연구 자료	7
2. 연구 대상	10
3. 연구 설계 및 시간 구조	13
4. 변수 정의	15
5. 분석 방법	22

III . 연구 결과

1. 연구대상자의 일반적 특성	31
1.1. 전체 연구대상자의 IPTW 적용 전, 일반적 특성	31
1.2. 전체 연구대상자의 IPTW 적용 후, 일반적 특성	35
1.3. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 일반적 특성	39
1.4. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 일반적 특성	43
1.5. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 일반적 특성	47
1.6. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 일반적 특성	51

2. 뇌졸중 발생 여부에 따른 생존분석	55
2.1. 전체 연구대상자의 IPTW 적용 전, 생존분석	55
2.2. 전체 연구대상자의 IPTW 적용 후, 생존분석	60
2.3. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 생존분석	66
2.4. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 생존분석	69
2.5. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 생존분석	72
2.6. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 생존분석	75
 IV. 고찰	
1. 연구 결과에 대한 고찰	78
2. 연구의 강점과 제한점	83
 V. 결론	86
 참고문헌	88
 Abstract	93

표 차 례

표 1. 2020년 36개 암종 및 전체 암의 신규 발생 및 사망 현황(상위 5개)	1
표 2. 2020년 대한민국 주요 암종별 발생 및 사망 현황(상위 5개)	3
표 3. 폐암 진단 정의를 위한 Concept ID 목록 및 설정값	11
표 4. 뇌졸중(뇌경색) 진단 정의를 위한 Concept ID 목록 및 설정값	17
표 5. 뇌졸중(뇌출혈) 진단 정의를 위한 Concept ID 목록 및 설정값	18
표 6. 성별 정의를 위한 Concept ID 목록 및 설정값	19
표 7. 만성질환 진단 정의를 위한 Concept ID 목록 및 설정값	20
표 8. 치료 관련 요인 정의를 위한 Concept ID 목록 및 설정값	21
표 9. 전체 연구대상자의 IPTW 적용 전, 일반적 특성	34
표 10. 전체 연구대상자의 IPTW 적용 후, 일반적 특성	38
표 11. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 일반적 특성	42
표 12. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 일반적 특성	46
표 13. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 일반적 특성	50
표 14. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 일반적 특성	54
표 15. 전체 연구대상자의 IPTW 적용 전, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	58
표 16. 전체 연구대상자의 IPTW 적용 전, Time-Dependent Cox Regression을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	59
표 17. 전체 연구대상자의 IPTW 적용 후, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	63
표 18. 전체 연구대상자의 IPTW 적용 후, Time-Dependent Cox Regression을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	64

표 19. 전체 연구대상자의 IPTW 적용 후, Marginal Structural Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	65
표 20. Truncation 비율별 Marginal Structural Model 분석 결과	65
표 21. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	68
표 22. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	71
표 23. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	74
표 24. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석	77

그림 차례

그림 1. 2020년 전 세계 암 발생률 및 사망률 분포(상위 암종 기준)	2
그림 2. OMOP-CDM 테이블 구성	9
그림 3. 연구대상자 선정 및 분석 흐름 개요	12
그림 4. 시간 흐름에 따른 연구 관찰 구조 및 사건 정의 예시	14
그림 5. 전체 연구대상자의 뇌졸중 진단 여부에 따른 추적관찰 기간 분포 ...	33
그림 6. 전체 연구대상자의 IPTW 적용 전후 공변량 간 Standardized Mean Difference	37
그림 7. 랜드마크 기간 2년 연구대상자의 뇌졸중 진단 여부에 따른 추적관찰 기간 분포	41
그림 8. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전후 공변량 간 Standardized Mean Difference	45
그림 9. 랜드마크 기간 4년 연구대상자의 뇌졸중 진단 여부에 따른 추적관찰 기간 분포	49
그림 10. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전후 공변량 간 Standardized Mean Difference	53
그림 11. 전체 연구대상자의 IPTW 적용 전, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test	57
그림 12. 전체 연구대상자의 IPTW 적용 후, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test	62
그림 13. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test	67

그림 14. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test	70
그림 15. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test	73
그림 16. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test	76

국 문 요 약

폐암 환자에서 뇌졸중 발생 여부에 따른 사망률 비교 :
병원 OMOP-CDM 데이터를 이용한 IPTW 기반 랜드마크 분석

연구 배경

폐암은 국내 암 사망률 1위(2020년 22.7 %)이며 5년 상대생존율이 38.5 %로 다른 주요 암종보다 현저히 낮다. 뇌졸중은 폐암 환자에서 빈번히 발생해 예후를 악화시킬 수 있으나, 뇌졸중 발생 ‘시점’이 장기 생존에 미치는 인과적 영향은 충분히 규명되지 않았다. 본 연구는 폐암 환자에서 뇌졸중 발생 여부와 시점이 전체 사망 위험에 미치는 효과를 평가하고자 하였다.

연구 대상 및 방법

본 연구는 세브란스병원 OMOP-CDM 데이터를 활용한 후향적 코호트 연구로, 2006~2018년 폐암 진단 환자 20,696명을 분석에 포함하였다. 주요 관심 변수는 뇌졸중 발생 여부 및 시점이며, 종속 변수는 사망 여부와 생존 기간이다. 공변량은 연령, 성별, 고혈압·당뇨·고지혈증 진단 여부, 수술 및 방사선 치료 여부이다. 성향점수 기반 안정화된 역처리확률가중법(IPTW)을 적용한 후, Cox 비례위험모형, 시간 종속 Cox 회귀모형, IPTW 기반 주변구조모형(MSM), 2년 및 4년 랜드마크 분석을 수행하였다.

연구 결과

전체 연구대상자에 대해 IPTW를 적용하고 다양한 공변량을 고정 변수로 통제한 Cox 비례위험모형에서는, 뇌졸중 진단 여부와 사망 위험 간에 유의한 차이가 없었다(HR=1.13, 95% CI:0.97-1.31). 반면, 뇌졸중 발생 시점과 공변량의 시간 변화를 반영한 시간 종속 Cox 회귀모형에서는 사망 위험이 약 2.4배 높았고(HR=2.43, 95% CI:1.99-2.96), IPTW 기반 MSM 분석에서도 유사한 결과가 나타났다(HR=2.62, 95% CI:2.13-3.22). 극단 가중치를 절단한 민감도 분석에서도 결과는 일관되었다(HR=3.02-3.05). 또한, 2년 및 4년 이상 생존한 환자군을 대상으로 한 IPTW 기반 Cox 비례위험모형에서도, 뇌졸중 진단은 두 시점 모두에서 사망 위험 증가와 유의한 관련이 있었다(2년 HR=1.56, 4년 HR=1.89).

결론

본 연구는 세브란스병원 OMOP-CDM 데이터를 활용하여, 폐암 환자에서 뇌졸중 발생 여부 및 시점이 사망 위험에 미치는 영향을 평가하였다. IPTW를 적용한 Cox 비례위험모형에서는 유의한 차이가 없었으나, 시간 종속 Cox 회귀모형과 MSM에서는 뇌졸중이 사망 위험을 유의하게 증가시키는 요인으로 나타났다. 2년 및 4년 시점의 랜드마크 분석에서도 유사한 경향이 확인되었으며, 이는 시간 의존적 구조를 반영한 분석이 실제 위험을 정밀하게 추정하는 데 중요함을 시사한다.

키워드: 폐암, 뇌졸중, 사망 위험, OMOP-CDM, 랜드마크 분석, IPTW, MSM

I . 서 론

1. 연구 배경

의학의 발전에 따라 질병에 대한 치료 기술이 진화하고 있으며, 인간의 수명을 점차 연장되고 있다. 그러나 현대 사회의 다양한 원인으로 인해 암 발생률은 점차 증가하고 있다. 그중 폐암(Lung Cancer)은 2020년 전 세계 암 통계에 따르면 신규 환자 수가 약 220만 명으로 전체 암 중 발생률 2위, 사망률은 1위를 기록하고 있다(Sung H et al., 2021).

표 1. 2020년 36개 암종 및 전체 암의 신규 발생 및 사망 현황(상위 5개)

암 종류	신규 발생 수 (건)	전체 발생 중 비율(%)	사망 수 (건)	전체 사망 중 비율(%)
유방암	2,261,419	11.7	684,996	6.9
폐암	2,206,771	11.4	1,796,144	18.0
전립선암	1,414,259	7.3	375,304	3.8
피부암	1,198,073	6.2	63,731	0.6
대장암	1,148,515	6.0	576,858	5.8

※ 출처: GLOBOCAN 2020, CA Cancer J Clin, Sung H et al., 2021

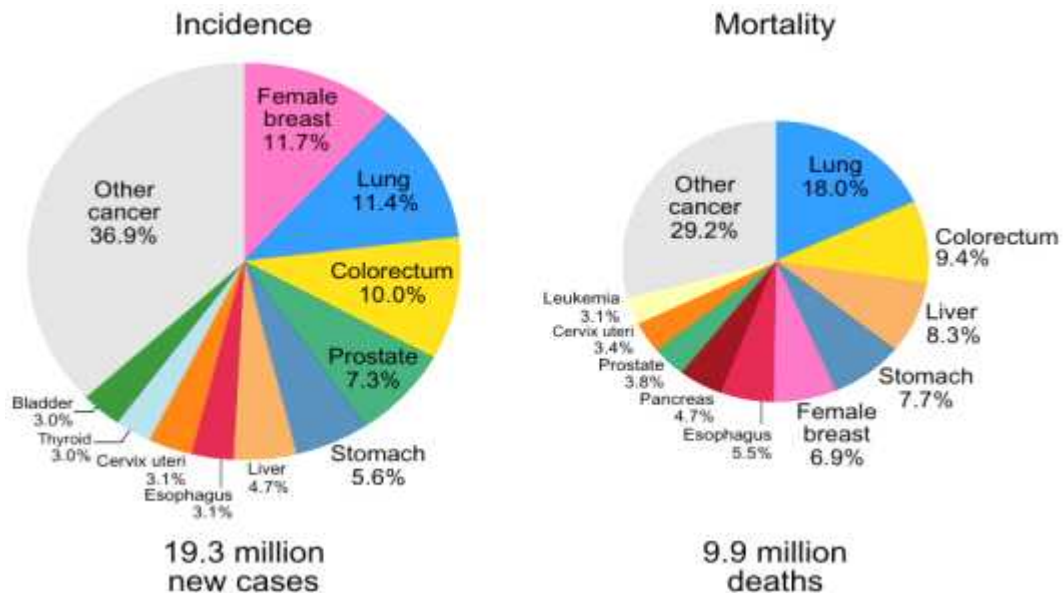


그림 1. 2020년 전 세계 암 발생률 및 사망률 분포(상위 암종 기준)

※ 출처: GLOBOCAN 2020, CA Cancer J Clin, Sung H et al., 2021

대한민국 2020년 국가암등록통계 자료에 의하면, 2020년 우리나라에서는 총 247,952명의 암 발생자가 있었고, 이 중 28,949명이 폐암 환자였다. 이는 전체 암 발생의 11.7%에 해당하며, 암종별로는 갑상선암 다음으로 두 번째로 높은 발생률을 보였다. 또한 폐암은 전체 암 사망자 중 22.7%인 18,673명이 사망해, 가장 높은 사망률을 보였다(Kang MJ et al., 2023).

표 2. 2020년 대한민국 주요 암종별 발생 및 사망 현황(상위 5개)

암 종류	신규 발생 수 (건)	전체 발생 중 비율(%)	사망 수 (건)	전체 사망 중 비율(%)
갑상선암	29,180	11.8	365	0.4
폐암	28,949	11.7	18,673	22.7
대장암	27,877	11.2	8,869	10.8
위암	26,662	10.8	7,510	9.1
유방암	24,923	10.1	2,745	3.3

※ 출처: 국가암등록통계, Kang MJ et al., 2023

폐암 환자의 5년 생존율은 38.5%로 대장암, 위암, 유방암, 전립선암 및 신장암의 생존율 70~90%와 비교하여 상당히 낮은 수준이다(National Cancer Information Center, 2021). 폐암 환자의 5년 생존율이 낮은 이유는 폐암의 초기 증상에 대해 관찰이 어려워 진단이 쉽지 않을 뿐만 아니라 재발과 전이 비율이 타종 암에 비해 월등히 높기 때문이다(Lee YE & Ryu E, 2021). 암과 관련된 다양한 합병증 중 뇌혈관 합병증(출혈 및 경색)은 전이(Metastases)에 이어 중추신경계(Central Nervous System, CNS) 병변에서 두 번째로 높은 빈도를 보이는 것으로 보고되었다(Graus F, Rogers LR and Posner JB, 1985). 대만의 국민건강보험의 인구 기반 청구 데이터 분석한 연구에 따르면 비암 환자 대비 폐암 환자에서 허혈성 뇌졸중 발생이 약 1.43배, 출혈성 뇌졸중이 약 1.78배 높게 발생하였다고 보고되었다(Chen PC et al., 2011).

뇌졸중(Stroke)은 중추신경계 병변으로 인해 뇌혈관이 막히거나 터짐으로써 발생하는 혈액 순환 장애로, 의식 장애와 신체 마비를 초래하는 심각한 뇌혈관 질환이다. 국내에서 뇌졸중은 암, 심장질환과 함께 주요 사망 원인으로 분류되며, 인구 10만 명당 45.8명이 사망하는 높은 사망률을 보인다(Statistics Korea, 2017). 뇌졸중은 크게 혈관이 막혀서 생기는 허혈성 뇌졸중(Ischemic Stroke)과 혈관이 터져서 생기는 출혈성 뇌졸중(Hemorrhagic Stroke)으로 나뉘며, 인지 기능 저하, 근력 약화, 비정상적인 근 긴장도, 비정상적인 자세와 움직임 감소, 협응성 감소 등 다양한 영구적 장애를 동반한다(Chang WH et al., 2016). 뇌졸중의 원인은 다양하지만, 암과 관련하여 뇌졸중 발생 위험이 증가하는 것으로 보고되고 있다. 암은 고 응고 상태(hypercoagulable state)를 유발하여 비세균성 혈전성 심내막염(nonbacterial thrombotic endocarditis)이나 전신 혈전증을 통해 허혈성 뇌졸중을 초래할 수 있으며, 암 전이로 인해 혈관이 물리적으로 압박되거나 침범되어 혈류가 차단되고, 암 치료(수술, 방사선 치료 등)로 인해 혈관 손상이나 응고 장애가 발생할 수 있다. 또한, 폐암 환자의 경우 응고 장애와 혈관 취약성이 동반되면서 허혈성 뇌졸중 이후 출혈성 전환(hemorrhagic transformation)의 위험이 증가할 수 있으며, 뇌전이(brain metastases)나 항응고 치료와 같은 요인으로 인해 자발적인 두개내 출혈(intracranial hemorrhage, ICH) 위험도 상승할 수 있다(Sands JM, Daly ME and Lee EQ, 2020). 특히, 암 환자의 뇌졸중은 일반 환자보다 더 복잡한 기전을 가지며, 더 나쁜 예후를 보이는 경향이 있다(Graus F, Rogers LR and Posner JB, 1985).

선행연구들에 따르면, 암 환자는 전반적으로 뇌졸중 위험도가 증가하는 경향을 보이며, 특히 폐암은 50대 이상에서 발병률이 높은 동시에 뇌졸중 위험도가 높은 암으로 알려져 있다(Jang HS et al., 2019). 또한, 폐암 환자 중 뇌졸중이 발생한 경우, 그렇지 않은 환자에 비해 사망 위험이 약 1.73배 높으며, 이러한 사망 위험을 증가시키는 주요 요인으로 연령, 성별, 만성질환 등이 영향을 미친다는 연구 결과가 보고된 바 있다(Chen L, Zhao X and Wang S, 2021). 그러나 국내 연구에서는 폐암 환자의 뇌졸중 발생에 대한 연구가 일부 진행되고 있음에도 불구하고, 이러한 위험 요인이 국내 폐암 환자에서 동일하게 작용하는지, 나아가 뇌졸중 발생이 사망률에 미치는 직접적인 영향을 분석한 연구는 부족한 실정이다. 따라서 본 연구에서는 세브란스병원 CDM 데이터를 활용하여 폐암 환자의 뇌졸중 발생과 사망률 간의 관계를 보다 체계적으로 분석하여, 국내 데이터를 기반으로 한 근거를 마련하고자 한다.

2. 연구 목적

본 연구는 2006년부터 2023년까지의 세브란스병원 OMOP-CDM 데이터를 활용하여 폐암으로 최초 진단받은 환자를 대상으로 최대 18년간 추적 관찰하고, 뇌졸중 발생 여부 및 발생 시점이 사망 위험에 미치는 영향을 평가하고자 한다. 인구사회학적 요인, 만성질환, 치료 관련 요인 등 다양한 공변량을 통제하여 사망 위험 요인을 통합적으로 분석하며, 본 연구 결과는 폐암 환자의 뇌졸중 예방과 사망률 감소를 위한 효율적인 관리 및 치료 전략 수립에 기여할 수 있을 것으로 기대된다.

구체적인 연구 목적은 다음과 같다.

첫째, 2006년부터 2023년까지의 세브란스병원 OMOP-CDM 데이터를 활용하여 폐암 환자 중 뇌졸중이 발생한 군과 발생하지 않은 군의 일반적 특성을 분석한다.

둘째, 전체 연구대상자에 대해 안정화 가중치를 적용한 역확률치료가중법(IPTW)을 활용하여 군 간의 특성 차이를 보정한 후, 가중치 적용 전·후 Kaplan-Meier 생존 곡선과 Log-Rank 검정을 통해 생존 양상의 차이를 비교하고, Cox 비례위험모형, 시간 종속 Cox 회귀모형 및 주변구조모형(MSM)을 활용하여 뇌졸중 발생 여부에 따른 사망 위험을 분석한다.

셋째, 2년 및 4년의 랜드마크 시점을 기준으로, IPTW 적용 전·후의 Kaplan-Meier 생존 곡선 및 Log-Rank 검정을 통해 생존 양상의 차이를 비교하고, Cox 비례위험모형을 활용하여 뇌졸중 발생 여부에 따른 사망 위험을 분석한다.

II. 연구 방법

1. 연구 자료

본 연구는 세브란스병원 OMOP-CDM 데이터를 이용하여 폐암 환자에서 뇌졸중 발생이 사망률에 미치는 영향을 파악하고자 하였다. 2006년부터 2023년까지 최대 18년 동안 적재된 데이터를 활용하였다.

의료 데이터 공유는 데이터 구조의 다양성, 품질 문제, 법적·윤리적 제약으로 인해 어려움이 존재한다(Hripcsak G et al., 2015). 기존 연구는 일부 데이터를 연구 주도 기관과 공유하는 방식이었으나, 이는 높은 비용과 시간이 소요되며, 개인정보 보호 문제도 있다. 이를 해결하기 위해 공통 데이터 모델(Common Data Model, CDM) 기반의 분산 연구망(Distributed Research Network)이 도입되었다. CDM은 데이터를 표준화하여 연구자가 원본 데이터를 직접 공유하지 않고도 연구를 수행할 수 있도록 하며, 개인정보 보호와 연구 효율성을 동시에 보장한다. 대표적인 사례인 OMOP-CDM(Observational Medical Outcomes Partnership Common Data Model)은 병원 간 비표준 용어의 차이를 해결하고 데이터 호환성을 확보하여 다기관 협력 연구를 지원한다(Park RW, 2017). OMOP-CDM은 미국 정부의 OMOP 프로젝트(2008)에서 시작되었으며, 이후 OHDSI(Observational Health Data Sciences and Informatics)가 이를 국제 협력 연구 모델로 발전시켰다(Hripcsak et al., 2015). OHDSI는 CDM을 활용해 약물 안전성 평가, 의료 질 평가, 인공지능 기반 예측 모델 개발 등 다양한 연구를 지원한다(Park RW, 2017; Observational Health Data Sciences and Informatics, 2021).

OMOP-CDM은 임상 진료 데이터(Standardized Clinical Data), 의료 시스템 정보(Standardized Health System Data), 건강 경제 데이터(Standardized Health Economics), 파생 요소(Standardized Derived Elements), 결과 스키마(Results Schema), 표준 용어체계(Standardized Vocabularies), 메타데이터 등으로 구성되어 있으며, 각 구성 요소는 그림 3에 제시하였다(Observational Health Data Sciences and Informatics, 2021). 그림 3은 CDM의 테이블 구조를 시각적으로 보여주는 도식으로, Person, Visit, Condition, Drug, Measurement 등 환자의 진료 이력을 담은 표준화된 테이블을 중심으로, 이를 기반으로 파생되는 요소(Condition Era, Drug Era 등)와 분석 결과를 저장하는 결과 테이블(Cohort, Cohort Definition) 등이 포함되어 있다. 또한, 표준 용어 계층은 SNOMED, RxNorm 등 국제 표준을 기반으로 개념(Class), 관계(Relationship), 상위·하위 개념(Ancessor) 간 연결 구조를 제공하여, 진단 및 치료 정보를 정밀하게 정의하고 연산할 수 있도록 구성되어 있다. OMOP-CDM에서 ‘Concept’은 각각의 진단, 처치, 약제, 검사 등 임상 항목을 표준화된 개념 ID로 정의한 단위를 의미하며, ‘Concept Set’은 이와 같은 개념들을 조합하여 특정 질환이나 조건을 정의하는 데 사용되는 개념 묶음이다. 연구자는 Concept Set을 구성할 때 상위 개념뿐 아니라 관련된 자식 개념까지 포함하여 포괄적인 질병 정의를 수행할 수 있다. 또한 ‘Cohort’는 특정 조건에 부합하는 환자 집단을 의미하며, ‘Cohort Definition’은 이 집단을 생성하기 위한 조건과 알고리즘을 정의한 구조이다. 연구자는 Concept Set을 기반으로 질환 또는 사건 정의를 설정하고, 이 정의를 바탕으로 분석 대상군(Cohort)을 생성하여 연구를 수행하게 된다(Park RW, 2017; Observational Health Data Sciences and Informatics, 2021).

본 연구에서는 이러한 구조를 바탕으로 폐암 및 뇌졸중을 정의하기 위해 SNOMED 기반의 진단 개념을 포함한 Concept Set을 구성하였고, 자식 개념까지 포함하는 하위 매핑을 적용하였다. 또한, 성별, 연령, 고혈압, 당뇨병, 고지혈증, 흡연 여부 등 주요 공변량 역시 모두 SNOMED 기반의 표준 개념을 활용하여 각각의 Concept 또는 Concept Set으로 정의하고 분석에 사용하였다. 이는 폐암 및 뇌졸중 환자를 보다 정확히 식별하고, 공변량의 측정 정확도와 임상적 포착력을 높이는 데 중요한 역할을 하였다.

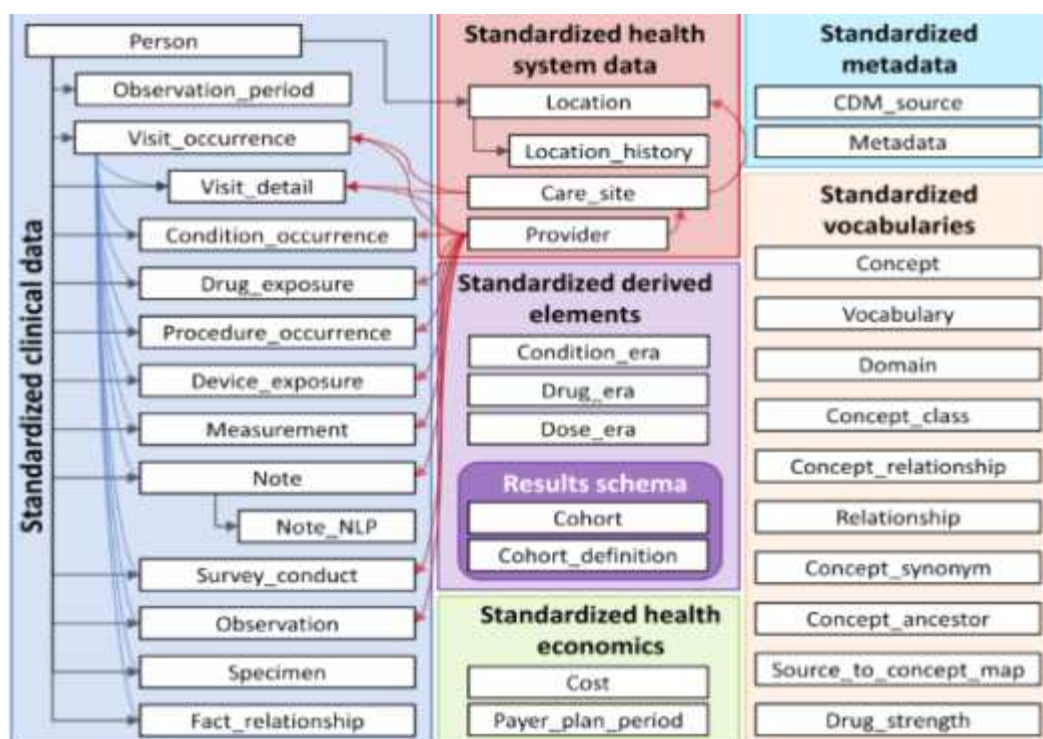


그림 2. OMOP-CDM 테이블 구성

2. 연구 대상

본 연구는 연세대학교 세브란스병원의 OMOP-CDM 데이터를 기반으로 수행되었으며, 전체 데이터베이스의 대상자 수는 총 6,401,087명이었다. 이 중 2006년부터 2018년까지 최초로 폐암 진단을 받은 환자는 21,044명이었다. 폐암 진단은 OMOP-CDM의 Standard Vocabulary인 SNOMED를 기반으로 하였으며, Concept Class는 Clinical Finding으로 설정하였다. 이 정의는 Ranitidine 복용이 암 발생에 미치는 영향을 평가한 선행연구에서 사용된 폐암 정의 기준을 참고하여 구성하였다(You SC et al., 2023). ‘Secondary malignant neoplasm of lung’ 및 ‘Kaposi's sarcoma of lung’의 하위 개념은 모두 제외하였다. 대부분의 Concept Set은 상위 개념의 하위 개념을 포함하는 방식으로 구성되었으며, 해당 Concept Set에 포함된 항목 중 하나라도 해당될 경우 폐암 진단으로 간주하였다(표 3).

선별된 환자 중 최초 폐암 진단 시점에 만 19세 미만인 24명, 폐암 진단 이전에 뇌졸중 진단을 받은 270명, 생존일이 0일 이하인 54명을 제외한 총 20,696명이 최종 연구대상자로 선정되었다. 이 중 뇌졸중이 발생한 환자는 437명, 발생하지 않은 환자는 20,259명이었다. 또한 본 연구에서는 폐암 진단 이후 2년 및 4년 시점을 기준으로 랜드마크 분석을 실시하였다. 2년 기준 분석에는 총 14,964명이 포함되었으며, 이 중 뇌졸중 발생 환자는 156명, 비발생 환자는 14,808명이었다. 4년 기준 분석에는 총 13,704명이 포함되었으며, 이 중 뇌졸중 발생 환자는 181명, 비발생 환자는 13,523명이었다. 선정된 대상자에 대해서는 역확률치료가중법(IPTW)을 적용하여 분석의 타당성을 확보한 후, 전체 대상자 및 랜드마크 시점 기준 대상자 각각에 대해 Kaplan-Meier 생존곡선 분석과 Log-Rank 검정, Cox 비례위험모형을 수행하였다. 이와

별도로, 전체 대상자에 대해서는 시간 종속 Cox 회귀모형과 주변구조모형(MSM)을 추가로 적용하여 시간 의존적 효과와 인과 추정을 보완하였다. 분석의 구체적인 방법은 ‘5. 분석 방법’ 절에서 상세히 기술하였다(그림 4).

본 연구는 연세대학교 세브란스병원 연구윤리심의위원회(Institutional Review Board, IRB)로부터 심의 면제를 승인받았다(IRB 승인번호: 4-2024-1523).

표 3. 폐암 진단 정의를 위한 Concept ID 목록 및 설정값

Concept ID	Concept Name	Exclude	Descendants
254583	Kaposi's sarcoma of lung	●	●
254591	Secondary malignant neoplasm of lung	●	●
443388	Malignant tumor of lung		●
4157333	Malignant neoplasm of main bronchus		●
4177112	Malignant tumor of trachea		●
4311499	Primary malignant neoplasm of respiratory tract		

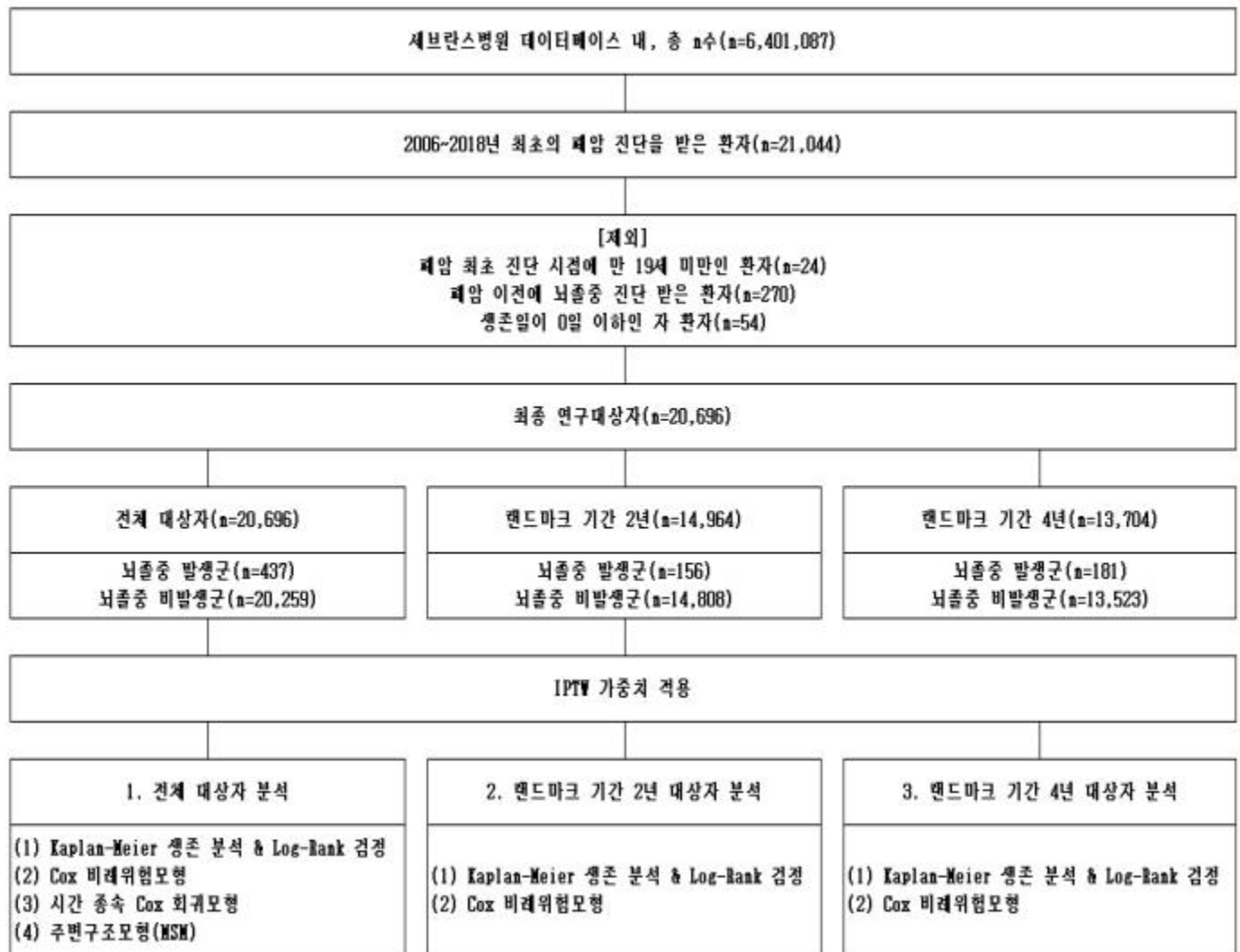


그림 3. 연구대상자 선정 및 분석 흐름 개요

3. 연구 설계 및 시간 구조

본 연구는 후향적 코호트 연구로, 폐암 최초 진단 시점을 기준으로 환자들을 추적 관찰하였다. 연구대상자의 폐암 진단 시점은 2006년부터 2018년 사이로 정의되었으며, 관찰 종료 시점은 2023년 12월 31일로 설정하였다. 관찰 기간 동안 사망한 환자는 사건(event)으로 간주되었고, 종료 시점까지 생존한 환자는 생존자로 정의하였다. 또한 종료 시점까지 사건이 발생하지 않은 경우는 중도절단(censored)으로 처리하였다. 그림 4는 이러한 연구의 시간 구조와 사건 정의 방식을 대표 환자 사례(ex.1 ~ ex.3)를 통해 시각적으로 설명한 것으로, 예를 들어 ex.1은 2010년에 뇌졸중을 진단받고 2016년에 사망한 환자로 사건 발생으로 간주된다. ex.2는 2008년 뇌졸중 진단 이후 2020년에 사망한 경우이며, 이 역시 사건으로 처리된다. 반면 ex.3은 2012년 뇌졸중 진단 후 관찰 종료일인 2023년 12월 31일까지 생존한 환자로, 이 경우는 중도절단으로 분류된다. 이와 같이 개별 환자의 시간 경과에 따른 사건 처리 방식은 본 연구의 생존 분석 구조를 이해하는 데 중요한 기초를 제공한다.

특히 본 연구에서는 폐암 생존자의 중장기 예후를 정밀하게 평가하기 위해 두 가지 분석 접근법을 병행하였다. 첫째, 전체 대상자에 대해 역확률치료가중법을 적용하여 뇌졸중 발생군과 비발생군 간의 공변량을 균형화한 뒤, 가중치를 반영한 Kaplan-Meier 생존 곡선 & Log-rank 검정, Cox 비례위험모형을 수행하였다. 또한 시간에 따라 변화하는 노출 변수인 뇌졸중의 영향을 보다 정확히 추정하기 위해 시간 종속 Cox 회귀모형과 주변구조모형도 함께 적용하였다. 둘째, 조기 사망 편향을 최소화하고 일정 기간 생존한 환자군을 기반으로 뇌졸중 발생 이후의 영향을 평가하기 위해

랜드마크 분석을 도입하였다. 폐암 진단 후 2년 및 4년 시점을 기준으로 각각 생존한 환자 집단을 설정한 후, 동일하게 역확률치료가중법을 적용하여 Kaplan-Meier 생존 곡선 & Log-rank 검정, Cox 비례위험모형을 단계적으로 수행하였다. 이처럼 분석 설계 전반에 걸쳐 환자의 시간 흐름에 따른 관찰 구조를 정교하게 반영하고 분석 기법을 계층적으로 적용함으로써, 본 연구는 생존 분석의 정밀도와 해석의 타당성을 동시에 확보하고자 하였다.

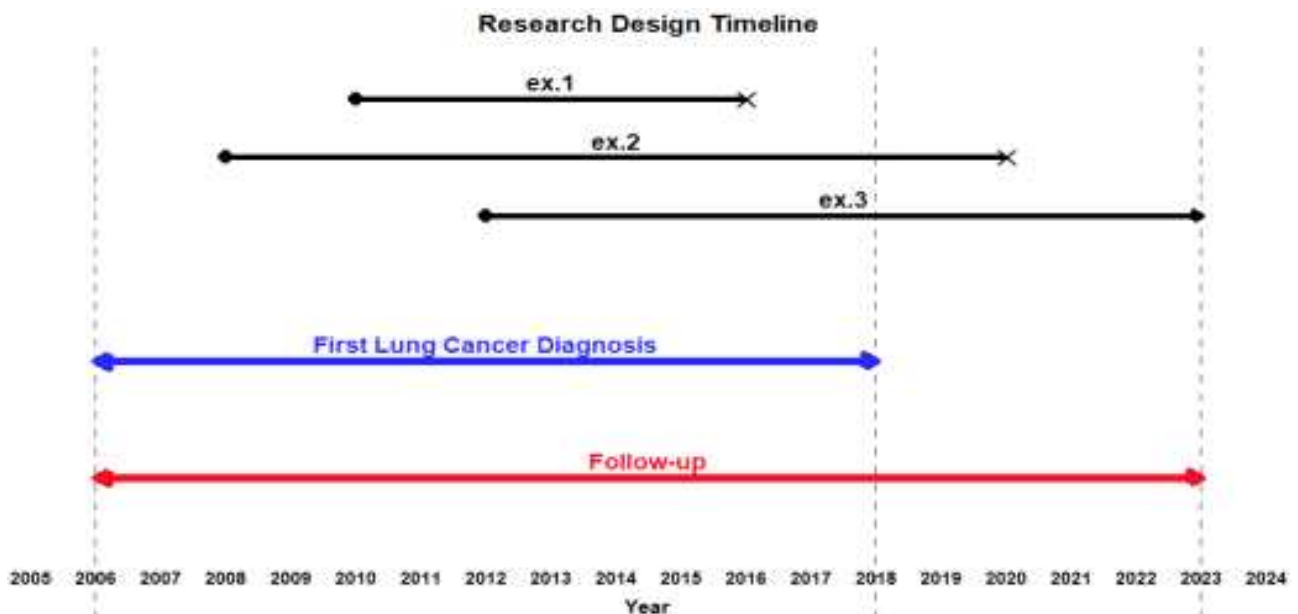


그림 4. 시간 흐름에 따른 연구 관찰 구조 및 사건 정의 예시

4. 변수 정의

가. 종속변수

1) 사망 여부

사망 여부는 OMOP-CDM의 `death` 테이블을 기반으로 정의하였으며, 사망일(`death_date`)이 존재하는 경우는 사망한 것으로 간주하여 값을 1로, 사망일이 존재하지 않는 경우는 생존으로 간주하여 값을 0으로 부여하였다. 이러한 이진 변수는 생존분석에서 종속변수(event indicator)로 사용되며, 각 랜드마크 시점 이후의 추적관찰 기간 동안 사망 발생 여부를 판단하는 데 활용되었다.

2) 생존 시간(Survival time)

모든 연구대상자의 관찰 시작 시점은 폐암으로 최초 진단된 날짜로 설정하였다. 생존 시간은 해당 진단일부터 사망이 발생한 날짜까지의 기간으로 정의하였다. 추적관찰 종료일인 2023년 12월 31일까지 사망이 발생한 경우에는 폐암 진단일부터 사망일까지의 기간을 생존 시간으로 하였고, 사망이 발생하지 않은 경우에는 폐암 진단일부터 추적관찰 종료일까지의 기간을 생존 시간으로 정의하였으며, 해당 사례는 생존분석에서 중도절단(censored)하였다.

나. 관심변수

뇌졸중 진단 여부는 OMOP-CDM의 Standard Vocabulary인 SNOMED를 기반으로 하였으며, Concept Class는 Clinical Finding으로 설정하였다. 본 연구에서는 허혈성 뇌졸중 및 출혈성 뇌졸중을 각각 별도의 Concept Set으로 구성하였으며, 연구대상자가 두 분류 Concept Set에 포함된 항목 중 하나라도 해당될 경우 뇌졸중 진단으로 간주하였다. 급성관상동맥증후군 환자에서 Ticagrelor와 Clopidogrel의 임상 결과를 비교한 선행연구에서 사용된 기준을 참고하여 구성하였다(You SC et al., 2020).

허혈성 뇌졸중은 ‘Cerebral infarction’, ‘Infarction - precerebral’ 등의 Concept을 포함하고, 해당 Concept들의 하위 개념까지 포함하여 Concept Set을 구성하였다(표 4). 출혈성 뇌졸중은 ‘Cerebral hemorrhage’, ‘Subarachnoid hemorrhage’, ‘Subdural hemorrhage’ 등의 Concept을 포함하여 개별 Concept ID를 기반으로 Concept Set을 구성하였다(표 5).

뇌졸중은 분석 목적에 따라 고정 변수(fixed covariate)로도 활용되었으며, 시간에 따른 노출 변화가 중요한 분석에서는 시간 가변 변수(time-varying covariate)로 처리되었다. 특히 뇌졸중 발생 시점은 폐암 진단일 이후 어느 시점에서든 발생 가능한 중간 사건(intermediate event)으로 간주하였으며, 해당 시점을 기준으로 분석 대상자의 노출 여부가 동적으로 반영되었다. 즉, 뇌졸중이 발생하기 전까지는 ‘비노출(0)’, 발생한 이후부터는 ‘노출(1)’로 간주하여 시간 종속 Cox 회귀모형(Time-dependent Cox Regression) 및 주변구조모형(Marginal Structural Model, MSM)에 반영하였다. 이와 같은 변수 구성은 중간 사건으로서의 뇌졸중이 사망에 미치는 인과적 영향을 보다 현실적으로 추정하기 위한 것이다.

표 4. 뇌졸중(뇌경색) 진단 정의를 위한 Concept ID 목록 및 설정값

Concept ID	Concept Name	Exclude	Descendants
443454	Cerebral infarction		●
4043731	Infarction - precerebral		●

표 5. 뇌졸중(뇌출혈) 진단 정의를 위한 Concept ID 목록 및 설정값

Concept ID	Concept Name	Exclude	Descendants
376713	Cerebral hemorrhage		
432923	Subarachnoid hemorrhage		
436430	Nontraumatic extradural hemorrhage		
439040	Subdural hemorrhage		
439847	Intracranial hemorrhage		
4049659	Subcortical hemorrhage		
4108952	Subarachnoid hemorrhage from carotid siphon and bifurcation		
4110185	Intracerebral hemorrhage, intraventricular		
4111708	Subarachnoid hemorrhage from vertebral artery		
4111709	Non-traumatic subdural hemorrhage		
4176892	Cortical hemorrhage		
4218781	Cerebral hemisphere hemorrhage		
4226021	Spontaneous hemorrhage		
4319328	Brain stem hemorrhage		
4326561	Cerebellar hemorrhage		

다. 공변량

본 연구에서는 인구사회학적 요인, 만성질환 요인, 치료 관련 요인을 통제변수로 선정하였다. 인구사회학적 요인은 연령과 성별을 포함하였으며, 만성질환 요인은 고혈압, 당뇨, 고지혈증 진단 여부를 포함하였다. 치료 관련 요인은 수술 여부와 방사선 치료 여부를 포함하였다.

1) 인구사회학적 요인

연령은 연구대상자가 최초로 폐암 진단을 받은 시점을 기준으로 산출하였다. 해당 정보는 OMOP-CDM의 person 테이블에 수록된 출생연도(year_of_birth)를 기반으로 계산하였다. 19세 미만 대상자는 제외하였으며, 연령은 50세 미만, 50-59세, 60-69세, 70-79세, 80세 이상의 총 5개 범주로 구분하여 변수로 구성하였으며, 본 연구에서는 고정 변수로 사용하였다.

성별은 OMOP-CDM의 person 테이블의 gender_concept_id 컬럼을 기반으로 하여, 남자(MALE: Concept ID 8507)와 여자(FEMALE: Concept ID 8532)로 구분하였으며, 이 역시 고정 변수로 활용하였다(표 6).

표 6. 성별 정의를 위한 Concept ID 목록 및 설정값

Concept ID	Concept Name	Exclude	Descendants
8507	MALE		
8532	FEMALE		

2) 만성질환 요인

고혈압, 당뇨, 고지혈증의 진단 여부는 OMOP-CDM의 Standard Vocabulary인 SNOMED를 기반으로 하였으며, Concept Class는 Clinical Finding으로 설정하였다. 각 질환에 대해 주요 Concept을 중심으로 하위 개념까지 포함하여 Concept Set을 구성하였고, 연구대상자가 해당 Concept Set에 포함된 항목 중 하나라도 해당될 경우 해당 질환이 있는 것으로 간주하였다. 해당 변수들은 분석 목적에 따라 고정 변수로도 사용되었으며, 시간에 따른 노출 상태 변화를 고려해야 하는 분석에서는 시간 가변 변수로 처리되었다(표 7).

표 7. 만성질환 진단 정의를 위한 Concept ID 목록 및 설정값

Concept ID	Concept Name	Exclude	Descendants
316866	Hypertensive disorder		●
201826	Type 2 diabetes mellitus		●
432867	Hyperlipidemia		●

3) 치료 관련 요인

수술 여부와 방사선 치료 여부는 OMOP-CDM의 Standard Vocabulary인 SNOMED를 기반으로 하였으며, Concept Class는 Clinical Procedure로 설정하였다. 각 항목에 대해 주요 Concept을 중심으로 하위 개념까지 포함하여 Concept Set을 구성하였고, 연구대상자가 해당 Concept Set에 포함된 항목 중 하나라도 해당될 경우 해당 치료를 받은 것으로 간주하였다. 해당 변수들은 분석 목적에 따라 고정 변수로도 사용되었으며, 시간에 따른 노출 상태 변화를 고려해야 하는 분석에서는 시간 가변 변수로 처리되었다(표 8).

표 8. 치료 관련 요인 정의를 위한 Concept ID 목록 및 설정값

Concept ID	Concept Name	Exclude	Descendants
4141937	Lung and/or mediastinum operations		●
4029715	Radiation oncology AND/OR radiotherapy		●

5. 분석 방법

본 연구에서는 폐암 환자의 생존에 영향을 미치는 요인을 분석하고자 다양한 통계 기법을 적용하였다. 각 기법은 연구 설계와 목적에 따라 적절히 선택되었으며, 그 개념과 본 연구에서의 활용 방식은 다음과 같다. 모든 가중 생존 모형(역확률치료가중법 적용 Cox 비례위험모형, 역확률치료가중법 적용 시간 종속 Cox, 주변구조모형)에는 Lin DY & Wei LJ(1989) 방식의 robust(샌드위치) 분산을 적용하여 95 % 신뢰구간을 계산하였고, 동일 개인을 cluster = person_id로 지정하였다. 반면, 가중치를 사용하지 않은 기본 Cox 모형은 partial-likelihood 기반 표준 오차를 사용하였다.

1) 카이제곱 검정(Chi-square test)

카이제곱 검정은 범주형 변수 간의 독립성 또는 적합성(goodness-of-fit)을 평가하는 데 사용되는 대표적인 비모수적(non-parametric) 통계 방법이다. 이 검정은 관측 빈도(observed frequency)와 기대 빈도(expected frequency) 간의 차이를 기반으로 통계량을 산출하며, 두 변수 간에 유의한 연관성이 있는지를 검정한다. 특히 범주형 데이터(예: 성별, 질환 유무 등)에 적합하며, 연속형 변수에는 적용할 수 없다. 카이제곱 검정은 Karl Pearson에 의해 고안되었으며, 표본이 정규분포를 따른다는 가정 없이도 가설 검정이 가능하다는 점에서 통계학적으로 중요한 의미를 갖는다(Rana RK & Singhal R, 2015).

본 연구에서는 뇌졸중을 진단받은 군과 진단받지 않은 군 간의 인구사회학적 특성(성별, 연령), 만성질환 요인(고혈압, 당뇨, 고지혈증), 치료 관련 요인(수술, 방사선 치료 여부)의 분포를 비교하였다. 모든 변수를 범주형으로 설정하여 빈도와 백분율로 제시하였고, 두 군 간의 차이는 chi-square 검정을 통해 통계적으로 유의한지를 평가하였다.

2) Kaplan-Meier 생존 분석(Kaplan-Meier Survival Analysis)

Kaplan-Meier 분석은 생존 곡선을 추정하기 위한 비모수적(non-parametric) 방법으로, 시간의 흐름에 따라 사건(예: 사망, 질병 발생 등)의 누적 확률을 추정하는 데 활용된다. 이 기법은 자료에 중도 절단(censoring)이 존재하는 경우에도 생존률을 안정적으로 추정할 수 있다는 장점이 있으며, 생존 곡선은 사건이 발생할 때마다 계단식으로 변화하는 특징을 갖는다. 모집단의 생존 함수를 특정 분포에 따라 가정하지 않아, 실제 임상 연구나 관찰연구에 널리 사용된다(Lee SW, 2023).

본 연구에서는, 뇌졸중을 진단받은 군과 진단받지 않은 군 간의 생존률 변화를 시간의 흐름에 따라 비교하고자 Kaplan-Meier 생존 곡선을 작성하였다. 이를 통해 두 군 간 생존 양상의 시각적 차이를 확인하였다.

3) Log-Rank 검정(Log-Rank Test)

Log-Rank 검정은 Kaplan-Meier 생존 곡선을 기반으로, 두 개 이상의 집단 간 생존 곡선의 차이가 통계적으로 유의한지를 평가하는 데 사용된다. 이 검정은 각 시간 구간에서의 사건 발생률을 기대값과 비교하여 전체 생존 곡선의 차이를 누적 통계량으로 산출하며, 비례위험 가정(proportional hazards assumption)을 전제로 한다. 로그-랭크 검정은 분석이 간단하고 널리 사용되지만, 집단 간 생존 곡선이 교차하거나 위험비가 시간에 따라 변할 경우에는 가정이 위배될 수 있으므로 해석 시 주의가 필요하다(Lee SW, 2023).

본 연구에서는, Kaplan-Meier 곡선에서 관찰된 뇌졸중을 진단받은 군과 진단받지 않은 군 간 생존률 차이가 통계적으로 유의한지를 평가하기 위해 Log-Rank 검정을 수행하였다.

4) Cox 비례위험모형(Cox Proportional Hazards Model)

Cox 비례위험모형은 준모수적 생존 회귀모형으로, 시간에 따라 위험비가 일정하다는 비례위험 가정(proportional hazards assumption)을 전제로 한다. 이 모형은 생존 시간에 영향을 미치는 다양한 공변량을 통제한 상태에서 각 변수의 상대적 효과를 추정할 수 있으며, 결과로 산출되는 위험비(Hazard Ratio)는 사건 발생에 대한 상대적 위험도를 의미한다. Cox 모형은 데이터에 특정 분포를 가정하지 않아 유연하게 활용될 수 있으며, 특히 시간 고정형 변수의 효과를 평가하는 데 적합하다. 다만, 분석 대상자가 시간 경과에 따라 노출 상태가 변할 수 있는 연구에서는 보장 시간 편향(guarantee-time bias)이 발생할 수 있다. 예를 들어 중간 시점에 치료나 질병이 발생하는 경우, 해당 사건 이전에 생존한 사람만이 노출군에 포함되기 때문에, 생존 편향이 결과를 왜곡할 수 있다(Cho IS et al., 2017).

본 연구에서는 Cox 비례위험모형을 활용하여, 폐암 환자에서 뇌졸중 발생 여부가 전체 추적 관찰 기간 동안 사망 위험에 미치는 영향을 평가하였다. 이 분석에서는 시간에 따라 변하지 않는 다양한 공변량을 고정된 변수로 포함하였으며, 뇌졸중 발생 여부 역시 고정된 노출 변수로 간주하여 분석을 수행하였다. 본 모형은 시간 경과에 따른 노출 변화는 반영하지 않지만, 뇌졸중 발생 자체가 장기 생존에 미치는 전반적인 경향을 파악하기 위한 목적으로 활용되었다. 본 단락에서는 역확률치료가중법을 적용하지 않은 기본 Cox 비례위험모형(unweighted)을 제시하고, 표준 오차는 partial-likelihood 방법으로 계산하였다. IPTW가 적용된 가중 Cox 결과는 7) 역확률치료가중법 단락에 요약하였다.

5) 시간 종속 Cox 회귀모형(Time-dependent Cox Regression)

시간 종속 Cox 회귀모형은 기존 Cox 비례위험모형을 확장하여, 시간에 따라 상태가 변하는 공변량을 모델에 포함할 수 있도록 설계된 생존분석 기법이다. 특히, 연구대상자가 분석 기간 중 치료를 시작하거나, 새로운 질병에 이환되거나, 상태가 변화하는 경우와 같이 노출 시점이 개인마다 상이한 상황에서 효과적으로 활용될 수 있다. 시간 종속 Cox 모형은 보장 시간 편향을 방지할 수 있다는 점에서 관찰연구에서 매우 중요한 도구이며, 중간 사건을 적절히 반영함으로써 보다 정밀한 위험 추정이 가능하다. 시간 종속 공변량은 일반적으로 “사건 발생 이전은 0(비노출), 이후는 1(노출)”로 정의되어, 분석 대상자의 노출 상태가 시간 축에 따라 동적으로 반영된다(Cho IS et al., 2017).

본 연구에서는 뇌졸중 발생 여부를 시간 종속 공변량으로 반영하고, 이와 함께 일부 공변량 또한 시간의 흐름에 따라 변화하는 형태로 모형에 포함하여, 시간 종속 Cox 회귀모형을 구성하였다. 이를 통해 뇌졸중 발생 시점을 기준으로 한 사망 위험 변화를 동적으로 분석하였으며, 시간에 따른 노출 상태 및 공변량 구조를 보다 현실적으로 반영한 분석을 수행하였다. 시간 종속 Cox 분석은 두 단계로 수행하였다. 첫째, 역활률치료가중법을 적용하지 않은 기본(unweighted) 모델(표준 오차: partial-likelihood). 둘째, 각 시간 구간별 역활률치료가중법을 부여한 가중(weighted) 모델(표준 오차: Lin DY & Wei LJ robust, cluster = person_id). 두 모델의 결과는 모두 7) 역활률치료가중법 단락에 제시하였다.

6) 랜드마크 분석(Landmark Analysis)

랜드마크 분석은 시간 종속 공변량을 분석에 포함하기 어렵거나, 시간에 따른 노출 상태를 고정된 변수로 재정의하고자 할 때 활용되는 분석 기법이다. 분석 기준 시점(landmark time)을 사전에 설정하고, 해당 시점까지 생존한 환자들을 대상으로 이후의 생존율 또는 사건 발생 위험을 비교한다. 이 방법은 조기 사망 편향(immortal time bias)을 줄일 수 있다는 장점이 있으며, 특정 시점 이후의 치료 효과나 질병 영향 등을 조건부로 평가하는 데 적합하다. 랜드마크 분석에서는 기준 시점 이후만을 분석에 포함하기 때문에, 기준 시점 이전의 사건이나 사망은 분석 대상에서 제외된다. 이로 인해 표본 수 감소에 따른 통계적 검정력 저하가 발생할 수 있으며, 기준 시점 설정이 분석 결과에 영향을 줄 수 있으므로 임상적·통계적으로 적절한 시점 선택이 중요하다(Cho IS et al., 2017).

본 연구에서는 시간-종속 공변량을 고정된 변수로 재정의하기 위해 랜드마크 설계를 적용하였다. 최초 폐암 진단일(index date)로부터 정확히 2 년과 4 년을 각각 랜드마크 시점으로 설정하고, 해당 시점까지 생존한 환자를 분석 대상으로 포함하였다. 랜드마크 시점 이전(0 - 2 년 · 0 - 4 년) 동안 뇌졸중이 한 차례라도 발생한 환자는 뇌졸중군, 발생하지 않은 환자는 비뇌졸중군으로 분류하였다. 랜드마크 이전에 사망하거나 추적이 종료된 환자는 조기 사망 편향을 방지하기 위해 제외하였다. 뇌졸중 발생 여부와 주요 기저 공변량은 모두 랜드마크 이전 정보로만 정의하여 이후 기간 자료에는 반영하지 않았다. 노출 및 공변량의 불균형을 보정하기 위해 안정화된 역활률치료가중법을 산출한 뒤, 2 년 · 4 년 랜드마크 집단 각각에서 IPTW-가중 Cox 비례위험모형을 수행하였다.

7) 역확률치료가중법(Inverse Probability of Treatment Weighting, IPTW)

역확률치료가중법(Inverse Probability of Treatment Weighting, IPTW)은 관찰연구에서 비무작위적으로 형성된 처치군과 대조군 간의 공변량 불균형을 보정하기 위한 준실험적 분석 기법이다. 각 대상자가 특정 처치를 받을 확률(성향점수, propensity score)의 역수를 가중치로 적용함으로써, 공변량이 균형을 이루는 가상의 무작위 배정 집단을 구성하는 원리에 기반한다. 특히 IPTW는 성향점수 매칭(Propensity Score Matching, PSM)과 달리 전체 표본을 유지할 수 있어 분석의 효율성과 외적 타당성(일반화 가능성)을 동시에 확보할 수 있다는 장점을 가진다(Robins JM, Hernán MA and Brumback B, 2000; Austin PC & Stuart EA, 2015).

본 연구에서는 폐암 환자에서 뇌졸중 발생 여부가 전체 생존에 미치는 영향을 평가하기 위해 IPTW를 적용하였다. 성향점수는 기준 시점에서의 고정된 베이스라인 변수(연령, 성별, 만성질환 이력, 치료 변수 등)를 기반으로 로지스틱 회귀모형을 이용해 추정하였으며, 이를 바탕으로 안정화된 가중치(stabilized weight)를 계산하였다. 이후 Kaplan-Meier 생존 곡선, Log-Rank 검정, Cox 비례위험모형, 시간 종속 Cox 모형, 주변구조모형(MSM) 등 다양한 생존 분석에 이 가중치를 적용하였다. 또한 조기 진단으로 인한 시간 편향을 보정하기 위해 2년 및 4년 기준의 랜드마크 분석에서도 동일한 가중치를 활용하였다. IPTW 가중치는 다음과 같은 식으로 정의된다:

$$W_i(t) = \frac{P(A_i(t) = a \mid \bar{A}_i(t-1))}{P(A_i(t) = a \mid \bar{A}_i(t-1), \bar{L}_i(t))}$$

여기서 $A_i(t)$ 는 시점 t 에서 개인 i 가 뇌졸중 진단 등 노출을 받았는지를 나타내는 변수이며, $\bar{A}_i(t-1)$ 은 시점 $t-1$ 까지의 과거 노출 이력을 의미한다. 또한 $\bar{L}_i(t)$ 는 시점 t 까지의 시간 종속 공변량을 의미한다.

이러한 정의를 바탕으로 계산된 안정화된 가중치를 분석에 적용함으로써, 시간에 따라 변화하는 교란 요인의 영향을 효과적으로 최소화할 수 있다. 가중치를 적용한 이후에는 공변량 간의 균형 상태를 표준화 평균차(Standardized Mean Difference, SMD)를 통해 확인하였으며, 모든 변수에서 SMD 값이 0.1 미만으로 나타나 공변량 균형이 적절히 확보된 것으로 평가되었다. 일반적으로 IPTW 분석에서는 극단적인 가중치 값으로 인해 추정 결과가 불안정해지는 것을 방지하고자 trimming(가중치의 상·하위 일부 절단)을 적용하기도 하나, 본 연구에서는 trimming 적용 시 전체 표본 수가 급격히 감소하고 특정 주요 집단이 과소 대표될 가능성이 있어 trimming 없이 안정화된 가중치만을 사용하였다. 생존 분석에서는 R의 survival 패키지 내 coxph() 함수를 활용하였으며, robust = TRUE 및 cluster = person_id 옵션을 적용하여 Lin DY & Wei LJ의 robust 분산 추정법을 사용해 표준 오차를 산출하였다. 이러한 분석 절차를 통해 IPTW 기반의 비교 가능한 집단을 구성함으로써, 관찰연구에서 발생할 수 있는 선택 편향을 최소화하고, 뇌졸중 발생 여부가 사망에 미치는 인과적 영향을 보다 신뢰성 있게 추정할 수 있었다.

8) 주변구조모형(Marginal Structural Model, MSM)

주변구조모형은 시간에 따라 변화하는 노출 변수와 공변량이 존재하는 관찰연구에서, 기존의 회귀모형으로는 적절히 추정하기 어려운 인과효과를 추정하기 위해 고안된 분석 방법이다. 특히 노출 변수와 시간 종속 공변량 사이에 교호작용이나 시간-의존성 편향이 존재할 경우, 전통적인 회귀모형은 이러한 편향을 효과적으로 보정하지 못해 왜곡된 추론에 이를 수 있다. MSM은 이와 같은 문제를 해결하기 위해, 시간 종속 공변량에 대한 영향을 통제하는 대신, 성향점수를 기반으로 한 역처리확률가중치법을 적용함으로써, 관찰자료를 마치 무작위 배정된 자료처럼 재구성하는 방식으로 접근한다. 즉, 시간 종속 공변량이 노출과 결과 모두에 영향을 주는 경우에도, IPTW를 통해 해당 공변량의 영향을 간접적으로 조정하고, 이를 기반으로 한 모형에서 노출 효과를 추정할 수 있다(Robins JM, Hernán MA and Brumback B, 2000).

본 연구에서는 뇌졸중 발생 이후의 사망 위험을 시간 흐름에 따라 보다 정밀하게 평가하고자 MSM을 활용하였다. 안정화된 IPTW를 기반으로 Cox 회귀모형의 로그 우도 함수를 시간 종속 형태로 구성하였으며, 다음과 같이 정의된다:

$$\mathcal{L}(\beta) = \prod_{i=1}^N \prod_t \left\{ \frac{W_i(t) \exp(\beta A_i(t))}{\sum_{j=1}^N W_j(t) \exp(\beta A_j(t))} \right\}^{dN_i(t)}$$

여기서 $A_i(t)$ 는 시점 t 에서 개인 i 가 노출(예: 뇌졸중 진단)을 받았는지를 나타내는 변수이며, $dN_i(t)$ 는 시점 t 에서 사건(예: 사망)이 발생했는지를 나타내는 지시 변수이다. 또한 $W_i(t)$ 는 해당 시점에서의 안정화된 IPTW 가중치를 의미한다. 본 연구의 생존 분석은 R의 survival 패키지 내 `coxph()` 함수에서 `robust = TRUE` 및 `cluster = person_id` 옵션을 적용하여, Lin DY & Wei LJ의 robust 분산 추정법을 사용해 표준 오차를 계산하였다. 이러한 분석

절차를 통해 IPTW 기반의 비교 가능한 집단을 구성함으로써, 관찰연구에서 발생할 수 있는 선택 편향을 최소화하고, 뇌졸중 발생 여부가 사망에 미치는 인과적 영향을 보다 신뢰성 있게 추정할 수 있었다.

본 연구의 모든 통계 분석은 R 소프트웨어(version 4.4.1, R Foundation for Statistical Computing, Vienna, Austria)를 사용하여 수행하였다. R은 다양한 통계 기법을 유연하게 구현할 수 있는 오픈소스 프로그래밍 언어로, 재현성과 확장성이 뛰어나 연구 및 학술분야에서 널리 활용되고 있다. 모든 분석은 유의수준 0.05에서 시행하였으며, p-value가 0.05 미만인 경우를 통계적으로 유의한 결과로 해석하였다.

Ⅲ. 연구 결과

1. 연구대상자의 일반적 특성

1.1. 전체 연구대상자의 IPTW 적용 전, 일반적 특성

연구대상자는 총 20,696명이었으며, 이 중 뇌졸중을 진단받지 않은 군은 20,259명(97.89%), 뇌졸중을 진단받은 군은 437명(2.11%)이었다. 연구대상자의 평균 추적관찰 기간은 7.04 ± 5.40 년이었으며, 뇌졸중을 진단받지 않은 군은 7.05 ± 5.40 년, 뇌졸중을 진단받은 군은 6.89 ± 5.30 년이었다(그림 5).

연령 분포는 뇌졸중을 진단받지 않은 군에서 50세 미만은 1,862명(9.19%), 50세 이상 60세 미만은 4,059명(20.04%), 60세 이상 70세 미만은 6,595명(32.55%), 70세 이상 80세 미만은 6,060명(29.91%), 80세 이상은 1,683명(8.31%)으로 나타났다. 뇌졸중을 진단받은 군에서는 50세 미만은 35명(8.01%), 50세 이상 60세 미만은 87명(19.91%), 60세 이상 70세 미만은 131명(29.98%), 70세 이상 80세 미만은 144명(32.95%), 80세 미만은 40명(9.15%)으로 나타났으며 이는 통계적으로 유의한 차이는 없었다($p=0.5225$).

성별은 뇌졸중을 진단받지 않은 군에서 남자는 13,309명(65.69%), 여자는 6,950명(34.31%)이었고 뇌졸중을 진단받은 군에서 남자는 295명(67.51%), 여자는 142명(32.49%)으로 통계적으로 유의한 차이는 없었다($p=0.4602$).

뇌졸중을 진단받지 않은 군에서 고혈압 여부에 대해 없음은 13,595명(67.11%), 고혈압 여부에 대해 있음은 6,664명(32.89%)이었고 뇌졸중을 진단받은 군에서 고혈압 여부에 대해 없음은 199명(45.54%), 고혈압 여부에 대해 있음은 238명(54.46%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$).

뇌졸중을 진단받지 않은 군에서 당뇨 여부에 대해 없음은 18,225명(89.96%), 당뇨 여부에 대해 있음은 2,034명(10.04%)이었고 뇌졸중을 진단받은 군에서 당뇨 여부에 대해 없음은 349명(79.86%), 당뇨 여부에 대해 있음은 88명(20.14%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$).

뇌졸중을 진단받지 않은 군에서 고지혈증 여부에 대해 없음은 18,900명(93.29%), 고지혈증 여부에 대해 있음은 1,359명(6.71%)이었고 뇌졸중을 진단받은 군에서 고지혈증 여부에 대해 없음은 369명(84.44%), 고지혈증 여부에 대해 있음은 68명(15.56%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$).

뇌졸중을 진단받지 않은 군에서 수술 여부에 대해 없음은 14,456명(71.36%), 수술 여부에 대해 있음은 5,803명(28.64%)이었고 뇌졸중 진단을 받은 군에서 수술 여부에 대해 없음은 272명(62.24%), 수술 여부에 대해 있음은 165명(37.76%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$).

뇌졸중을 진단받지 않은 군에서 방사선 치료 여부에 대해 없음은 12,838명(63.37%), 방사선 치료 여부에 대해 있음은 7,421명(36.63%)이었고 뇌졸중 진단을 받은 군에서 방사선 치료 여부에 대해 없음은 200명(45.77%), 방사선 치료 여부에 대해 있음은 237명(54.23%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$, 표 9).

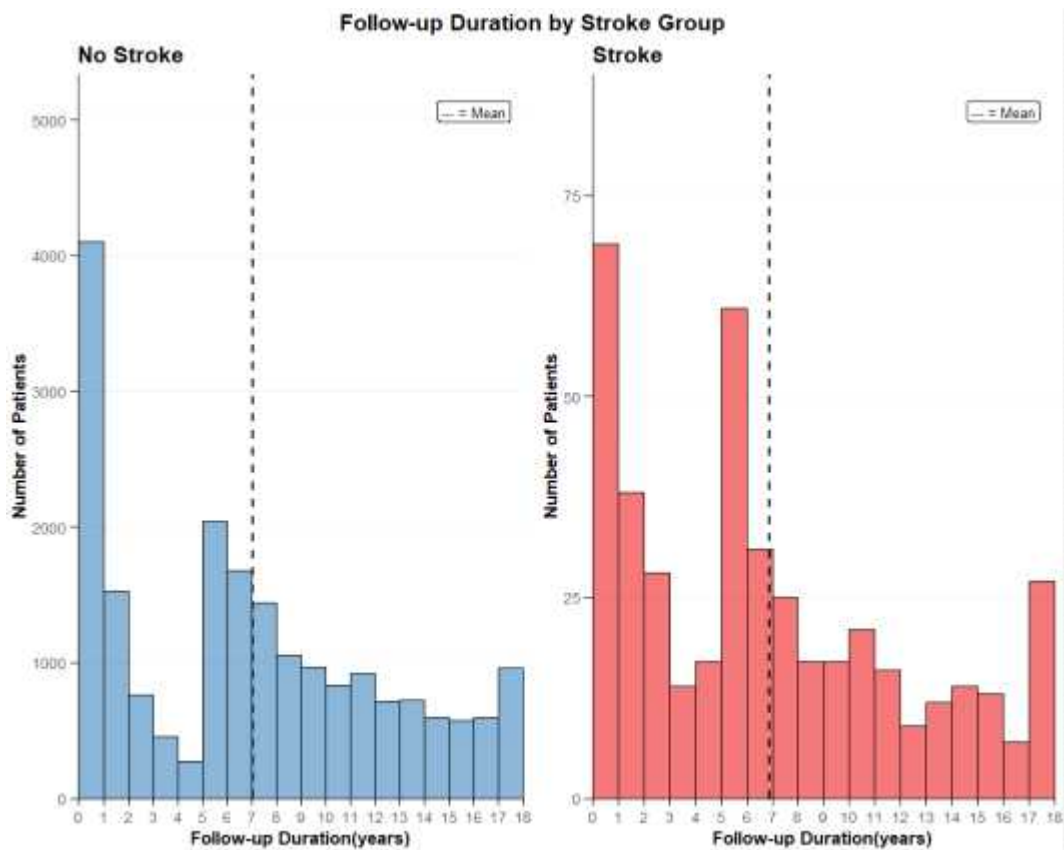


그림 5. 전체 연구대상자의 뇌졸중 진단 여부에 따른 추적관찰 기간 분포

표 9. 전체 연구대상자의 IPTW 적용 전, 일반적 특성

	뇌졸중을 진단받지 않은 군 N = 20,259 (97.89%) N(%)	뇌졸중을 진단받은 군 N = 437 (2.11%) N(%)	p-value	SMD
연령				
50세 미만	1,862(9.19)	35(8.01)	0.5225	0.030
50세 이상 60세 미만	4,059(20.04)	87(19.91)		
60세 이상 70세 미만	6,595(32.55)	131(29.98)		
70세 이상 80세 미만	6,060(29.91)	144(32.95)		
80세 이상	1,683(8.31)	40(9.15)		
성별				
남	13,309(65.69)	295(67.51)	0.4602	0.018
여	6,950(34.31)	142(32.49)		
고혈압 유무				
없음	13,595(67.11)	199(45.54)	<.0001	0.216
있음	6,664(32.89)	238(54.46)		
당뇨 유무				
없음	18,225(89.96)	349(79.86)	<.0001	0.101
있음	2,034(10.04)	88(20.14)		
고지혈증 유무				
없음	18,900(93.29)	369(84.44)	<.0001	0.089
있음	1,359(6.71)	68(15.56)		
수술 여부				
없음	14,456(71.36)	272(62.24)	<.0001	0.091
있음	5,803(28.64)	165(37.76)		
방사선 치료 여부				
없음	12,838(63.37)	200(45.77)	<.0001	0.176
있음	7,421(36.63)	237(54.23)		

1.2. 전체 연구대상자의 IPTW 적용 후, 일반적 특성

IPTW를 적용하여 연구대상자 20,696명을 분석하였다. 공변량 균형 상태를 평가하기 위해 변수 간 표준화 평균차(Standardized Mean Difference, SMD)를 확인한 결과, 모든 변수의 SMD 값이 0.1 미만으로 나타나 뇌졸중 진단 유무에 따른 두 집단 간 기저 특성이 잘 균형을 이룬 것으로 확인되었다. 이러한 공변량 균형은 러브플롯(Love plot)을 통해 시각적으로 제시하였으며, IPTW 적용 전후의 SMD 변화가 명확히 나타났다(그림 6).

연령 분포는 뇌졸중을 진단받지 않은 군에서 50세 미만은 1,862명(9.19%), 50세 이상 60세 미만은 4,059명(20.04%), 60세 이상 70세 미만은 6,595명(32.55%), 70세 이상 80세 미만은 6,060명(29.91%), 80세 이상은 1,683명(8.31%)으로 나타났다. 뇌졸중을 진단받은 군에서는 50세 미만은 35명(8.01%), 50세 이상 60세 미만은 87명(19.91%), 60세 이상 70세 미만은 131명(29.98%), 70세 이상 80세 미만은 144명(32.95%), 80세 미만은 40명(9.15%)으로 나타났으며 이는 통계적으로 유의한 차이는 없었다($p=0.5418$).

성별은 뇌졸중을 진단받지 않은 군에서 남자는 13,309명(65.69%), 여자는 6,950명(34.31%)이었고 뇌졸중을 진단받은 군에서 남자는 295명(67.51%), 여자는 142명(32.49%)으로 통계적으로 유의한 차이는 없었다($p=0.2872$).

뇌졸중을 진단받지 않은 군에서 고혈압 여부에 대해 없음은 13,595명(67.11%), 고혈압 여부에 대해 있음은 6,664명(32.89%)이었고 뇌졸중을 진단받은 군에서 고혈압 여부에 대해 없음은 199명(45.54%), 고혈압 여부에 대해 있음은 238명(54.46%)으로 통계적으로 유의한 차이를 보였다($p=0.0377$).

뇌졸중을 진단받지 않은 군에서 당뇨 여부에 대해 없음은 18,225명(89.96%), 당뇨 여부에 대해 있음은 2,034명(10.04%)이었고 뇌졸중을 진단받은 군에서 당뇨 여부에 대해 없음은 349명(79.86%), 당뇨 여부에 대해 있음은 88명(20.14%)으로 통계적으로 유의한 차이는 없었다($p=0.1357$).

뇌졸중을 진단받지 않은 군에서 고지혈증 여부에 대해 없음은 18,900명(93.29%), 고지혈증 여부에 대해 있음은 1,359명(6.71%)이었고 뇌졸중을 진단받은 군에서 고지혈증 여부에 대해 없음은 369명(84.44%), 고지혈증 여부에 대해 있음은 68명(15.56%)으로 통계적으로 유의한 차이는 없었다($p=0.5689$).

뇌졸중을 진단받지 않은 군에서 수술 여부에 대해 없음은 14,456명(71.36%), 수술 여부에 대해 있음은 5,803명(28.64%)이었고 뇌졸중 진단을 받은 군에서 수술 여부에 대해 없음은 272명(62.24%), 수술 여부에 대해 있음은 165명(37.76%)으로 통계적으로 유의한 차이는 없었다($p=0.4575$).

뇌졸중을 진단받지 않은 군에서 방사선 치료 여부에 대해 없음은 12,838명(63.37%), 방사선 치료 여부에 대해 있음은 7,421명(36.63%)이었고 뇌졸중 진단을 받은 군에서 방사선 치료 여부에 대해 없음은 200명(45.77%), 방사선 치료 여부에 대해 있음은 237명(54.23%)으로 통계적으로 유의한 차이는 없었다($p=0.0565$, 표 10).

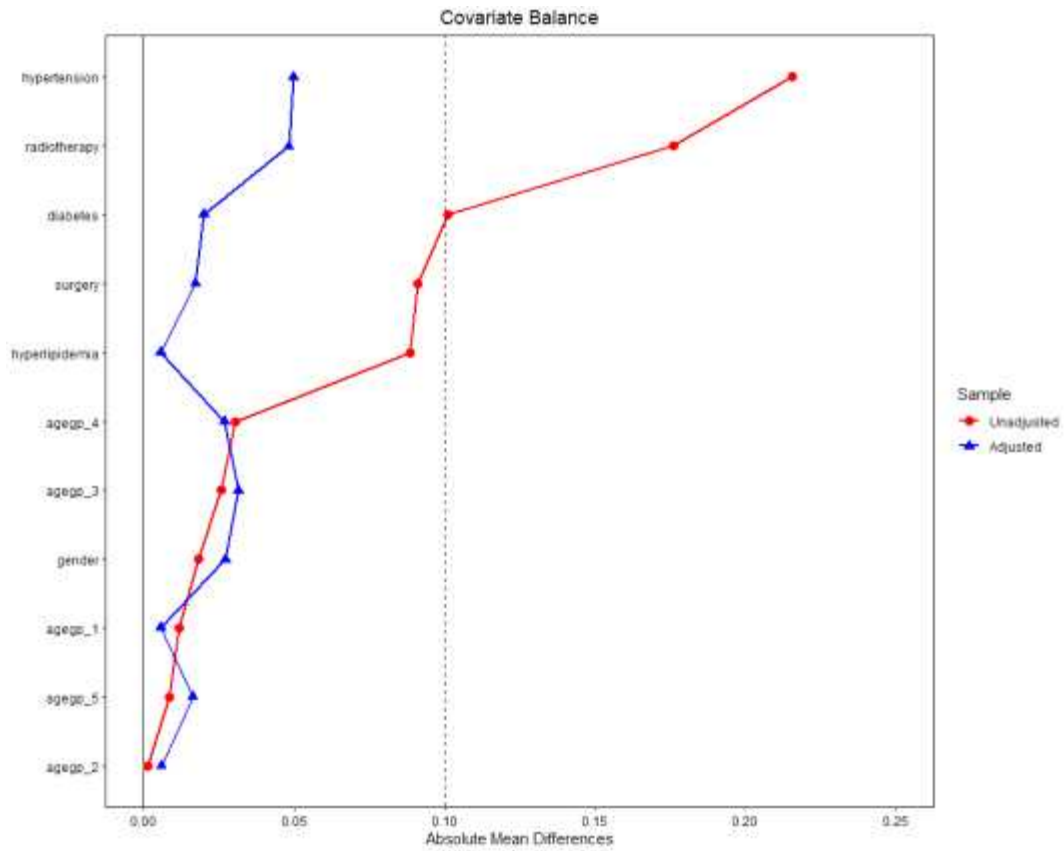


그림 6. 전체 연구대상자의 IPTW 적용 전후 공변량 간 Standardized Mean Difference

표 10. 전체 연구대상자의 IPTW 적용 후, 일반적 특성

	뇌졸중을 진단받지 않은 군 N = 20,259 (97.89%) N(%)	뇌졸중을 진단받은 군 N = 437 (2.11%) N(%)	p-value	SMD
연령				
50세 미만	1,862(9.19)	35(8.01)	0.5418	0.031
50세 이상 60세 미만	4,059(20.04)	87(19.91)		
60세 이상 70세 미만	6,595(32.55)	131(29.98)		
70세 이상 80세 미만	6,060(29.91)	144(32.95)		
80세 이상	1,683(8.31)	40(9.15)		
성별				
남	13,309(65.69)	295(67.51)	0.2872	0.027
여	6,950(34.31)	142(32.49)		
고혈압 유무				
없음	13,595(67.11)	199(45.54)	0.0377	0.050
있음	6,664(32.89)	238(54.46)		
당뇨 유무				
없음	18,225(89.96)	349(79.86)	0.1357	0.020
있음	2,034(10.04)	88(20.14)		
고지혈증 유무				
없음	18,900(93.29)	369(84.44)	0.5689	0.006
있음	1,359(6.71)	68(15.56)		
수술 여부				
없음	14,456(71.36)	272(62.24)	0.4575	0.017
있음	5,803(28.64)	165(37.76)		
방사선 치료 여부				
없음	12,838(63.37)	200(45.77)	0.0565	0.048
있음	7,421(36.63)	237(54.23)		

1.3. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 일반적 특성

연구대상자는 총 14,964명이었으며, 이 중 뇌졸중을 진단받지 않은 군은 14,808명(98.96%), 뇌졸중을 진단받은 군은 156명(1.04%)이었다. 연구대상자의 평균 추적관찰 기간은 9.48 ± 4.33 년이었으며, 뇌졸중을 진단받지 않은 군은 9.49 ± 4.32 년, 뇌졸중을 진단받은 군은 8.31 ± 4.54 년이였다(그림 7).

연령 분포는 뇌졸중을 진단받지 않은 군에서 50세 미만은 1,487명(10.04%), 50세 이상 60세 미만은 3,218명(21.73%), 60세 이상 70세 미만은 4,883명(32.98%), 70세 이상 80세 미만은 4,135명(27.92%), 80세 이상은 1,085명(7.33%)으로 나타났다. 뇌졸중을 진단받은 군에서는 50세 미만은 10명(6.41%), 50세 이상 60세 미만은 27명(17.31%), 60세 이상 70세 미만은 49명(31.41%), 70세 이상 80세 미만은 52명(33.33%), 80세 미만은 18명(11.54%)으로 나타났으며 이는 통계적으로 유의한 차이는 없었다($p=0.0644$).

성별은 뇌졸중을 진단받지 않은 군에서 남자는 9,152명(61.80%), 여자는 5,656명(38.20%)이었고 뇌졸중을 진단받은 군에서 남자는 111명(71.15%), 여자는 45명(28.85%)으로 통계적으로 유의한 차이를 보였다($p=0.0209$).

뇌졸중을 진단받지 않은 군에서 고혈압 여부에 대해 없음은 10,755명(72.63%), 고혈압 여부에 대해 있음은 4,053명(27.37%)이었고 뇌졸중을 진단받은 군에서 고혈압 여부에 대해 없음은 80명(51.28%), 고혈압 여부에 대해 있음은 76명(48.72%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$).

뇌졸중을 진단받지 않은 군에서 당뇨 여부에 대해 없음은 13,647명(92.16%), 당뇨 여부에 대해 있음은 1,161명(7.84%)이었고 뇌졸중을 진단받은 군에서 당뇨 여부에 대해 없음은 134명(85.90%), 당뇨 여부에 대해 있음은 22명(14.10%)으로 통계적으로 유의한 차이를 보였다($p=0.0062$).

뇌졸중을 진단받지 않은 군에서 고지혈증 여부에 대해 없음은 14,012명(94.62%), 고지혈증 여부에 대해 있음은 796명(5.38%)이었고 뇌졸중을 진단받은 군에서 고지혈증 여부에 대해 없음은 140명(89.74%), 고지혈증 여부에 대해 있음은 16명(10.26%)으로 통계적으로 유의한 차이를 보였다($p=0.0124$).

뇌졸중을 진단받지 않은 군에서 수술 여부에 대해 없음은 9,764명(65.94%), 수술 여부에 대해 있음은 5,044명(34.06%)이었고 뇌졸중 진단을 받은 군에서 수술 여부에 대해 없음은 104명(66.67%), 수술 여부에 대해 있음은 52명(33.33%)으로 통계적으로 유의한 차이는 없었다($p=0.9153$).

뇌졸중을 진단받지 않은 군에서 방사선 치료 여부에 대해 없음은 11,010명(74.35%), 방사선 치료 여부에 대해 있음은 3,798명(25.65%)이었고 뇌졸중 진단을 받은 군에서 방사선 치료 여부에 대해 없음은 84명(53.85%), 방사선 치료 여부에 대해 있음은 72명(46.15%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$, 표 11).

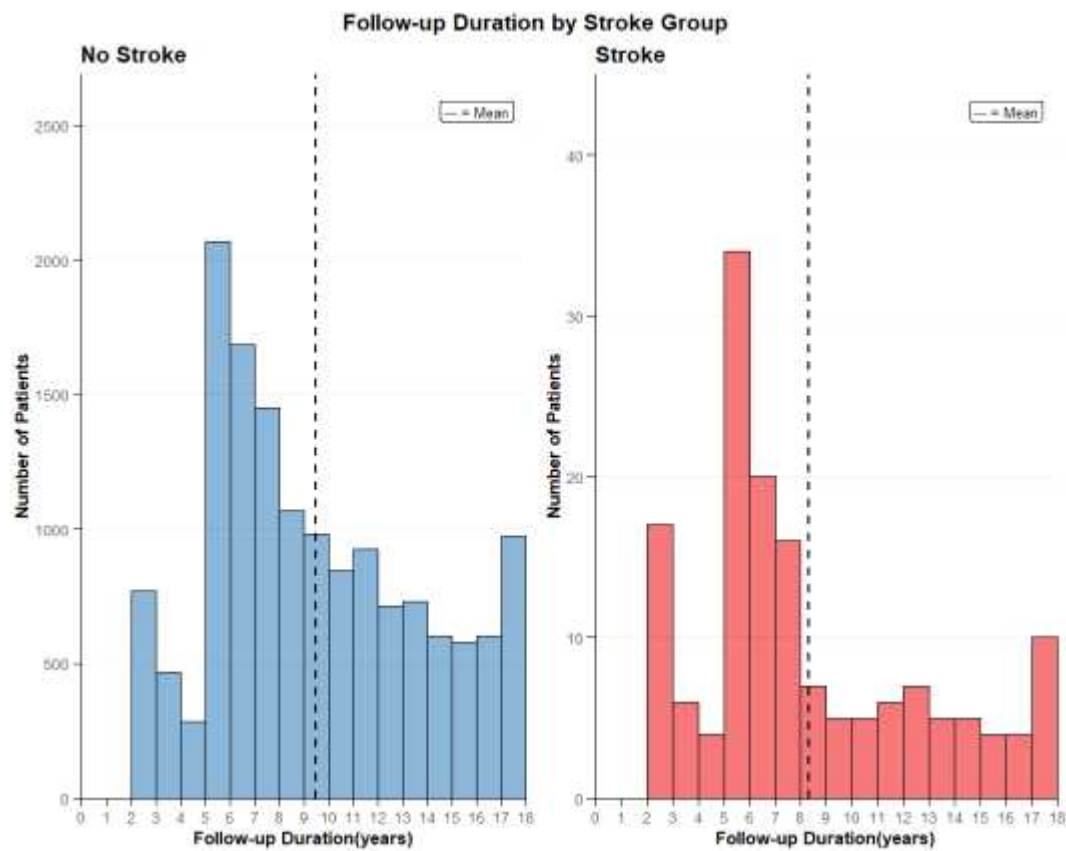


그림 7. 랜드마크 기간 2년 연구대상자의 뇌졸중 진단 여부에 따른 추적관찰 기간 분포

표 11. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 일반적 특성

	뇌졸중을		뇌졸중을		p-value	SMD
	진단받지 않은 군		진단받은 군			
	N = 14,808		N = 156			
	(98.96%)		(1.04%)			
	N(%)		N(%)			
연령						
50세 미만	1,487(10.04)		10(6.41)		0.0644	0.054
50세 이상 60세 미만	3,218(21.73)		27(17.31)			
60세 이상 70세 미만	4,883(32.98)		49(31.41)			
70세 이상 80세 미만	4,135(27.92)		52(33.33)			
80세 이상	1,085(7.33)		18(11.54)			
성별						
남	9,152(61.80)		111(71.15)		0.0209	0.093
여	5,656(38.20)		45(28.85)			
고혈압 유무						
없음	10,755(72.63)		80(51.28)		<.0001	0.213
있음	4,053(27.37)		76(48.72)			
당뇨 유무						
없음	13,647(92.16)		134(85.90)		0.0062	0.063
있음	1,161(7.84)		22(14.10)			
고지혈증 유무						
없음	14,012(94.62)		140(89.74)		0.0124	0.049
있음	796(5.38)		16(10.26)			
수술 여부						
없음	9,764(65.94)		104(66.67)		0.9153	0.007
있음	5,044(34.06)		52(33.33)			
방사선 치료 여부						
없음	11,010(74.35)		84(53.85)		<.0001	0.205
있음	3,798(25.65)		72(46.15)			

1.4. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 일반적 특성

IPTW를 적용하여 연구대상자 14,964명을 분석하였다. 공변량 균형 상태를 평가하기 위해 변수 간 SMD를 확인한 결과, 모든 변수의 SMD 값이 0.1 미만으로 나타나 뇌졸중 진단 유무에 따른 두 집단 간 기저 특성이 잘 균형을 이룬 것으로 확인되었다. 이러한 공변량 균형은 러브플롯을 통해 시각적으로 제시하였으며, IPTW 적용 전후의 SMD 변화가 명확히 나타났다(그림 8).

연령 분포는 뇌졸중을 진단받지 않은 군에서 50세 미만은 1,487명(10.04%), 50세 이상 60세 미만은 3,218명(21.73%), 60세 이상 70세 미만은 4,883명(32.98%), 70세 이상 80세 미만은 4,135명(27.92%), 80세 이상은 1,085명(7.33%)으로 나타났다. 뇌졸중을 진단받은 군에서는 50세 미만은 10명(6.41%), 50세 이상 60세 미만은 27명(17.31%), 60세 이상 70세 미만은 49명(31.41%), 70세 이상 80세 미만은 52명(33.33%), 80세 미만은 18명(11.54%)으로 나타났으며 이는 통계적으로 유의한 차이는 없었다($p=0.5306$).

성별은 뇌졸중을 진단받지 않은 군에서 남자는 9,152명(61.80%), 여자는 5,656명(38.20%)이었고 뇌졸중을 진단받은 군에서 남자는 111명(71.15%), 여자는 45명(28.85%)으로 통계적으로 유의한 차이는 없었다($p=0.2405$).

뇌졸중을 진단받지 않은 군에서 고혈압 여부에 대해 없음은 10,755명(72.63%), 고혈압 여부에 대해 있음은 4,053명(27.37%)이었고 뇌졸중을 진단받은 군에서 고혈압 여부에 대해 없음은 80명(51.28%), 고혈압 여부에 대해 있음은 76명(48.72%)으로 통계적으로 유의한 차이는 없었다($p=0.0737$).

뇌졸중을 진단받지 않은 군에서 당뇨 여부에 대해 없음은 13,647명(92.16%), 당뇨 여부에 대해 있음은 1,161명(7.84%)이었고 뇌졸중을 진단받은 군에서 당뇨 여부에 대해 없음은 134명(85.90%), 당뇨 여부에 대해 있음은 22명(14.10%)으로 통계적으로 유의한 차이는 없었다($p=0.0736$).

뇌졸중을 진단받지 않은 군에서 고지혈증 여부에 대해 없음은 14,012명(94.62%), 고지혈증 여부에 대해 있음은 796명(5.38%)이었고 뇌졸중을 진단받은 군에서 고지혈증 여부에 대해 없음은 140명(89.74%), 고지혈증 여부에 대해 있음은 16명(10.26%)으로 통계적으로 유의한 차이는 없었다($p=0.4997$).

뇌졸중을 진단받지 않은 군에서 수술 여부에 대해 없음은 9,764명(65.94%), 수술 여부에 대해 있음은 5,044명(34.06%)이었고 뇌졸중 진단을 받은 군에서 수술 여부에 대해 없음은 104명(66.67%), 수술 여부에 대해 있음은 52명(33.33%)으로 통계적으로 유의한 차이는 없었다($p=0.4806$).

뇌졸중을 진단받지 않은 군에서 방사선 치료 여부에 대해 없음은 11,010명(74.35%), 방사선 치료 여부에 대해 있음은 3,798명(25.65%)이었고 뇌졸중 진단을 받은 군에서 방사선 치료 여부에 대해 없음은 84명(53.85%), 방사선 치료 여부에 대해 있음은 72명(46.15%)으로 통계적으로 유의한 차이를 보였다($p=0.0244$, 표 12).

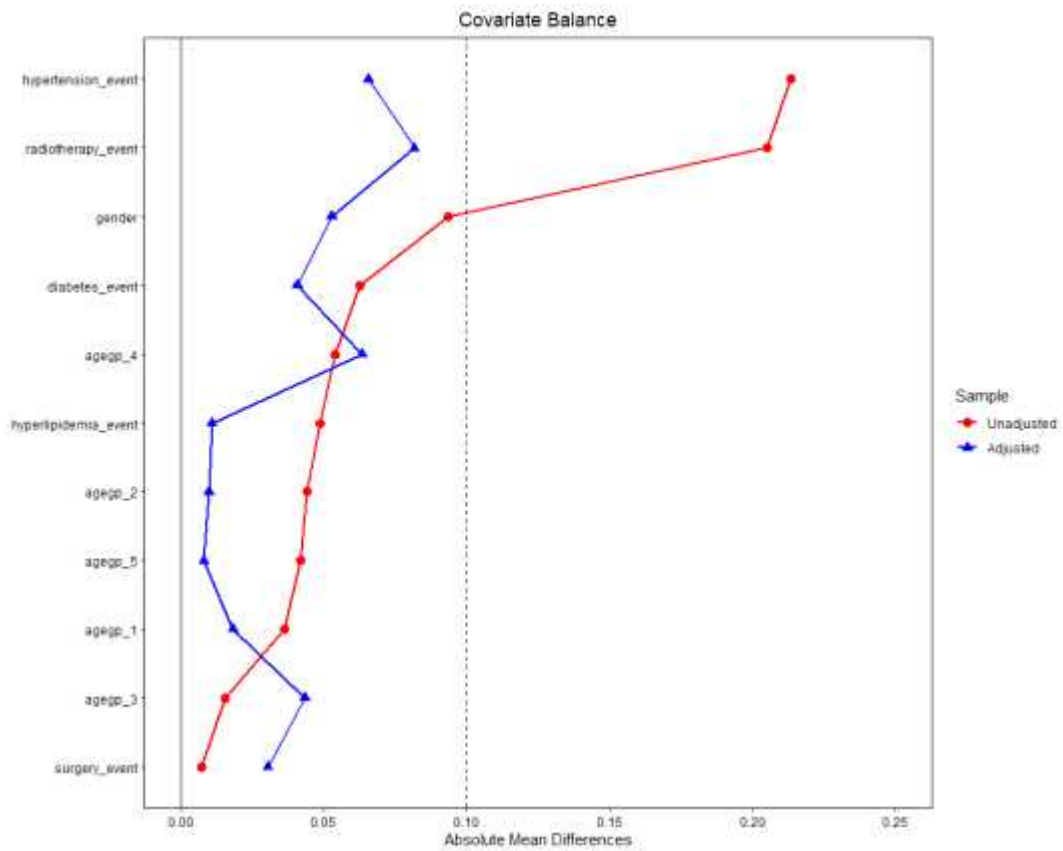


그림 8. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전후 공변량 간 Standardized Mean Difference

표 12. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 일반적 특성

	뇌졸중을 진단받지 않은 군 N = 14,808 (98.96%) N(%)	뇌졸중을 진단받은 군 N = 156 (1.04%) N(%)	p-value	SMD
연령				
50세 미만	1,487(10.04)	10(6.41)	0.5306	0.064
50세 이상 60세 미만	3,218(21.73)	27(17.31)		
60세 이상 70세 미만	4,883(32.98)	49(31.41)		
70세 이상 80세 미만	4,135(27.92)	52(33.33)		
80세 이상	1,085(7.33)	18(11.54)		
성별				
남	9,152(61.80)	111(71.15)	0.2405	0.053
여	5,656(38.20)	45(28.85)		
고혈압 유무				
없음	10,755(72.63)	80(51.28)	0.0737	0.066
있음	4,053(27.37)	76(48.72)		
당뇨 유무				
없음	13,647(92.16)	134(85.90)	0.0736	0.041
있음	1,161(7.84)	22(14.10)		
고지혈증 유무				
없음	14,012(94.62)	140(89.74)	0.4997	0.011
있음	796(5.38)	16(10.26)		
수술 여부				
없음	9,764(65.94)	104(66.67)	0.4806	0.031
있음	5,044(34.06)	52(33.33)		
방사선 치료 여부				
없음	11,010(74.35)	84(53.85)	0.0244	0.082
있음	3,798(25.65)	72(46.15)		

1.5. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 일반적 특성

연구대상자는 총 13,704명이었으며, 이 중 뇌졸중을 진단받지 않은 군은 13,523명(98.68%), 뇌졸중을 진단받은 군은 181명(1.32%)이었다. 연구대상자의 평균 추적관찰 기간은 10.09 ± 4.00 년이었으며, 뇌졸중을 진단받지 않은 군은 10.01 ± 4.00 년, 뇌졸중을 진단받은 군은 9.19 ± 4.17 년이었다(그림 9).

연령 분포는 뇌졸중을 진단받지 않은 군에서 50세 미만은 1,328명(9.82%), 50세 이상 60세 미만은 2,901명(21.45%), 60세 이상 70세 미만은 4,461명(32.99%), 70세 이상 80세 미만은 3,822명(28.26%), 80세 이상은 1,011명(7.48%)으로 나타났다. 뇌졸중을 진단받은 군에서는 50세 미만은 14명(7.74%), 50세 이상 60세 미만은 38명(20.99%), 60세 이상 70세 미만은 53명(29.28%), 70세 이상 80세 미만은 57명(31.49%), 80세 미만은 19명(10.50%)으로 나타났으며 이는 통계적으로 유의한 차이는 없었다($p=0.3567$).

성별은 뇌졸중을 진단받지 않은 군에서 남자는 8,332명(61.61%), 여자는 5,191명(38.39%)이었고 뇌졸중을 진단받은 군에서 남자는 129명(71.27%), 여자는 52명(28.73%)으로 통계적으로 유의한 차이를 보였다($p=0.0099$).

뇌졸중을 진단받지 않은 군에서 고혈압 여부에 대해 없음은 9,699명(71.72%), 고혈압 여부에 대해 있음은 3,824명(28.28%)이었고 뇌졸중을 진단받은 군에서 고혈압 여부에 대해 없음은 85명(46.96%), 고혈압 여부에 대해 있음은 96명(53.04%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$).

뇌졸중을 진단받지 않은 군에서 당뇨 여부에 대해 없음은 12,376명(91.52%), 당뇨 여부에 대해 있음은 1,147명(8.48%)이었고 뇌졸중을 진단받은 군에서 당뇨 여부에 대해 없음은 154명(85.08%), 당뇨 여부에 대해 있음은 27명(14.92%)으로 통계적으로 유의한 차이를 보였다($p=0.0033$).

뇌졸중을 진단받지 않은 군에서 고지혈증 여부에 대해 없음은 12,685명(93.80%), 고지혈증 여부에 대해 있음은 838명(6.20%)이었고 뇌졸중을 진단받은 군에서 고지혈증 여부에 대해 없음은 157명(86.74%), 고지혈증 여부에 대해 있음은 24명(13.26%)으로 통계적으로 유의한 차이를 보였다($p=0.0002$).

뇌졸중을 진단받지 않은 군에서 수술 여부에 대해 없음은 8,762명(64.79%), 수술 여부에 대해 있음은 4,761명(35.21%)이었고 뇌졸중 진단을 받은 군에서 수술 여부에 대해 없음은 117명(64.64%), 수술 여부에 대해 있음은 64명(35.36%)으로 통계적으로 유의한 차이는 없었다($p=0.9978$).

뇌졸중을 진단받지 않은 군에서 방사선 치료 여부에 대해 없음은 10,080명(74.54%), 방사선 치료 여부에 대해 있음은 3,443명(25.46%)이었고 뇌졸중 진단을 받은 군에서 방사선 치료 여부에 대해 없음은 97명(53.59%), 방사선 치료 여부에 대해 있음은 84명(46.41%)으로 통계적으로 유의한 차이를 보였다($p<0.0001$, 표 13).

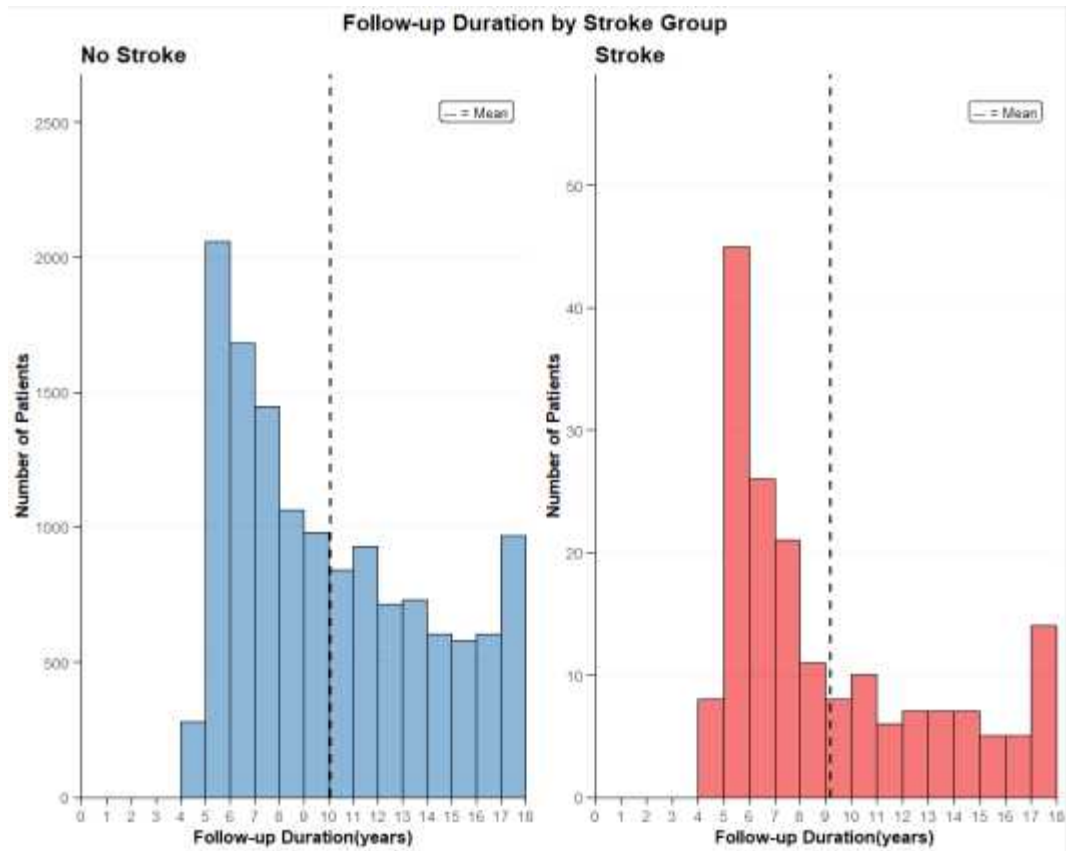


그림 9. 랜드마크 기간 4년 연구대상자의 뇌졸중 진단 여부에 따른 추적관찰 기간 분포

표 13. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 일반적 특성

	뇌졸중을		p-value	SMD
	진단받지 않은 군			
	N = 13,523			
	(98.68%)			
	N(%)	N(%)		
연령				
50세 미만	1,328(9.82)	14(7.74)	0.3567	0.037
50세 이상 60세 미만	2,901(21.45)	38(20.99)		
60세 이상 70세 미만	4,461(32.99)	53(29.28)		
70세 이상 80세 미만	3,822(28.26)	57(31.49)		
80세 이상	1,011(7.48)	19(10.50)		
성별				
남	8,332(61.61)	129(71.27)	0.0099	0.097
여	5,191(38.39)	52(28.73)		
고혈압 유무				
없음	9,699(71.72)	85(46.96)	<.0001	0.248
있음	3,824(28.28)	96(53.04)		
당뇨 유무				
없음	12,376(91.52)	154(85.08)	0.0033	0.064
있음	1,147(8.48)	27(14.92)		
고지혈증 유무				
없음	12,685(93.80)	157(86.74)	0.0002	0.071
있음	838(6.20)	24(13.26)		
수술 여부				
없음	8,762(64.79)	117(64.64)	0.9978	0.002
있음	4,761(35.21)	64(35.36)		
방사선 치료 여부				
없음	10,080(74.54)	97(53.59)	<.0001	0.209
있음	3,443(25.46)	84(46.41)		

1.6. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 일반적 특성

IPTW를 적용하여 연구대상자 13,704명을 분석하였다. 공변량 균형 상태를 평가하기 위해 변수 간 SMD를 확인한 결과, 모든 변수의 SMD 값이 0.1 미만으로 나타나 뇌졸중 진단 유무에 따른 두 집단 간 기저 특성이 잘 균형을 이룬 것으로 확인되었다. 이러한 공변량 균형은 러브플롯을 통해 시각적으로 제시하였으며, IPTW 적용 전후의 SMD 변화가 명확히 나타났다(그림 10).

연령 분포는 뇌졸중을 진단받지 않은 군에서 50세 미만은 1,328명(9.82%), 50세 이상 60세 미만은 2,901명(21.45%), 60세 이상 70세 미만은 4,461명(32.99%), 70세 이상 80세 미만은 3,822명(28.26%), 80세 이상은 1,011명(7.48%)으로 나타났다. 뇌졸중을 진단받은 군에서는 50세 미만은 14명(7.74%), 50세 이상 60세 미만은 38명(20.99%), 60세 이상 70세 미만은 53명(29.28%), 70세 이상 80세 미만은 57명(31.49%), 80세 미만은 19명(10.50%)으로 나타났으며 이는 통계적으로 유의한 차이는 없었다($p=0.5545$).

성별은 뇌졸중을 진단받지 않은 군에서 남자는 8,332명(61.61%), 여자는 5,191명(38.39%)이었고 뇌졸중을 진단받은 군에서 남자는 129명(71.27%), 여자는 52명(28.73%)으로 통계적으로 유의한 차이는 없었다($p=0.2094$).

뇌졸중을 진단받지 않은 군에서 고혈압 여부에 대해 없음은 9,699명(71.72%), 고혈압 여부에 대해 있음은 3,824명(28.28%)이었고 뇌졸중을 진단받은 군에서 고혈압 여부에 대해 없음은 85명(46.96%), 고혈압 여부에 대해 있음은 96명(53.04%)으로 통계적으로 유의한 차이는 없었다($p=0.0857$).

뇌졸중을 진단받지 않은 군에서 당뇨 여부에 대해 없음은 12,376명(91.52%), 당뇨 여부에 대해 있음은 1,147명(8.48%)이었고 뇌졸중을 진단받은 군에서 당뇨 여부에 대해 없음은 154명(85.08%), 당뇨 여부에 대해 있음은 27명(14.92%)으로 통계적으로 유의한 차이를 보였다($p=0.0303$).

뇌졸중을 진단받지 않은 군에서 고지혈증 여부에 대해 없음은 12,685명(93.80%), 고지혈증 여부에 대해 있음은 838명(6.20%)이었고 뇌졸중을 진단받은 군에서 고지혈증 여부에 대해 없음은 157명(86.74%), 고지혈증 여부에 대해 있음은 24명(13.26%)으로 통계적으로 유의한 차이는 없었다($p=0.5468$).

뇌졸중을 진단받지 않은 군에서 수술 여부에 대해 없음은 8,762명(64.79%), 수술 여부에 대해 있음은 4,761명(35.21%)이었고 뇌졸중 진단을 받은 군에서 수술 여부에 대해 없음은 117명(64.64%), 수술 여부에 대해 있음은 64명(35.36%)으로 통계적으로 유의한 차이는 없었다($p=0.3557$).

뇌졸중을 진단받지 않은 군에서 방사선 치료 여부에 대해 없음은 10,080명(74.54%), 방사선 치료 여부에 대해 있음은 3,443명(25.46%)이었고 뇌졸중 진단을 받은 군에서 방사선 치료 여부에 대해 없음은 97명(53.59%), 방사선 치료 여부에 대해 있음은 84명(46.41%)으로 통계적으로 유의한 차이는 없었다($p=0.0869$, 표 14).

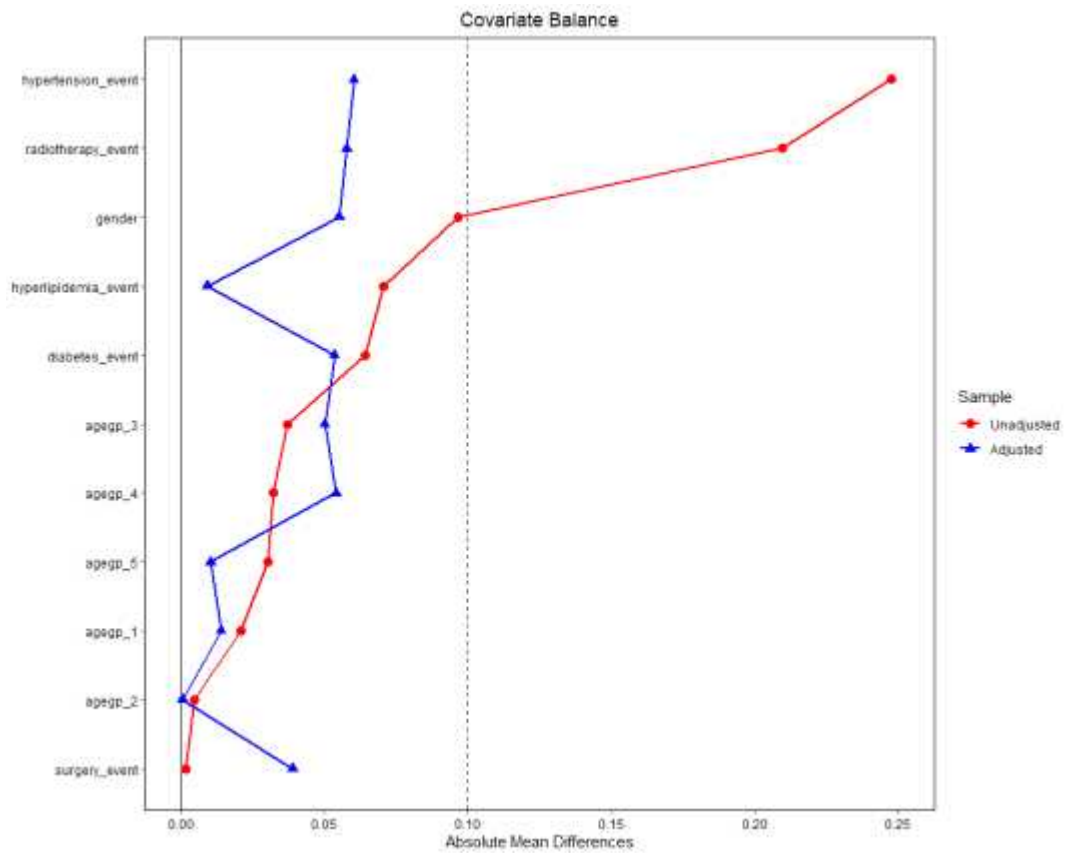


그림 10. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전후 공변량 간
 Standardized Mean Difference

표 14. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 일반적 특성

	뇌졸중을		p-value	SMD
	진단받지 않은 군			
	진단받은 군			
	N = 13,523	N = 181		
	(98.68%)	(1.32%)		
	N(%)	N(%)		
연령				
50세 미만	1,328(9.82)	14(7.74)	0.5545	0.054
50세 이상 60세 미만	2,901(21.45)	38(20.99)		
60세 이상 70세 미만	4,461(32.99)	53(29.28)		
70세 이상 80세 미만	3,822(28.26)	57(31.49)		
80세 이상	1,011(7.48)	19(10.50)		
성별				
남	8,332(61.61)	129(71.27)	0.2094	0.055
여	5,191(38.39)	52(28.73)		
고혈압 유무				
없음	9,699(71.72)	85(46.96)	0.0857	0.060
있음	3,824(28.28)	96(53.04)		
당뇨 유무				
없음	12,376(91.52)	154(85.08)	0.0303	0.053
있음	1,147(8.48)	27(14.92)		
고지혈증 유무				
없음	12,685(93.80)	157(86.74)	0.5468	0.009
있음	838(6.20)	24(13.26)		
수술 여부				
없음	8,762(64.79)	117(64.64)	0.3557	0.039
있음	4,761(35.21)	64(35.36)		
방사선 치료 여부				
없음	10,080(74.54)	97(53.59)	0.0869	0.058
있음	3,443(25.46)	84(46.41)		

2. 뇌졸중 발생 여부에 따른 생존분석

2.1. 전체 연구대상자의 IPTW 적용 전, 생존분석

연구대상자 20,696명을 대상으로 뇌졸중 발생 여부에 따른 사망 위험을 분석하기 위해 Kaplan-Meier 생존 곡선을 작성하였고, 교란변수를 통제된 상태에서 사망 위험도를 추정하기 위해 Cox 비례위험모형을 수행하였다. 또한 뇌졸중 발생 시점이 사망 위험에 미치는 영향을 반영하기 위해 시간 종속 Cox 회귀모형을 수행하였다.

Kaplan-Meier 생존 곡선 및 Log-Rank 검정 결과, 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 생존 곡선이 더 가파르게 감소하였으며, 두 집단 간 생존 곡선 차이는 통계적으로 유의하였다($p=0.027$, 그림 11).

Cox 비례위험모형의 분석 결과, 다른 요인을 통제된 상태에서 뇌졸중이 발생한 집단은 비발생 집단에 비해 사망 위험이 4% 높았으나 통계적으로 유의하지는 않았다(HR 1.04, 95% CI:0.90-1.20). 연령이 증가할수록 사망 위험은 유의하게 증가하였으며, 60세 이상 70세 미만은 50세 미만보다 9% 증가하였고(HR 1.09, 95% CI:1.01-1.19), 70세 이상 80세 미만은 27%, 80세 이상은 31% 증가하였다(HR 1.27, 95% CI:1.17-1.39; HR 1.31, 95% CI:1.18-1.47). 여성은 남성보다 사망 위험이 28% 낮았고(HR 0.72, 95% CI:0.68-0.75), 고혈압이 있는 경우는 20% 높았다(HR 1.20, 95% CI:1.14-1.26). 고지혈증이 있는 경우는 사망 위험이 21% 낮았으며(HR 0.79, 95% CI:0.72-0.87), 수술을 받은 경우는 48% 낮았다(HR 0.52, 95% CI:0.49-0.55). 방사선 치료를 받은 환자는 사망 위험이 2.14배로 가장 크게 증가하였다(HR 2.14, 95% CI:2.04-2.24, 표 15).

시간 종속 Cox 회귀모형의 분석 결과, 다른 요인을 통제한 상태에서 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 2.20배로 유의하게 증가하였다(HR 2.20, 95% CI:1.90-2.55). 연령이 증가할수록 사망 위험은 유의하게 증가하였으며, 70세 이상 80세 미만은 50세 미만보다 사망 위험이 19% 증가하였고(HR 1.19, 95% CI:1.09-1.30), 80세 이상은 20% 증가하였다(HR 1.20, 95% CI:1.07-1.34). 여성은 남성보다 사망 위험이 26% 낮았으며(HR 0.74, 95% CI:0.70-0.78), 고혈압이 있는 경우는 39% 높았다(HR 1.39, 95% CI:1.32-1.46). 당뇨가 있는 경우 사망 위험은 22% 증가하였고(HR 1.22, 95% CI:1.13-1.31), 수술을 받은 경우는 37% 낮았다(HR 0.63, 95% CI:0.59-0.66). 방사선 치료를 받은 환자는 사망 위험이 3.34배로 가장 크게 증가하였다(HR 3.34, 95% CI:3.19-3.50, 표 16).

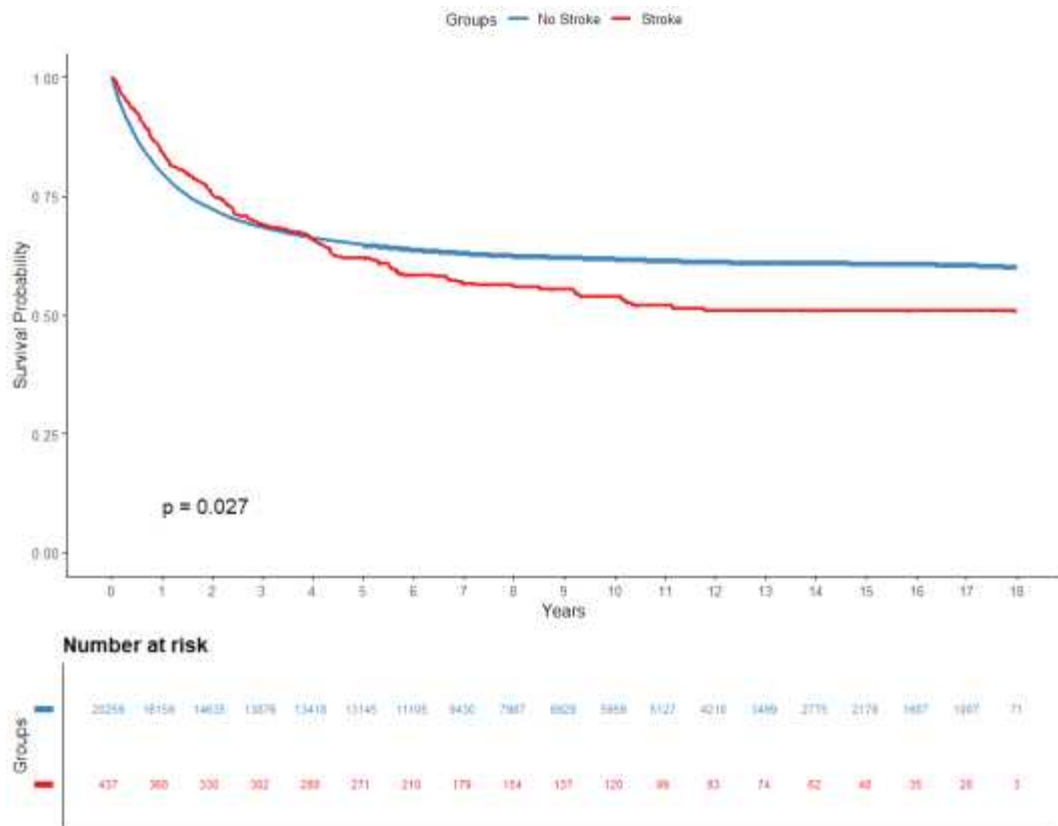


그림 11. 전체 연구대상자의 IPTW 적용 전, 뇌졸중 발생 여부에 따른
 Kaplan-Meier survival curve 및 Log-Rank test

표 15. 전체 연구대상자의 IPTW 적용 전, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	1.04	0.90	1.20	0.5930
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.97	0.88	1.06	0.4823
60세 이상 70세 미만	1.09	1.01	1.19	0.0454
70세 이상 80세 미만	1.27	1.17	1.39	<.0001
80세 이상	1.31	1.18	1.47	<.0001
성별				
남	1.00	-	-	-
여	0.72	0.68	0.75	<.0001
고혈압 유무				
없음	1.00	-	-	-
있음	1.20	1.14	1.26	<.0001
당뇨 유무				
없음	1.00	-	-	-
있음	0.98	0.91	1.06	0.6779
고지혈증 유무				
없음	1.00	-	-	-
있음	0.79	0.72	0.87	<.0001
수술 여부				
없음	1.00	-	-	-
있음	0.52	0.49	0.55	<.0001
방사선 치료 여부				
없음	1.00	-	-	-
있음	2.14	2.04	2.24	<.0001

표 16. 전체 연구대상자의 IPTW 적용 전, Time-Dependent Cox Regression을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	2.20	1.90	2.55	<.0001
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.94	0.86	1.03	0.2025
60세 이상 70세 미만	1.04	0.96	1.14	0.3442
70세 이상 80세 미만	1.19	1.09	1.30	0.0001
80세 이상	1.20	1.07	1.34	0.0013
성별				
남	1.00	-	-	-
여	0.74	0.70	0.78	<.0001
고혈압 유무				
없음	1.00	-	-	-
있음	1.39	1.32	1.46	<.0001
당뇨 유무				
없음	1.00	-	-	-
있음	1.22	1.13	1.31	<.0001
고지혈증 유무				
없음	1.00	-	-	-
있음	1.02	0.93	1.12	0.6966
수술 여부				
없음	1.00	-	-	-
있음	0.63	0.59	0.66	<.0001
방사선 치료 여부				
없음	1.00	-	-	-
있음	3.34	3.19	3.50	<.0001

2.2. 전체 연구대상자의 IPTW 적용 후, 생존분석

IPTW를 적용한 연구대상자 20,696명을 대상으로, 뇌졸중 발생 여부에 따른 사망 위험을 분석하기 위해 Kaplan-Meier 생존 곡선을 작성하였고, 교란변수를 통제한 상태에서 사망 위험도를 추정하기 위해 Cox 비례위험모형을 수행하였다. 또한 시간에 따라 변화하는 혼란변수를 적절히 보정하고 인과 추론의 정확도를 높이기 위해 stabilized IPTW를 적용한 MSM 분석을 수행하였다.

Kaplan-Meier 생존 곡선 및 Log-Rank 검정 결과, 두 집단 간 생존 곡선 차이는 통계적으로 유의하지 않았다($p=0.116$, 그림 12).

Cox 비례위험모형의 분석 결과, 다른 요인을 통제한 상태에서 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 13% 높았으나 통계적으로 유의하지는 않았다(HR 1.13, 95% CI:0.97-1.31). 연령이 증가할수록 사망 위험은 유의하게 증가하였으며, 60세 이상 70세 미만 집단은 50세 미만 집단에 비해 사망 위험이 10% 높았고(HR 1.10, 95% CI:1.01-1.19), 70세 이상 80세 미만은 28%(HR 1.28, 95% CI:1.17-1.39), 80세 이상은 32% 증가하였다(HR 1.32, 95% CI:1.18-1.47). 여성은 남성보다 사망 위험이 28% 낮았으며(HR 0.72, 95% CI:0.69-0.76), 고혈압이 있는 경우에는 사망 위험이 20% 높았다(HR 1.20, 95% CI:1.14-1.27). 고지혈증이 있는 경우에는 사망 위험이 21% 낮았고(HR 0.79, 95% CI:0.72-0.87), 수술을 받은 경우는 49% 낮았다(HR 0.51, 95% CI:0.49-0.54). 방사선 치료를 받은 환자는 사망 위험이 2.14배로 가장 크게 증가하였다(HR 2.14, 95% CI:2.04-2.24, 표 17).

시간 종속 Cox 회귀모형의 분석 결과, 다른 요인을 통제한 상태에서 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 2.43배로 유의하게 증가하였다(HR 2.43, 95% CI:1.99-2.96). 연령이 증가할수록 사망 위험은 유의하게 증가하였으며, 70세 이상 80세 미만은 50세 미만에 비해 사망 위험이 19% 증가하였고(HR 1.19, 95% CI:1.09-1.30), 80세 이상은 19% 증가하였다(HR 1.19, 95% CI:1.06-1.34). 여성은 남성보다 사망 위험이 25% 낮았으며(HR 0.75, 95% CI:0.71-0.78), 고혈압이 있는 경우에는 사망 위험이 40% 높았다(HR 1.40, 95% CI:1.32-1.47). 당뇨가 있는 경우 사망 위험은 22% 증가하였고(HR 1.22, 95% CI:1.12-1.32), 수술을 받은 경우는 37% 낮았다(HR 0.63, 95% CI:0.59-0.67). 방사선 치료를 받은 환자는 사망 위험이 3.35배로 가장 크게 증가하였다(HR 3.35, 95% CI:3.19-3.52, 표 18).

MSM 분석 결과, 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 2.62배 높았고, 이는 통계적으로 유의하였다(HR 2.62, 95% CI:2.13-3.22, 표 19). 이어서 IPTW의 극단값을 제한하기 위해 트렁케이션(truncation)을 각각 1%, 5%, 10% 수준으로 적용한 MSM 분석을 수행한 결과, 1% 트렁케이션에서는 사망 위험이 3.02배로 나타났으며(HR 3.02, 95% CI:2.54-3.60), 5% 트렁케이션 시에는 3.04배(HR 3.04, 95% CI:2.55-3.62), 10% 트렁케이션 시에는 3.05배(HR 3.05, 95% CI:2.56-3.62)로 모두 통계적으로 유의한 결과를 보였다(표 20).

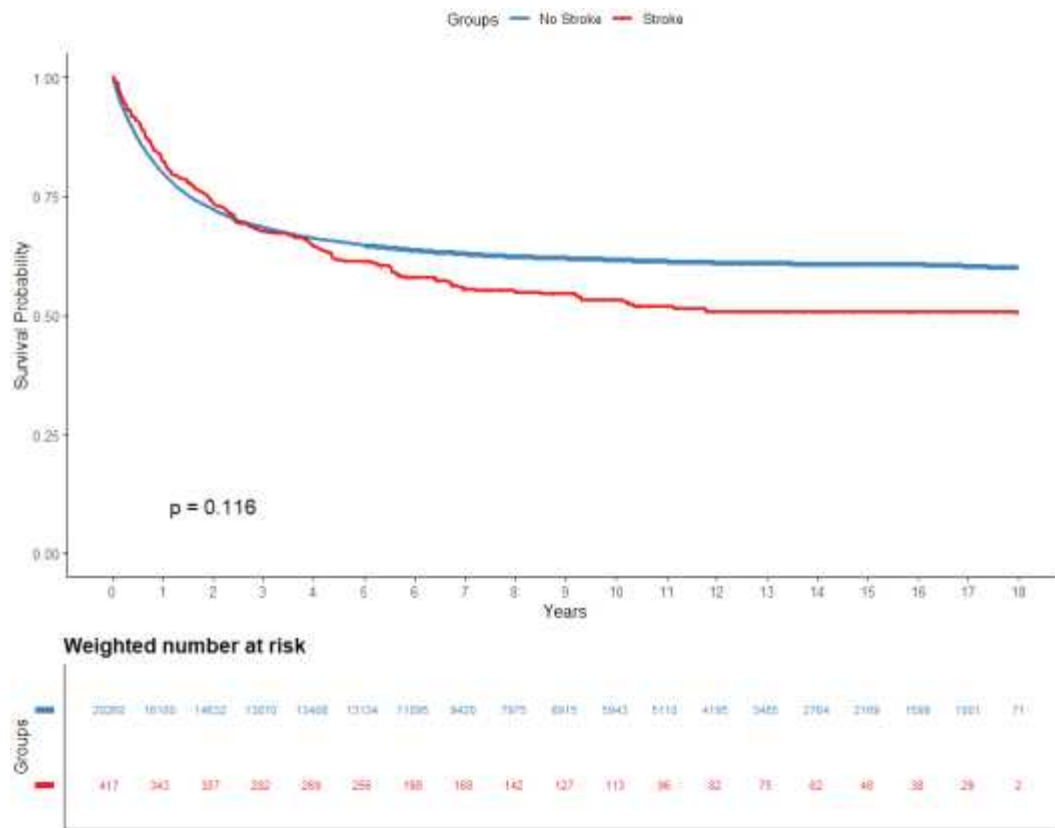


그림 12. 전체 연구대상자의 IPTW 적용 후, 뇌졸중 발생 여부에 따른
 Kaplan-Meier survival curve 및 Log-Rank test

표 17. 전체 연구대상자의 IPTW 적용 후, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	1.13	0.97	1.31	0.1160
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.97	0.88	1.06	0.4542
60세 이상 70세 미만	1.10	1.01	1.19	0.0312
70세 이상 80세 미만	1.28	1.17	1.39	<.0001
80세 이상	1.32	1.18	1.47	<.0001
성별				
남	1.00	-	-	-
여	0.72	0.69	0.76	<.0001
고혈압 유무				
없음	1.00	-	-	-
있음	1.20	1.14	1.27	<.0001
당뇨 유무				
없음	1.00	-	-	-
있음	0.98	0.92	1.06	0.6737
고지혈증 유무				
없음	1.00	-	-	-
있음	0.79	0.72	0.87	<.0001
수술 여부				
없음	1.00	-	-	-
있음	0.51	0.49	0.54	<.0001
방사선 치료 여부				
없음	1.00	-	-	-
있음	2.14	2.04	2.24	<.0001

표 18. 전체 연구대상자의 IPTW 적용 후, Time-Dependent Cox Regression을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	2.43	1.99	2.96	<.0001
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.94	0.86	1.03	0.1869
60세 이상 70세 미만	1.04	0.96	1.14	0.3237
70세 이상 80세 미만	1.19	1.09	1.30	<.0001
80세 이상	1.19	1.06	1.34	0.0029
성별				
남	1.00	-	-	-
여	0.75	0.71	0.78	<.0001
고혈압 유무				
없음	1.00	-	-	-
있음	1.40	1.32	1.47	<.0001
당뇨 유무				
없음	1.00	-	-	-
있음	1.22	1.12	1.32	<.0001
고지혈증 유무				
없음	1.00	-	-	-
있음	1.02	0.91	1.14	0.7476
수술 여부				
없음	1.00	-	-	-
있음	0.63	0.59	0.67	<.0001
방사선 치료 여부				
없음	1.00	-	-	-
있음	3.35	3.19	3.52	<.0001

표 19. 전체 연구대상자의 IPTW 적용 후, Marginal Structural Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	2.62	2.13	3.22	<.0001

표 20. Truncation 비율별 Marginal Structural Model 분석 결과

	IPTW Mean	HR
	(min-max)	(95% CI)
No truncation	1.00	2.62
	(0.14 - 2.60)	(2.13 - 3.22)
1% truncation	1.00	3.02
	(0.99 - 1.07)	(2.54 - 3.60)
5% truncation	1.00	3.04
	(0.99 - 1.03)	(2.55 - 3.62)
10% truncation	1.00	3.05
	(0.99 - 1.02)	(2.56 - 3.62)

2.3. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 생존분석

연구대상자 14,964명을 대상으로 뇌졸중 발생 여부에 따른 사망 위험을 분석하기 위해 Kaplan-Meier 생존 곡선을 작성하였고, 교란변수를 통제된 상태에서 사망 위험도를 추정하기 위해 Cox 비례위험모형을 수행하였다.

Kaplan-Meier 생존 곡선 및 Log-Rank 검정 결과, 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 생존 곡선이 더 가파르게 감소하였으며, 두 집단 간 생존 곡선 차이는 통계적으로 유의하였다($p=0.0003$, 그림 13).

Cox 비례위험모형의 분석 결과, 다른 요인을 통제된 상태에서 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 1.43배로 유의하게 증가하였다(HR 1.43, 95% CI:1.03-1.99). 연령이 증가할수록 사망 위험은 유의하게 감소하였으며, 70세 이상 80세 미만은 50세 미만에 비해 사망 위험이 22% 낮았고(HR 0.78, 95% CI:0.67-0.92), 80세 이상은 42% 낮았다(HR 0.58, 95% CI:0.46-0.74). 고혈압이 있는 경우 사망 위험은 22% 높았으며(HR 1.22, 95% CI:1.10-1.35), 방사선 치료를 받은 환자는 사망 위험이 2.67배로 가장 크게 증가하였다(HR 2.67, 95% CI:2.44-2.92, 표 21).

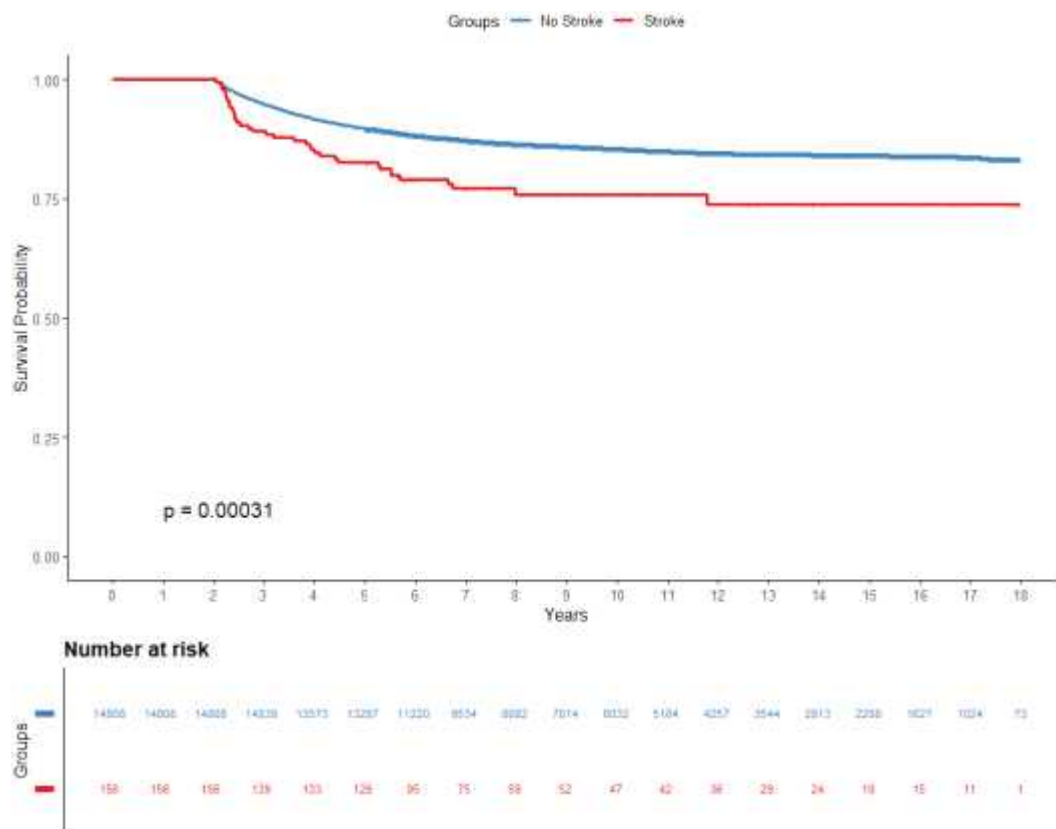


그림 13. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test

표 21. 랜드마크 기간 2년 연구대상자의 IPTW 적용 전, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	1.43	1.03	1.99	0.0351
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.91	0.78	1.06	0.2408
60세 이상 70세 미만	0.89	0.77	1.03	0.1277
70세 이상 80세 미만	0.78	0.67	0.92	0.0025
80세 이상	0.58	0.46	0.74	<.0001
성별				
남	1.00	-	-	-
여	0.94	0.86	1.03	0.1653
고혈압 유무				
없음	1.00	-	-	-
있음	1.22	1.10	1.35	<.0001
당뇨 유무				
없음	1.00	-	-	-
있음	1.12	0.96	1.30	0.1519
고지혈증 유무				
없음	1.00	-	-	-
있음	1.10	0.92	1.31	0.3103
수술 여부				
없음	1.00	-	-	-
있음	0.99	0.91	1.10	0.9708
방사선 치료 여부				
없음	1.00	-	-	-
있음	2.67	2.44	2.92	<.0001

2.4. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 생존분석

IPTW를 적용한 연구대상자 14,964명을 대상으로, 뇌졸중 발생 여부에 따른 사망 위험을 분석하기 위해 Kaplan-Meier 생존 곡선을 작성하였고, 교란변수를 통제한 상태에서 사망 위험도를 추정하기 위해 Cox 비례위험모형을 수행하였다.

Kaplan-Meier 생존 곡선 및 Log-Rank 검정 결과, 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 생존 곡선이 더 가파르게 감소하였으며, 두 집단 간 생존 곡선 차이는 통계적으로 유의하였다($p=0.014$, 그림 14).

Cox 비례위험모형의 분석 결과, 다른 요인을 통제한 상태에서 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 1.56배로 유의하게 증가하였다(HR 1.56, 95% CI:1.09-2.22). 연령이 증가할수록 사망 위험은 유의하게 감소하였으며, 70세 이상 80세 미만은 50세 미만에 비해 사망 위험이 21% 낮았고(HR 0.79, 95% CI:0.67-0.92), 80세 이상은 41% 낮았다(HR 0.59, 95% CI:0.46-0.74). 고혈압이 있는 경우 사망 위험은 22% 높았으며(HR 1.22, 95% CI:1.10-1.35), 방사선 치료를 받은 환자는 사망 위험이 2.68배로 가장 크게 증가하였다(HR 2.68, 95% CI:2.46-2.93, 표 22).

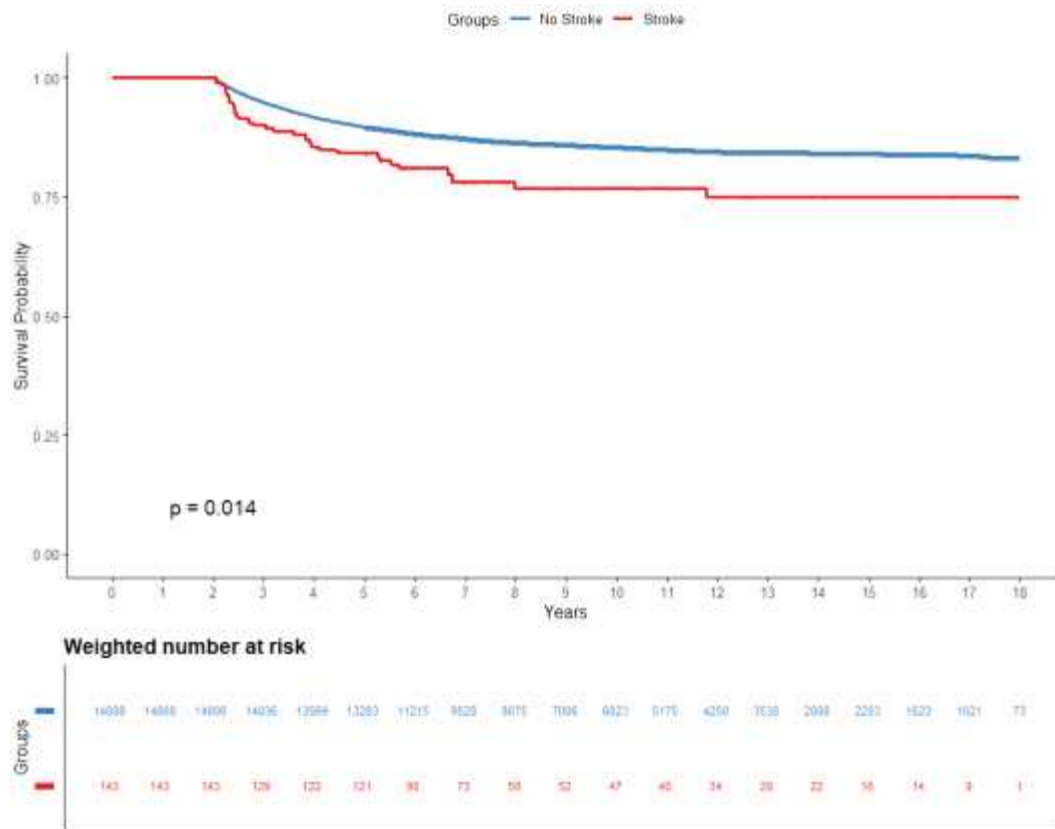


그림 14. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test

표 22. 랜드마크 기간 2년 연구대상자의 IPTW 적용 후, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	1.56	1.09	2.22	0.0140
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.91	0.78	1.07	0.2527
60세 이상 70세 미만	0.89	0.77	1.04	0.1421
70세 이상 80세 미만	0.79	0.67	0.92	0.0030
80세 이상	0.59	0.46	0.74	<.0001
성별				
남	1.00	-	-	-
여	0.94	0.86	1.03	0.1687
고혈압 유무				
없음	1.00	-	-	-
있음	1.22	1.10	1.35	0.0001
당뇨 유무				
없음	1.00	-	-	-
있음	1.10	0.95	1.29	0.2076
고지혈증 유무				
없음	1.00	-	-	-
있음	1.10	0.91	1.32	0.3305
수술 여부				
없음	1.00	-	-	-
있음	0.99	0.91	1.10	0.9905
방사선 치료 여부				
없음	1.00	-	-	-
있음	2.68	2.46	2.93	<.0001

2.5. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 생존분석

연구대상자 13,704명을 대상으로 뇌졸중 발생 여부에 따른 사망 위험을 분석하기 위해 Kaplan-Meier 생존 곡선을 작성하였고, 교란변수를 통제된 상태에서 사망 위험도를 추정하기 위해 Cox 비례위험모형을 수행하였다.

Kaplan-Meier 생존 곡선 및 Log-Rank 검정 결과, 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 생존 곡선이 더 가파르게 감소하였으며, 두 집단 간 생존 곡선 차이는 통계적으로 유의하였다($p=0.0015$, 그림 15).

Cox 비례위험모형의 분석 결과, 다른 요인을 통제된 상태에서 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 1.67배로 유의하게 증가하였다(HR 1.67, 95% CI:1.07-2.61). 80세 이상 집단은 50세 미만 집단에 비해 사망 위험이 37% 낮았고(HR 0.63, 95% CI:0.42-0.95), 여성은 남성보다 사망 위험이 14% 낮았다(HR 0.86, 95% CI:0.75-0.99). 고혈압이 있는 경우 사망 위험은 20% 증가하였으며(HR 1.20, 95% CI:1.03-1.40), 수술을 받은 경우 사망 위험은 59% 증가하였다(HR 1.59, 95% CI:1.38-1.83). 방사선 치료를 받은 환자는 사망 위험이 2.45배로 가장 크게 증가하였다(HR 2.45, 95% CI:2.13-2.81, 표 23).

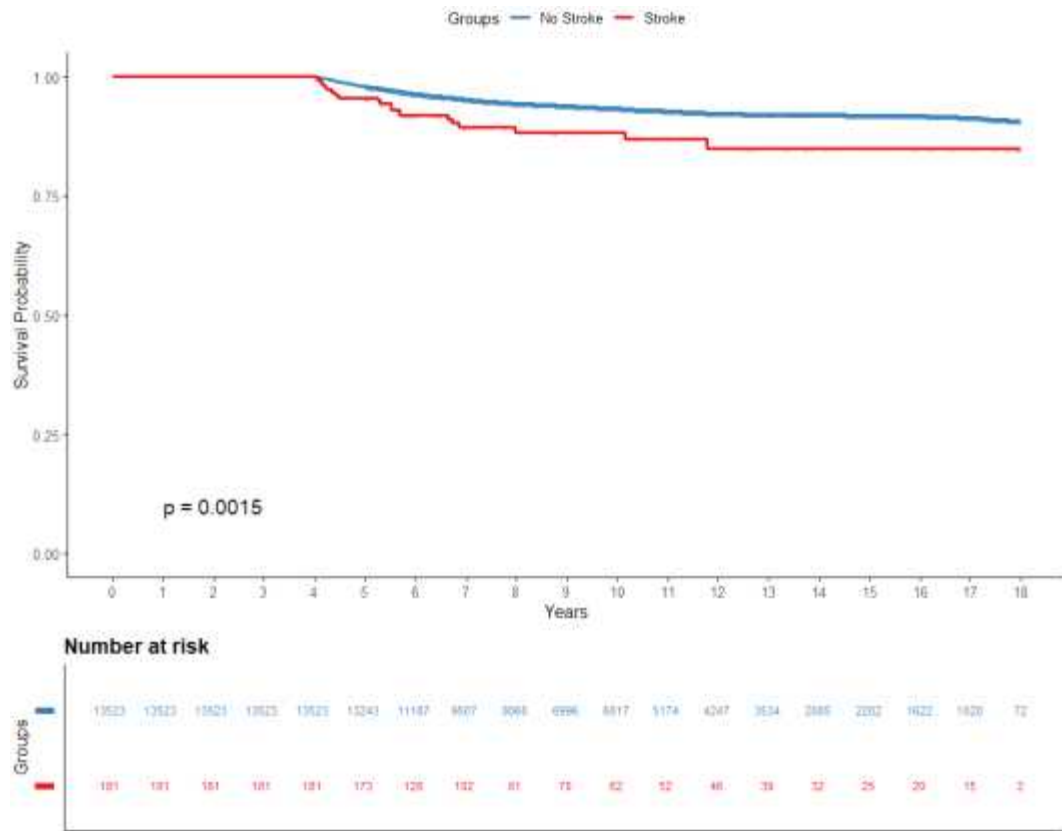


그림 15. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test

표 23. 랜드마크 기간 4년 연구대상자의 IPTW 적용 전, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	1.67	1.07	2.61	0.0234
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.91	0.70	1.18	0.4743
60세 이상 70세 미만	1.09	0.85	1.39	0.4899
70세 이상 80세 미만	0.99	0.77	1.28	0.9386
80세 이상	0.63	0.42	0.95	0.0269
성별				
남	1.00	-	-	-
여	0.86	0.75	0.99	0.0447
고혈압 유무				
없음	1.00	-	-	-
있음	1.20	1.03	1.40	0.0199
당뇨 유무				
없음	1.00	-	-	-
있음	1.06	0.84	1.33	0.6247
고지혈증 유무				
없음	1.00	-	-	-
있음	1.24	0.97	1.59	0.0888
수술 여부				
없음	1.00	-	-	-
있음	1.59	1.38	1.83	<.0001
방사선 치료 여부				
없음	1.00	-	-	-
있음	2.45	2.13	2.81	<.0001

2.6. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 생존분석

IPTW를 적용한 연구대상자 13,704명을 대상으로, 뇌졸중 발생 여부에 따른 사망 위험을 분석하기 위해 Kaplan-Meier 생존 곡선을 작성하였고, 교란변수를 통제된 상태에서 사망 위험도를 추정하기 위해 Cox 비례위험모형을 수행하였다.

Kaplan-Meier 생존 곡선 및 Log-Rank 검정 결과, 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 생존 곡선이 더 가파르게 감소하였으며, 두 집단 간 생존 곡선 차이는 통계적으로 유의하였다($p=0.013$, 그림 16).

Cox 비례위험모형의 분석 결과, 다른 요인을 통제된 상태에서 뇌졸중이 발생하지 않은 집단에 비해 뇌졸중이 발생한 집단의 사망 위험은 1.89배로 유의하게 증가하였다(HR 1.89, 95% CI:1.14-3.12). 80세 이상 집단은 50세 미만 집단에 비해 사망 위험이 36% 낮았고(HR 0.64, 95% CI:0.42-0.97), 고혈압이 있는 경우 사망 위험은 21% 증가하였다(HR 1.21, 95% CI:1.04-1.42). 수술을 받은 경우 사망 위험은 59% 증가하였으며(HR 1.59, 95% CI:1.37-1.84), 방사선 치료를 받은 환자는 사망 위험이 2.44배로 가장 크게 증가하였다(HR 2.44, 95% CI:2.13-2.80, 표 24).

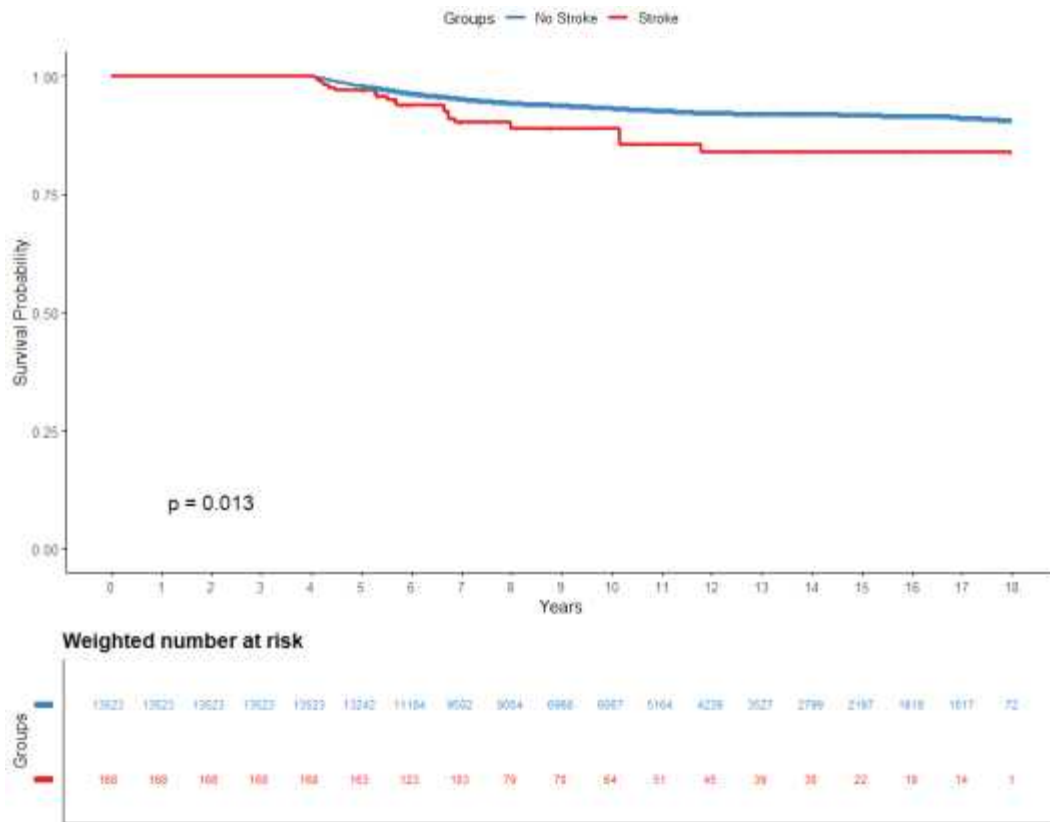


그림 16. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, 뇌졸중 발생 여부에 따른 Kaplan-Meier survival curve 및 Log-Rank test

표 24. 랜드마크 기간 4년 연구대상자의 IPTW 적용 후, Cox Proportional Hazards Model을 이용한 뇌졸중 발생 여부에 따른 사망 위험 분석

	HR	95% CI		p-value
		Lower	Upper	
뇌졸중 유무				
없음	1.00	-	-	-
있음	1.89	1.14	3.12	0.0130
연령				
50세 미만	1.00	-	-	-
50세 이상 60세 미만	0.92	0.71	1.20	0.5411
60세 이상 70세 미만	1.09	0.86	1.40	0.4648
70세 이상 80세 미만	1.01	0.78	1.29	0.9880
80세 이상	0.64	0.42	0.97	0.0352
성별				
남	1.00	-	-	-
여	0.87	0.75	1.01	0.0615
고혈압 유무				
없음	1.00	-	-	-
있음	1.21	1.04	1.42	0.0165
당뇨 유무				
없음	1.00	-	-	-
있음	1.06	0.84	1.34	0.6370
고지혈증 유무				
없음	1.00	-	-	-
있음	1.22	0.95	1.57	0.1206
수술 여부				
없음	1.00	-	-	-
있음	1.59	1.37	1.84	<.0001
방사선 치료 여부				
없음	1.00	-	-	-
있음	2.44	2.13	2.80	<.0001

V. 고 찰

1. 연구 결과에 대한 고찰

본 연구는 폐암 환자에서 뇌졸중 발생 여부가 전체 생존에 미치는 영향을 분석하고자 수행되었다. 비교군 간의 기저 특성 차이를 최소화하기 위해 IPTW를 적용하였으며, 가중치 부여 전·후 두 군 간 기저 특성의 균형 여부는 SMD를 통해 평가하였다. SMD 기준치인 0.1을 기준으로 공변량 균형을 판단하였으며, 모든 변수에서 해당 기준을 만족하였다. 다만 일부 변수에서는 IPTW 적용 후 오히려 SMD 값이 소폭 증가하는 양상이 관찰되었는데, 이는 변수 간 상호 상관관계나 비대칭적인 분포 특성에 기인한 것으로 보인다. 그럼에도 불구하고 전반적인 공변량 균형은 유지되었고, 분석의 타당성에는 큰 영향을 미치지 않은 것으로 해석된다. 이와 동일한 방식으로, 2년 및 4년 랜드마크 분석에서도 각 기준 시점까지 생존한 환자를 대상으로 IPTW를 적용한 후, 군 간 기저 특성의 균형을 SMD를 통해 검토하였다. 대부분의 변수에서 SMD 값이 0.1 이하로 유지되어 공변량의 적절한 균형이 확보되었으며, 분석 결과의 타당성을 뒷받침하였다. 앞서 전체 분석에서와 마찬가지로, 일부 변수에서는 IPTW 적용 이후에도 SMD 값이 소폭 증가하거나 비대칭적인 분포를 보이는 경향이 나타났으나, 이는 변수 간 상호 상관관계나 생존자 특성에 따른 선택 편향 등으로 해석될 수 있다. 그럼에도 불구하고 전반적인 공변량 균형은 유지되었고, 랜드마크 분석 결과에 대한 신뢰도를 높이는 기반이 되었다.

전체 연구대상자에 대해서는 보조 분석의 일환으로 Kaplan-Meier 생존 곡선, Log-Rank 검정, Cox 비례위험모형을 수행하였다. 시간에 따른 위험 변화까지 고려하기 위해 시간 종속 Cox 회귀모형을 추가로 수행하였으며, 시간-변이 공변량을 체계적으로 보정하기 위해 안정화된 IPTW 기반 MSM을 별도로 구축하였다. 또한, 조기 발견에 따른 시간 편향을 보정하고 장기 생존자의 조건부 생존율을 평가하기 위해, 최초 폐암 진단일로부터 2년 및 4년간 생존한 환자를 기준으로 한 랜드마크 분석을 실시하였다. 해당 분석에서도 IPTW를 적용한 후 군 간 기저 특성을 카이제곱 검정과 SMD로 비교하였고, 생존율 및 사망 위험 비교를 위해 Kaplan-Meier 생존 곡선, Log-Rank 검정, Cox 비례위험모형을 활용하였다.

전체 연구대상자 및 랜드마크 기간별 연구대상자 분석 결과, IPTW 적용 전에는 뇌졸중 진단 여부에 따라 여러 변수에서 통계적으로 유의한 차이가 나타났다. 전체 연구대상자의 경우, 고혈압, 당뇨, 고지혈증 유무, 수술 여부, 방사선 치료 여부에서 유의한 차이가 나타나 두 군 간 기저 특성의 불균형이 뚜렷하게 확인되었으며, 반면 연령과 성별은 유의한 차이를 보이지 않았다. 랜드마크 기간 2년 분석에서도 성별, 고혈압, 당뇨, 고지혈증, 방사선 치료 여부에서 통계적으로 유의한 차이가 나타났고, 4년 분석에서도 이와 유사하게 고혈압, 당뇨, 고지혈증, 성별, 방사선 치료 여부에서 두 군 간 유의한 차이가 관찰되었다. 이러한 결과는 뇌졸중 진단 여부에 따라 두 집단 간 기저 특성이 상이함을 시사한다. 이에 따라 세 분석 모두에서 IPTW를 적용하였으며, 적용 후에는 변수 간 SMD가 모두 0.1 미만으로 유지되어, 공변량 균형이 적절히 확보된 것을 확인할 수 있었다. 이는 이후 생존 분석에서 군 간 비교의 타당성을 확보하는 데 중요한 기반이 되었다.

전체 연구대상자에 대해 IPTW 적용하고 다양한 공변량을 고정된 변수로 통제된 상태에서 Cox 비례위험모형을 수행한 결과, 뇌졸중을 진단받은 군과 진단받지 않은 군 간의 사망 위험에는 통계적으로 유의한 차이가 없었다(HR=1.13, 95% CI:0.97-1.31). 반면, 시간 종속 Cox 회귀모형에서는 뇌졸중 발생 시점을 시간 종속 공변량으로 반영하고 일부 공변량 또한 시간 변화에 따라 반영한 상태에서 분석을 수행한 결과, 뇌졸중을 진단받은 군은 진단받지 않은 군에 비해 약 2.4배 높은 사망 위험을 나타냈으며, 이는 통계적으로 유의하였다(HR=2.43, 95% CI:1.99-2.96). 이러한 결과는 뇌졸중의 단순한 발생 유무만으로는 사망 위험을 충분히 설명하기 어렵고, 뇌졸중이 질병 경과 중 언제 발생했는지를 고려하는 시간의존적 분석이 보다 정밀한 위험 추정을 가능하게 한다는 점을 시사한다. 즉, 뇌졸중은 폐암 생존 과정에서 시기적으로 중요한 전환점이 될 수 있으며, 발생 시점에 따라 생존에 미치는 영향이 크게 달라질 수 있음을 보여준다. 추가로, IPTW를 이용한 MSM 분석에서도 뇌졸중을 진단받은 군은 진단받지 않은 군 대비 사망 위험이 유의하게 증가하였다(HR = 2.62, 95% CI:2.13-3.22). 극단 가중치를 1%, 5%, 10% 수준으로 절단한 민감도 분석에서도 HR이 3.02-3.05 범위로 일관되게 유의성을 유지해, 가중치 분포에 대한 의존성이 크지 않음을 확인하였다. 이러한 일관된 결과는 뇌졸중 발생이 폐암 환자의 생존에 미치는 부정적 영향이 가중치 선택이나 모델 사양 변화에도 견고함을 뒷받침하며, 뇌졸중 예방과 조기 개입의 임상적 중요성을 더욱 부각시킨다. 한편, 기존 선행연구들에서는 랜드마크 분석, IPTW, 시간 종속 Cox 회귀모형, MSM 같은 고도화된 분석 기법을 적용하지 않았음에도 불구하고, 폐암 환자에서 뇌혈관질환(또는 뇌졸중) 이후 전체 사망 위험이 증가하는 경향이 보고되었다. Liu et al.(2023)은 미국 SEER 데이터를 이용해 암 진단 이후 뇌혈관질환 관련 사망 위험을 평가한 결과, 폐암 환자의 표준화사망비(SMR)가 1.47(95 % CI:

1.42-1.52)로 일반 인구 대비 47 % 높다고 보고하였다. 그러나 이 연구는 전체 암 환자를 대상으로 한 인구 기반 단면 분석(population-level analysis)으로 SMR과 AER에 주로 의존하였으며, 개별 위험 요인이나 시간의존적 변화를 반영하지 못했다. Chen L et al.(2022) 또한 SEER 데이터를 활용해 폐암 환자 82,454명을 분석한 결과, 4,821명(5.85 %)이 뇌졸중으로 사망했고, SMR은 1.73(95 % CI:1.69-1.78)로 나타났다. Cox 비례위험모형에서 고령, 남성, 흑인, 고등급 암, 원격 전이, 진행된 병기(TNM) 등이 유의한 위험 요인으로 제시되었으나, 뇌졸중 발생 시점의 시간의존성이나 개별 시간가변 교란은 고려되지 않았다. 두 선행연구 모두 시간의존적 구조를 충분히 반영하지 못했음에도 불구하고 뇌졸중 이후 사망 위험이 증가한다는 결과를 보여, 본 연구와 비교 가능성이 있다. 본 연구는 이러한 선행연구의 한계를 보완하고자 시간 구조를 반영한 분석 설계로 2년 및 4년 랜드마크 시점을 설정하고, IPTW를 적용하여 보다 정밀한 예후 분석을 수행하였다.

랜드마크 시점을 기준으로 2년 및 4년 이상 생존한 폐암 환자 집단에 대해 IPTW 가중치를 적용하고 모든 공변량을 고정 변수로 통제된 상태에서 Cox 비례위험모형을 적용한 결과, 두 시점 모두 뇌졸중 진단은 사망 위험을 유의하게 증가시키는 요인으로 나타났다. 2년 랜드마크 시점에서 뇌졸중을 진단받은 군은 진단받지 않은 군에 비해 사망 위험이 약 1.6배 높았으며(HR = 1.56, 95% CI:1.09-2.22), 4년 랜드마크 시점에서도 약 1.9배 높은 위험을 보였다(HR = 1.89, 95% CI:1.14-3.12). 이 같은 결과는 뇌졸중이 장기 생존 폐암 환자의 예후를 지속적으로 악화시키는 주요 요인임을 보여주며, 랜드마크 분석이 이러한 장기적 영향을 평가하는 데 효과적인 분석 전략임을 시사한다. 특히, 기준 시점 이전 사망자를 제외함으로써 조기 사망에 따른 편향을 줄이고, 일정 기간 이상 생존한 환자 집단을 대상으로 뇌졸중 이후의 생존 영향을 보다 명확하게 관찰할 수 있다는 점에서 이 방법의 의의는 크다.

실제로 가장 치명적인 초기 사망군을 제외한 이후에도 뇌졸중 진단군의 사망 위험이 지속적으로 상승했다는 점은, 뇌졸중이 폐암 생존자의 장기 예후에 미치는 영향이 상당히 견고함을 뒷받침한다.

본 연구는 IPTW를 통해 공변량 균형을 확보한 후, Cox 비례위험모형, 시간 종속 Cox 회귀모형, MSM, 그리고 2년 및 4년 랜드마크 분석을 단계적으로 수행하였다. 고정된 기준 시점만을 가정한 Cox 모형에서는 뇌졸중과 사망 위험 간 유의한 연관성이 확인되지 않았으나, 뇌졸중 발생 시점과 시간 경과에 따른 위험 구조를 반영한 시간 종속 모형 및 MSM, 초기 사망 편향을 제거한 랜드마크 분석에서는 모두 통계적으로 유의한 위험 증가가 관찰되었다. 이는 사망 위험 추정 시 뇌졸중의 발생 시점과 시간-변이 공변량을 고려하는 것이 매우 중요함을 의미한다. 이러한 일관된 결과는 뇌졸중이 폐암 생존 과정에서 급성기뿐 아니라 장기적으로도 임상적으로 중요한 영향을 미치는 결정적인 요인임을 입증하며, 폐암 환자 관리에 있어 뇌졸중 예방과 조기 대응, 그리고 생존기 집중 관리 전략의 필요성을 강하게 시사한다.

2. 연구의 강점과 제한점

본 연구는 국내에서 2006년부터 2023년까지 18년간 축적된 세브란스병원 OMOP-CDM 데이터를 활용하여, 폐암 환자를 대상으로 뇌졸중 발생이 사망 위험에 미치는 영향을 분석한 국내 최초의 대규모 후향적 코호트 연구라는 점에서 의의가 있다. 서론에서도 언급했듯이, 폐암 환자의 뇌졸중 발생에 대한 연구는 일부 수행된 바 있으나, 뇌졸중 발생이 사망률에 미치는 직접적인 영향을 정량적으로 분석한 연구는 국내에서는 보고된 바 없으며, 기존 문헌은 대부분 국외 사례에 국한되어 있다. 특히 본 연구는 연세대학교 보건대학원에서 세브란스병원 OMOP-CDM 데이터를 분석 자료원으로 활용한 첫 번째 논문으로, 국내 의료데이터 기반 보건학 연구의 선도적 사례로 평가될 수 있다. 세브란스병원 OMOP-CDM 데이터는 국내 대형병원의 실 진료 데이터를 기반으로 구축된 표준화된 대규모 임상 데이터베이스로, 진단, 처방, 시술, 검사, 사망 등 다양한 의료 정보를 국제적으로 통용 가능한 공통 데이터 구조(Common Data Model)에 맞춰 정비한 것이 특징이다. 이러한 데이터 구조는 국내 실정을 반영한 연구임에도 불구하고 국제적 비교 가능성을 제공하며, 향후 다기관·다국가 공동연구로의 확장 가능성이라는 중요한 장점을 지닌다. 또한, 표준화된 코딩 체계와 구조 덕분에 분석 결과의 재현 가능성(reproducibility)과 신뢰도(reliability)를 높일 수 있으며, 장기적으로는 정밀의료, 다기관 메타분석, 의료 질 평가 등 다양한 분야에서 핵심 인프라로 활용될 수 있다. 분석 측면에서도 본 연구는 단순 생존분석을 넘어 고도화된 통계 기법을 단계적으로 적용해 보다 정밀한 추정과 해석을 도모하였다. 먼저 기저 특성의 불균형을 해소하기 위해 안정화된 IPTW를 사용해 실험군과 대조군 간 공변량 균형을 확보하였고, 뇌졸중이라는 시간

의존적 노출의 특성을 고려하여 시간 종속 Cox 회귀모형으로 시간 변화까지 반영한 분석을 수행하였다. 이어 노출과 시간 가변 교란을 동시에 보정하기 위해 MSM을 적용해 인과 추정의 견고성을 한층 강화하였다. 마지막으로 2년 및 4년 랜드마크 분석을 도입해 기준 시점 이전 조기 사망자를 제외함으로써 조기 사망 편향을 최소화하고, 해당 시점 이후 생존자 집단에서 뇌졸중의 순수한 장기 효과를 평가하였다. 이러한 다층적 분석 전략은 단일 방법론에 의존하지 않고 연구 설계와 통계적 접근 모두에서 타당성을 확보하려는 시도로서, 국내 보건의료 빅데이터 연구에서 활용 가능한 모범적 분석 프레임워크를 제시한다.

반면, 본 연구는 세브란스병원 OMOP-CDM 데이터를 활용한 후향적 코호트 연구라는 설계적 특성상 몇 가지 한계점을 지닌다. 첫째, OMOP-CDM 기반 분석에서는 특정 개념(Concept)이 데이터에 존재하면 ‘존재함’으로 간주되지만, 데이터에 존재하지 않을 경우에는 ‘존재하지 않음’으로 처리되는 구조적 특성을 갖는다. 이는 실제로 해당 정보가 누락된 것인지, 혹은 진료 기록 자체가 존재하지 않는 것인지를 구분할 수 없어, 결측(missing)과 부재(absence)를 식별할 수 없는 구조적 한계를 내포한다. 이러한 데이터 구조는 일부 변수의 해석에 있어 잠재적인 정보 왜곡 또는 오분류의 위험을 수반할 수 있다. 둘째, 특정 개념이 존재하더라도 병원 자체의 매핑(Mapping) 작업이 충분히 완료되지 않은 경우, 해당 개념이 CDM에 제대로 반영되지 않아 누락되는 문제가 발생할 수 있다. 실제로 본 연구에서는 흡연, 음주, 암 병기(TNM stage), 화학요법(Chemotherapy) 등 임상적으로 중요한 변수들이 CDM 데이터에 충분히 매핑되지 않아 분석에 포함시키지 못하였으며, 이는 연구 해석에 있어 주요 제약으로 작용하였다. 이러한 문제를 해결하기 위해서는 향후 OMOP-CDM 내 주요 임상 변수들에 대한 정밀한 매핑 강화 및 표준화 작업이 선행되어야 하며, 연구 목적에 따라 CDM

기반 분석의 한계를 인식하고 보완할 수 있는 병렬적 데이터 접근 전략도 함께 고려할 필요가 있다. 향후 후속 연구에서는 보다 명확한 변수 정의와 정보 누락을 최소화할 수 있는 대규모 전향적 코호트 연구 또는 임상시험 기반의 연구 설계를 도입함으로써, 폐암 환자에서 뇌졸중 발생이 사망에 미치는 영향을 보다 신뢰도 높고 정밀하게 추정할 수 있을 것으로 기대된다.

VI. 결 론

본 연구는 세브란스병원 OMOP-CDM 데이터를 활용하여, 폐암 환자에서 뇌졸중 발생 여부가 사망 위험에 미치는 영향을 분석한 후향적 코호트 연구이며, 2006년부터 2023년까지의 장기 추적 데이터를 기반으로 전체 연구대상자를 분석하였다. 군 간 기저 특성의 불균형을 최소화하기 위해 IPTW를 적용하여 연령, 성별, 고혈압, 당뇨, 고지혈증, 수술 여부, 방사선 치료 여부 등 주요 공변량을 균형화하였으며, 조기 사망 편향을 줄이기 위해 2년 및 4년 생존 시점을 기준으로 한 랜드마크 분석을 병행하였다. 전체 연구대상자에 대해 고정된 공변량을 포함한 일반 Cox 비례위험모형을 수행한 결과, 뇌졸중 진단 여부가 사망 위험에 유의한 영향을 미치지 않았으나, 뇌졸중 발생 시점과 일부 공변량을 시간 종속 공변량으로 반영한 시간 종속 Cox 회귀모형에서는 뇌졸중이 사망 위험을 유의하게 증가시키는 요인으로 나타났다. MSM에서도 같은 방향의 위험 증가가 재확인되어 분석 결과의 견고성을 강화하였다. 한편, 2년 및 4년 랜드마크 분석에서 동일 공변량을 고정 변수로 통제된 Cox 비례위험모형을 수행한 결과 뇌졸중을 진단받은 군의 사망 위험이 통계적으로 유의하게 상승함을 보여주었다. 이러한 차이는 시간 구조를 반영하는 방법에 따라 추정치가 달라질 수 있음을 시사하지만, 모든 고도화된 분석에서 뇌졸중의 부정적 영향이 일관되게 확인되었다는 점이 특징적이다.

종합적으로, 본 연구는 다양한 통계 기법과 연구 설계를 접목해 폐암 환자에서 뇌졸중 발생이 사망 위험에 미치는 영향을 다각도로 규명함으로써 학술적·실용적 가치를 지녔다. 특히 시간 종속 Cox 회귀모형, MSM, 그리고 랜드마크 분석을 병행함으로써 시간 구조와 조기 사망 편향을 동시에 고려한

정밀 추정이 가능했으며, 이러한 접근은 향후 보건의료 빅데이터 기반 생존 연구의 방법론적 방향성을 제시한다. 본 결과는 임상 현장에서 폐암 환자의 뇌졸중 예방·조기 대응 및 고위험군 모니터링 체계 구축의 필요성을 뒷받침하며, 향후 다기관·전향적 연구를 통한 인과 관계 규명이 더욱 중요함을 시사한다.

참 고 문 헌

Austin PC, Mamdani MM, Stukel TA, Anderson GM, Tu JV. The use of the propensity score for estimating treatment effects: administrative versus clinical data. Stat Med. 2005;24(10):1563-78.

Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat. 2011;10(2):150-61.

Austin PC. A comparison of 12 algorithms for matching on the propensity score. Stat Med. 2014;33(6):1057-69.

Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011;46(3):399-424.

Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med. 2015;34(28):3661-79.

Braitman LE, Rosenbaum PR. Rare outcomes, common treatments: analytic strategies using propensity scores. Ann Intern Med. 2002;137(8):693-5.

Chang WH, Sohn MK, Lee J, Kim DY, Lee SG, Shin YI, Oh GJ, Lee YS, Joo MC, Han EY, Kang C, Kim YH. Predictors of functional level and quality of life at 6 months after a first-ever stroke: the KOSCO study. J Neurol. 2016 Jun;263(6):1166-77.

Chen PC, Muo CH, Lee YT, Yu YH, Sung FC. Lung cancer and incidence of stroke: a population-based cohort study. Stroke 2011;42(11):3034-9.

Chen L, Zhao X, Wang S. Factors leading to the risk of stroke mortality: a cross-sectional study with lung cancer patient-based large sample. Eur J Cancer Prev 2021;31(1):14-8.

Cho IS, Chae YR, Kim JH, Yoo HR, Jang SY, Kim GR, Nam CM. Statistical methods for elimination of guarantee-time bias in cohort studies: a simulation study. BMC Med Res Methodol. 2017;17(1):126.

D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med. 1998;17(19):2265-81.

Graus F, Rogers LR, Posner JB. Cerebrovascular complications in patients with cancer. Medicine(Baltimore) 1985;64(1):16-35.

Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong IC, Rijnbeek PR, van der Lei J, Pratt N, Norén GN, Li YC, Stang PE, Madigan D, Ryan PB. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Stud Health Technol Inform* 2015;216:574-8.

Jang HS, Choi J, Shin JW, Chung JW, Bang OY, Kim GM, Seo WK, Lee JY. The Long-term effect of cancer on incident stroke: a nationwide population-based cohort study in Korea. *Front Neurol* 2019;10:52.

Kang MJ, Jung KW, Bang SH, Choi SH, Park EH, Yun EH, Kim HJ, Kong HJ, Im JS, Seo HG; Community of Population-Based Regional Cancer Registries*. Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2020. *Cancer Res Treat*. 2023;55(2):385-99.

Lee SW. Kaplan-Meier and Cox proportional hazards regression in survival analysis: statistical standard and guideline of Life Cycle Committee. *Life Cycle*. 2023;3:e8.

Lee YE, Ryu E. Moderating effect of inner strength between mood status and quality of life in female patients with lung cancer. *Asian Oncol Nurs*. 2021;21:24-32.

Lin DY, Wei LJ. The Robust Inference for the Cox Proportional Hazards Model. *Journal of the American Statistical Association*. 1989;84(408):1074-1078.

Liu Q, Zhang W, Li Q, Chen L, Wang X, Wang Z, Wang Y, Wang J. Risk of cerebrovascular disease after cancer diagnosis in the United States. *iScience*. 2023;26(7):107165.

National Cancer Information Center. Cancer Statistics in 2021. Goyang, 2021.

Observational Health Data Sciences and Informatics. The book of OHDSI. New York: OHDSI, 2021.

Park RW. Sharing Clinical Big Data While Protecting Confidentiality and Security: Observational Health Data Sciences and Informatics. *Healthc Inform Res*. 2017;23(1):1-3.

Rana RK, Singhal R. Chi-square test and its application in hypothesis testing. *Journal of the Practice of Cardiovascular Sciences*. 2015;1(1):69-71.

Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550-60.

Sands JM, Daly ME, Lee EQ. Neurologic complications of lung cancer. *Cancer*. 2020;126(20):4455-65.

Statistics Korea. Causes of Death Statistics in 2016. Daejeon, 2017.

Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71(3):209-249.

You SC, Rho Y, Bikdeli B, Kim J, Siapos A, Weaver J, Londhe A, Cho J, Park J, Schuemie M, Suchard MA, Madigan D, Hripcsak G, Gupta A, Reich CG, Ryan PB, Park RW, Krumholz HM. Association of Ticagrelor vs Clopidogrel With Net Adverse Clinical Events in Patients With Acute Coronary Syndrome Undergoing Percutaneous Coronary Intervention. *JAMA*. 2020;324(16):1640-50.

You SC, Seo SI, Falconer T, Yanover C, Duarte-Salles T, Seager S, Posada JD, Shah NH, Nguyen PA, Kim Y, Hsu JC, Van Zandt M, Hsu MH, Lee HL, Ko H, Shin WG, Pratt N, Park RW, Reich CG, Suchard MA, Hripcsak G, Park CH, Prieto-Alhambra D. Ranitidine Use and Incident Cancer in a Multinational Cohort. *JAMA Netw Open*. 2023;6(9):e2333495.

Abstract

Mortality Risk Associated with Stroke in Lung Cancer Patients: An IPTW-Weighted Landmark Analysis Using Hospital-Based OMOP-CDM Data

Changjun Ko

Department of Health Informatics & Biostatistics

Graduate School of Public Health

Yonsei University

(Directed by Professor Sohee Park, PhD)

Background

Lung cancer remains the leading cause of cancer-related deaths in Korea (22.7% in 2020), with a 5-year relative survival rate of only 38.5%, significantly lower than that of other major cancers. Stroke frequently occurs in lung cancer patients and may worsen prognosis; however, the causal impact of stroke timing on long-term survival has not been sufficiently investigated. This study aimed to evaluate the effect of stroke occurrence and its timing on all-cause mortality in patients with lung cancer.

Methods

This retrospective cohort study utilized OMOP-CDM data from Severance Hospital. A total of 20,696 patients diagnosed with lung cancer between 2006 and 2018 were included. The primary exposure was the occurrence and timing of stroke, and the outcome was all-cause mortality and survival duration. Covariates included age, sex, diagnosis of hypertension, diabetes, and hyperlipidemia, as well as history of surgery and radiotherapy. To adjust for confounding, stabilized inverse probability of treatment weighting (IPTW) based on propensity scores was applied. Subsequently, Cox proportional hazards models, time-dependent Cox regression, IPTW-based marginal structural models (MSM), and landmark analyses at 2- and 4-year timepoints were performed.

Results

In the Cox proportional hazards model adjusted with IPTW and fixed covariates, stroke diagnosis was not significantly associated with mortality risk (HR=1.13, 95% CI: 0.97-1.31). However, in the time-dependent Cox model accounting for the timing of stroke and changes in covariates over time, the mortality risk increased by approximately 2.4 times (HR=2.43, 95% CI: 1.99-2.96). The IPTW-based MSM yielded similar results (HR=2.62, 95% CI: 2.13-3.22). Sensitivity analyses using truncated weights (1%, 5%, 10%) also showed consistent findings (HR=3.02-3.05). In the landmark analyses of patients who survived ≥ 2 or ≥ 4 years after diagnosis, stroke diagnosis remained significantly associated with increased mortality risk (2-year HR=1.56; 4-year HR=1.89).

Conclusion

This study used OMOP-CDM data from Severance Hospital to assess the impact of stroke occurrence and timing on mortality in patients with lung cancer. While the IPTW-adjusted Cox model showed no significant association, both the time-dependent Cox model and MSM revealed a significantly increased risk of death following stroke. Consistent patterns were also observed in the 2- and 4-year landmark analyses, highlighting the importance of incorporating time-dependent structures in survival analysis to more accurately estimate clinical risks.

Keywords: Lung cancer, Stroke, Mortality risk, OMOP-CDM, Landmark analysis, Inverse Probability of Treatment Weighting (IPTW), Marginal Structural Model (MSM)