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

간호학석사 학위논문

전자의무기록자료 기반 심장수술 후 섬망 예측모델 개발 및 평가

2023 년 8 월

서울대학교 대학원

간호학과 간호정보학 전공

이 지 연

전자의무기록자료 기반 심장수술 후 섬망 예측모델 개발 및 평가

지도 교수 김 현 의

이 논문을 간호학 석사 학위논문으로 제출함
2023 년 6 월

서울대학교 대학원
간호학과 간호정보학 전공
이 지 연

이지연의 간호학석사 학위논문을 인준함
2023 년 6 월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

국 문 초 록

본 연구의 목적은 전자의무기록자료를 활용하여 심장수술 후 중환자 섬망 예측모델을 개발하고 평가하는 것이다. 서울대학교병원 Clinical Data Warehouse (CDW) 연구 검색 시스템인 Seoul National University Hospital Patient Research Environment (SUPREME)를 통해 2017년부터 2021년까지의 5년간 병원 심폐기계중환자실에 입실한 성인 2,083명의 원시자료를 추출한 뒤, R 프로그램을 이용하여 분석하였다.

최대한 누락 없이 선별하기 위해 임상관찰기록의 The Confusion Assessment Method for the ICU (CAM-ICU) 평가 결과, 간호일지 표준 진술문, 진단명, 정신건강의학과 타과의뢰 및 haloperidol 처방내역의 5 가지 자료 중 한 가지 이상에서 섬망 여부가 확인될 경우 섬망 환자로 정의했다. 섬망 발생 건의 수는 전체 2,083건 중 417건(20.0%)이었고, CAM-ICU, 타과의뢰 및 haloperidol만으로 정의된 수와 큰 차이가 없었다(410건).

문헌고찰을 바탕으로 중환자실 입실 직후 시점에 결정되는 46개의 위험요인을 선정한 후, 카이제곱 검정 및 T 검정으로 단변량 분석을 시행하여 p-value가 0.05보다 작았던 최종 35개의 위험요인으로 섬망 예측모델을 개발하였다. 모델 개발에 사용된 최종 위험요인 변수 중 수술 전 변수로 대상자의 성별, 연령, 체질량지수, 수술 전 재실 여부 및 수술 전 입원 기간 등이 포함되었고, 수술 자체 관련 변수로는 대동맥 수술 여부, 관상동맥 수술 여부, 판막 수술 여부, 예상 실혈량, 수술 시간, 심폐기 사용 여부 및 시간이 포함되었으며, 수술 후 변수로는 산소포화도 이상 여부, 체온, 심박수, 평균 동맥 혈압, 동맥혈 산도(pH), 동맥혈 산소 분압(pO_2), 동맥혈 이산화탄소 분압(pCO_2), 혈당, 젖산, 헤모글로빈, 혈소판, 섬유소원(Fibrinogen), 활성화 부분 트롬보플라스틴 시간(activated Partial Thromboplastin Time, aPTT), 프로트롬빈 시간(Prothrombin Time, PT), 나트륨, 칼슘, 마그네슘, 혈중요소질소(Blood Urea Nitrogen, BUN), 크레아티닌, 총 빌리루빈(Total Bilirubin), 알부민, 그리고 C-

반응성 단백질(C-Reactive Protein, CRP)이 사용 되었다.

예측모델은 로지스틱 회귀(Logistic Regression), 의사결정나무(Decision Tree), 랜덤포레스트(Random Forest), 그리고 eXtreme Gradient Boosting (XGBoost)의 기계학습 방법을 이용하여 개발하였고 각각의 성능을 비교 평가했다. 섬망군이 추출된 전체 환자군의 약 20%로 비섬망군 대비 균형이 맞지 않아 Synthetic Minority Oversampling Technique (SMOTE) 기법을 사용하여 섬망군과 비섬망군의 비를 1:1로 맞춘 뒤에 모델을 학습시켰다.

예측모델 성능 비교를 위해 반복적 K-겹 교차검증(Repeated K-fold cross-validation) 방법으로 5-겹 교차검증을 10회 반복하였으며 각각의 검증에서 산출된 민감도(Sensitivity), 특이도(Specificity), 양성 예측도(Positive Predictive Value), 음성 예측도(Negative Predictive Value) 및 AUROC (Area Under Receiver Operating Characteristic Curve)의 평균치를 최종 성능 지표로 사용하였다. 네 개의 기계학습 방법 중에서 SMOTE를 적용 후 랜덤포레스트 기법으로 개발한 예측모델의 성능이 민감도 0.9050, 특이도 0.8650, 양성 예측도 0.8702, 음성 예측도 0.9011, AUROC 0.9575로 가장 뛰어났지만, 일반화 가능성에 대한 후속 연구가 필요하다.

본 연구는 섬망의 세부 유형 및 발생 시각을 고려하지 못했다는 제한점을 가진다. 추후 전자의무기록의 data quality를 보완함과 동시에 전자의무기록의 자유형식 기록을 활용할 수 있는 연구도 계속 되어야 할 것이다.

주요어 : 전자의무기록, 예측모델, 섬망, 기계학습, 중환자, CDW

학 번 : 2021-22756

목 차

제 1 장 서 론 1

제 1 절 연구 필요성 1
제 2 절 연구 목적 4
제 3 절 용어 정의 5

제 2 장 문헌고찰 6

제 1 절 섬망 정의 6
제 2 절 섬망 위험요인 8
제 3 절 섬망 예측모델 12

제 3 장 연구의 틀 16

제 1 절 연구 의의 16
제 2 절 연구 과정 17

제 4 장 연구 방법 18

제 1 절 연구 설계 18
제 2 절 자료 수집 19
제 3 절 대상자 선정 20
제 4 절 섬망 환자 정의 25
제 5 절 섬망 위험요인 정의 26
제 6 절 자료 전처리 31
제 7 절 자료 분석 33
제 8 절 윤리적 고려 34

제 5 장 연구 결과 35

제 1 절 대상자 일반적 특성 35
제 2 절 섬망 환자 선별 38
제 3 절 섬망 위험요인 결정 41
제 4 절 섬망 예측모델 성능 비교 49

제 6 장 논 의 51
제 1 절 섬망 환자 선별 51
제 2 절 섬망 위험요인 결정 54
제 3 절 섬망 예측모델 성능 비교 56
제 4 절 연구 제한점 58

제 7 장 결 론 및 제 언 60
제 1 절 결론 60
제 2 절 제언 62

참고문헌 64

부 록 76

Abstract 88

List of Tables

[Table 1] Classification of Major Cardiac Surgeries with Examples 22
[Table 2] Inclusion and Exclusion Criteria 24
[Table 3] List of Risk Factors for Delirium 28
[Table 4] General Characteristics of Patients 37
[Table 5] Delirium Rate with Subject Selection Process 38
[Table 6] Delirium Rate using Different Operational Definitions 39
[Table 7] Inclusion Relations of Different Operational Definitions 40
[Table 8] Properties of 11 Discrete Variables 41
[Table 9] Properties of 35 Continuous Variables 42
[Table 10] Datasets before and after Applying SMOTE 44
[Table 11] Univariate Analysis of 11 Discrete Risk Factor Variables 46
[Table 12] Univariate Analysis of 35 Continuous Risk Factor Variables 47
[Table 13] Evaluation of Delirium Prediction Models 49

List of Figures

[Figure 1] Data-driven Care System 16
[Figure 2] Data Preprocessing Example of Surgery Name Variables 30
[Figure 3] Subject Selection Process 35
[Figure 4] Inclusion Relations of Different Operational Definitions 39
[Figure 5] Change in Boxplot Distribution after Treating Outliers (ALT) 44
[Figure 6] Comparison of ROC curves between Algorithms50
[Figure 7] Comparison of ROC curves between Datasets... 50

부록 목차

[부록 1] 의학연구윤리심의위원회(IRB) 심의 결과 통보서 76
[부록 2] 주요 심장수술 분류 및 해당 세부수술명 목록 78
[부록 3] 연속변수 이상치 기준, 개수 및 비율 80
[부록 4] R 코드 예시: 수술명 관련 변수 81
[부록 5] R 코드 예시: Estimated Blood Loss (EBL) 변수 83
[부록 6] R 코드 예시: 연속변수 단변량 분석 85
[부록 7] R 코드 예시: 랜덤포레스트 예측모델 87

제 1 장 서 론

제 1 절 연구 필요성

섬망은 단기간에 발생하고 변동하는 의식수준, 주의력 및 인지능력의 장애이다(American Psychiatric Association, 2013). 그 발생률은 중환자실 및 완화의료 현장에서 가장 높은 것으로 알려져 있다(Hshieh et al., 2017). 이 중에서 특히 중환자실에서의 섬망 발생은 중대한 환자안전사고로 이어질 가능성이 높기 때문에, 현장 근무자들이 주목하게 되는 사안이다. 섬망 증세를 보이는 중환자가 치료적 목적으로 삽입된 인공기도관 및 중심정맥관 등을 스스로 제거해버리는 사고가 빈번히 발생한다(Tilouche et al., 2018). 섬망 여부를 예측할 수 있으면 사전에 고위험군으로 설정하여 사고를 방지하기 위한 예방적 중재를 추가할 수 있다. 장기적으로도 섬망의 발생은 평균 재원 기간 증가 및 퇴원 시 사망률 증가와 관련이 있고 결과적으로 의료비 증가의 원인이 된다(Vasilevskis et al., 2018).

한편, 내과계 중환자와 외과계 중환자는 서로 다른 섬망 발생 양상을 보이기 때문에 섬망 문제를 다룰 때 이를 구별하는 것이 필요하다. Kim 등(2018)의 연구에 따르면, 섬망 증세를 보였던 중환자의 섬망 지속 시간 및 퇴원 시점 섬망 회복 여부는 수술 여부에 따라 통계적으로 유의한 차이를 나타냈다. 보통 내과계 중환자가 더 많은 기저질환을 보유한다는 사실도 참고해야 한다. 나아가 외과계 수술 중에서도 심장수술은 심폐기(Cardiopulmonary bypass, CPB)의 사용과 같은 특징적인 절차를 포함하기에 비심장수술과 분리되어 다뤄져야 한다(Silva et al., 2021). 심장수술의 특징적인 심폐기의 사용, 흉골 절개(median sternotomy), 색전 위험(embolic load) 및 긴 수술/마취 시간은 섬망의 위험요인으로 작용한다(Berger et al., 2018).

구체적인 기전은 명확하게 밝혀지지 않았지만, 섬망의 발생에 대해서는 수많은 요인들이 유의한 상관관계를 보임이 알려져 있다. Marcantonio 등(2001)은 무작위 대조군 실험연구(Randomized controlled trial)를 통해 섬망 위험요인에 대해 선제적으로 조기 중재를 시행한 군에서 섬망 발생률이 유의미하게 낮게 나타나는 것을 입증하였다. 위험요인 자료를 이용하여 환자별 섬망 발생을 사전에 예측할 수 있다면, 고위험군을 집중적으로 관리할 수 있게 되므로 효율적인 의료자원 활용이 가능해질 것이다. 그러나 아직까지 섬망 발생을 효과적으로 예측할 수 있는 수단은 부재하다. 임상현장에서는 환자의 현재 섬망 여부를 판단할 수 있는 도구만 활용되고 있다.

중환자실 섬망의 사정 도구로는 The Confusion Assessment Method for the ICU (CAM-ICU)와 The Intensive Care Delirium Screening Checklist (ICDSC)가 널리 사용되고 있다(Krewulak et al., 2018). 국내에서는 번역을 거친 한국어판 CAM-ICU가 가장 보편적으로 사용되고 있다(강지연 등, 2021). CAM-ICU는 섬망의 진단 기준인 Diagnostic and Statistical Manual of Mental Disorders (DSM)에 충실하게 근거하여 개발된 섬망 사정 도구인 Confusion Assessment Method (CAM)를 중환자용으로 수정한 도구이다(Ely, Margolin, et al., 2001). 개발 당시에는 471명의 환자를 대상으로 한, 연구 간호사의 평가 결과가 96.5%의 민감도 및 99%의 특이도를 보였다(Ely, Inouye, et al., 2001).

그러나 실제 현장에서는 해당 도구가 적극적으로 활용되지 못하고 있다. Swan (2014)의 연구에서는 CAM-ICU 사용에 대한 교육을 받지 않은 간호사들의 CAM-ICU 기록 423 건 중 87 건(21%)이 잘못되었음을 발견했는데, 그 중 80 건은 환자가 의식이 있어 평가가 가능한 상태였음에도 'Unable to access'로 기록을 한 경우였다. Awan 등 (2021)의 연구에서도 의식이 명료하거나 약간의 진정 상태인 환자의 CAM-ICU 기록 중 16%가 'Unable to assess'로 입력되어 있었는데, 특히 환자가 80 세가 넘고, 기관 내 삽관이 되어있고, 억제대가 있는 등의 경우가 유의하게 많았다.

CAM-ICU를 사용 중인 국내 일개 병원 간호사 150명을 대상으로 한 질적 연구에서 간호사들은 해당 섬망 도구를 제대로 사용하지 못하는 이유로 바쁜 업무, 사정 도구의 복잡성, 진정 상태 혹은 기관내관 유지 환자에 적용 어려움 등을 꼽았다(공경희 등, 2015). CAM-ICU 를 사용하지 않는 병원에서 간호사들을 새롭게 교육한 뒤 실시했던 질적 연구(Jung et al., 2013)에서도 유사한 결과를 보였으며 해당 도구를 반복적으로 사용할 때 환자들이 비협조적인 모습을 보였음을 보고하고 있다. 이렇듯 현장에서는 섬망 예측 도구는 차치하고 섬망 사정 도구마저 제대로 활용되고 있지 못하고 있는 실정인 것이다.

이에 본 연구는 전자의무기록자료를 이용하여 간호사들에게 업무 과중의 부담을 주지 않고도 심장수술 중환자의 섬망 발생 위험을 자동으로 예측할 수 있는 기계학습 기반의 알고리즘을 개발하고자 한다. 전자의무기록의 자료가 실시간으로 연동되는 모델을 통해 섬망 발생 고위험군을 예측할 수 있다면, 중환자 간호 업무의 효율 및 환자 안전 향상에 기여할 수 있을 것이다. 본 연구의 결과는 근거 기반의 맞춤형 간호를 제공하는 기반이 될 수 있다.

제 2 절 연구 목적

본 연구의 목적은 일개 대학병원의 전자의무기록자료를 활용하여 최적의 심장수술 중환자 대상 섬망 예측모델을 개발하기 위함이다. 본 연구의 구체적인 목적은 다음과 같다.

1) 연구 대상자 중 심장수술 후 섬망 환자를 선별한다.
2) 문헌고찰 및 임상경험을 기반으로 섬망 위험요인을 선정한 뒤, 단변량 분석을 수행해 유의한 변수를 선별한다.
3) 로지스틱 회귀, 의사결정나무, 랜덤포레스트 및 eXtreme Gradient Boosting (XGBoost)의 기계학습 방법을 이용하여 예측모델을 개발하고 예측 타당도를 비교 평가한다.

제 3 절 용어 정의

1) 섬망

섬망은 일시적으로 발생하고 양상이 변화하는 급성 혼돈 상태를 의미하고, 주의 집중력 및 인지능력의 저하를 포함하며 정신운동성(psychomotor) 행동 문제를 동반한다(American Psychiatric Association, 2013).

본 연구에서는 섬망 환자를, 중환자실 재실 기간 중 다음의 5 가지 조건 중 하나를 만족하는 환자로 조작적으로 정의하였다.

조건 1. 임상관찰기록지 내 CAM-ICU 평가 결과가 양성임
조건 2. 간호일지 내 섬망이 있다는 진술문이 존재함
조건 3. 진단명에 섬망이 추가됨
조건 4. 정신건강의학과에 섬망 관련 타과의뢰가 작성됨
조건 5. 투약 오더 중 Haloperidol 주사제제가 처방됨

2) 심장수술

심장수술은 심장 및 대동맥과 관련된 심혈관질환에 대한 심장혈관흉부외과 심장 파트의 수술을 의미한다.

본 연구에서는 다루는 심장수술의 목록은 [부록 2]와 같다.

제 2 장 문헌고찰

제 1 절 섬망의 정의

섬망의 이론적 정의라고 할 수 있는, 미국 정신의학 협회(American Psychiatric Association, APA)가 2013 년에 발표한 정신질환의 진단 및 통계 편람(The Diagnostic and Statistical Manual of Mental Disorders, Fifth edition, DSM-5)에 따른 섬망의 세부 진단 기준은 다음과 같다.

(a) 주의 집중력, 지남력 및 의식수준의 장애

(b) 몇 시간 또는 며칠 등의 단기간에 발생한 변화이며 그 정도는 하룻동안에도 큰 변동이 있음

(c) 기억력, 언어능력 등 인지능력의 장애가 추가적으로 있음

(d) 앞의 장애가 다른 신경인지적 원인으로는 설명이 되지 않으며 혼수 등의 심각한 각성 저하 상태가 아님

(e) 앞의 장애가 다른 질병 상태 및 약물 사용 등이 직접 작용한 생리학적 결과라는 증거가 있음

Inouye 등(1990)은 DSM 진단 기준에 기초하여, 정신건강의학과 전문의가 아닌 일반 의료진들도 현장에서 간편하게 사용할 수 있는 섬망 사정 도구인 Confusion Assessment Method (CAM)를 개발하였다. 대상 환자에게 특성 1(급격히 발생했으며 변동을 보임), 특성 2(주의력 결핍), 특성 3(정리되지 않은 사고), 특성 4(의식수준의 변화)와 관련된 간단한 질문들을 던지며 5 분 이내 결과를 낼 수 있게 하는 도구이다. 특성 1 과 특성 2 를 모두 가지고 있으면서, 특성 3 또는 특성 4 를 나타내면 양성으로 판별된다. CAM-ICU 는 인공호흡기 사용이 빈번한 중환자실 환경을 고려하여 대상 환자가 육성을 내지 않고도 답할 수 있는 질문으로 기존의 CAM 을 수정한 버전이다.

그러나 실제 임상현장에서는 섬망이 발생하였음에도 그에 대한 평가가 누락되는 경우가 빈번하다(Awan et al., 2021). 연구 간호사가 아닌 일반 간호사들에 의해 CAM-ICU가 사용될 때 47%의 민감도, 98%의 특이도, 95%의 양성 예측도, 72%의 음성 예측도를 보였다는 연구 결과도 있다(van Eijk et al., 2011). 민감도가 특히 낮다는 점에서, 상당히 많은 수의 섬망 환자들이 CAM-ICU 결과만으로 선별될 수 없다는 것을 알 수 있다.

일반적으로 섬망과 관련된 연구들에서 섬망을 조작적으로 정의할 때는 주로 중환자실 간호사가 작성한 CAM-ICU 결과를 바탕으로 정신건강의학과 의사의 판단을 추가하는 방식을 사용한다(손정현 등, 2014). 여기에 섬망 치료에 주로 사용되는 약물인 haloperidol과 quetiapine의 처방 여부를 더해 섬망 환자를 선별하기도 한다. Lee 등(2017)의 연구에서는 간호사들이 평가한 CAM-ICU 결과가 양성이거나 섬망 치료 목적의 haloperidol이 처방된 환자를 섬망 환자라고 정의했다. 후향적 연구에서 연구 간호사가 직접 의무기록 전반을 검토하여 환자별 섬망 발생 여부를 확인하거나(노은영 등, 2019), 전향적 연구에서 연구 시작 전에 섬망 사정을 수행할 의료진들을 대상으로 관련 교육을 여러 차례 상세히 진행한 경우도 있다(Menzenbach et al., 2022).

본 연구에서는 전자의무기록 내의 다양한 자료원을 활용하여 섬망을 조작적으로 정의하는 방법을 시도하였다. 서울대학교병원에서 중환자 섬망 여부 사정 도구로 사용하고 있는 CAM-ICU와 더불어 haloperidol 처방 여부, 섬망 관련 정신건강의학과 타과의뢰 작성 여부, 간호일지 섬망 진술문 유무, 그리고 섬망 진단명 유무를 활용하여 섬망 환자를 선별할 것이다.

제 2 절 섬망의 위험요인

흔히 언급되는 섬망의 위험요인으로는 연령, 진정제 및 진통제의 사용 여부, 인공호흡기의 사용 여부, 대사성 산증 여부, 응급 입원 여부, 수술 여부, 중환자실 재실 기간, 혈중요소질소/크레아티닌(BUN/Cr) 수치 등이 있다(Ruppert et al., 2020). 그 외에 입원 전 알코올 섭취량, 동반질환 여부, 유치도뇨관 및 억제대 사용 여부, 감염 여부 등도 섬망과 관련이 있으며 이들의 복합적인 작용으로 인해 섬망이 유발된다(Hshieh et al., 2017).

본 연구에서는 여러 가지 위험요인을 통해 환자가 수술 직후 중환자실에 입실하는 시점에서 남은 재실 기간 중 섬망 증세를 보일 지를 예측하는 모델을 개발하고자 한다. 따라서 위험요인 변수를 선정할 때 중환자실 재실 기간 내의 기록으로 산출되는 변수보다는 중환자실 입실 전에 값이 결정되는 변수를 중심으로 포함시켰다. 수술 전 결정되는 변수, 수술 자체와 관련된 변수와 더불어 수술 후 중환자실 입실 시점에 측정되었던 활력징후 및 검사결과까지만 고려하였다.

섬망의 명확한 기전은 밝히기 어렵지만 섬망 증세의 발생은 혈액뇌장벽(Blood-brain barrier, BBB) 투과성의 변화를 동반하는 급성 염증 반응과 관련이 있는 것으로 알려졌다(Maldonado, 2013). 이 과정에는 섬유소원(Fibrinogen)도 중요하게 관여를 한다(Terrando et al., 2022). 염증 반응과 관련된 지표로는 C-반응성 단백질(C-reactive protein, CRP)이나 백혈구 수치 등을 볼 수 있다.

수술과 관련된 여러 요인들은 서로 복합적으로 작용하여 급성 염증 반응에 큰 원인을 제공하는데 (Cerejeira et al., 2010), 특히 심폐기의 사용은 대규모의 전신적 염증 반응을 초래하기에 심장수술 환자는 높은 섬망 발생 위험을 가진다 (Jarvela et al., 2018). 따라서 심폐기 작동 여부 및 시간을 섬망 위험요인으로 포함시킬 수 있으며, 심폐기 사용에는 체온을 낮추는 과정이 포함되어있기 때문에 수술 중 및 수술 직후 체온도 관련된 변수로 생각할 수 있다(Smulter et al., 2013). 심폐기의 사용은 혈액 내

응고 관련 인자들의 수치에도 큰 영향을 주는데, 특히 프로트롬빈 시간(Prothrombine Time, PT), 섬유소원과 혈소판 수치가 영향을 받는 것으로 밝혀졌다(Hofer et al., 2016).

Rudiger 등(2016)은 심장수술 자체와 관련된(intraoperative) 변수들에 대해 단변량 분석을 시행했을 때 심폐기 사용 시간 ($p<0.001$), 최저 평균 동맥혈압($p=0.004$), 최저 헤모글로빈 수치($p=0.005$), 최저 체온($p=0.035$) 등이 수술 후 섬망 발생과 유의한 관련을 가짐을 보였다. 또한 혈액검사결과를 입실 직후, 입실 1 일차, 입실 2 일차의 세 시점에서 비교하였을 때, 입실 직후의 수치들이 섬망군과 비섬망군 사이에 가장 통계적으로 유의한 차이를 나타냄을 보였다. 입실 직후의 젖산($p<0.001$), Alanine aminotransferase ($p=0.024$), CRP ($p=0.025$), 백혈구 ($p=0.035$) 및 헤모글로빈($p=0.047$)이 해당된다. 이는 심장수술 후 환자의 혈액검사 수치를 변수로 사용할 때 여러 개의 검사 시점 중에서 입실 직후의 수치를 선택하는 것이 합당함을 뒷받침한다.

섬망의 발생 기전에 대한 가설 중에는 조직손상, 저산소증, 감염 등으로 인한 산소의 공급 대비 수요 증가를 대뇌 기능 저하 등의 원인으로 주장하는 것도 있다(Maldonado, 2013). 최저 평균 동맥혈압 및 최저 헤모글로빈 수치가 이와 관련 있다. 섬망의 발생은 수술 중 실혈량 및 수술 직후 헤모글로빈 수치(Marcantonio et al., 1998), 수술 후 대뇌 산소포화도(Eertmans et al., 2020)와도 관련이 있다고 밝혀졌다.

한편, 관상동맥 우회 수술 환자를 대상으로 전향적으로 시행했던 Zhang 등(2015)의 연구에서 수술 후 전해질 불균형은 섬망군에서 1.5 배 더 많이 나타났는데($p=0.0003$), 칼륨, 나트륨, 칼슘 및 마그네슘 수치는 심장기능과 관련이 크기 때문에 심장수술 후 불균형이 발생하는 경우가 많다. 혈청 전해질 불균형 및 신장기능 저하는 신경전달물질의 전달에 영향을 미쳐 섬망 증세 발생에 영향을 주는 것으로 추정된다(Marcantonio et al., 2006).

성별이나 체질량지수는 연구에 따라 섬망 발생과의 관련성 정도가 다르게 보고된다. Bakker 등(2012)의 심장수술 후 섬망 연구에서는 관련성이

낮은 것으로 밝혀졌지만, Kim 등(2023)의 연구에서는 반대의 결과를 시사한다. 여성일수록, 체질량지수가 낮을수록 섬망 발생 위험이 높다는 주장이 있다. 결과는 다를 수 있지만, 성별과 체질량지수는 대상자의 일반적인 특성과 관련된 기본정보로서 대부분의 연구가 다루고 있다.

Bakker 등의 연구에서 prehospitalization 여부가 섬망군(19%)과 비섬망군(16%) 간에 유의한 차이를 보였고(p=0.482), 이전 심장수술 여부 또한 섬망군(16%)과 비섬망군(5%) 간 차이가 있었다(p=0.014).

환자들의 영양상태와 관련된 지표 중에는 혈청 알부민 수치가 심장수술 후 섬망 발생의 중요한 예측 인자로 작용한다는 연구가 있다(Kim et al., 2023). 같은 연구에서 총 콜레스테롤 지수나 이상적 몸무게 대비 실제 몸무게 비율은 섬망의 발생과 관련이 적었다.

본 연구에서는 국내외의 문헌고찰을 바탕으로 섬망의 위험요인으로 가능한 중요 변수들을 선별하였다. 섬망의 위험요인 중 수면의 질(Van Rompaey, 2012) 등 연구 병원의 전자의무기록에서 관련 데이터를 추출할 수 없는 요인은 대상에서 제외하였다. 추후 개발될 모델의 활용성을 고려하여 연구 병동에 수술 후 입실하는 모든 환자들을 대상으로 일상적으로 시행되어 결측치 개수가 적은 검사들의 수치를 위험요인 변수로 포함시키고자 했다.

본 연구에서 수술 후 섬망 예측모델을 개발하는 데 있어서 크게 세 종류의 변수를 사용했다. 첫째로, 수술 전에 결정되는 변수이다. 성별, 연령, 체질량지수 등의 기본정보와 수술 전 입원 기간, 수술 전 입실 여부를 변수로 선정하였다. 두 번째로는 수술 자체와 관련된 변수이다. 수술의 종류, 응급수술 여부 및 재수술 여부, 수술 시간, 심폐기 사용 여부 및 사용 시간, 그리고 예상 실혈량(Estimated Blood Loss, EBL)의 값을 선정했다. 마지막으로는 수술 직후 획득되는 변수로 수술 직후 중환자실에 입실하는 시점에 측정된 활력징후와 채혈한 혈액의 검사결과를 포함했다. 활력징후는 본래 체온, 호흡수, 심박수, 혈압의 네 가지를 의미하는데, 전신마취 수술 직후에는 환자가 인공호흡기에 의존하며 자가호흡을 못하기 때문에, 호흡수는 제외하고 산소포화도를 포함시켰다. 혈액검사 수치로는

전혈구 검사(Complete Blood Count)의 백혈구, 헤모글로빈, 혈소판, 응고검사의 섬유소원(Fibrinogen), 활성화 부분 트롬보 플라스틴 시간(activated Partial Thromboplastin Time, 이하 aPTT), 프로트롬빈 시간(Prothrombin Time, 이하 PT), 혈청검사의 칼륨, 나트륨, 칼슘, 마그네슘, 인, 혈중요소질소(Blood Urea Nitrogen, 이하 BUN), 크레아티닌, Aspartate aminotransferase(이하 AST), Alanine aminotransferase (이하 ALT), 총 빌리루빈(Total Bilirubin), Alkaline phosphatase (이하 Alk.phos), 알부민, C-반응성 단백질(C-Reactive Protein, 이하 CRP), 그리고 동맥혈가스 검사의 산도(pH), 산소 분압(pO_2), 이산화탄소 분압(pCO_2), 중탄산이온(HCO_3^-), 혈당 및 젖산을 포함시켰다. 이상으로 총 46 개의 변수를 선정하였다.

제 3 절 섬망의 예측모델

1) 예측모델 알고리즘

본 연구에서는 로지스틱 회귀, 의사결정나무, 랜덤포레스트, 그리고 XGBoost 알고리즘을 각각 사용하여 예측모델을 개발한 후 성능을 비교하고자 한다. 섬망의 병태생리가 명확하게 밝혀지지 않고 수많은 위험요인들만 알려진 상황에서는, 복잡한 연산에 능한 기계학습(Machine Learning) 방법이 예측모델 개발에 적합하다(Jauk et al., 2020). 로지스틱 회귀 방법을 이용하여 섬망 위험을 예측하려는 시도는 꾸준히 있어 왔는데, 최근 연구들에 따르면 최신의 기계학습 방법들이 로지스틱 회귀 보다 더 좋은 성능을 보이고 있다(Lindroth et al., 2018). Hur 등(2021)이 The Medical Information Mart for Intensive Care III (MIMIC-III) 데이터베이스를 기반으로 여러 가지 기계학습 방법으로 섬망 예측모델을 개발했을 때, 랜덤포레스트 및 XGBoost 등의 모델들이 로지스틱 회귀 모델보다 더 높은 민감도, 특이도, 양성 예측도, 음성 예측도 및 AUROC (Area Under Receiver Operating Characteristic Curve) 값을 보였다. 심장수술 후 섬망을 특정하여 기계학습 기법에 따른 예측모델 성능을 비교한 연구는 소수에 불과하다.

본 연구에서 개발하는 심장수술 후 섬망 예측모델의 종속변수는 '섬망 발생 여부' 이며, 1 (섬망 발생) 또는 0 (섬망 미발생)의 값을 가진다. 각 모델은 섬망 발생 확률값을 계산하는데, 그 확률이 0.5 보다 크면 1 (섬망 발생)의 값을 가진다고 처리한다.

중환자 섬망 예측모델에 대한 체계적 문헌고찰 연구에서 대부분의 모델이 로지스틱 회귀를 사용한 것을 확인할 수 있다(Ruppert et al., 2020). 로지스틱 회귀는 전통적인 통계학적 방법에 근거하나, 최근 컴퓨터의 연산 능력 향상으로 이전에 다루지 못했던 더 많은 변수에 대한 계산을 빠르게 할 수 있게 되었기 때문에 기계학습 방법 중의 하나로 다뤄지기도 한다. 결과변수가 0 또는 1 의 값을 가지는 이분형 변수일 때의 분류기(classifier) 개발에 사용되는 기법이다. 종속변수가 1 일 확률이 p

$(0 \leq p \leq 1)$일 때, $p/(1-p)$을 비차비(odds)라고 하는데, 이에 로그를 취하는 로짓(logit) 변환을 거치면 $(-\infty, \infty)$의 범위를 가지게 된다. 여기에 독립변수들을 선형 결합 형태로 나타내어 선형 회귀모델처럼 계산할 수 있다. 그러면 독립변수들의 값을 통해 종속변수가 1 일 확률을 계산할 수 있다. Nusinovici 등(2020)이 20 개 정도의 예측변수를 사용하여 심혈관질환, 만성 신질환, 당뇨 및 고혈압을 각각 예측하는 모델을 개발하는 연구에서, 적당한 개수의 서로 구별되는 변수들에 대해서는 로지스틱 회귀가 타 기계학습 방법보다 나은 성능을 보일 수 있다고 결론지었다.

한편, 대표적인 기계학습 방법인 의사결정나무는 말 그대로 나무에서 가지가 자라나듯, 예시들이 여러 분기점을 거치며 분류될 때 특정 종속변수에 대해 가장 순도(purity)가 높게 분류되게 하는 방법을 찾는다. 의사결정나무의 가장 대표적인 알고리즘으로는 C5.0 알고리즘이 있는데, 순도를 계산할 때 엔트로피를 사용한다. 분기점에서 분류의 기준으로 사용하는 특징을 고를 때, 전체 엔트로피 값이 많이 줄어들게 해서 분기점 이후의 그룹 각각의 내부 동질성이 증가하게끔 하는 최선의 특징을 선택한다. 분기점에 사용되는 특징은 명목 변수뿐만이 아닌 연속 변수 또한 가능하다. 연속변수일 때는 어떤 수치를 기준으로 나누어야 더 깔끔하게 분류할 수 있는지를 정해준다. 한편, 가지의 수가 너무 많으면 의사결정나무는 훈련 데이터에 대해 과적합되기 쉽다. C5.0 알고리즘은 우선 가지를 아주 많이 만든 후, 결과에 영향이 없는 분기를 제거하는 식의 사후 가지치기 방법을 사용한다. 기계학습 방법 중 의사결정나무는 높은 해석력(Perceptive interpretability)을 보이며 결과가 도출되기까지의 과정을 일반적인 언어로 설명할 수 있다는 장점을 가진다.

랜덤포레스트 와 XGBoost 방법은 여러 개의 의사결정나무를 각자 다른 원리로 조합하여 사용하는 앙상블 알고리즘이다. 의사결정나무는 다양한 기계학습 기법의 기초가 되는데, 이러한 기법들을 'tree based' 알고리즘이라 일컫는다. 랜덤포레스트와 XGBoost 는 의사결정나무보다 해석력은 떨어지지만, 쉽게 과적합되지 않고 높은 예측 성능을 보인다.

랜덤포레스트는 Breiman (2001)이 제안한 방법으로 간단한 의사결정나무를 수백 또는 수천 개 만들어 예측 결과를 다수결 투표로 정하는 방법이다. 사용하는 앙상블의 기본 원리는 'Bagging (Bootstrap Aggregating)'이다. 의사결정나무를 하나 생성할 때마다 원래의 훈련 데이터에서 복원추출 방법으로 같은 크기의 새로운 훈련 데이터를 생성한다. 이러한 방식을 Bootstrap 샘플링이라고 한다. 나무 생성과정에서 분기점에 사용할 특징을 선택할 때, 원래의 의사결정나무 기법과는 다르게 전체 특징을 후보로 두지 않고, 그 중 사전에 정해 둔 개수만큼의 특징을 랜덤으로 선택하여 후보로 둔다. 보통 전체 특징 수의 제곱근에 해당되는 개수만큼의 랜덤 특징을 고려한다. 생성된 모든 의사결정나무의 결과를 합산(aggregating)하여 평균치를 통해 최종 예측 확률을 예측한다. 사용하는 의사결정나무의 개수 및 분기점에서 선택할 특징 후보 수를 사용자가 지정할 수 있다. 임상 문제 관련 예측모델에 대한 체계적 문헌고찰에서 랜덤포레스트는 로지스틱 회귀와 더불어 가장 많이 사용되는 기계학습 방법임이 확인되었다(Christodoulou et al., 2019).

XGBoost (eXtreme Gradient Boosting)는 비교적 가장 최근에 소개된 기계학습 기법으로 많은 연구에서 그 예측 성능을 널리 인정받고 있다. 2015 년 국제 인공지능 개발 경진대회인 Kaggle 에서 29 개의 우승 솔루션 중 17 개가 XGBoost 를 사용한 것으로 공개되며 그 우수성이 주목받기 시작했다(Chen et al., 2016). 섬망과 관련된 연구도 여럿 진행되었는데, 그 중 Bishara 등(2022)의 수술 후 섬망 예측모델 개발 연구에서 XGBoost 모델과 로지스틱 회귀 모델의 AUROC 는 각각 0.851, 0.762 를 보였다.

XGBoost 는 의사결정나무를 'Gradient Boosting'이라는 앙상블 알고리즘으로 조합한다. 앙상블을 구성하며 순차적으로 새로운 기본 모델을 추가할 때, 기존의 모델들로 산출한 예측 결과보다 결과가 더 정확해지는 방향으로 구성한다. 최종 결과값을 산출하기 위해서는 시작 모델의 예측값에 추가되는 모델들에서 계산되는 예측값과 실젯값 사이의 차이가 순차적으로 더해지는데, 보통 0.02 에서 0.2 사이의 값으로

지정되는 '학습률(Learning rate)'을 곱하여 더한다. 사용하는 의사결정나무의 수나 크기, 학습률 등을 사용자가 지정할 수 있다. 학습률이 낮고 사용하는 의사결정나무의 수가 많을수록 과적합을 방지하고 더 정확한 XGBoost 모델이 되는 것으로 알려져 있다.

2) 데이터 불균형 처리

예측모델을 개발할 때 소수의 예시들에 대한 학습이 충분히 이루어지게 하도록 Synthetic Minority Oversampling Technique (SMOTE) 기법을 이용할 수 있다. 기계학습 모델 분류기에 데이터를 학습시킬 때, 예측하려는 클래스의 비율이 너무 작으면, 테스트 데이터에서 낮은 민감도를 보이게 된다. 소수의 클래스에 대해 충분한 학습이 이루어지지 못하기 때문이다. SMOTE 는 2002 년에 처음 소개된 이후 이러한 불균형 데이터를 문제를 처리하는 방법으로 널리 활용되고 있다. 단순히 oversampling 하는 방법을 넘어서, 선형 보간법을 통해 소수의 클래스에 속하는 예시를 새롭게 합성하여 추가하는 방법이다. 소수 클래스에 속하는 예시와 그와 가장 유사한 다른 예시 사이에서 랜덤으로 새롭게 예시를 생성하는데, 가장 유사한 예시를 고를 때는 사전에 정해놓은 개수만큼의 가까운 예시 중에서 랜덤으로 하나를 선택한다. Chawla 등(2002)은 다양한 사례를 통해 SMOTE 의 우수성을 입증하였고, Fernández 등(2018)은 그 후 오랫동안 SMOTE 가 다양하게 변형되며 널리 쓰였음을 확인하였다.

제 3 장 연구의 틀

제 1 절 연구 의의

본 연구 수행의 의의는 Celi 등(2014)이 제시한 'Data-driven Care System' 모델로 설명할 수 있다. 이 모델은 실질적인 'Material energy flow' 이외 'Information flow'를 통해 임상현장에서의 데이터의 역할을 설명한다.

기본적으로 섬망과 같은 임상 문제(Clinical issue)와 의료진의 중재(Clinical practice)의 영향을 받는 환자 상태(State of patient)가 있고, 환자의 상태는 센서를 통해 각종 수치들로 컴퓨터 데이터베이스에 저장된다. 중재의 결과로서의 환자의 상태는 피드백(Feed back) 과정을 통해 또다시 의료진의 중재에 영향을 미친다. 이는 'Material energy flow'로 설명된다.

그런데 예측모델과 같은 데이터 기반 의사결정 지원이 갖춰져 있으면, 'Information flow'를 통해 임상 문제(Clinical issue)가 환자 상태(State of patient)에 영향을 미치기 전에 피드포워드(Feed forward) 과정을 통한 예방적 중재(Clinical practice)가 가능해진다. 이 모델을 통해 임상에서 예측모델의 중요성을 설명할 수 있다.

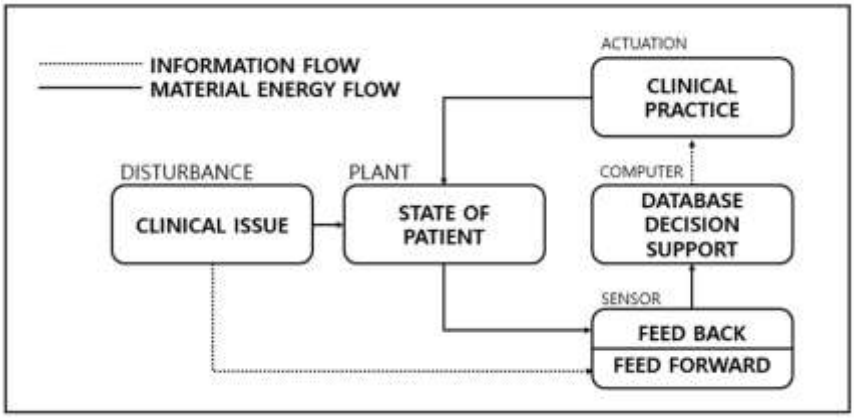

Figure 1. Data-driven Care system

제 2 절 연구 과정

본 연구의 전체적인 흐름은 Cross Industry Standard Process for Data Mining in Medicine (CRISP-MED-DM) 모델을 기반으로 설계하였다. 이 모델은 2000 년에 개발된 뒤 현재까지 데이터 마이닝 작업에 이론적 틀을 제공하는 데 가장 많이 사용되고 있는(Schröer et al., 2021) Cross Industry Standard Process for Data Mining (CRISP-DM)을 Olegas (2015)가 의학 분야에 맞춰 변형한 모델이다. 원 모델과 마찬가지로 CRISP-MED-DM 은 총 6 단계로 구성이 되어있다. 이 단계에 맞춰 본 연구의 흐름을 정리하면 다음과 같다.

첫 번째 문제 이해(Problem understanding) 단계는 해결하고자 하는 문제를 이해하고 데이터 분석의 목표를 분명히 하는 단계로, 본 연구에서는 연구의 필요성을 서술하고 관련된 문헌을 고찰하는 부분이 해당된다. 두 번째로 데이터 이해(Data understanding) 단계에서는 기관의 승인을 받아 사전에 정의한 연구 대상자들의 전자의무기록에 접근한 뒤 원시자료의 특성을 이해한다. 다음으로 세 번째는 데이터 준비(Data preparation) 단계로, 데이터 분석을 위해 원시자료를 가공한다. 결측치 및 이상치를 정제하고 최종 분석이 가능하도록 전처리 하는 과정을 포함한다. 네 번째인 모델링(Modeling) 단계에서는 가공된 데이터세트를 기반으로 예측모델을 개발하고, 민감도, 특이도, 양성 예측도, 음성 예측도 및 AUROC (Area Under Receiver Operating Characteristic Curve) 등의 수치를 계산한다. 개발된 모델들의 성능을 비교하며 최적의 모델을 선정한다. 다섯 번째로 평가(Evaluation) 단계에서는 본 연구의 목적이 이루어졌는지 점검하고 전체 연구과정에서 얻은 교훈 및 한계점에 대해 논의한다. 마지막 여섯 번째인 배치(Deployment) 단계에서는 선정된 최종 모델이 연구병원의 전자의무기록시스템에 적용될 수 있는 방법을 제시하는 것인데, 본 연구에서는 생략하였다.

제 4 장 연구 방법

제 1 절 연구 설계

본 연구는 전자의무기록을 활용한 후향적 관찰 연구(Retrospective observational study)이다. 기존에 입원했던 환자들의 기록을 사용하여 특정 결과변수의 확률을 예측하는, 기계학습 분류기(Machine learning classifier)를 개발한다. 이를 통해 추후 새로운 환자의 섬망 위험을 예측하는 것을 목적으로 한다. 이러한 측면에서는 예측적 상관관계 연구(Predictive correlational design)라고 할 수 있다.

본 연구의 진행 방법은 다음과 같이 요약할 수 있다. 우선 연구 대상자 조건에 맞는 환자들의 데이터를 선별한다. 그 후 결과변수인 섬망 발생 여부를 파악하여 데이터를 라벨링 한다. 여기에 섬망의 위험요인 변수들의 데이터를 추가하여 분석할 데이터세트을 완성한다. 이후 결측치 처리, 이상치 처리 및 연속변수 표준화 등의 데이터 전처리 과정을 수행한다. 라벨링과 전처리가 완료된 최종 데이터세트를 가지고, 지도학습(Supervised learning) 방법의 기계학습 방법으로 예측모델을 개발한다. 이 때 위험요인들에 대해 단변량 분석을 시행하여 섬망군과 비섬망군 간에 통계적으로 유의한 차이를 보이지 않는 요인은 포함시키지 않는다. 반복적 K-겹 교차검증을 적용하여 모델의 성능을 기술한다. 기법의 종류 등에 따라 달라지는 결과변수 예측력을 비교하고 평가하여, 최적의 예측모델을 선정한다

제 2 절 자료 수집

본 연구에서 사용한 전자의무기록자료는 대한민국 서울특별시 종로구에 위치한 서울대학교병원의 자료이다. 해당 병원의 Hospital Information System (이하 HIS)은 이지케어텍 사의 'Best Care 2.0'을 사용한다. 연구에 사용할 변수를 선정할 때 HIS 를 통해 접근할 수 있는 전자의무기록지인 '임상관찰기록지', '간호일지', '수술기록', '마취기록' 등에서 관련 자료를 수집할 수 있는지를 우선적으로 고려하였다. 서울대학교병원의 Clinical Data Warehouse (CDW) 연구 검색 시스템인 Seoul National University Hospital Patient Research Environment (이하 SUPREME)를 통해 사전에 승인 받은 연구자가 직접 원하는 HIS 상의 전자의무기록자료를 엑셀 형식으로 추출할 수 있다.

본 연구에 활용한 모든 전자의무기록자료는 SUPREME 1.0 (2018. 04.30~) 및 SUPREME 2.0 (2023.02.10~)에서 연구자가 추출하였다. 추출한 원시자료는 R (version 4.2.3; R foundation for Statistical Computing, Vienna, Austria)을 통해 가공했다.

제 3 절 대상자 선정

1) 연구 대상 병동 특성

본 연구는 연구 기간 내 서울대학교병원의 심폐기계중환자실에 입실한 환자의 수술 후 섬망 문제를 다룬다. 해당 중환자실은 총 8 병상이며, 성인 심장혈관흉부외과의 심장 파트 수술 후 관리를 담당한다. 심장혈관흉부외과는 심장 파트와 폐 파트 두 개로 나누어 운영되며, 전자는 심장 및 대동맥 질환 수술을 맡는다. 년간 약 700 명 이상의 환자가 입실하며, 평균적으로 약 4 일정도 머문다. 사망률 및 재입실률은 각각 약 2%정도이다. 해당 심폐기계중환자실은 단일 진료과가 운영하는 소수 병상의 중환자실로, 수술 후 환자에 대한 오더 세트 및 검사 루틴이 장기간 유지되는 특징을 가지고 있다.

일반적으로 입실하는 환자는 대개 기관삽관이 이루어진 상태로, 내경정맥 중심정맥관, 유치도뇨관, 1~2 개의 동맥관, 1 개 이상의 말초정맥관을 모두 가지고 있다. 보통 1~2 일 이내에 기관발관을 하는 경우가 대부분인데, 기관내관이 있는 상태에는 모든 환자에게 양 손목에 신체보호대가 적용되며 Remifentanil 50mcg/cc, 5~10cc/hr 및 Dexmedetomidine 4mcg/cc, 0.2~0.4 mcg/kg/hr 정도가 소량 투약된다. 기존의 많은 연구들이 연장된 인공호흡기 사용 기간, 억제대 사용, 진통제 및 진정제 사용 등이 섬망의 위험을 증가시키는 것을 입증하였다. 내과계 중환자들과는 다르게 심폐기계중환자실의 환자들은 대부분 유사한 인공호흡기 이탈 과정을 경험하기 때문에 위 변수들이 섬망 발생 여부 차이에 미치는 영향을 어느 정도 배제한 상태에서 다른 변수들의 영향을 살필 수 있을 것으로 기대된다.

또한 해당 중환자실은 연구자가 2018 년 4 월부터 2022 년 7 월까지 일반 간호사로서 근무하며 실제로 많은 환자의 임상 정보를 수집하고 전산에 기록했던 곳이다. 따라서 임상 전문가의 입장으로서 실제 현장에서 HIS 를 활용하는 실정을 고려할 수 있다.

2) 주요 심장수술 정의

연구 대상자 선정 과정 중 주요 심장수술을 분리하여 정의하는 과정이 필요했다. SUPREME을 이용해 연구 기간 내 심장혈관흉부외과 심장파트의 집도의들의 성명을 특정하여 4,050건의 수술기록을 추출하였다. 그러나 마취종류를 '전신', 퇴실장소를 '중환자실'이라고 특정하였음에도 'Mediastinal irrigation', 'Debridement of wound', 'Bleeding control', 'Sternal Wound closure', 'VA ECMO weaning' 등 주요 수술이라고는 간주하기 어려운 수술에 대한 기록이 무수하였다. 연구에서 다룰 수술을 관상동맥 수술이나 심장판막 수술 등의 대표적인 심장수술 유형 몇 가지로 한정하면 이러한 상황을 피하기 수 있다. 그러나 본 연구에서는 최대한 많은 수술 후 입실 건을 다루기 위해 기간 내 작성된 수술 기록지의 수술명 전수를 살피며 주요 수술에 포함시킬 수술명을 분리하였다.

본 연구에서는 수술기록지의 수술명을 기준으로 수술을 분류하였다. 수술기록지에는 Figure 2 에서와 같이 자유형식으로 입력된 수술명('수술명 추가 입력')과 함께 표준화된 수술명이 최대 5 개까지('1st 수술명', '2nd 수술명' 등) 선택되어 기록된다. 이를 차례로 자유수술명과 세부수술명으로 칭하겠다. 세부수술명에는 작성자 선택에 따라 구체적인 수술 부위나 사용한 재료 등의 속성을 함께 입력할 수 있다. 쉼표 뒤나 괄호 안에 표현되는 수술 속성 정보 고려하지 않았을 때, 앞서 추출한 4,050 건의 수술기록에 대해 총 213 개의 세부수술명이 확인되었다. 연구자가 직접 검토하여 58 개의 세부수술명을 주요 수술로 인정하였다(부록 2). 수술 속성 정보를 포함한 세부수술명 전체로는 총 379 건이 확인되었으며 이 중 195 건을 주요 수술로 간주하였다. 추가적으로 수술 부위 및 수술 빈도를 고려하여 앞서 인정한 주요 수술들을 대동맥 수술, 관상동맥 수술, 심장판막 수술, 부정맥 수술, 그리고 그 외의 수술이라는 다섯 종류로 나누었다(Table 1).

Table 1. Classification of Major Cardiac Surgeries with Examples

Surgery type		Number	(when including all details)	Examples of specific surgery names
AORTA	Replacement	2	(7)	'Replacement of aorta (descending thoracic)', 'Replacement of thoracoabdominal aortic aneurysm'
	Bentall op/ David op	2	(2)	'Composite graft interposition for annuloaortic ectasia (Bentall operation)'
	Repair/ Resection	5	(12)	'Repair of aneurysm (ascending aorta)', 'Resection of aneurysm, descending thoracic aorta'
	Stent	1	(1)	'Insertion of aortic stent'
	Bypass	4	(11)	'Aorta to other artery bypass (subclavian/ left/ artificial vessel graft)', 'Thoracic aorta to abdominal aorta bypass'
CORONARY	Off pump CABG	1	(5)	'Off pump CABG (coronary artery bypass graft) surgery (LAD/ LITA)'
	CABG	4	(4)	'Aortocoronary bypass of one coronary artery', 'Redo coronary artery bypass graft'
VALVE	Repair	4	(27)	'Open heart valvuloplasty without replacement (aortic valve)', 'Repair of heart valve (tricuspid valve/ valve)'
	Replacement	2	(26)	'Replacement of heart valve (aortic valve/ tissue valve)', 'Replacement of redo-valve (unspecified valve)'
	else	2	(2)	'Revision of prosthesis of heart valve', 'Operation of heart valve'
MAZE	Maze op	2	(3)	'Cox-Maze procedure', 'Maze operation (bilateral)'
ELSE	ASD	1	(12)	'Repair of atrial septal defect (secundum/ patch closure/ bioprosthesis)'
	VSD	1	(14)	'Repair of ventricular septal defect (primary closure)'
	Mass	12	(19)	'Excision of cardiac tumor (partial)', 'Excision of mediastinal mass (midline/ sternotomy)'
	Window op	1	(7)	'Pericardial window operation (bilateral/ VATS)'
	Vessel	4	(21)	'Thrombectomy (bilateral/ full sternotomy/ partial)', 'Operation for thromboembolism'
	Thymectomy	2	(13)	'Thymectomy (full sternotomy)', 'Subxiphoid approach thymectomy (bilateral / robotic/ total)'
	L-VAD	3	(4)	'Insertion of VAD (Ventricular assist device)', 'Implantation of heart assist device'
	Heart TPL	2	(2)	'Transplantation of heart', 'Transplantation of heart-lung'
	else	3	(3)	'Total correction of tetralogy of Fallot', 'Repair of heart'
	Total	58	(195)	op=operation; CABG= Coronary Artery Bypass Graft; LITA= Left Internal Thoracic Artery; LAD= Left Anterior Descending coronary artery; ASD= Atrial Septal Defect; VSD= Ventricular Septal Defect; VATS= Video-Assisted Thoracic Surgery; L-VAD= Left Ventricular Assist Device; TPL= Transplant

3) 연구 대상자 선정

본 연구는 2017 년 1 월 1 일부터 2021 년 12 월 31 일까지, 총 5 년 동안의 서울대학교병원의 심폐기계중환자실 입실 건을 대상으로 했다. 연구 대상자 선정 조건(inclusion criteria)과 제외 조건(exclusion criteria)은 Table 2 와 같다.

선정 조건으로는 입실 당시 기준 만 18 세 이상의 성인으로 주요 심장수술 직후 심폐기계중환자실에 입실한 환자로 설정하였다. 앞서 정한 [부록 2]의 주요 심장수술 목록을 기준으로 하였다. 주요 수술을 받기 전부터 중환자실에 입실해 있던 경우는 제외하지 않고 별도의 변수를 생성하여 처리하였다. 하나의 입실 건에 주요 수술 2 개 이상이 시행된 경우는 대상자에서 제외하였다. 추가로 재실 기간 내내 RASS (Richmond Agitation−Sedation Scale) 점수가 −4~−5 였던 경우, 최근 2 년 이내 해당 중환자실에 입실한 이력이 있는 경우, 그리고 재실 기간이 24 시간이 되지 않는 경우도 제외하였다.

중환자실 재실 기간 내내 RASS 점수가 −4 (물리/통증 자극에만 반응하며 눈을 뜨지 않는 경우) 또는 -5 (자극에 아무 반응하지 않는 경우)인 환자는 뇌손상 및 진정치료 등의 이유로 섬망 평가가 불가능한 상태이다. 임상관찰기록지의 RASS 점수 자료를 추출하여, RASS 점수 −4 또는 −5 에는 ‘0’, 그 외 RASS 점수 −3~+5 에는 ‘1’이 대응하는 변수인 ‘rass2’를 새로 생성한 뒤, 각 입실 건에 대해서 존재하는 모든 rass2 값의 합이 0 인 입실 건을 분석에서 제외했다.

또한 섬망과 관련된 외생변수의 영향을 최소화하기 위해, 최근 2 년 이내 재입실한 경우를 연구 대상에서 제외했다. 2015 년부터의 중환자실 입퇴실시간 자료를 확보한 뒤, 각 환자별로 각 입실 건에 대해 최근 2 년 이내 입실한 횟수를 코딩하여 처리하였다.

마지막으로 중환자실 입실시각과 퇴실시각의 차이가 24 시간 이내인 자료를 분석에서 제외하였다. 본 연구는 재실 기간 내 섬망 관련 기록이 하나라도 있으면 섬망 환자라고 조작적으로 정의하고 있는데, 24 시간은 실제 섬망 환자가 섬망이라고 기록되거나 관찰되기에 부족한 시간이라

판단했다. Ko 등(2019), Simeone (2018) 과 이현심(2013) 등 중환자실 섬망과 관련된 다른 많은 연구들에서도 재실 기간이 24 시간 이내인 경우는 연구 대상에서 제외하고 있다.

Table 2. Inclusion and Exclusion Criteria

Inclusion criteria	Exclusion criteria
· Age ≥ 18 years · Admitted to the Intensive Care Unit (ICU) after a major cardiac surgery	· Received 2 or more major cardiac surgeries before discharged · RASS −4~−5 throughout the whole ICU stay · History of admission in the same ICU within last 2 years · ICU stay less than 24 hours

제 4 절 섬망 환자 정의

본 연구에서 정의한 섬망 환자는 중환자실에 있던 동안 섬망에 대한 다섯 가지 조건(조건 1~5) 중 하나를 만족시킨 환자를 의미한다. 조건 1 에 대해서는, 중환자실 재실 기간 중 작성된 임상관찰기록 중 최종 CAM-ICU 결과가 양성인 것이 하나라도 존재하는 경우로 구체화했다. 조건 2 에 대해서는, 간호일지의 표준 진술문 중에 '섬망 있음' 이나 '수술 후 섬망 보임'이 사용되었으면서 그 작성 일시가 입실 날짜와 퇴실 날짜 사이인 경우로 구체화했다. 조건 3 에 대해서는, 환자의 진단명 리스트 중에서 단어 'Delirium'이 포함된 진단명이 존재하면서 그 진단 일자가 입실 날짜와 퇴실 날짜 사이인 경우로 구체화했다. 조건 4 는 재실 기간 중 심장혈관흉부외과 주치의가 정신건강의학과에 섬망과 관련된 타과의뢰를 작성한 경우를 의미한다. 조건 5 에 대해서는, 약 처방기록 중 약품명(성분명)에 'Haloperidol 5mg/1ml inj'의 기록이 존재하며, 그 처방일이 입실 날짜와 퇴실 날짜 사이인 경우로 정의했다.

조건 4 에 해당하는 입실 건을 가려내기 위하여 섬망과 관련된 타과의뢰를 작성한 경우를 정성적으로 판별해야 했다. 이에 비해 다른 네 개의 조건들은 추가적인 작업 없이 섬망 여부를 명확하게 가려낼 수 있었다. SUPREME 에서 전체 대상자들의 '병원공통_타과회신_FREE TEXT' 서식 관련 자료를 전부 추출한 뒤, 심폐기계중환자실 재실 기간 내에 정신건강의학과 의사가 작성한 회신을 일차적으로 선별하였다. 그런데 정신건강의학과 회신은 정신질환의 증상 정도를 나타내는 척도인 Clinical Global Impression (CGI)을 포함한 '정신건강의학과_타과회신_CGI' 서식을 사용하는 경우도 있어서 두 가지 모두를 고려해야 했다. 이차적으로는 서식의 내용에 '섬망', 'delirium', 'delirious', 또는 'haloperidol' 단어가 포함된 것들의 내용을 검토하였는데, 그 중 섬망 관련 과거 history 혹은 관련 증세 없음에 대한 내용이었던 경우를 제외하였다.

제 5 절 섬망 위험요인 정의

문헌고찰 이후 섬망 예측모델에 사용될 예측변수 후보로 선정한 46 개 변수 목록은 Table 3 와 같다. 총 11 개의 이산변수(이분형)와 35 개의 연속변수가 해당된다. 각 변수를 정의하고 관련 자료를 추출할 수 있는 전자의무기록 내 자료원을 확인하였다.

수술 전 결정되는 변수 중 연령은 중환자실 입퇴실정보의 입실 날짜와 기본정보의 생년월일을 조합하여 입실 당시의 만 나이로 반영하였고, 체질량지수의 계산에는 임상관찰기록지의 수술 직전의 몸무게를 사용하였다. 수술 전 입원 기간은 중환자실 입퇴실정보의 입실날짜와 수진정보의 입원날짜와의 차이를 계산하여 생성하였다. 수술 전 재실 여부는 수술 전 심폐기계중환자실에 입실해있던 경우에 1의 값을 가진다. 원래는 입실날짜와 수술목록의 수술날짜의 차이를 계산하여 연속변수로 사용하려 하였으나 수술 전 중환자실에 입실해 있던 대상자의 수가 88명에 지나지 않아 이분형 변수로 바꾸어 처리하였다.

수술 자체 관련 변수 중 수술명과 관련된 다섯 개의 변수(AORTA, CORONARY, VALVE, MAZE, ELSE)는 Table 1의 분류를 활용하였다. 입실 직전에 시행된 수술의 세부진단명 중 하나라도 위 다섯 분류 중 하나에 속한다면, 해당 변수의 값은 1로 처리했다. SUPREME에서 추출한 수술기록 원시자료를 최종 데이터세트의 수술명 변수로 가공하는 과정을 Figure 2에서 하나의 예시를 들어 부연하였다. 대표 수술명 1개가 아닌 세부수술명 5개를 모두 활용하는 것은 복합수술이 많은 심장혈관흉부외과의 특징을 적절하게 반영한다는 장점을 가지고 있다.

재수술 여부는 수술기록지 내 표준화된 세부수술명 또는 자유형식 수술명에 ‘cardiac redo operation’ 또는 ‘replacement of redo-valve’와 같이 ‘redo’ 단어를 포함하거나, 자유형식 수술명에 ‘3rd OHS’에서처럼 개흉술(Open Heart Surgery)을 의미하는 ‘OHS’ 단어를 포함한 경우에 1의 값이 부여되었다. ‘응급수술 여부’는 전산상 수술목록에 기록된 수술 유형이 ‘정규’가 아닌 ‘응급’이거나 수술기록지에서 자유형식으로 기록되는

수술명에 ‘emergency’, ‘em’, ‘urgent’ 또는 ‘urgency’의 단어를 포함한 경우에 1의 값이 부여되었다. 실제로 의무기록을 살펴보다 보면 자유형식으로 기록되는 항목에서 더 정확하고 자세한 정보가 담겨 있는 경우가 많다. 특정 단어를 포함하는지를 판단하는 수준의 간단한 자연어 처리를 통해 주어진 자료를 최대한 활용하고자 했다.

심폐기 사용 여부 및 사용 시간은 체외순환기록지를, 예상 실혈량은 마취기록지를 활용하였으며, 수술 시간은 수술목록의 수술 시작 시간과 수술 종료 시간의 차이를 계산하여 구하였다.

수술 직후 중환자실에 입실하여 측정되는 체중 및 활력징후는 임상관찰기록지의 자료를 사용하였다. 수술기록지의 수술 종료 시각과 임상관찰기록지의 측정 시각을 비교하여, 수술 종료 후 가장 먼저 측정된 체중 값과 체온 값을 선택하였다. 평균 동맥혈압, 심박수 및 산소 포화도의 경우에는 가장 먼저 측정된 값을 포함하여 측정 시간이 빠른 순서대로 세 개를 추출 한 뒤 중앙값을 선택하였고, 특히 평균 동맥혈압은 자료 추출 시 추가적으로 범위를 20이상 200이하로 설정하여 이상치의 영향을 최소화했다. 체중 및 체온은 간호사가 직접 입력한 값인 반면 나머지 세 변수는 측정 장비에서 자동으로 연동되는 수치들이기 때문이다. 환자가 수술 후 입실하여 중환자실의 모니터링 장비가 새롭게 연결되고 몸이 닦이고 의복이 입혀지는 과정에서 발생하는 움직임으로 인해 비정상적인 측정값이 빈번하게 연동된다. 심지어 평균 동맥혈압의 경우에는, 환자가 수술 후 입실한 직후 바로 말초동맥관에서 채혈이 이루어지기 때문에 채혈하는 동안 비정상적으로 높은 수치가 전산에 연동되는 경우가 많다. 이러한 변수별 특성을 고려하여 자료를 수집하였다.

산소포화도 이상 여부는 수술 직후 측정된 산소포화도가 98% 미만인 경우에 1의 값을 가진다. 수술 직후 산소포화도 측정값은 보통 100%의 값을 가지기 때문에 연속변수보다 이산변수로 활용하는 것이 적절하다.

Table 3. List of Risk Factors for Delirium

Variable	Variable type	Data source
1. Preoperative variable		
Sex (1='Male')	discrete (binary)	basic info.
Age (years)	continuous	basic info., admission data
Body Mass Index (kg/㎡)	continuous	clinical observation sheet
Pre ICU admission (1='Yes')	discrete (binary)	surgery list, admission data
Hospitalization before the surgery (days)	continuous	basic info., admission data
2. Intraoperative variable		
Surgery type		
Aorta op. (1='Yes')	discrete (binary)	surgical record
Coronary artery op. (1='Yes')	discrete (binary)	surgical record
Valve op. (1='Yes')	discrete (binary)	surgical record
Maze op. (1='Yes')	discrete (binary)	surgical record
else (1='Yes')	discrete (binary)	surgical record
Redo op. (1='Yes')	discrete (binary)	surgical record
Emergency op. (1='Yes')	discrete (binary)	surgical record, surgery list
CPB use (1='Yes')	discrete (binary)	ECC record
CPB time (min)	continuous	ECC record
Operation time (min)	continuous	surgery list
EBL (ml)	continuous	anesthesia record
3. Immediate postoperative variable		
Body weight gain (kg)	continuous	clinical observation sheet
Vital sign		
SpO_2 abnormality (1='<98%')	discrete (binary)	clinical observation sheet
Body temperature (℃)	continuous	clinical observation sheet
Heart rate (bpm)	continuous	clinical observation sheet
mean ABP (mmHg)	continuous	clinical observation sheet
Arterial Blood Gas analysis		
pH	continuous	laboratory result
pO_2 (mmHg)	continuous	laboratory result
pCO_2 (mmHg)	continuous	laboratory result
HCO_3^- (mmol/L)	continuous	laboratory result
Complete Blood Count		
WBC ($\times 10^3$/ul)	continuous	laboratory result
Hemoglobin (g/dL)	continuous	laboratory result
Platelet ($\times 10^3$/ul)	continuous	laboratory result

Coagulation test		
Fibrinogen (mg/dL)	continuous	laboratory result
aPTT (sec)	continuous	laboratory result
PT (INR)	continuous	laboratory result
Other blood serum test		
Potassium (mmol/L)	continuous	laboratory result
Sodium (mmol/L)	continuous	laboratory result
Calcium (mg/dL)	continuous	laboratory result
Magnesium (mEq/L)	continuous	laboratory result
Phosphorus (mg/dL)	continuous	laboratory result
BUN (mg/dL)	continuous	laboratory result
Creatinine (mg/dL)	continuous	laboratory result
AST (IU/L)	continuous	laboratory result
ALT (IU/L)	continuous	laboratory result
Total Bilirubin (mg/dL)	continuous	laboratory result
Alk.phos (IU/L)	continuous	laboratory result
Albumin (g/dL)	continuous	laboratory result
CRP (mg/dL)	continuous	laboratory result
Glucose (mg/dL)	continuous	laboratory result
Lactate (mmol/L)	continuous	laboratory result

Info.=information; ECC= Extracorporeal Circulation; op.=operation; CPB= Cardiopulmonary Bypass; EBL= Estimated Blood Loss; ABP= Arterial Blood Pressure; pO_2= partial pressure of oxygen; pCO_2= partial pressure of carbon dioxide; HCO_3^- = Bicarbonate ion; WBC= White Blood Cell; aPTT = activated Partial Thromboplastin Time; PT= Prothrombin Time; BUN= Blood Urea Nitrogen; AST= Aspartate aminotransferase; ALT= Alanine aminotransferase; Alk.phos= Alkaline phosphatase; CRP= C-reactive protein

Raw data format (surgical records)

(환자 ID)	(기록시간)	수술명 추가입력 <- 병원공통_수술기록_FREE TEXT	TAP with Tri Ad ring 28mm under De Vega annuloplasty, MV repair (secondary chordae resection in A1, A2), Cox-maze IV op. (Cryocath) with LAA ligation
(환자 ID)	(기록시간)	5th 수술명 <- 병원공통_수술기록_FREE TEXT	Cardioplegia
(환자 ID)	(기록시간)	4th 수술명 <- 병원공통_수술기록_FREE TEXT	Extracorporeal circulation auxiliary to open heart surgery
(환자 ID)	(기록시간)	3rd 수술명 <- 병원공통_수술기록_FREE TEXT	Cox-Maze procedure
(환자 ID)	(기록시간)	2nd 수술명 <- 병원공통_수술기록_FREE TEXT	Open heart valvuloplasty without replacement (mitral valve / valve)
(환자 ID)	(기록시간)	1st 수술명 <- 병원공통_수술기록_FREE TEXT	Open heart annuloplasty of tricuspid valve

		op.free	op1	op2	op3	op4	op5
(환자 ID)	(기록시간)	TAP with Tri Ad ring 28mm under De Vega annuloplasty, MV repair (secondary chordae resection in A1, A2), Cox-maze IV op. (Cryocath) with LAA ligation	Open heart annuloplasty of tricuspid valve	Open heart valvuloplasty without replacement (mitral valve / valve)	Cox-Maze procedure	Extracorporeal circulation auxiliary to open heart surgery	Cardioplegia

Final dataset format

	AORTA	CORONARY	MAZE	VALVE	ELSE
(환자 ID)	0	0	1	1	0

Figure 2. Data Preprocessing Example of Surgery Name Variables

제 6 절 자료 전처리

1) 결측치 처리

결측치 처리에는 단순히 각 변수의 전체 평균값으로 대체하는 방법(imputation)을 택했다. 다만 입실 직후 검사결과 수치인 경우 전체의 평균값으로 대체하기 이전에, 기록을 검토하였을 때 입실 직전 측정된 수치가 있었으면 그 값으로 대체하였다.

2) 이상치 처리

본 연구에서는 35개의 연속변수에 대해 4분위수를 구한 뒤 '1사분위수 −1.5*사분위범위'를 하한으로, '3사분위수+1.5*사분위범위'를 상한으로 하여 이상치를 정의했다. 이상치는 가까운 상한 또는 하한의 경곗값으로 대체했다. 당뇨 여부를 예측하는 기계학습 모델을 개발한 Roy 등(2021)이나 Hasan 등(2020)의 연구에서도 같은 기준 및 방식을 사용했다. 의사결정나무 방법은 이상치의 영향을 적게 받지만, 로지스틱 회귀 등은 이상치 처리 여부가 최종 성능에 큰 차이를 줄 수 있다(Nyitrai et al., 2019). 이상치 처리는 기계학습 모델이 좀 더 일반화된 패턴을 학습할 수 있도록 한다.

$$\text{upper limit} = 3^{rd}\,\text{quartile} + 1.5 * \text{interquartile range}$$
$$\text{lower limit} = 1^{st}\,\text{quartile} - 1.5 * \text{interquartile range}$$

3) 연속변수 표준화

연속변수마다 다양한 단위를 사용하고 가지는 값의 범위가 다르기 때문에, 표준화하여 그 차이로 인한 영향력을 감소시키는 과정이 필요하다. 모든 연속변수에서 분포가 정규성을 띄는 것을 검정한 후, 평균 0 과 표준편차 1을 가지도록 표준화한 수치를 분석에 사용했다.

4) 데이터 불균형 처리

예측모델을 개발하기 직전에, 최종 데이터세트에 Synthetic Minority Oversampling Technique-Nominal Continuous (SMOTE-NC) 기법을 적용하여 섬망 환자와 비섬망 환자의 비를 1:1 정도로 맞춘 뒤 모델을 학습시켰다. 비섬망 환자의 데이터는 그대로 유지하면서 합성된 섬망 환자의 데이터를 추가하였다. SMOTE와 같은 기법을 활용하면 예측모델의 민감도를 증가시킬 수 있다. SMOTE-NC 기법은 소수 클래스에 대한 새로운 데이터를 합성할 때 명목변수의 형태를 그대로 유지할 수 있도록 고안된 방법으로, 초기에 연속변수만을 대상으로 개발되었던 SMOTE의 응용 형태이다(Chawla et al., 2002). 예를 들어 0 또는 1의 값을 가지는 성별 변수에 일반적인 SMOTE 기법을 적용할 경우 0.734 등의 값으로 합성이 되지만, SMOTE-NC 기법을 적용하면 0 또는 1의 값으로 합성이 된다.

제 7 절 자료 분석

예측모델 개발에 사용한 다양한 통계적, 기계학습 기법은 R 프로그램의 내장 함수 외 다양한 패키지를 활용하여 구현하였다. 연구 목적에 따른 구체적인 분석 방법은 다음과 같다.

1) 대상자의 일반적 특성에 대한 이산변수(빈도, 백분율) 및 연속변수(평균, 표준편차, 중앙값, 최솟값 및 최댓값)는 기술통계를 사용하여 제시하였다.

2) 섬망의 위험요인 변수 각각에 대해 단변량 분석을 시행하여 섬망군과 비섬망군 간에 통계적으로 유의한 차이를 보인 변수를 선별하였다(양측검정, 유의수준 0.05). 이산변수에 대해서는 카이제곱 검정을, 연속변수에 대해서는 정규성 검정 및 등분산 검정 후 이표본 T 검정을 시행했다.

3) R 의 caret 패키지를 이용하여 로지스틱 회귀(method='glm', family=bionomial()), 의사결정나무(method='C5.0'), 랜덤포레스트(method='rf') 및 XGBoost (method='XgbTree')의 기계학습 기법을 각각 구현하였다. 반복적 K-겹 교차검증(Repeated K-fold cross-validation) 방법으로 5-겹 교차검증을 10 회 반복한 후의 평균치로 모델별 일반적인 예측 성능을 추정하였다.

4) 개발된 모델들 간에 예측 타당도를 비교하는 데에는 민감도(Sensitivity), 특이도(Specificity), 양성 예측도(Positive Predictive Value), 음성 예측도(Negative Predictive Value) 및 Receiver Operating Characteristic curve (ROC Curve)의 Area under ROC curve (AUROC) 지표를 사용하였다. R 의 caret 패키지와 pROC 패키지를 이용하였다.

제 8 절 윤리적 고려

본 연구는 전자의무기록을 이용한 후향적 관찰 연구로, 연구 병원인 서울대학교병원의 기관생명윤리위원회(Institutional Review Board, IRB)에서 승인을 받을 때 피험자 사전동의를 면제받을 수 있었다(IRB No. H-2201-074-1291). 이후 간호본부의 자료수집에 대한 승인을 거쳐, 정보화실로부터 SUPREME 시스템 사용 권한을 부여 받았다. 간호 관리자가 아닌 일반 간호사는 IRB 승인 없이 SUPREME 에 접속 할 수 없다. SUPREME 시스템 접속 후에는 IRB 심의번호를 입력하여 필요한 자료를 추출할 수 있다.

수집된 모든 대상자의 정보는 잠금장치가 있는 연구실에 비밀번호가 설정된 별도의 파일로 보관하여 타인의 접근이 불가능하도록 관리하였다. 연구와 관련된 기록은 생명윤리법에 따라 연구가 종료된 시점 이후 3 년까지 보안을 유지하며 보관한다. 대상자 개인정보에 관한 사항은 보관기간이 끝나면 파기한다.

제 5 장 연구 결과

제 1 절 대상자 일반적 특성

1) 대상자 선정 결과

앞서 정한 선정 조건과 제외 조건에 따른 연구 대상자 선정 과정을 나타내면 Figure 3와 같다. 수술 시간 등의 기록이 부족한 5건과 수술 기록은 있으나 준비 중 아나필락틱 쇼크로 수술이 취소된 1건을 추가로 제외하였다. 최종 연구 대상으로 선정된 입실 건수는 2,083건이다. 실제 환자 수는 2,075명인데, 편의를 위하여 이후로는 매 입실 건마다 다른 환자가 입실 하였다고 가정할 것이다. 선정된 2,083건 중에 단 8건만이 재입실한 경우이고 이전 입실과의 시간 간격은 2년이 넘는다. 따라서 최종 연구 대상자 수가 2,083 '명' 이라고 표'현함에 무리가 없다.

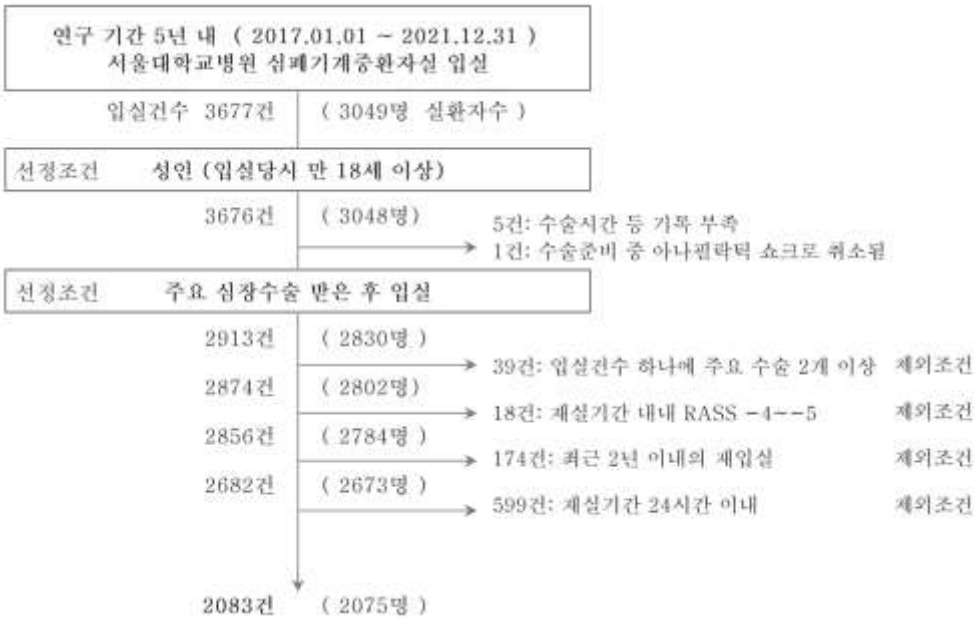

Figure 3. Subject Selection Process

2) 일반적 특성

최종 연구 대상으로 선정된 2,038명의 일반적 특성은 Table 4와 같다. 남성의 비율은 60.5%이고, 만 65세 이상의 비율은 59.6%이다. 평균 연령은 65.8±11.9세로, 가장 나이가 많은 대상자는 96세이다. 주요 수술 중 262건(12.6%)이 재수술이고, 205건(9.8%)이 응급수술이다. 1,489건(71.5%)의 수술에서 심폐기가 사용되었으며, 심장판막 수술(1,015건, 48.7%), 관상동맥 수술(660건, 31.7%), 대동맥 수술(376건, 18.1%), 부정맥 수술(214건, 10.3%) 순으로 건수가 많았다.

대상자의 95.8%의 경우가 수술이 끝남과 동시에 심폐기계중환자실에 처음 입실하는데, 수술 전 병원 입원 기간의 중앙값과 중환자실 재실 기간의 중앙값은 각각 5일과 2일이다. 입실한 환자의 98.2%가 이후 일반병실로 전동을 가며, 병원 총 입원 기간의 중앙값은 18일이다. 병원 퇴원 시의 사망률은 3.3%이다. 대상자 중 마지막으로 퇴원한 자는 2023년 4월 7일에 퇴원했기 때문에 병원 입원 기간을 포함한 Table 4의 모든 변수에서 결측치는 없다.

Table 4. General Characteristics of Patients (N=2,083)

Variable	Categories	n	(%)
Sex (Male)		1,260	(60.5%)
Age group	< 65	840	(40.3%)
	≥65, <80	1,048	(50.3%)
	≥80	195	(9.3%)
Age (years)	M±SD	65	± 11.9
	median(min, max)	67	(19, 96)
Redo op. (Yes)		262	(12.6%)
EM op. (Yes)		205	(9.8%)
CPB use (Yes)		1,489	(71.5%)
Surgery type	Aorta op.	376	(18.1%)
	Coronary artery op.	660	(31.7%)
	Valve op.	1,015	(48.7%)
	Maze op.	214	(10.3%)
Hospitalization before the surgery (days)	M±SD	7.1	± 14.9
	median(min, max)	5	(0, 363)
Pre ICU admission (Yes)		88	(4.2%)
ICU length of stay (days)	M±SD	3.7	± 5.4
	median(min, max)	2.1	(1, 121)
ICU discharge type	Death	34	(1.6%)
	Transfer	2,045	(98.2%)
	Discharge	4	(0.2%)
Length of hospitalization (days)	M±SD	28.6	± 62.5
	median(min, max)	18	(5, 1,532)
In-hospital death (Yes)		68	(3.3%)

op.=operation ; EM=emergency ; CPB = Cardiopulmonary bypass

제 2 절 섬망 환자 선별

본 연구에서 정한 섬망의 조작적 정의에 따라 섬망 환자를 가려냈을 때, 총 2,083명의 대상자 중 417명(20.0%)이 해당되었다. Table 5를 보면 대상자 선정에 대한 조건이 추가됨에 따라 섬망 발생 비율이 크게 달라지지 않는 것을 확인할 수 있다. 주요 심장수술을 받은 조건이 추가되었을 때에도 크게 변화가 없었다. 다만 중환자실 재실 기간 24시간 이상 조건을 추가하였을 때 섬망 발생 비율이 16.0%에서 20.0%로 증가하였다.

Table 5. Delirium Rate with Subject Selection Process

	Subject selection process criteria	Delirium	Total subject	Delirium rate
1	· admitted to the ICU	607	3,677	16.5%
2	· age ≥ 18 years	607	3,676	16.5%
3	· received a major cardiac surgery right before the admission and no more before discharged	474	2,874	16.5%
4	· at least one RASS score > −4	474	2,856	16.6%
5	· no history of admission in the same ICU within last 2 years	429	2,682	16.0%
6	· ICU length of stay ≥ 24 hours	417	2,083	20.0%

RASS = Richmond Agitation−Sedation Scale

전체 대상자 중 417 명이 본 연구에서 설정한 섬망의 조작적 정의에 대한 다섯 개의 조건 중 하나 이상을 만족시켰다. 각각의 세부 조건을 만족시킨 대상자의 수는 Table 6 에 정리하였다. Haloperidol IM 투약으로 정의되는 섬망 환자 수가 341 명으로 가장 많았고, 간호일지 섬망 진술문으로 정의되는 수가 13 명으로 가장 적었다. Figure 4 와 Table 7 을 통해 세부 조건들 간의 포함관계를 나타내었다.

Table 6. Delirium Rate using Different Operational Definitions

Operational definition	Yes	No	Yes%
1. CAM-ICU positive	71	2,012	3.4%
2. Nursing notes about delirium	13	2,070	0.6%
3. Diagnosed as delirium	78	2,005	3.7%
4. Neuropsychiatry consultation	227	1,856	10.9%
5. Prescription of haloperidol IM	341	1,742	16.4%
Delirium (meets more than one conditions in 1~5)	417	1,666	20.0%
Total	2,083		-

CAM-ICU= The Confusion Assessment Method for the Intensive Care Unit; IM= intamuscular injection

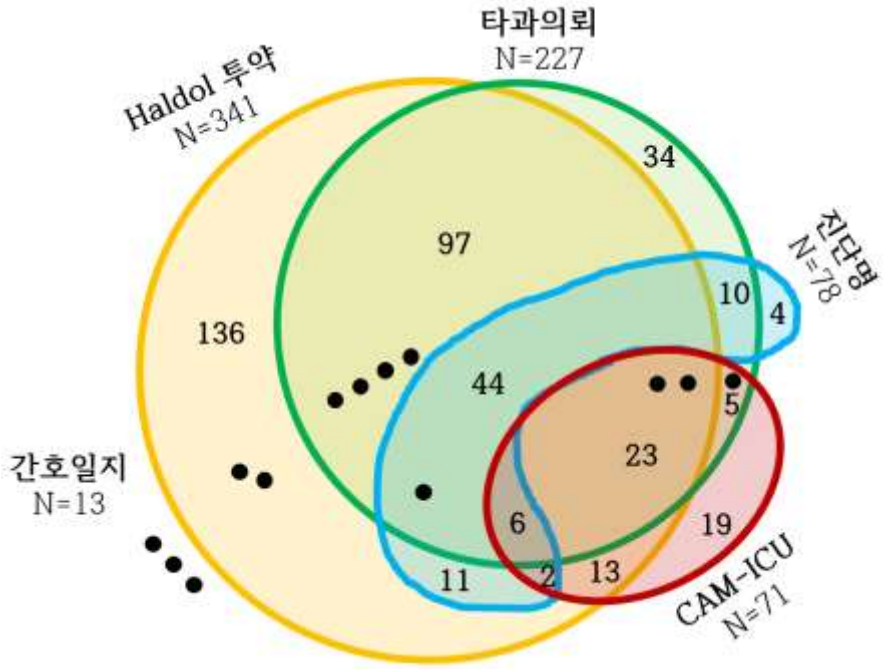

Figure 4. Inclusion Relations of Different Operational Definitions

Table 7. Inclusion Relations of Different Operational Definitions

Type	Sum	Delirium definition ('1' = Yes, '0' = No)					Number of cases		
		CAM	note	dg	cst	hdl			
01		0	1	1	1	1	1		
02	4	1	0	1	1	1	6	9	
03		1	1	0	1	1	2		
04		0	0	1	1	1	44		
05		0	1	0	1	1	4		
06	3	1	0	0	1	1	23	74	
07		1	0	1	0	1	2		
08		1	1	0	1	0	1		
09		0	0	0	1	1	97		
10		0	0	1	0	1	11		417
11	2	0	0	1	1	0	10	138	
12		0	1	0	0	1	2		
13		1	0	0	0	1	13		
14		1	0	0	1	0	5		
15		0	0	0	0	1	136		
16		0	0	0	1	0	34		
17	1	0	0	1	0	0	4	196	
18		0	1	0	0	0	3		
19		1	0	0	0	0	19		
20	0	0	0	0	0	0	1,666		
Total							2,083		

CAM= Confusion Assessment Method for Intensive Care Unit;
note= nursing note; dg= diagnosis; cst= consultation; hdl= haloperidol

제 3 절 섬망 위험요인 결정

1) 섬망 위험요인 자료 특성

섬망 예측모델에 사용될 예측변수 후보로 선정된 11개의 이산변수(이분형)와 35개의 연속변수의 데이터는 각각 Table 8, Table 9와 같은 특성을 나타낸다.

산소포화도 측정값은 전체 대상자 중 516명(75.2%)이 100%의 값을 가지고, 표준편차가 3.1이기 때문에 연속변수보다 이산변수로 활용하는 것이 적절함을 확인했다.

Table 8. Properties of 11 Discrete Variables

Variable	Number of '1'(%)	Number of '0'(%)	missing
1. Preoperative variable			
Sex (1='Male')	1,260(60.5%)	823(39.5%)	0
Pre ICU admission (1='Yes')	88(4.2%)	1,995(95.8%)	0
2. Intraoperative variable			
Surgery type			
Aorta op. (1='Yes')	376(18.1%)	1,707(81.9%)	0
Coronary artery op. (1='Yes')	660(31.7%)	1,423(68.3%)	0
Valve op. (1='Yes')	1,015(48.7%)	1,068(51.3%)	0
Maze op. (1='Yes')	214(10.3%)	1,869(89.7%)	0
else (1='Yes')	223(10.7%)	1,860(89.3%)	0
Redo op. (1='Yes')	262(12.6%)	1,821(87.4%)	0
Emergency op. (1='Yes')	205(9.8%)	1,878(90.2%)	0
CPB use (1='Yes')	1,489(71.5%)	594(28.5%)	0
3. Immediate postoperative variable			
SpO_2 abnormality(1='<98%')	265(12.7%)	1,818(87.3%)	0

Table 9. Properties of 35 Continuous Variables

Variable	min	median	M±SD	max	missing
1. Preoperative variable					
Age (years)	19	67	65.8 ±11.9	96	0
Body Mass Index (kg/m^2)	11.7	23.4	23.5 ±3.4	38.6	13
Hospitalization before the surgery (days)	0	5	7.1 ±14.9	363	0
2. Intraoperative variable					
Estimated Blood Loss (ml)	0	400	599.5 ±800.5	11,700	0
Operation time (min)	105	310	323.9 ±99	865	0
CPB time (min)	17	162	169.8 ±65.4	487	594
3. Immediate postoperative variable					
Body weight gain (kg)	−9.8	1.5	1.52 ±1.98	11.9	13
Vital Sign					
Body temperature (℃)	33	36.1	36 ±0.72	38.6	0
Heart rate (bpm)	35	79	79.9 ±15	160	0
Mean ABP (mmHg)	24	79	80.1 ±18.1	152	0
Arterial Blood Gas analysis					
pH	6.99	7.42	7.42 ±0.07	7.68	0
pO_2 (mmHg)	46	194	201.3 ±88.3	667	0
pCO_2 (mmHg)	10	38	38.9 ±6.3	77	0
HCO_3^- (mmol/L)	11.3	25.1	25.1 ±2.7	36.5	0
Complete Blood Count					
WBC ($\times 10^3$/ul)	0.7	9.6	10.3 ±4.5	56.5	0
Hemoglobin (g/dL)	5.9	11.3	11.4 ±1.5	17	0
Platelet ($\times 10^3$/ul)	11	108	114.4 ±49.5	608	0
Coagulation test					
Fibrinogen (mg/dL)	46	201	211.5 ±66.5	565	0
aPTT (sec)	20.1	33.8	36.9 ±12.7	201.1	8
PT (INR)	0.85	1.24	1.28 ±0.2	3.59	0
Other blood serum test					
Potassium (mmol/L)	2.2	4.2	4.2 ±0.5	7.5	0
Sodium (mmol/L)	127	139	139.3 ±3.7	154	0
Calcium (mg/dL)	4	8.6	8.8 ±1.4	26.6	0
Magnesium (mEq/L)	0.8	1.8	1.9 ±0.4	3.8	2
Phosphorus (mg/dL)	0.5	3	3 ±0.9	7.3	0
BUN (mg/dL)	4	15	16.8 ±8.3	82	0
Creatinine (mg/dL)	0.23	0.79	1.08 ±1.2	11.75	0
AST (IU/L)	6	40	72.5 ±139.3	4,355	0
ALT (IU/L)	1	18	27.5 ±49.8	1,063	0
Total Bilirubin (mg/dL)	0.2	1.3	1.5 ±1.1	33.2	0
Alk.phos (IU/L)	14	42	45.6 ±18.6	235	1
Albumin (g/dL)	1.2	3	3 ±0.5	4.7	0
CRP (mg/dL)	0.01	0.94	1.6 ±2.2	20.2	3
Glucose (mg/dL)	24	132	138.3 ±39.7	488	0
Lactate (mmol/L)	0.3	1.4	1.9 ±1.6	13.3	4

2) 결측치 처리

초기에 위험요인 변수를 선정할 때부터 연구 대상자들에게서 일상적으로 획득되는 변수들을 대상으로 선정했기 때문에, 최종 데이터세트에서 결측치의 수는 매우 적은 편이었다. 체질량지수와 체중 증가량 13건, 심폐기 사용 시간 594건, 젖산 4건, aPTT 8건, 마그네슘 2건, Alk.phos 1건, 그리고 CRP 3건의 결측치가 있었다.

체질량지수와 체중 증가량에 대한 13건의 결측치는 공통적으로 수술 직전 몸무게 측정값의 부재로 인해 나타났다. 대동맥 박리 등으로 응급실을 통해 입원하여 곧바로 수술이 이루어지는 경우에는 종종 수술 직전의 몸무게를 측정하지 못하는 상황이 발생한다. 실제로, 결측치를 보인 13건 중 7건에서 앞서 응급수술 여부를 반영하기 위해 생성했던 EM 변수의 값이 1인 것을 확인하였다. 수술 직후 몸무게와 수술 직전 몸무게의 차이인 체중 증가량 변수의 평균값이 1.5kg인 것을 고려하여 13건의 수술 직전 몸무게는 수술 직후 몸무게에서 1.5kg을 더한 값으로 대체하였다. 이에 따라 수술 직전의 체질량지수도 계산하여 대체하였다.

혈액검사 변수 중에 각 2건, 1건, 3건의 결측치가 있었던 마그네슘, Alk.phos, CRP 변수에 대해서도 평균값으로 각 결측치를 대체하였다. 젖산 변수 결측치 4건 중 1건에 대해서는, 해당 환자가 수술장에서 퇴실하기 1시간 전에 시행한 검사 수치가 있어 그 값으로 대체하였다.

한편, 젖산 수치의 결측치 중 나머지 3건은 검사장비의 한계로 결과 수치가 ‘>15’로 기록된 경우이다. 이와 유사하게 aPTT 8건도 ‘>400’로 기록되었다. 두 경우에 대해 일차적으로 각각 15, 400으로 대체한 뒤 이상치 처리를 적용했다.

기간 내 작성된 체외순환기록지가 없어서 심폐기 사용 시간 데이터의 결측치가 있던 594건에 대해서는 심폐기 사용 시간에 0의 값을 부여했다. 심폐기를 작동하였음에도 체외순환기록이 누락된 경우가 있을 가능성이 있으나 반영하지 않았다.

3) 이상치 처리

46개의 섬망 위험요인 변수에서 평균적으로 3.8%의 이상치가 발견되었다. 이상치가 가장 많았던 변수는 ALT이고, 전체 2,083건 중 258건(12.4%)의 값이 해당되었다. Figure 5는 이상치를 처리하기 전과 후의 ALT 변수의 수치 분포에 대한 boxplot이다.

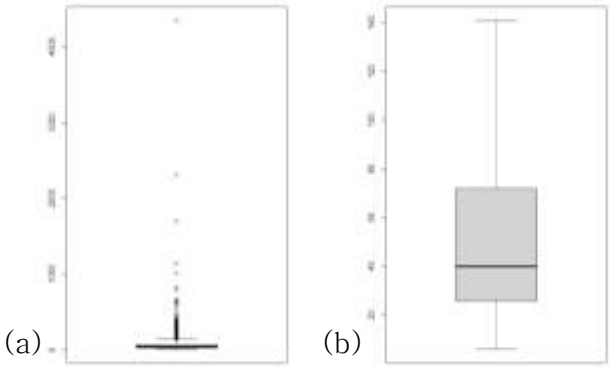

Figure 5. Change in Boxplot Distribution after Treating Outliers (ALT)
(a)before, (b)after

4) 데이터 불균형 처리

Table 10은 최종 데이터세트에 SMOTE 의 일종인 SMOTE-NC를 적용한 결과이다. 최종 데이터세트의 섬망 환자수는 417명이고 섬망 발생률은 20%인데, SMOTE 로 기존의 데이터를 기반으로 섬망 환자의 데이터를 1,249건 합성하여 추가하였다. 그 결과 만들어진 섬망 발생률 50%의 데이터세트를 예측모델 개발에 사용하였다.

Table 10. Datasets before and after Applying SMOTE

	Delirium	No delirium	Total	Delirium Rate
Final dataset	417	1,666	2,083	20%
Final dataset + SMOTE	1,666	1,666	3,332	50%

SMOTE = Synthetic Minority Oversampling TEchnique

5) 단변량 분석

예측모델을 개발하기 전에, 단변량 분석을 시행하여 섬망군과 비섬망군 간에 46개의 섬망 위험요인 수치값이 통계적으로 유의한 차이를 나타내는지 확인했다. 이산변수 11개에 대해서는 카이제곱 검정을 시행했다(Table 11). 연속변수 35개에 대해서는 정규성 검정 및 등분산 검정을 거쳐 이표본 T 검정을 시행했다(Table 12).

단변량 분석 결과 심장수술 후 섬망이 발생한 환자는 그렇지 않은 환자에 비해 다음의 특징을 가진다. 먼저 이산변수에 대해서는, 섬망 환자는 수술에 심폐기를 사용한 경우가 많고, 대동맥 수술이나 심장판막 수술 받은 경우가 많고, 관상동맥 수술을 받은 경우가 적고, 수술 전부터 입실해있던 경우가 많고, 응급수술이나 재수술인 여자 환자의 경우가 많다. 그리고 수술 직후 산소포화도가 98% 미만인 경우가 많다. 연속변수에 대해서는, 섬망 환자는 연령이 높고, 체질량지수는 작은 편이며, 수술 시간 및 심폐기 사용 시간이 길고, 수술 중 예상 실혈량이 많으며, 수술 직후 체온이 낮고 평균 동맥혈압이 낮으며 심박수가 높다. 검사 수치를 살펴보았을 때 섬망 환자는 헤모글로빈, 혈소판, 섬유소원, 동맥혈 산소 분압, 알부민, 동맥혈 pH, 나트륨 수치가 낮았고 PT, 혈당, 젖산, 칼슘, 마그네슘, BUN, 총 빌리루빈, CRP, aPTT, 동맥혈 이산화탄소 분압, 크레아티닌 수치가 높았다.

한편, 섬망군과 비섬망군 간에 별 차이를 보이지 않았던 변수들도 있었다. 일반적인 유의수준 0.05를 기준으로 p-value가 컸던 이산변수 2개(기타 주요 수술 여부, 부정맥 수술 여부)와 연속변수 9개(ALT, AST, 수술 전 입원 기간, 백혈구, 중탄산이온, 인, 칼륨, 체중 증가량, Alk.phos)를 최종 예측변수 후보에서 제외하였다. 최종 46개 변수 중 35개의 변수를 예측모델 개발에 사용했다.

Table 11. Univariate Analysis of 11 Discrete Risk Factor Variables

(N=2,083)

Variable	Delirium N=417 n (%)	No delirium N=1,666 n (%)	*p*
CPB use (Yes)	358 (85.9%)	1,131 (67.9%)	0000***
Aorta op.[a] (Yes)	126 (30.2%)	250 (15%)	.0000***
Coronary artery op.[a] (Yes)	87 (20.9%)	573 (34.4%)	.0000***
SpO_2 abnormality (Yes)	77 (18.5%)	177 (11.3%)	.0001***
Pre ICU admission (Yes)	32 (7.7%)	56 (3.4%)	.0001***
Emergency op. (Yes)	62 (14.9%)	143 (8.6%)	.0001***
Valve op.[a] (Yes)	237 (56.8%)	778 (46.7%)	.0002***
Redo op. (Yes)	67 (16.1%)	195 (11.7%)	.0163*
Sex (Male)	234 (56.1%)	1,026 (61.6%)	.0410*
Other op.[a] (Yes)	39 (9.4%)	184 (11%)	.3176
Maze op.[a] (Yes)	39 (9.4%)	175 (10.5%)	.4885

Variables are arranged in ascending order by *p*−value;
* $p<.05$; ** $p<.01$; *** $p<.001$;
[a] Variables derived from surgery name data;
CPB= Cardiopulmonary bypass; op.= operation

Table 12. Univariate Analysis of 35 Continuous Risk Factor Variables (N=2,083)

Variable	Delirium N=417 M±SD	No delirium N=1,666 M±SD	*p*
Age (years)	69.3±10.6	64.9±12.1	.0000***
Operation time (min)	367.1±107.8	313.14±93.6	.0000***
CPB time (min)	168.4±92.5	109.6±91.4	.0000***
Body temperature (℃)	35.87±0.7	36.1±0.7	.0000***
mean ABP (mmHg)	76.1±17.8	81.1±18.1	.0000***
Hemoglobin (g/dL)	11.0±1.5	11.5±1.5	.0000***
Platelet ($\times 10^3$/ul)	101.0±43.9	117.7±50.3	.0000***
Fibrinogen (mg/dL)	193.6±66.0	216.0±65.9	.0000***
PT (INR)	1.36±0.25	1.26±0.18	.0000***
pO_2 (mmHg)	181.2±86.9	206.3±87.9	.0000***
Glucose (mg/dL)	148.8±41.5	135.7±38.8	.0000***
Lactate (mmol/L)	2.6±2.2	1.8±1.5	.0000***
Calcium (mg/dL)	9.1±1.5	8.7±1.3	.0000***
Magnesium (mEq/L)	2.0±0.4	1.8±0.4	.0000***
BUN (mg/dL)	18.7±9.2	16.4±8.0	.0000***
Albumin (g/dL)	2.9±0.6	3.1±0.5	.0000***
pH	7.41±0.07	7.42±0.06	.0001***
Body Mass Index (kg/㎡)	22.97±3.55	23.68±3.4	.0002***
Total Bilirubin (mg/dL)	1.69±1.11	1.46±1.11	.0002***
CRP (mg/dL)	1.99±2.7	1.5±1.98	.0004***
EBL (ml)	724.3±902.5	568.2±770.0	.0012**
aPTT (sec)	39.9±18.1	36.9±15.7	.0019**

pCO_2 (mmHg)	39.8±6.8	38.7±6.2	.0021**
Creatinine (mg/dL)	1.25±1.33	1.03±1.12	.0021**
Sodium (mmol/L)	138.8±4.4	139.5±3.5	.0044**
Heart rate (bpm)	81.5±14.9	79.5±15.0	.0154*
ALT (IU/L)	31.5±59.8	26.5±47.0	.1119
AST (IU/L)	79.2±101.5	70.8±147.2	.1682
Hospitalization before the surgery (days)	8.1±19.2	6.9±13.6	.2208
WBC ($\times 10^3$/ul)	10.5±4.8	10.2±4.5	.3331
HCO_3^- (mmol/L)	25.0±2.9	25.1±2.6	.3790
Phosphorous (mg/dL)	3.1±1.0	3.0±0.9	.4096
Potassium (mmol/L)	4.2±0.6	4.2±0.5	.5798
Body weight gain (kg)	1.5±2.3	1.5±1.9	.8368
Alk.phos (IU/L)	45.6±20.8	45.6±17.9	.9642

Variables are arranged in ascending order by p-value;
$^*p<.05$; $^{**}p<.01$; $^{***}p<.001$;
CPB= Cardiopulmonary Bypass; ABP= Arterial Blood Pressure; PT= Prothrombin Time; pO_2= partial pressure of oxygen; BUN= Blood Urea Nitrogen; CRP= C-reactive protein; EBL= Estimated Blood Loss; aPTT = activated Partial Thromboplastin Time; pCO_2= partial pressure of carbon dioxide; ALT= Alanine aminotransferase; AST= Aspartate aminotransferase; WBC= White Blood Cell; HCO_3^- = Bicarbonate ion; Alk.phos= Alkaline phosphatase

제 4 절 섬망 예측모델 성능 비교

35개의 섬망 위험요인 변수를 사용하여 로지스틱 회귀, 의사결정나무, 랜덤포레스트 및 XGBoost 각각의 알고리즘으로 예측모델을 개발하였다. 5-겹 교차검증을 10회 반복하여 모델별 성능을 추정하였고, 그 결과는 Table 13과 같다. 전처리로 SMOTE를 적용하지 않았을 때 전반적으로 모델들의 민감도가 낮았다(0.0628~0.2302). SMOTE를 적용하여 섬망과 비섬망군의 비를 1:1 로 맞춘 뒤에 모델을 개발했을 때는 전반적으로 예측 성능이 향상되었다. 특히 민감도 및 양성 예측도가 많이 향상되었다.

알고리즘별로 살펴보았을 때, SMOTE를 적용하지 않은 데이터에서는 로지스틱 회귀가 가장 높은 AUROC (0.7564) 값을 나타냈지만, SMOTE 적용 후에는 랜덤포레스트가 가장 높은 AUROC (0.9575) 값을 보였다. SMOTE 적용 후 랜덤포레스트로 개발된 예측모델의 성능이 가장 뛰어나서 최종 모델로 선택했다. 이 때 민감도, 특이도, 양성 예측도, 및 음성 예측도는 각각 0.9050, 0.8650, 0.8702 및 0.9011 이었다. Figure 6와 Figure 7에서는 최종 모델의 ROC curve를 다른 모델과 비교했다.

Table 13. Evaluation of Delirium Prediction Models

Algorithm	Sensitivity	Specificity	PPV	NPV	AUROC
Dataset : imbalanced dataset (20% delirium)					
Logistic Regression	**0.2302**	0.9537	**0.5546**	**0.8319**	**0.7564**
Decision Tree	0.1874	0.9490	0.4789	0.8235	0.7204
Random Forest	0.0628	**0.9827**	0.4755	0.8073	0.7453
XGBoost	0.1760	0.9487	0.4621	0.8214	0.6933
Dataset : balanced dataset (50% delirium) after applying SMOTE					
Logistic Regression	0.7186	0.6994	0.7051	0.7131	0.7860
Decision Tree	0.8852	**0.8657**	0.8683	0.8829	0.9464
Random Forest	**0.9050**	0.8650	**0.8702**	**0.9011**	**0.9575**
XGBoost	0.8367	0.8031	0.8095	0.8311	0.9014

PPV= Positive Predictive Value ; NPV= Negative Predictive Value; AUROC= Area Under Receiver Operating Characteristic Curve; XGBoost= eXtreme Gradient Boosting; SMOTE= Synthetic Minority Oversampling Technique

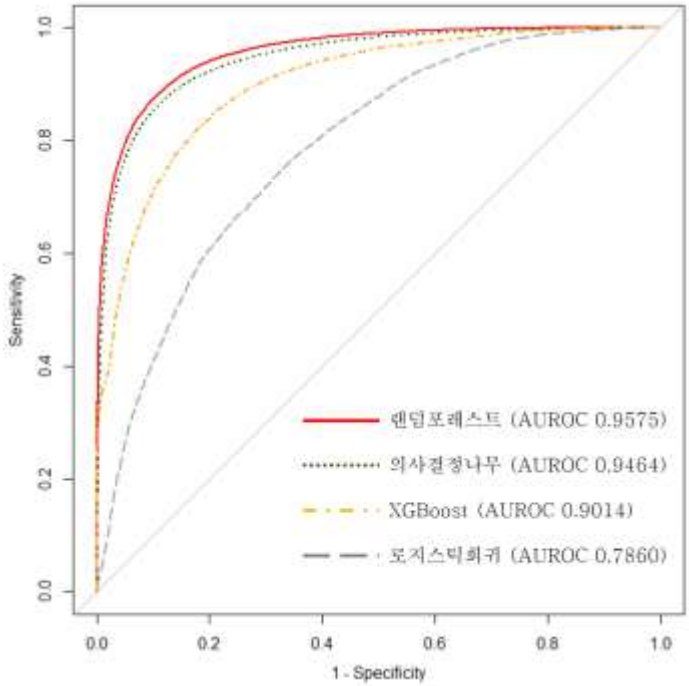

Figure 6. Comparison of ROC curves between Algorithms

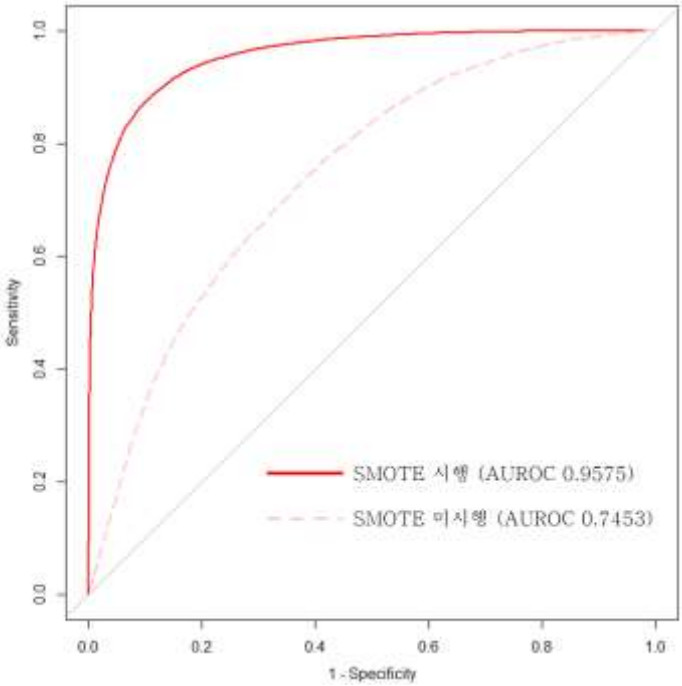

Figure 7. Comparison of ROC curves between Datasets
(Random Forest)

제 6 장 논 의

제 1 절 섬망 환자 선별

본 연구에서의 섬망 발생률은 20.0%이다. 이는 본 연구에서와 같은 중환자실을 대상으로 섬망을 주제로 다뤘던 노은영과 박연환(2019)의 연구 결과와 일치한다. 해당 연구는 2016년 9월부터 1년동안 심장수술을 받고 입실한 421명의 환자의 15%에서 섬망이 발생하였음을 보고하였는데, 연구 대상자를 설정할 때 재실 기간에 대한 제한은 두지 않았었다. Table 5를 통해 확인할 수 있듯이, 본 연구에서도 24시간 이상 재실이라는 대상자 선정 조건을 제외하면 섬망 발생률은 16.0%로 유사한 수치를 보인다. 노은영과 박연환의 연구는 연구 간호사가 대상자들의 의무기록을 검토한 뒤 Korean Nursing Delirium Screening Scale (Korean Nu-DESC) 도구를 사용하여 섬망 환자를 선별한 연구였는데, 사용한 도구가 달라도 선별 결과가 크게 차이 나지 않음을 짐작할 수 있다. 이와 관련하여 Rood 등(2018)이 체계적 문헌고찰을 통해 섬망 사정 도구의 종류나 사용 빈도에 따른 섬망 발생률의 차이가 유의하지 않음을 확인한 연구가 있다.

심장수술 환자에게서 섬망 발생률은 4.1~54.9% 정도로 다양하게 보고되기 때문에(Chen et al., 2021) 다른 연구에서의 심장수술 후 섬망 발생률은 본 연구에서의 섬망 발생률과 차이가 있을 수 있다. 연구마다 대상자의 일반적 특성이 다르기 때문임은 물론이고, 대상 중환자실에서 행해지는 루틴이 다른 것도 그 원인일 수 있다. 예를 들어, Likhvantsev 등(2021)은 무작위 대조군 실험연구를 통하여 심폐기를 사용한 관상동맥 수술 혹은 판막 수술 환자들에게서 Dexmedetomidine 투약 여부가 섬망 발생률에 유의한 영향을 미침을 밝혔다. 본 연구의 중환자실에서는 수술 후 입실한 환자들에게 Dexmedetomidine 투약을 0.2mcg/kg/hr 의 용량으로 시작하여 기관내관 발관 시까지 유지하는 루틴이 있는데, 해당 약물을 사용하지 않는 중환자실에서는 섬망 발생률이 다르게 나타날 수 있음을

시사한다.

그 뿐만이 아니라 중환자실에서의 빛 노출 정도나 보호자 면회 허용 여부 및 정도 등과 같이 기록상으로 드러나지 않는 요소들도 섬망 발생률 차이의 원인일 수 있다. 섬망 환자의 대부분은 수면 문제를 가지고 있는데, 자연광에 노출되는 것은 일주기(circadian) 리듬과 수면 패턴을 정상적으로 유지하는 데 있어 가장 중요한 요소이다(Luetz et al., 2019). 또한 Eghbali-Babdi 등(2017)은 실험연구를 통해 정기적이고 구조화된 가족 면회가 심장수술 후 중환자의 섬망 발생 위험을 약 21% 낮춤을 증명하였다. 본 연구의 중환자실에서는 외부와 연결된 창문이 없고, 중증 환자 치료로 인해 야간에도 조명을 소등하지 못하는 경우가 많으며 보호자 면회가 매우 제한적으로 이루어지기 때문에, 그렇지 않은 중환자실에서는 본 연구에서보다 섬망 발생률이 더 낮을 수 있다. 한편, 본 연구에서는 기존의 일반적인 섬망 연구들과 다르게 섬망 환자를 조작적으로 정의할 때 haloperidol 처방 내역, CAM-ICU 평가, 정신건강의학과 타과의뢰뿐만이 아니라 간호일지나 진단명 기록도 추가하여 살폈다. 이는 실제 임상현장에서는 CAM-ICU 평가가 잘 활용되지 않으며, 전자의무기록의 다양한 기록지에 섬망 여부가 산발적으로 기록되고 있는 점을 고려한 조치이다. 그런데 본 연구의 대상자 2,083명 중에 섬망 환자는 417건이었는데, haloperidol, CAM-ICU, 그리고 타과의뢰 조건만으로 410건을 가려낼 수 있었다.

예상했던 것과 다르게 간호일지나 진단명 조건을 추가함으로써 더 선별된 섬망 환자의 수는 매우 적었는데 이는 섬망 발생에 대한 기록이 여전히 누락되었음을 의심해 볼 수 있으나, 추후 연구를 통해 기록의 질이 섬망 예측의 정확성에 미치는 영향을 검토해 볼 필요가 있다. 또한 이와 같은 결과는 섬망과 관련된 기록이 누락되는 것을 방지하기 위해서는, 섬망의 조작적 정의를 여러 가지 사용하는 것보다는 Menzenbach 등(2022)의 연구에서와 같이 현장의 의료진들을 대상으로 섬망 기록 방법에 대한 교육을 철저하게 하는 것이 더 효과적일 수 있음을 시사한다.

Dubiel 등(2022)은 The Society of Thoracic Surgery National

Database (STS-ND)의 섬망 정의를 사용했을 때보다 CAM-ICU 섬망 평가 결과를 사용했을 때 더 많은 섬망 환자를 선별할 수 있었고, 심장수술 일 년 후의 기능 정도도 더 정확하게 예측할 수 있었음을 밝혔다. 이는 환자의 건강성과 혹은 임상 예후의 예측에 정확한 섬망의 정의가 무엇보다도 중요함을 보여준다. 따라서 섬망 환자를 누락 없이 정확하게 선별하는 방법에 대한 연구가 계속 되어야 한다.

섬망의 조작적 정의 조건들 간의 포함관계를 비교하였을 때 특정한 경향성을 발견하기는 어려웠다. 이는 조작적 정의의 조건들이 중복 없이 효과적으로 선별되었음을 의미한다고 볼 수도 있다. 기존의 섬망 정의에 사용된 자료원에 따른 포함력 사이의 관계는 연구된 바가 없다. 따라서 본 연구 결과는 추후 다른 연구들에서 섬망을 조작적으로 정의할 때 기반이 될 수 있을 것으로 기대한다.

제 2 절 섬망 위험요인 결정

본 연구에서 단변량 분석을 통해 결정된 섬망의 위험요인들은 유사한 타 연구에서 사용한 위험요인들과 차이가 있었다. 예를 들어 노은영과 박연환(2019)이 421명의 심장수술 후 중환자를 대상으로 수행했던 연구에서 성별, Aortic aneurysm 여부, pH 이상 여부, 혈당 이상 여부, 나트륨 이상 여부, 그리고 수술 시간이 섬망 발생 여부와 유의한 관계를 보이지 않았음을 보고한 반면 본 연구에서는 성별, 대동맥 수술 여부, pH(연속변수), 나트륨(연속변수), 수술 시간(연속변수)이 유의한 변수로 나타났다. 이러한 차이가 나타난 이유는 첫째로 본 연구의 표본 수가 5배 정도 많아서 귀무가설을 기각하기에 유리했기 때문으로 사료된다. p-value가 충분히 작지 않았지만, 섬망 환자군 중 여성 비율과 대동맥 수술 비율이 높고, 평균 동맥혈 pH가 낮으며 수술 시간이 긴 경향은 본 연구에서와 동일하게 확인되었다.

두 번째 이유로는 노은영과 박연환의 연구에서는 연속변수를 이산변수로 변환하여 사용하면서 정보의 소실이 있었을 가능성이 있다. 그리고 무엇보다 본 연구에서는 중환자실 입실 직후의 검사결과만 가지고 분석했으나, 위 연구는 중환자실 재실 기간 중 가장 정상범위를 벗어난 수치를 사용했다는 중요한 차이가 있다. 그럼에도 불구하고 연령(연속변수), 알부민 이상 여부(이산변수), 크레아티닌 이상 여부(이산변수)나 동맥혈 산소 분압 이상 여부(이산변수)는 본 연구에서와 마찬가지로 유의한 변수인 것으로 확인되었으며 섬망군은 비섬망군에 비해 연령이 높고, 알부민 수치가 낮으며, 크레아티닌 수치가 높고, 산소 분압 수치가 낮은 경향이 있음을 보여주었다.

본 연구의 결과는 삼성서울병원 심장중환자실(Cardiac Intensive Care Unit, CICU)에서 행해졌던 Kim 등(2023)의 연구 결과와도 비교해볼 수 있다. 2012년 9월부터 2018년 12월 사이 해당 중환자실에 24시간 이상 재실한 성인 환자 2,783명을 대상으로 한 연구이며, 섬망 발생률은 24.3%이다. 본 연구에서와 마찬가지로 섬망 발생과 관련하여 심장수술 후

중환자실 입실 직후의 다양한 혈액검사 항목에 대해 단변량 분석을 시행하였는데 섬망군이 비섬망군에 비해 총 빌리루빈, BUN, 크레아티닌, 혈당, CRP 가 높고, 헤모글로빈, 혈소판, 알부민이 낮은 것이 본 연구의 결과와 동일하게 확인되었다. 다만 본 연구에서는 백혈구 수치가 섬망군이 비섬망군보다 높았지만 통계적으로 유의하게 차이가 나는 수준은 아니었던 반면, Kim 등의 연구에서는 유의한 차이를 보였다. 또한 본 연구에서 다루지 않았던 NT-pro BNP, Troponin I, Total/ LDL/ HDL Cholesterol, Triglycerides 의 수치도 다루고 있었는데 모두 유의한 차이를 보였다. 이는 본 연구에서 다루지 않은 많은 혈액검사들도 섬망군 예측요인으로 포함할 필요성을 시사하나 이들 혈액검사들은 서울대학교병원 심폐기계중환자실 환자들에게 루틴으로 시행되지 않기 때문에 활용에 한계가 있다. 심장수술 후 섬망의 위험요인들은 심폐기의 사용과 밀접한 관련을 가지고 있는 경우가 많다. 대동맥 수술이나 판막 수술은 거의 심폐기를 사용하는 수술인 반면 연구 병원의 관상동맥 수술에서는 off pump CABG의 비율이 높아 18.3%만이 심폐기를 사용했기 때문에, 수술 종류도 섬망 발생과 관련이 있었을 것이다. 또한 심폐기 사용 시간이 길면 자연히 수술 시간도 길어지고, 체온을 낮추는 과정이 포함되기 때문에 수술 직후 체온은 낮아지고, 항응고 요법을 동반한 체외순환으로 인해 응고 관련 검사결과도 이상치를 보이게 되는 등 다양한 예측변수에 연쇄 효과를 가져오는 중요한 변수임을 확인할 수 있었다.

제 3 절 섬망 예측모델 성능 비교

본 연구에서는 최종 데이터세트에 SMOTE 기법을 적용하여 데이터 불균형을 해소한 뒤에 개발한 예측모델의 성능이 그렇지 않은 경우보다 우수했으며 특히 양성 예측도와 민감도가 모두 크게 증가했다. 양성 예측도가 낮은 예측모델을 현장에 적용하게 되면 간호사들의 alarm fatigue을 유발할 것이다. 민감도가 낮은 모델을 사용하면 실제 섬망 고위험군을 조기에 가려내지 못하기 때문에 예측모델을 개발하는 의도와 맞지 않는다. 따라서 데이터 불균형 해소는 예측모델 성능의 유의미한 개선을 가져왔다고 볼 수 있다.

본 연구에서 예측모델 개발에 사용한 로지스틱 회귀, 의사결정나무, 랜덤포레스트, 및 XGBoost 중 랜덤포레스트가 민감도, 양성 예측도, 음성 예측도, AUROC의 항목에서 가장 우수한 성능을 보였다. 특이도도 1위와 거의 유사한 수치를 나타냈다. Hur 등(2021)의 섬망 예측 연구에서는 로지스틱 회귀(AUROC 0.875)와 랜덤포레스트(AUROC 0.916) 보다 XGBoost (AUROC 0.919)의 예측 타당도가 더 높았다. 이 연구에서는 외과계와 내과계 중환자를 모두 대상으로 만들었고 더 많은 예측변수를 사용했다는 차이는 있다.

심장수술 후 중환자를 대상으로 섬망 예측모델을 개발할 때 여러 가지 머신러닝 기법을 비교 평가한 연구는 소수에 불과하다. 그 중 Mufti 등(2019)은 5,584명의 자료를 이용하여 로지스틱 회귀, 의사결정나무, 랜덤포레스트, 인공신경망, 서포트 벡터 머신, 그리고 베이지안 네트워크의 방법을 비교하였다. 이 연구에서의 섬망 발생률은 11.4%(507명)였는데, 비섬망 환자의 예시를 무작위로 선택하는 undersampling 방법을 채택하여 섬망 발생률을 50%로 교정한 뒤 예측모델을 개발하였다. 본 연구에서와 마찬가지로, 데이터의 불균형을 해소 했을 때 양성 예측도와 민감도가 향상되었다. 해당 연구에서 가장 성능의 좋았던 모델은 서포트 벡터 머신 기법을 활용한 모델(AUROC 0.811)인데, 본 연구에서의 최종 모델(AUROC 0.9575)보다 낮은 성능을 보였다.

본 연구에서 개발한 예측모델은 중환자실 입실 직후의 시점에서 결정되는 변수만으로는 좋은 예측성능을 보였다. 이는 중환자 입실 기간 내 섬망 발생 여부를, 중환자 입실 직후까지의 시점에서 결정되는 변수들만으로 예측할 수 있음을 확인했다는 의의를 가진다. 보통 외과계 환자의 수술 후 섬망의 경우 환자가 전신마취에서 깨어나는 시점에서 섬망 증세를 보이고, 카테터를 제거하는 등의 위험한 상황이 발생하는 상황이 잦기 때문에 섬망 위험을 조기에 예측할 필요가 있는데, 본 연구는 섬망 발생 위험을 충분히 조기에 파악하는 것이 가능함을 보여주었다.

제 4 절 연구 제한점

본 연구는 다음의 제한점을 가진다.

첫째, 섬망의 종류에 대한 고려를 하지 않았다. 섬망은 정신운동성 행동(psychomotor activity) 양상에 따라 과활동성(hyperactive), 저활동성(hypoactive) 및 혼재형(mixed)의 유형으로 분류된다(American Psychiatric Association, 2013). CAM-ICU 최종 결과가 양성일 때, RASS 점수가 양수이면 과활동성, 음수이면 저활동형, 그리고 양수와 음수 모두의 값을 가지면 혼재형으로 정의하는 방법이 일반적이다(Peterson et al., 2006). 위 섬망 유형에 따라 재원일수 및 사망률 등의 결과가 유의하게 다른 것이 알려져 있다(la Cour et al., 2022).

수술 후 섬망을 발생 시각에 따라 24시간 이내이면 Emergence Delirium, 24시간 이상 72시간 이내 이면 post-op delirium, 그리고 몇 주나 몇 달 이후이면 post-op cognitive dysfunction으로 분류하기도 한다(Berger et al., 2018). 그 외에 Girard 등(2018)은 섬망의 유형을 임상 양상에 따라 hypoxic, septic, sedative, metabolic, 및 unclassified의 다섯 가지로 나누어 연구할 필요성을 제기하기도 했다. 본 연구에서는 이러한 섬망의 세부 유형을 나누어 살펴보지 않았다는 한계를 가진다.

둘째, 본 연구에서는 섬망 발생 시각에 대한 분석을 포함하지 못했다. 섬망 환자를 정의할 때 중환자실 재실 기간 내에 한번이라도 관련 기록이 있었으면 섬망 환자에 해당되는 것으로 여겼다. 섬망 정의에 활용된 정신건강의학과 타과의뢰는 보통 정규시간에 주치의가 시간이 날 때 작성되기 때문에 그 작성 시각은 섬망의 발생 시각과 차이가 날 수밖에 없다. 타과의뢰 내용에도 구체적인 섬망 발생 시각 정보는 부재한 경우가 대부분이며. 진단명 작성 일시 또한 마찬가지이다. 그리고 중환자실에서 행동문제를 동반한 섬망 문제가 발생하여 환자에게 haloperidol 주사제를 빠르게 투약해야 하는 상황에서는 보통 담당의에게 구두처방을 받아 병동 비품약을 먼저 사용한 뒤 시간 차를 두고 전산상의 오더를 받기 때문에,

약물 오더 발행 시각 역시 섬망 발생 시각을 추정을 위한 믿을 만한 정보로 보기 어렵다. 전자의무기록만을 활용한 후향적 연구에서 섬망의 발생 시각 정보를 획득하기 어려운 게 현 실정이다.

셋째, 섬망의 위험요인으로 사용한 변수를 더욱 세분화해서 반영하거나 영향력이 있다고 알려진 변수를 모두 예측모델에 포함시키지 않았다. 심장혈관흉부외과의 수술을 받은 환자에 대해 연구할 때, 수술 대상 심장판막의 종류(mitral, aortic, tricuspid 등)나 판막 성형/치환 여부를 구별할 필요가 있는데, 본 연구에서는 그렇게 하지 못했다는 한계가 있다. 2,083건의 연구대상 입실 건 중 48.7%에 해당하는 1,015건이 심장판막과 관련된 수술을 한 것으로 드러났는데, 좀 더 그룹을 세분화할 필요가 있는 것으로 사료된다. 이는 판막수술에 대한 표준화된 세부수술명에서 상세한 수술 판막 부위 및 사용한 재료 등이 함께 작성된 경우가 많지 않았기 때문에 발생한 한계이다. 자유형식으로 작성된 수술명이 있기에, 이를 하나씩 살펴본다면 필요한 추가 정보를 얻을 수 있을 것으로 기대된다. 혹은 결측치를 감수하고 전산 상의 세부수술명을 이용하여 심장판막 수술의 종류를 더 상세히 반영한 뒤 이를 통해 다른 의미 있는 결과를 도출할 수도 있을 것이다. 그리고 입원간호기록의 기저질환(당뇨, 고혈압, 고지혈증 여부 등)이나 음주력을 변수로 활용하지 못하였는데 이들은 중환자실 입실하기 전에 결정되는 변수들이기 때문에, 본 연구에서 쓰인 예측모델의 위험요인 후보로 적합함에도 포함하지 못했다.

예측모델의 학습에 사용된 데이터가 아닌 임상현장에서 새롭게 수집된 자료를 사용하여 예측모델의 성능 검증이 이뤄지지 않은 점도 제한점으로 꼽을 수 있다. 또한 일개 병원 중환자실 데이터만을 활용한 연구이기에 본 연구의 결과만으로는 예측모델의 성능을 타 병원 환자들에게 일반화하기는 어렵다.

제 7 장 결 론 및 제 언

제 1 절 결 론

본 연구는 일개 병원 중환자실에서 2017년과 2021년 사이 입실한 환자 2,083명의 전자의무기록자료를 바탕으로 심장수술 후 섬망의 예측모델을 개발하고 평가한 연구이다.

연구의 방법과 관련하여, 본 연구는 연구자가 간호사로서 수년간 근무하면서 수집 및 기록 과정에 참여하였던 중환자실의 전자의무기록 자료를 직접 추출하여 가공하고 분석하였다는 데 의의를 가진다. 연구 병원의 전산실을 거치지 않고 CDW 연구검색 시스템에서 필요한 모든 자료를 추출하여 이용하였다. 연구자가 임상에 대해 충분히 이해하고 있기 때문에 수술명과 관련된 변수를 생성할 때, 복합수술이 많은 심장혈관 흉부외과의 특성을 반영하여 대표 수술명이 아닌 세부수술명 다섯 개를 모두 분석 대상으로 삼을 수 있었다. 재수술 여부나 응급수술 여부를 정의할 때는 자유형식으로 작성된 수술명까지 고려하였는데, 이것도 기록 누락이 잦은 현장의 특성을 파악하고 있기에 가능했던 접근이다. 또한 수술 전 재실 여부 및 체중 변화와 같이 기존 연구들에는 잘 다뤄지지 않았지만, 연구자가 경험적으로 중요하다고 느꼈던 변수들을 생성하고 실제 관련성 여부를 확인해볼 수 있었다. 그리고 평균 동맥혈압 변수 자료를 다룰 때처럼, CDW를 통해 추출한 데이터에 발견되는 결측치 및 이상치에 대한 이해가 높기 때문에 그런 상황에서 추가 추출 조건을 더해 더 실제에 가까운 자료를 확보할 수 있었다.

한편 본 연구의 결과는 다음의 세 가지로 요약된다. 첫째, 심장수술 후 중환자실에 입실한 환자의 20%가 재실 기간 중에 섬망 증세를 나타냈다. 전자의무기록의 다섯 가지 자료원을 활용하여 섬망 환자를 최대한 누락 없이 선별하려고 시도했으나, 기존의 연구 방법들과 차이가 있는 결과는 얻지 못했다. 둘째, 섬망군과 비섬망군 간에 유의한 차이를 나타낸 35개의

변수를 확인했다. 이 중 수술 전 결정되는 변수로 대상자의 성별, 연령, 체질량지수, 수술 전 재실 여부 및 수술 전 입원 기간 변수가 있었다. 수술 자체와 관련된 변수로는 대동맥 수술 여부, 관상동맥 수술 여부, 판막 수술 여부, 예상 실혈량, 수술 시간, 및 심폐기 사용 시간이 있었다. 마지막으로 수술직후 획득된 변수로는 산소포화도 이상 여부, 체온, 심박수, 평균 동맥혈압, 동맥혈 pH, 동맥혈 산소 분압, 동맥혈 이산화탄소 분압, 혈당, 젖산, 헤모글로빈, 혈소판, 섬유소원, aPTT, PT, 나트륨, 칼슘, 마그네슘, BUN, 크레아티닌, 총 빌리루빈, 알부민, 그리고 CRP가 있었다. 셋째, 최종 데이터세트에 SMOTE를 적용하여 섬망군과 비섬망군의 비를 1:1로 조정한 뒤 랜덤포레스트 알고리즘을 기반으로 개발한 예측모델의 성능이 가장 우수한 것으로 확인되었다. 랜덤포레스트 알고리즘이 로지스틱 회귀, 의사결정나무, 및 XGBoost 보다 높은 민감도, 양성 예측도, 음성 예측도 및 AUROC 를 나타냈다. 가장 성능이 좋았던 이 모델을 최종 모델로 선택하였다.

본 연구는 최종 모델의 일반화 가능성을 확인하지 못했다는 제한점을 가진다. 그러나 본 연구의 결과를 통해 환자가 심장수술 후 중환자실에 입실했을 때 남은 재실 기간 동안의 섬망 발생 여부가 우수한 성능으로 예측될 수 있음을 확인했다. 최종 예측모델에 이용되는 변수들은 심장수술 후 중환자실에 입실하는 거의 모든 환자들에게서 결측치 없이 획득되며 그 수가 지나치게 많지도 않기 때문에 이는 실용적인 예측모델을 개발할 수 있음을 의미한다. 후속 연구를 통해 본 연구에서와 같은 중환자실의 추가 자료나 타 병원 자료를 사용하여 최종 예측모델의 성능을 검증할 필요가 있다.

제 2 절 제 언

본 연구를 마무리하며 향후 섬망 연구 및 전자의무기록을 활용한 연구의 방향에 대해 다음과 같이 추가적으로 제언하고자 한다.

첫째, 전자의무기록 임상관찰기록지의 'memo' 항목을 활용해 섬망 환자를 선별하는 연구를 제언한다. 중환자실 간호사들은 환자와 관련된 특별한 이벤트가 발생했을 때 그 내용을 발생 시각에 맞추어 memo 항목에 우선적으로 남긴다. 간호일지 등에서와는 달리 임상관찰기록지에서는 당시 환자의 활력징후 수치와 함께 내용을 조회할 수 있다는 점이 편리하기 때문이다. 즉, 중환자에 대한 기록에서 임상관찰기록지는 비교적 가장 정확한 정보를 담고 있고 신뢰할 만한 자료원인 것이다. 그럼에도 입력 데이터가 정형화 되어있지 않아 그 활용 가능성은 아직 미지수이다. '환자가 밤새 잠을 자지 못하고 헛소리함', 'Haldol 0.5@ IM inj' 등의 기록이 실시간으로 남겨지는데, 이를 활용할 수 있다면 섬망 증세 발현 시각에 대한 의미 있는 정보를 얻을 수 있을 것이다. 최근 Jung 등(2022)의 연구에서 환자의 섬망 증세를 시사하는 69개의 단어를 지정한 뒤, 의무기록에 포함되는지 여부를 파악하여 진단에 참고한 바 있다. 앞으로 전자의무기록에 대한 자연어 처리 연구가 더 중요해질 것으로 생각된다.

둘째, 중환자실 재실 기간 중 획득되는 다양한 변수들을 예측변수에 포함시켜 재실 중인 중환자의 섬망 발생 확률이 계속 업데이트되는 예측모델 개발에 대한 연구를 제언한다. 본 연구를 비롯하여 현재 개발되어 있는 섬망 예측모델의 대부분은 입실 이전 혹은 입실 후 24시간 이내와 같이 특정 시간 대로 고정된 예측변수를 사용하는 정적인 모델이다(Rupert et al., 2020). 그러나 매일 환자의 상태가 변화하는 중환자실에서 예측력을 높이기 위해서는 정적인 모델로는 한계가 있다. 중환자실 재실 기간 중의 변수를 활용하는 방향은 내과계 중환자들의 섬망 예측모델 개발 방향과도 유사할 것이다.

셋째, 전자의무기록자료의 data quality 실황과 관련된 질적 연구가 필요하다. 이를 근거로 data quality를 높이기 위해 자료를 입력하는 단계에서부터 정확성을 보장할 수 있도록 구체적인 방안을 마련해야 한다. 본 연구에서 전자의무기록자료를 검토하는 중에 동일한 항목임에도 기록지마다 다른 값으로 기록되는 경우를 빈번하게 마주했다. 예를 들어서, 마취기록과 체외순환기록지 간에 심폐기 작동 시작 및 종료 시간이 일치하지 않는 경우가 있었다. 같은 항목에 대해서는 기록지가 달라도 서로 일치되는 값을 가질 수 있도록 전산 기반을 마련해야 한다. 이러한 연구가 누적된다면 추후 전자의무기록자료를 활용한 많은 연구들이 큰 도움을 받을 것으로 기대한다.

참 고 문 헌

강지연, & 최민정. (2021). 국내 중환자실 섬망 예방 중재에 관한 체계적 고찰 및 메타분석. *중환자간호학회지*, *14*(3), 141-156. https://doi.org/10.34250/jkccn.2021.14.3.141

공경희, 하이경, & 강인순. (2015). 섬망사정도구를 사용하는 중환자실 간호사의 섬망인식, 섬망사정의 장애요인 및 섬망사정 수행의 중요도. *중환자간호학회지*, *8*(2), 33-42.

노은영, & 박연환. (2019). 중환자실 내 성인 심장수술 환자의 섬망발생률 및 위험요인: 후향적 조사연구. *성인간호학회지*, *31*(2), 146-155. https://doi.org/10.7475/kjan.2019.31.2.146

손정현, 박진영, 라세희, 신증수, 손인정, 오주영, 안지선, 안석균, & 김재진. (2014). 국내 일 병원 중환자실에서 관찰한 섬망 환자의 예후. *신경정신의학*, *53*(6), 418-425. https://doi.org/10.4306/jknpa.2014.53.6.418

이현심, & 김소선. (2013). 내과계 중환자 섬망발생 선별모형 개발. *임상간호연구*, 19(3), 357-368.

American Psychiatric Association. (2013). Delirium. In *Diagnostic and statistical manual of mental disorders* (5th ed.), 596-598. https://doi.org/10.1176/appi.books.9780890425596

Awan, O. M., Buhr, R. G., & Kamdar, B. B. (2021). Factors influencing CAM-ICU documentation and inappropriate "Unable to Assess" responses. *American Journal of Critical Care*, *30*(6), 99-107. https://doi.org/10.4037/ajcc2021599

Bakker, R. C., Osse, R. J., Tulen, J. H., Kappetein, A. P., & Bogers, A. J. (2012). Preoperative and operative predictors of delirium after cardiac surgery in elderly patients. *European Journal of Cardio-thoracic Surgery*, *41*(3), 544-549. https://doi.org/10.1093/ejcts/ezr031

Berger, M., Terrando, N., Smith, S. K., Browndyke, J. N., Newman, M. F., & Mathew, J. P. (2018). Neurocognitive function after cardiac surgery: From phenotypes to mechanisms. *Anesthesiology*, *129*(4), 829–851. https://doi.org/10.1097/ALN.0000000000002194

Bishara, A., Chiu, C., Whitlock, E. L., Douglas, V. C., Lee, S., Butte, A. J., Leung, J. M., & Donovan, A. L. (2022). Postoperative delirium prediction using machine learning models and preoperative electronic health record data. *BMC Anesthesiology*, *22*(1). https://doi.org/10.1186/s12871-021-01543-y

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Celi, L., Csete, M., & Stone, D. (2014). Optimal data systems: The future of clinical predictions and decision support. *Current Opinion in Critical Care*, *20*(5), 573–580. https://doi.org/10.1097/MCC.0000000000000137

Cerejeira, J., Firmino, H., Vaz-Serra, A., & Mukaetova-Ladinska, E. B. (2010). The neuroinflammatory hypothesis of delirium. *Acta Neuropathologica*, *119*(6), 737–754. https://doi.org/10.1007/s00401-010-0674-1

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence research*, *16*, 321–357.

Chen, H., Mo, L., Hu, H., Ou, Y., & Luo, J. (2021). Risk factors of postoperative delirium after cardiac surgery: A meta-analysis. *Journal of Cardiothoracic Surgery*, *16*(1), 113. https://doi.org/10.1186/s13019-021-01496-w

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International*

Conference on Knowledge Discovery and Data Mining, USA. https://doi.org/10.1145/2939672.2939785

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, *110*, 12–22. https://doi.org/10.1016/j.jclinepi.2019.02.004

Corradi, J. P., Thompson, S., Mather, J. F., Waszynski, C. M., & Dicks, R. S.. (2018). Prediction of incident delirium using a random forest classifier. *Journal of Medical Systems*, *42*(12). https://doi.org/10.1007/s10916-018-1109-0

Dubiel, C., Hiebert, B. M., Stammers, A. N., Sanjanwala, R. M., Tangri, N., Singal, R. K., Manji, R. A., Rudolph, J. L., & Arora, R. C. (2022). Delirium definition influences prediction of functional survival in patients one-year postcardiac surgery. *The Journal of Thoracic and Cardiovascular Surgery*, *163*(2), 725–734. https://doi.org/10.1016/j.jtcvs.2020.07.028

Eertmans, W., De Deyne, C., Genbrugge, C., Marcus, B., Bouneb, S., Beran, M., Fret, T., Gutermann, H., Boer, W., Vander Laenen, M., Heylen, R., Mesotten, D., Vanelderen, P., & Jans, F. (2020). Association between postoperative delirium and postoperative cerebral oxygen desaturation in older patients after cardiac surgery. *British Journal of Anaesthesia*, *124*(2), 146–153. https://doi.org/10.1016/j.bja.2019.09.042

Ely, E. W., Inouye, S. K., Bernard, G. R., Gordon, S., Francis, J., May, L., Truman, B., Speroff, T., Gautam, S., Margolin, R., Hart, R. P., & Dittus, R. (2001). Delirium in mechanically ventilated patients: Validity and reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *JAMA: Journal of the*

American Medical Association, *286*(21), 2703–2710. https://doi.org/10.1001/jama.286.21.2703

Ely, E. W., Margolin, R., Francis, J., May, L., Truman, B., Dittus, R., Speroff, T., Gautam, S., Bernard, G. R., & Inouye, S. K. (2001). Evaluation of delirium in critically ill patients: Validation of the Confusion Assessment Method for the Intensive Care Unit (CAM–ICU). *Critical Care Medicine*, *29*(7), 1370–1379. https://doi.org/10.1097/00003246–200107000–00012

Fernandez, A., Garcia, S., Herrera, F., & Chawla, N. (2018). SMOTE for learning from imbalanced data: Progress and challenges, marking the 15–year anniversary. *Journal of Artificial Intelligence Research*, *61*, 863–905. https://doi.org/10.1613/jair.1.11192

Girard, T. D., Thompson, J. L., Pandharipande, P. P., Brummel, N. E., Jackson, J. C., Patel, M. B., Hughes, C. G., Chandrasekhar, R., Pun, B. T., Boehm, L. M., Elstad, M. R., Goodman, R. B., Bernard, G. R., Dittus, R. S., & Ely, E. W. (2018). Clinical phenotypes of delirium during critical illness and severity of subsequent long–term cognitive impairment: A prospective cohort study. *The Lancet Respiratory Medicine*, *6*(3), 213–222. https://doi.org/10.1016/S2213–2600(18)30062–6

Hasan, M. K., Alam, M. A., Das, D., Hossain, E., & Hasan, M. (2020). Diabetes prediction using ensembling of different machine learning classifiers. *IEEE Access*, *8*, 76516–76531. https://doi.org/10.1109/access.2020.2989857

Hofer, J., Fries, D., Solomon, C., Velik–Salchner, C., & Ausserer, J. (2016). A snapshot of coagulopathy after cardiopulmonary bypass. *Clinical and Applied Thrombosis/Hemostasis*, *22*(6), 505–511. https://doi.org/10.1177/1076029616651146

Hshieh, T. T., Inouye, S. K, & Oh, E. S. (2017). Delirium in the elderly. *Psychiatric Clinics of North America*, *41*(1), 1-17. https://doi.org/10.1016/j.psc.2017.10.001

Hur, S., Ko, R.-E., Yoo, J., Ha, J., Cha, W. C., & Chung, C. R. (2021). A machine learning-based algorithm for the prediction of intensive care unit delirium (PRIDE): Retrospective study. *JMIR Medical Informatics*, *9*(7). https://doi.org/10.2196/23401

Inouye, S. K., Bogardus, S. T., Charpentier, P. A., Leo-Summers, L., Acampora, D., Holford, T. R., & Cooney, L. M.. (1999). A multicomponent intervention to prevent delirium in hospitalized older patients. *New England Journal of Medicine*, *340*(9), 669-676. https://doi.org/10.1056/nejm199903043400901

Inouye, S. K., Vandyck, C. H., Alessi, C. A., Balkin, S., Siegal, A. P., & Horwitz, R. I. (1990). Clarifying confusion - The Confusion Assessment Method - A new method for detection of delirium. *Annals of Internal Medicine*, *113*(12), 941-948. https://doi.org/10.7326/0003-4819-113-12-941

Jarvela, K., Porkkala, H., Karlsson, S., Martikainen, T., Selander, T., & Bendel, S. (2018). Postoperative delirium in cardiac surgery patients. *Journal of Cardiothoracic and Vascular Anesthesia*, *32*(4), 1597-1602. https://doi.org/10.1053/j.jvca.2017.12.030

Jauk, S., Kramer, D., Grossauer, B., Rienmuller, S., Avian, A., Berghold, A., Leodolter, W., & Schulz, S. (2020). Risk prediction of delirium in hospitalized patients using machine learning: An implementation and prospective evaluation study. *Journal of the American Medical Informatics Association*, *27*(9), 1383-1392. https://doi.org/10.1093/jamia/ocaa113

Jung, J. W., Hwang, S., Ko, S., Jo, C., Park, H. Y., Han, H. S., Lee, M. C., Park, J. E., & Ro, D. H. (2022). A machine-learning model to

predict postoperative delirium following knee arthroplasty using electronic health records. *BMC Psychiatry*, *22*(1). https://doi.org/10.1186/s12888-022-04067-y

Jung, J. H., Lim, J. H., Kim, E. J., An, H. C., Kang, M. K., Lee, J., Min, Y. K., Park, E. Z., Song, X. H., Kim, H. R., & Lee, S. M. (2013). The experience of delirium care and clinical feasibility of the CAM-ICU in a Korean ICU. *Clinical Nursing Research*, *22*(1), 95-111. https://doi.org/10.1177/1054773812447187

Kim, S., Kim, J. J., Oh, J., Park, J., & Park, J. Y. (2018). Delirium characteristics and outcomes in medical and surgical lnpatients: A subgroup analysis. *Journal of Critical Care*, *43*, 156-162. https://doi.org/10.1016/j.jcrc.2017.08.011

Kim, T. W., Ko, R. E., Na, S. J., Chung, C. R., Choi, K. H., Park, C. M., & Yang, J. H. (2023). Associations of albumin and nutritional index factors with delirium in patients admitted to the cardiac intensive care unit. *Frontiers in Cardiovascular Medicine*, *10*. https://doi.org/10.3389/fcvm.2023.1100160

Ko, C., Kim, J. J., Cho, D., Oh, J., & Park, J. Y. (2019). A study on clinical variables contributing to differentiation of delirium and non-delirium patients in the ICU. *Korean Journal of Psychosomatic Medicine*, *27*(2), 101–110. https://doi.org/10.22722/KJPM.2019.27.2.101

Krewulak, K. D., Stelfox, H. T., Leigh, J. P., Ely, E. W., & Fiest, K. M. (2018). Incidence and prevalence of delirium subtypes in an adult ICU: A systematic review and meta-analysis. *Critical Care Medicine*, *46*(12), 2029-2035. https://doi.org/10.1097/ccm.0000000000003402

la Cour, K. N., Andersen-Ranberg, N. C., Weihe, S., Poulsen, L. M., Mortensen, C. B., Kjer, C. K. W., Collet, M. O., Estrup, S., &

Mathiesen, O. (2022). Distribution of delirium motor subtypes in the intensive care unit: A systematic scoping review. *Critical Care*, *26*(1). https://doi.org/10.1186/s13054-022-03931-3

Lantz, B. (2020). *R을 활용한 머신러닝* (윤성진 외, Trans.; 3rd ed.). 에이콘출판. (Original work published 2019).

Lee, A., Mu, J. L., Joynt, G. M., Chiu, C. H., Lai, V. K. W., Gin, T., & Underwood, M. J. (2017). Risk prediction models for delirium in the intensive care unit after cardiac surgery: A systematic review and independent external validation. *British Journal of Anaesthesia*, *118*(3), 391-399. https://doi.org/10.1093/bja/aew476

Likhvantsev, V. V., Landoni, G., Grebenchikov, O. A., Ovezov, A. M., Skripkin, Y. V., Lembo, R., Gaevskiy, D. I., Tereshina, A. A., & Yavorovskiy, A. G. (2021). Perioperative dexmedetomidine supplement decreases delirium incidence after adult cardiac surgery: A randomized, double-blind, controlled study. *Journal of Cardiothoracic and Vascular Anesthesia*, *35*(2), 449-457. https://doi.org/10.1053/j.jvca.2020.02.035

Lindroth, H., Bratzke, L., Purvis, S., Brown, R., Coburn, M., Mrkobrada, M., Chan, M. T. V., Davis, D. H. J., Pandharipande, P., Carlsson, C. M., & Sanders, R. D. (2018). Systematic review of prediction models for delirium in the older adult inpatient. *BMJ Open*, 8(4), Article e019223. https://doi.org/10.1136/bmjopen-2017-019223

Maldonado, J. R. (2013). Neuropathogenesis of delirium: Review of current etiologic theories and common pathways. *American Journal of Geriatric Psychiatry*, *21*(12), 1190-1222. https://doi.org/10.1016/j.jagp.2013.09.005

Marcantonio, E. R., Flacker, J. M., Wright, R. J., & Resnick, N. M.

(2001). Reducing delirium after hip fracture: A randomized trial. *Journal of American Geriatrics Society, 49*(5), 516–522. https://doi.org/10.1046/j.1532-5415.2001.49108.x

Marcantonio, E. R., Goldman, L., Orav, E. J., Cook, E. F., & Lee, T. H. (1998). The association of intraoperative factors with the development of postoperative delirium. *The American Journal of Medicine*, *105*(5), 380–384. https://doi.org/10.1016/S0002-9343(98)00292-7

Marcantonio, E., Rudolph, J., Culley, D., Crosby, G., Alsop, D., & Inouye, S. (2006). Review article: serum biomarkers for delirium. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 61*(12), 1281–1286. https://doi.org/10.1093/gerona/61.12.1281

Menzenbach, J., Kirfel, A., Guttenthaler, V., Feggeler, J., Hilbert, T., Ricchiuto, A., Staerk, C., Mayr, A., Coburn, M., Wittmann, M., & Group, P. C. (2022). PRe-Operative Prediction of postoperative DElirium by appropriate SCreening (PROPDESC) development and validation of a pragmatic POD risk screening score based on routine preoperative data. *Journal of Clinical Anesthesia*, *78*. https://doi.org/10.1016/j.jclinane.2022.110684

Mufti, H., Hirsch, G., Abidi, S., & Abidi, S. (2019). Exploiting machine learning algorithms and methods for the prediction of agitated delirium after cardiac surgery: Models development and validation study. *JMIR Medical Informatics*, *7*(4). https://doi.org/10.2196/14993

Nusinovici, S., Tham, Y. C., Chak Yan, M. Y., Wei Ting, D. S., Li, J., Sabanayagam, C., Wong, T. Y., & Cheng, C. Y. (2020). Logistic regression was as good as machine learning for predicting major chronic diseases. *Journal of Clinical Epidemiology*, *122*, 56–69.

https://doi.org/10.1016/j.jclinepi.2020.03.002

Nyitrai, T., & Virág, M. (2019). The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences*, *67*, 34-42. https://doi.org/10.1016/j.seps.2018.08.004

Olegas, N. (2015). CRISP data mining methodology extension for medical domain. *Baltic Journal of Modern Computing*, *3*(2), 92.

Ouimet, S., Riker, R., Bergeon, N., Cossette, M., Kavanagh, B., & Skrobik, Y. (2007). Subsyndromal delirium in the ICU: Evidence for a disease spectrum. *Intensive Care Medicine*, *33*(6), 1007-1013. https://doi.org/10.1007/s00134-007-0618-y

Peterson, J., Pun, B., Dittus, R., Thomason, J., Jackson, J., Shintani, A., & Ely, E. (2006). Delirium and its motoric subtypes: A study of 614 critically ill patients. *Journal of the American Geriatrics Society*, *54*(3), 479-484. http://doi.org/10.1111/j.1532-5415.2005.00621.x

Rood, P., Huisman, G., Vermeulen, H., Schoonhoven, L., Pickkers, P., & van den Boogaard, M. (2018). Effect of organisational factors on the variation in incidence of delirium in intensive care unit patients: A systematic review and meta-regression analysis. *Australian Critical Care*, *31*(3), 180-187. https://doi.org/10.1016/j.aucc.2018.02.002

Roy, K., Ahmad, M., Waqar, K., Priyaah, K., Nebhen, J., Alshamrani, S. S., Raza, M. A., Ali, I., & Uddin, M. I. (2021). An enhanced machine learning framework for type 2 diabetes classification using imbalanced data with missing values. *Complexity*, *2021*, 1-21. https://doi.org/10.1155/2021/9953314

Rudiger, A., Begdeda, H., Babic, D., Krüger, B., Seifert, B., Schubert, M., Spahn, D. R., & Bettex, D. (2016). Intra-operative events

during cardiac surgery are risk factors for the development of delirium in the ICU. *Critical Care*, *20*(1), 264–264. https://doi.org/10.1186/s13054-016-1445-8

Ruppert, M. M., Lipori, J., Patel, S., Ingersent, E., Cupka, J., Ozrazgat-Baslanti, T., Loftus, T., Rashidi, P., & Bihorac, A. (2020). ICU delirium-prediction models: A systematic review. *Critical Care Explorations*, *2*(12). https://doi.org/10.1097/CCE.0000000000000296

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia computer science*, *181*, 526-534. https://doi.org/10.1016/j.procs.2021.01.199

Silva, A. R., Regueira, P., Albuquerque, E., Baldeiras, I., Cardoso, A. L., Santana, I., & Cerejeira, J. (2021). Estimates of geriatric delirium frequency in noncardiac surgeries and its evaluation across the years: A systematic review and meta-analysis. *Journal of the American Medical Directors Association*, *22*(3). https://doi.org/10.1016/j.jamda.2020.08.017

Simeone, S., Pucciarelli, G., Perrone, M., Teresa, R., Gargiulo, G., Guillari, A., Castellano, G., Tommaso, L. D., Niola, M., & Iannelli, G. (2018). Delirium in ICU patients following cardiac surgery: An observational study. Journal of clinical nursing, *27*(9-10), 1994–2002. https://doi.org/10.1111/jocn.14324

Smulter, N., Lingehall, H. C., Gustafson, Y., Olofsson, B., & Engstrom, K. G. (2013). Delirium after cardiac surgery: Incidence and risk factors. *Interactive Cardiovascular and Thoracic Surgery*, *17*(5), 790-796. https://doi.org/10.1093/icvts/ivt323

Swan, J. T. (2014). Decreasing inappropriate unable-to-assess

ratings for the confusion assessment method for the intensive care unit. *American Journal of Critical Care*, *23*(1), 60–69. https://doi.org/10.4037/ajcc2014567

Terrando, N., & Akassoglou, K. (2022). Breaking barriers in postoperative delirium. *British Journal of Anaesthesia*, *129*(2), 147–150. https://doi.org/10.1016/j.bja.2022.05.004

Tilouche, N., Hassen, M. F., Ali, H. B., Jaoued, O., Gharbi, R., & El Atrous, S. S. (2018). Delirium in the intensive care unit: Incidence, risk factors, and impact on outcome. *Indian Journal of Critical Care Medicine*, *22*(3), 144–149. https://doi.org/10.4103/ijccm.IJCCM_244_17

van Eijk, M. M., van den Boogaard, M., van Marum, R. J., Benner, P., Eikelenboom, P., Honing, M. L., van der Hoven, B., Horn, J., Izaks, G. J., Kalf, A., Karakus, A., Klijn, I. A., Kuiper, M. A., de Leeuw, F.-E., de Man, T., van der Mast, R. C., Osse, R.-J., de Rooij, S. E., Spronk, P. E., . . . Slooter, A. J. (2011). Routine use of the confusion assessment method for the intensive care unit: A multicenter study. *American Journal of Respiratory Critical Care Medicine*, *184*(3), 340–344. https://doi.org/10.1164/rccm.201101-0065OC

van Rompaey, B., Elseviers, M. M., van Drom, W., Fromont, V., & Jorens, P. G. (2012). The effect of earplugs during the night on the onset of delirium and sleep perception: A randomized controlled trial in intensive care patients. *Critical Care*, *16*(3). https://doi.org/10.1186/cc11330

Vasilevskis, E. E., Chandrasekhar, R., Holtze, C. H., Graves, J., Speroff, T., Girard, T. D., Patel, M. B., Hughes, C. G., Cao, A., Pandharipande, P. P., & Ely, E. W. (2018). The cost of ICU delirium and coma in the intensive care unit patient. *Med Care*,

56(10), 890-897.
https://doi.org/10.1097/MLR.0000000000000975

Zhang, W. Y., Wu, W. L., Gu, J. J., Sun, Y., Ye, X. F., Qiu, W. J., Su, C. Q., Zhang, S. Q., & Ye, W. Q. (2015). Risk factors for postoperative delirium in patients after coronary artery bypass grafting: A prospective cohort study. *Journal of Critical Care*, *30*(3), 606-612. https://doi.org/10.1016/j.jcrc.2015.02.003

부록

부록1. 의학연구윤리심의위원회(IRB) 심의 결과 통보서

서울대학교의과대학/서울대학교병원
의학연구윤리심의위원회

서울대학교의과대학/서울대학교병원 의학연구윤리심의위원회	
Tel : 82-02-2072-0694/2266	
FAX : 82-02-3675-6824	서울특별시 종로구 대학로 101번지 (우)03080

심의결과통보서

IRB No.	H-2201-074-1291			제출경로	서울대병원		
수신	책임연구자	이지연		소속		직위	간호직
	의뢰기관						
연구과제명	전자의무기록자료를 활용한 중환자 섬망 예측모델 개발 및 평가						
Protocol No.				Version No.	1.2		
생명 윤리법에 따른 분류	■ 인간대상연구 □ 인체유래물연구 □ 배아줄기세포주이용연구 □ 배아연구 □ 체세포복제배아연구 □ 단성생식배아연구 □ 배아생성의료기관 □ 인체유래물은행						
연구종류	□전향적 연구(Prospective Study) ■후향적 연구(Retrospective Study)						
	■비중재(관찰)연구	□사례(환자)등록연구(Registry study) □코호트 연구(Cohort study) □사례(환자)군연구(Case Series) □증례보고(Case Report) □단면연구(Cross-Sectional study) ■환자대조군연구(Case-Control study) □조사,설문,인터뷰연구 □사회행동과학 연구(Social, Behavioral & Research) □인체유래물 조사분석연구 □인체유래물저장소(Repository) □생태학적 연구(Ecological study) □시판후사용성적조사(PMS) □임상시험용의약품/의료기기 치료목적 사용 □기타					
	□의약품/의료기기 임상시험 등	□의약품/생물학적제제 임상시험	□예비연구(Pilot Study) □약동학·약력학 연구 □생물학적동등성 □제1상 □제1/2상 □제2상 □제2/3상 □제3상 □제4상				
		□첨단바이오의약품 임상시험	□세포치료제 □유전자치료제 □조직공학치료제 □융복합치료제 □장기추적조사				
		□의료기기 임상시험	□예비연구(Pilot Study) □탐색 임상시험 □확증 임상시험				
			분류번호/등급				
		□체외진단의료기기 임상적 성능시험	□탐색 임상시험 □확증 성능시험				
			분류번호/등급				

서울대학교의과대학/서울대학교병원
의학연구윤리심의위원회

		식약처 승인 여부	□식약처 승인 대상 □식약처 승인 제외 대상	
	□기타 중재연구	□기타 임상시험	□의료행위(수술법,마취법 등) □개인용 건강관리 중재 □기타	
		□인체적용시험	□화장품 인체적용시험 □건강기능식품 인체적용시험	
	□첨단재생임상연구	□세포치료 □유전자치료 □조직공학치료 □융복합치료 □장기추적조사 □인체세포등 채취 및 처리연구		
		첨단심의위원회 심의결과	위원회 승인일	
			위험도	□저위험 □중위험 □고위험 식약처 승인일( )
연구목적	□국내(MFDS)허가용 □해외허가용 □학술용			
연구계획서 승인일	2022년 02월 18일 (정기보고주기 : 12개월)			
승인유효 만료일	2023년 02월 17일			
심의대상	연구계획심의 의뢰서(수정후신속심의에 대한 답변)			
심의종류	신속심의		심의일자	2022년 02월 17일
접수일자	2022년 02월 09일		심의결과통보일	2022년 02월 18일
심의목록	1. 연구계획심의 의뢰서(수정후신속심의에 대한 답변) 2. 연구계획서 ver1.2 3. 중례기록서			
심의결과	승인			
연구의 위험도	최소위험 연구(minimal risk)			
심의의견	- 검토의견에 대한 답변 및 변경 사항을 확인하였고, IRB승인 기준에 따라 승인합니다.			

의 학 연 구 윤 리 심 의 위 원 회 위 원 장

본 통보서에 기재된 사항은 IRB의 기록된 내용과 일치 함을 증명합니다.
본 기관 IRB는 생명윤리 및 안전에 관한 법률, 약사법, 의료기기법 및 ICH-GCP 등 관련 연구를 준수합니다.
본 연구와 이해충돌(Conflict of Interest)가 있는 위원이 있을 경우 연구의 심의에서 배제하였습니다.

부록2. 주요 심장수술 분류 및 해당 세부수술명 목록

	Surgery Types	Specific Surgery names (details not included)
1	AORTA	Aorta to other artery bypass
2		Aortosubclavian-carotid artery bypass
3		Composite graft interposition for annuloaortic ectasia (Bentall operation)
4		Excision of vessel
5		Femoro-femoral bypass
6		Insertion of aortic stent
7		Repair of aneurysm
8		Repair of ruptured aneurysm of sinus valsalva
9		Replacement of aorta
10		Replacement of thoracoabdominal aortic aneurysm
11		Resection of aneurysm
12		Resection of vessel, abdominal aorta
13		Root remodeling operation for annuloaortic ectasia (David operation)
14		Thoracic aorta to abdominal aorta bypass
15	CORONARY	Aortocoronary bypass
16		Aortocoronary bypass of one coronary artery
17		Aortocoronary bypass of two or more coronary artery
18		Off pump CABG (coronary artery bypass graft) surgery
19		Redo coronary artery bypass graft
20	MAZE	Cox-Maze procedure
21		Maze operation
22	VALVE	Open heart annuloplasty of tricuspid valve
23		Open heart valvuloplasty without replacement
24		Operation on heart valve
25		Pulmonary valve valvuloplasty
26		Repair of heart valve
27		Replacement of heart valve
28		Replacement of redo-valve
29		Revision of prosthesis of heart valve
30	ELSE	Embolectomy or thrombectomy
31		Endarterectomy

32	Excision of benign mediastinal tumor
33	Excision of cardiac tumor
34	Excision of malignant soft tissue tumor
35	Excision of mediastinal lesion or tissue
36	Excision of mediastinal mass
37	Excision of myxoma
38	Excision of pericardial mass
39	Excision of pleural mass
40	Implantation of heart assist device
41	Insertion of VAD (Ventricular assist device)
42	Mechanical cardiopulmonary bypass assist (left ventricular assist device (L−VAD))
43	Operation for mediastinal tumor
44	Operation for thromboembolism
45	Pericardial window operation
46	Repair of atrial septal defect
47	Repair of heart
48	Repair of ventricular septal defect
49	Resection of intracardiac tumor or foreign body
50	Revision of vascular procedure with interposition graft
51	Robotic mass excision, mediastinum
52	Subxiphoid approach thymectomy
53	Thrombectomy
54	Thymectomy
55	Total correction of tetralogy of Fallot
56	Transplantation of heart
57	Transplantation of heart−lung
58	Wide excision of malignant soft tissue tumor

부록3. 연속변수 이상치 기준, 개수 및 비율

Variable	Q1	Q3	Q1− 1.5*IQR	Q3+ 1.5*IQR	Number of outliers	Percentage of outliers (%)
1. Preoperative variable						
Age (years)	59	74	36.5	96.5	55	2.6
Body Mass Index (kg/m^2)	21.3	25.6	14.7	32.2	27	1.3
Hospitalization before the surgery (days)	3	7	−3	13	157	7.5
2. Intraoperative variable						
Estimated Blood Loss (ml)	100	800	−950	1850	102	4.9
Operation time (min)	257	375	80	552	60	2.9
CPB time (min)	0	187	−280.5	467.5	2	0.1
3. Immediate postoperative variable						
Body weight gain (kg)	0.4	2.65	−2.98	6.03	68	3.3
Vital Sign						
Body temperature (℃)	35.6	36.5	34.25	37.85	42	2.0
Heart rate (bpm)	70	88	43	115	53	2.5
Mean ABP (mmHg)	67	92	29.5	129.5	16	0.8
Arterial Blood Gas analysis						
pH	7.38	7.47	7.24	7.6	25	1.2
pO_2 (mmHg)	131	259	−61	451	18	0.9
pCO_2 (mmHg)	35	42	24.5	52.5	75	3.6
HCO_3^- (mmol/L)	23.5	26.8	18.55	31.75	43	2.1
Complete Blood Count						
WBC ($\times 10^3$/ul)	7.28	12.38	−0.38	20.04	73	3.5
Hemoglobin (g/dL)	10.3	12.3	7.3	15.3	17	0.8
Platelet ($\times 10^3$/ul)	80.5	140	−8.75	229.25	44	2.1
Coagulation test						
Fibrinogen (mg/dL)	166	245	47.5	363.5	64	3.1
aPTT (sec)	30.6	38.8	18.3	51.1	158	7.6
PT (INR)	1.16	1.35	0.87	1.64	97	4.7
Other blood serum test						
Potassium (mmol/L)	3.9	4.5	3	5.4	40	1.9
Sodium (mmol/L)	137	142	129.5	149.5	25	1.2
Calcium (mg/dL)	8	9.3	6.05	11.25	76	3.6
Magnesium (mEq/L)	1.6	2.1	0.85	2.85	27	1.3
Phosphorus (mg/dL)	2.5	3.5	1	5	54	2.6
BUN (mg/dL)	12	19	1.5	29.5	148	7.1
Creatinine (mg/dL)	0.66	0.99	0.17	1.48	202	9.7
AST (IU/L)	26	72	−43	141	258	12.4
ALT (IU/L)	13	26	−6.5	45.5	186	8.9
Total Bilirubin (mg/dL)	1	1.8	−0.2	3	95	4.6
Alk.phos (IU/L)	34	52	7	79	84	4.0
Albumin (g/dL)	2.7	3.4	1.65	4.45	29	1.4
CRP (mg/dL)	0.52	1.74	−1.31	3.57	184	8.8
Glucose (mg/dL)	110	160	35	235	39	1.9
Lactate (mmol/L)	1	2.2	−0.8	4	161	7.7

부록4. R code 예시: 수술명 관련 변수

```
# 1. SUPREME 1.0, '병원공통_수술기록_FREE TEXT' 서식에서 추출한 원본 excel
파일을 불러온다. (해당 기록지에는 최대 16개의 항목에 대한 내용 있음)

OP <- readxl::read_excel("OP.xlsx")
View(OP)
colnames(OP)<-c('idid', 'ID', 'write.date', 'type', 'value') #변수명 지정


# 2. 행/열 전환 후 원하는 순서대로 기존 16개 열을 재배열 후 변수명을 지정한다.
## 이 중 op1~op5이 표준화된 세부수술명임 (재배열 후 9~13열에 해당됨)
## op1 은 NA 없음, 나머지 op2~op5는 NA 값 가지기도 함

OP <- tidyr::pivot_wider(OP, names_from = 'type', values_from = 'value')
OP <- OP [,c(1,2,3,4,5,6,7,8,12,11,10,9,14,13,15,16)]
colnames(OP) <- c('idid', 'ID', 'write.date', 'anes', 'date', 'dg.bef', 'dg.aft','op.free',
                  'op1', 'op2', 'op3', 'op4', 'op5', 'finding1', 'finding2', 'finding3')


# 3. 한편, op 1~op5에 등장하는 수술명 전수에 대한 검토를 시행하였고, 표준화된
세부수술명 195개(세부속성 포함; 원본 excel 파일에 기록된 형태 그대로; 포함하지 않을
경우 58개)를 다섯 종류로 나누어 분석에 포함시켰다. opname.xlsx 파일에 이와 같은
수술명 분류 결과를 저장한 뒤 활용하였다.

op_name_group <- readxl::read_excel("opname.xlsx")
View(op_name_group)

group_aorta<- op_name_group$ 'op name'[1:33] #33개
group_coronary<- op_name_group$ 'op name'[34:42] #9개
group_valve<- op_name_group$'op name' [43:97] #55개
group_maze<- op_name_group$'op name' [98:100] #3개
group_else <- op_name_group$'op name' [101:195] #95개


# 4. OP 에 수술명 관련 변수(이분형)에 대한 새로운 5열을 추가하고 (직전에는 16열
이었음), 앞서 시행한 수술명 분류에 따라 값을 지정한다.
## OP의 9~13열은 op1~op5이고, 새로 추가된 열은 17~21열임

OP[,17:21]<-0
colnames(OP)[,17:21]<-c('AORTA', 'CORONARY', 'VALVE','MAZE', 'ELSE')
for(i in 1:nrow(OP)){
```

```
  for(l in 9:13) {if(OP[i,l] %in% group_aorta) {OP$AORTA[i]<-1}}}
for(i in 1:nrow(OP)) {
  for(l in 9:13) {if(OP[i,l] %in% group_coronary) {OP$CORONARY[i]<-1}}}
for(i in 1:nrow(OP)) {
  for(l in 9:13) {if(OP[i,l] %in% group_valve) {OP$VALVE[i]<-1}}}
for(i in 1:nrow(OP)) {
  for(l in 9:13) {if(OP[i,l] %in% group_maze) {OP$MAZE[i]<-1}}}
for(i in 1:nrow(OP)) {
  for(l in 9:13) {if(OP[i,l] %in% group_else) {OP$ELSE[i]<-1}}}
```

5. OP의 date(수술기록지 상 수술일자에 해당된다)를 활용하여 각 행에 대해 입실 건 num 을 대응시킨다.
본 연구에서는 연구 기간 중 해당 중환자실 입실 건 전체를 입실순서로 나열 후 1부터 차례로 번호를 부여하였다. 각 입실 건에 대한 구분자로서의 num 변수를 활용하였다. 같은 환자가 여러 번 입실하는 경우가 있어 환자 ID는 사용에 제한이 있었다.
다른 변수의 가공에도 아래와 코드를 동일하게 혹은 변형하여 활용하였다. (그래서 편의를 위해 dataset 이나 BASE 등의 중간변수를 사용하였다.)

```
dataset <- OP
BASE <- base  #연구 대상자 ID, RID(Research ID), num, 중환자실 입퇴실시간
              #(in.date time, out.date.time) 등이 저장되어있음

dataset$num <-0
for(i in 1:length(group)) {
    id <- group[i]
    table1 <- dataset[which(dataset$ID == id),]
    table2 <- BASE[which(BASE$ID ==id),]
    for(l in 1:nrow(table1)) {
      for( m in 1:nrow(table2)) {
         if (table1$date[l] >= date(table2$in.date.time[m])-1 &
             table1$date[l] <= date(table2$out.date.time)[m] ) {
             table1$num[l] <- table2$num[m]
             dataset[which(dataset$ID == id), 'num'] <- table1$num}
     }}}

library(dplyr)
dataset <- dataset %>% filter(num != 0)
OP <- dataset
```

입실 건수 1개에 1개의 주요 수술이 대응되는 경우가 본 연구의 분석 대상이다.

부록5. R code 예시: Estimated Blood Loss (EBL) 변수

```
# 1. SUPREME 2.0 에서 추출한 원본 excel 파일을 불러온다.
## 마취기록지 하나당, 시간별로 기록된 EBL값을 합쳐 총량을 구해야 한다
## 마취기록의 '총 EBL' 항목이 별도로 있으나, 총 EBL 항목에는 기록 시간이 따로 없어
match 시키기 어렵기 때문이다 (주요 수술 외 Blood control 등의 추가 수술로 마취기록이
여러 개 일 수 있고, 마취기록지가 완성된 시간이 수술 시간과 차이가 날 수 있다.)

ebl <- readxl::read_excel("EBL.xlsx")
head(ebl)



# 2. 원본 엑셀파일에서 필요한 열을 추출한다.

ebl <- ebl[,c(1,5,7)]
colnames(ebl) <- c('RID', 'rec.time', 'value')
ebl$rec.time <- lubridate::ymd_hms(ebl$rec.time)
ebl$value <- as.numeric(ebl$value)
head(ebl)



# 3. 연구 대상자 자료만 선별한다. - 연구 대상자 ID, RID(Research ID), num(입실 건
구분자) 등이 저장된 'base' 데이터프레임 이용

ebl <- ebl[which(ebl$RID %in% base$RID),]
ebl <- ebl[!is.na(ebl$rec.time),]



# 4. 시간별로 작성된 EBL 수치에 대해 해당되는 입실 건(num)을 대응시킨다. (각 num에
대해 RID가 같으면서, 수술 입실 시간과 퇴실 시간 사이 기록된 EBL을 대응시킨다.)

base_ebl <- base[which(base$RID %in% ebl$RID),] #오류방지용
nrow(base) ; nrow(base_ebl)

dataset <-ebl
BASE <- base_ebl
head(dataset); head(BASE)

group <- unique(dataset$RID)
dataset$num <- NA

for (i in 1:length(group)) {
  id <- group[i]
  table1 <- dataset[which(dataset$RID == id),]
  table2 <- BASE[which(BASE$RID ==id),]
  {for (l in 1:nrow(table1))
```

```
  { for( m in 1:nrow(table2) )
    { if (table1$'rec.time'[l] >= table2$'in.time'[m]
        & table1$'rec.time'[l] <= table2$'out.time'[m])
    {T<-table1$'rec.time'[l]
     dataset[which(dataset$RID == id
                & dataset$rec.time == T), 'num'] <- table2$num[m]}
    }}}}

head(dataset)
dataset <- dataset[!is.na(dataset$num),]
dataset <- dataset[!is.na(dataset$value),]

ebl <- dataset
ebl <- ebl[,c(4,3)]



# 5. num 별 EBL의 총 값을 구한다(ebl2).

library(dplyr)
ebl2 <- ebl %>% group_by(num) %>% summarise('EBL'=sum(value))
View(ebl2)



# 6. base에 ebl2 정보를 통합시킨다. ebl2가 없는 num에는 0값을 대응시킨다. (실혈량이
적어서 수치가 기록 되지 않은 것으로 간주한다.)

ebl3 <- merge (base, ebl2, by='num', all=T)
ebl3[is.na(ebl3$EBL), 'EBL'] <-0
ebl3 <- ebl3 %>% arrange(num)
View(ebl3)
```

부록6. R code 예시: 연속변수 단변량 분석

```
# 1. T-test 결과를 담을 빈 데이터프레임을 생성하고 1행과 1열의 변수명을 채운다.
## FINAL 은 연속변수 표준화 처리 전의 최종 데이터세트로, 15열부터 49열까지가
연속변수 35개에 해당한다. 50열은 최종 섬망 여부를 나타내는'deli_1' 변수이다

test<- as.data.frame( matrix(NA, nrow=ncol(FINAL), ncol=19) )
colnames(test) <- c('variable', 'test', 'F.statistic', 'F.df1', 'F.df2', 'F.conf.int1',
                    'F.conf.int2', 'F.p.value', 'T.statistic', 'T.df', 'T.conf.int1',
                    'T.conf.int2', 'T.p.value', 'T.check', 'ave1', 'ave0', 'select',
                    'sd1', 'sd0')  #T-test 관련 통계량 이름 나열
test$variable[15:49] <- colnames(FINAL)[15:49]  #연속변수 이름 옮기기
View(test)



# 2. 섬망군과 비섬망군의 자료를 분리한 뒤, 등분산 검정과 T 검정(등분산 검정 결과에
따른)을 수행하고 그 결과를 test 데이터프레임에 표현한다.

deli1 <- FINAL[FINAL$deli_1==1,] #섬망 417건
deli0 <- FINAL[FINAL$deli_1==0,] #비섬망 1,666건

for (i in 15:49) {
 test[i,3]<- round(var.test(deli1[,i], deli0[,i])$statistic,2)
 test[i,4]<- var.test(deli1[,i], deli0[,i])$parameter[1]
 test[i,5]<- var.test(deli1[,i], deli0[,i])$parameter[2]
 test[i,6]<- var.test(deli1[,i], deli0[,i])$conf.int[1]
 test[i,7]<- var.test(deli1[,i], deli0[,i])$conf.int[2]
 test[i,8]<- round(var.test(deli1[,i], deli0[,i])$p.value, 4) }

test[which(!is.na(test$F.p.value)& test$F.p.value<0.05), 'test'] <-'이분산T검정'
test[which(!is.na(test$F.p.value)& test$F.p.value >=0.05), 'test'] <-'등분산T검정'

Equal <- test[which(test$test=='등분산T검정'),'variable']
notEqual <- test[which(test$test=='이분산 T검정'),'variable']

for (i in 15:49){
 if (test$variable[i] %in% Equal) {
  test[i,9]<- round(t.test(deli1[,i], deli0[,i], var.equal=TRUE)$statistic, 2)
  test[i,10]<- t.test(deli1[,i], deli0[,i], var.equal=TRUE)$parameter
  test[i,11]<- t.test(deli1[,i], deli0[,i], var.equal=TRUE)$conf.int[1]
  test[i,12]<- t.test(deli1[,i], deli0[,i], var.equal=TRUE)$conf.int[2]
  test[i,13]<- round(t.test(deli1[,i], deli0[,i], var.equal=TRUE)$p.value,4)
  test[i,15]<- round(t.test(deli1[,i], deli0[,i], var.equal=TRUE)$estimate[1],2)
  test[i,16]<- round(t.test(deli1[,i], deli0[,i], var.equal=TRUE)$estimate[2],2)
  test[i,18]<- round( sd( deli1[,i] )  ,2)
  test[i,19]<- round( sd( deli0[,i] )  ,2)
 }
 else{
  test[i,9]<- round(t.test(deli1[,i], deli0[,i], var.equal=FALSE)$statistic,2)
```

```
  test[i,10]<- t.test(deli1[,i], deli0[,i], var.equal=FALSE)$parameter
  test[i,11]<- t.test(deli1[,i], deli0[,i], var.equal=FALSE)$conf.int[1]
  test[i,12]<- t.test(deli1[,i], deli0[,i], var.equal=FALSE)$conf.int[2]
  test[i,13]<- round(t.test(deli1[,i], deli0[,i], var.equal=FALSE)$p.value,4)
  test[i,15]<- round(t.test(deli1[,i], deli0[,i], var.equal=FALSE)$estimate[1],2)
  test[i,16]<- round(t.test(deli1[,i], deli0[,i], var.equal=FALSE)$estimate[2],2)
  test[i,18]<- round( sd( deli1[,i] )  ,2)
  test[i,19]<- round( sd( deli0[,i] )  ,2)
 }}


# 3. *를 이용하여 p-value 범위를 시각적으로 표현하고, p-value 0.05를 기준으로
선택될 변수는 select=1로 표시한다.

test$T.check<-NA
test[which(!is.na(test$T.p.value)& test$T.p.value<0.05 & test$T.p.value>=0.01),
    'T.check'] <- '*'
test[which(!is.na(test$T.p.value)& test$T.p.value<0.01 &
              test$T.p.value>=0.001), 'T.check'] <- '**'
test[which(!is.na(test$T.p.value)& test$T.p.value<0.001), 'T.check'] <- '***'

test[which(!is.na(test$T.p.value)& test$T.p.value<0.05), 'select'] <- 1
test[which(!is.na(test$T.p.value)& test$T.p.value>=0.05), 'select'] <- 0


## 4. p-value 가 낮은 순으로 정렬한 뒤, 결과를 확인한다.

test1<-test[!is.na(test$variable),]
library(dplyr)
test1 <- test1 %>% arrange(T.p.value)
View(test1)
```

부록7. R code 예시: 랜덤포레스트 예측모델

```
# 1. SMOTE-NC 적용
## FINAL2는 SMOTE를 제외한 모든 전처리가 끝난 데이터세트임

levels(FINAL2$deli_1)<- c('no', 'yes') #오류방지용
library(recipes)
library(themis)
recipe_smotenc <- recipe(deli_1 ~., data=FINAL2) %>%
    step_impute_knn(all_predictors()) %>%
    step_smotenc(deli_1, over_ratio=1) %>%
    prep()
data_smoted <- step_smotenc(recipe_smotenc, neighbors=5)$template


# 2. 랜덤포레스트 예측모델 개발
## repeated K-fold cross validation (5-fold, repeated 10 times)
## mtry (분기점에서 고려할 랜덤 특징 수) 여러 개 시도하여 가장 우수한 성능 선택:
hyperparatmeter tuning

library(caret)
repeat_cv <- trainControl(method='repeatedcv',
                              number=5, repeats=10, classProbs = TRUE,
                              summaryFunction = twoClassSummary,
                              savePredictions=TRUE, verboseIter=TRUE)
grid_rf <- expand.grid(.mtry=c(2,5,6,19,35))

set.seed(123)
rf <- train( deli_1~., FINAL2, mdthod='rf', trControl=repeat_cv,
             tuneGrid=grid_rf, metric='ROC')
rf_smoted <- train( deli_1~., data_smoted, mdthod='rf', trControl=repeat_cv,
                        tuneGrid=grid_rf, metric='ROC')
rf$finalModel; rf_smoted$finalModel
confusionMatrix(rf$pred$pred, rf$pred$obs, positive='yes')
confusionMatrix(rf_smoted$pred$pred, rf_smoted$pred$obs, positive='yes')


# 3. ROC 커브 그리기 (SMOTE 적용 전후 ROC 커브 비교)

ibrary(pROC)
roc_rf <- roc(rf$pred$obs, rf$pred$yes)
roc_rf_smoted <- roc(rf_smoted$pred$obs, rf_smoted$pred$yes)
roc_rf; roc_rf_smoted
par(mfrow=c(1,1))
plot(roc_rf_smoted, col='red', lty=1, legacy.axes=TRUE)
plot(roc_rf, col='pink', lty=5, add=TRUE)
```

Abstract

Development and Evaluation of a Prediction Model for Delirium after Cardiac Surgery Using Electronic Medical Records

Ji Yun Lee

College of Nursing

The Graduate School

Seoul National University

The purpose of the study is to develop and evaluate a prediction model for delirium after cardiac surgery using electronic medical records. A retrospective analysis of 2,083 patients who were admitted to Cardiopulmonary Intensive Care Unit (CPICU) in Seoul National University Hospital between 2017 and 2021 was

performed. Raw data was retrieved from Clinical Data Warehouse (CDW) using 'Seoul National University Hospital Patient Research Environment (SUPREME)', and analyzed with R programming.

Five operational definitions of delirium were used in order to include as many delirious patients as possible. The Confusion Assessment Method for the ICU (CAM-ICU) results of clinical observation sheets, structured statements of narrative nursing records, diagnoses, neuropsychiatry consultations, and prescriptions of haloperidol were used. 417 (20.0%) out of 2,083 patients were identified as delirious patients and no significant difference was found when only used two operational definitions, which were based on consultation and prescription data (410).

46 risk factors were selected after literature review, and univariate analysis with either chi-square test or T-test was performed. The 35 risk factors with p-value less than 0.05 were filtered and included in prediction models. Preoperative variables include sex, age, body mass index, pre ICU admission, and length of hospitalization before the surgery. Intraoperative variables include types of surgery (aorta, coronary, heart valve), estimated blood loss (EBL), operation time, use of cardiopulmonary bypass (CPB), and CPB time. Postoperative variables include abnormality in pulse oximetry (SpO_2), body temperature, heart rate, mean Arterial Blood Pressure (mABP), arterial blood gas (pH, pO_2, pCO_2), glucose, lactate, hemoglobin, platelet, fibrinogen, activated Partial Thromboplastin Time (aPTT), Prothrombin Time (PT), sodium, calcium, magnesium, Blood Urea Nitrogen (BUN), creatinine, total

bilirubin, albumin, and C-reactive protein (CRP).

Delirium prediction models were developed using 4 different algorithms, which were logistic regression, decision tree, random forest, and eXtreme Gradient Boosting (XGBoost). When developing prediction models, balanced dataset with 50% delirium rate was used after applying Synthetic Minority Oversampling Technique (SMOTE).

Performance of each model was evaluated using repeated K-fold cross-validation method. 5-fold cross-validation was repeated 10 times and the mean values of each sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and AUROC (Area Under Receiver Operating Characteristic Curve) were compared. The prediction model based on random forest using balanced dataset performed best and its sensitivity, specificity, PPV, NPV, and AUROC were 0.9050, 0.8650, 0.8702, 0.9011, and 0.9575 respectively. Further research is needed to prove its generalizability.

One of the limitations of this study is that the study does not consider any subtype of delirium or time of delirium onset. More researches should be conducted on data quality of electronic medical records and utilizing unstructured data with natural language processing.

keywords : Electronic medical records, Prediction model, Delirium, Machine Learning, Intensive care unit, CDW

Student Number : 2021-22756