**11장. 공변량 분석 (Covariate Analysis)**

* 1. **서론**

모델링-시뮬레이션은 관찰한 현상을 일관된 모델로서 정립하여 현상 안에 담긴 의미를 일반화하는 것이며, 이렇게 일반화된 지식을 토대로 특정한 상황에서 약동-약력학적 결과를 예측하여, 의사 결정 등에 활용하고자 하는 것이다. 특히, 임상시험 수행 혹은 시판 후 상황 등에서 환자 요인, 약물 요인, 질병 요인에 따라 달라지는 약물의 치료 효과 등을 예측하는 것은 임상 개발 전략 또는 임상시험 설계 등을 결정하는 데에 핵심적인 정보가 된다. 예를 들어, 성인을 대상으로 수집한 약동-약력학 데이터로부터, 연령이 증가할수록 관심 대상 약물의 청소율이 감소하는 현상을 모델로서 구축할 수 있으며, 이러한 모델을 이용해 50세 이상 환자군과 50세 미만 환자군을 나누어 시뮬레이션함으로써, 특정 용량-용법 투여 시 약물 효과를 두 군에서 비교 예측해 볼 수 있다. 이 때, ‘연령’과 같이 약동-약력학 모델의 구조 또는 파라미터에 유의미한 영향을 끼치는 다양한 요인을 공변량(covariate)라 한다. 즉, 모델링-시뮬레이션의 작업 흐름(workflow) 속에서 공변량은 모델의 핵심적인 구성 요소 중 하나라 하겠다.

공변량이라 함은 단순히 데이터 수집 단계에서 모아진 정보의 항목들을 의미하는 것이 아니라, 그러한 항목들 중 통계학적인 분석을 통해 모델에 반영하는 것이 적절하다고 입증된 것만을 의미한다. 수집된 데이터 중 약동-약력학 분석에 종속변수로서 사용되는 약물의 농도 또는 작용에 대한 측정값(바이오마커로서 측정되는 경우가 많음)을 제외한 데이터 항목들(인구학적 정보, 질환 특성 등)은 실제 모델에 반영된 공변량과 구분하여, 잠재적 공변량(potential covariate)이라 지칭한다. 즉, 공변량 분석(covariate analysis)이란 잠재적 공변량에 대한 적절한 통계분석을 통해 모델링의 대상이 되는 약물의 약동-약력학 모델에 유의미한 영향을 미치는 공변량을 찾아내는 과정이라 할 수 있으며, 여기에는 데이터 항목 평가(variable evaluation), 스크리닝(covariate screening), 전진선택(forward selection), 후진제거(backward elimination), 검증(validation)의 과정이 포함된다.



< 그림 1 > 공변량 분석의 workflow

약동-약력학 모델에 적절한 공변량이 반영되기 위해서는 다음의 요건이 갖추어져야 한다.

1) 공변량 분석에 가치가 있는 항목들에 대한 정보가 충분히 수집될 것 (정보의 충분성)

2) 잠재적 공변량 간의 관계 파악을 통해 정보의 중복성을 최소화할 것 (관계의 독립성)

3) 적절한 통계학적 절차를 통해 필수적인 공변량만을 선택할 것 (절차의 정당성)

4) 약동-약력학적 지식을 통해 공변량의 의미를 해석할 수 있을 것 (임상적 유의성)

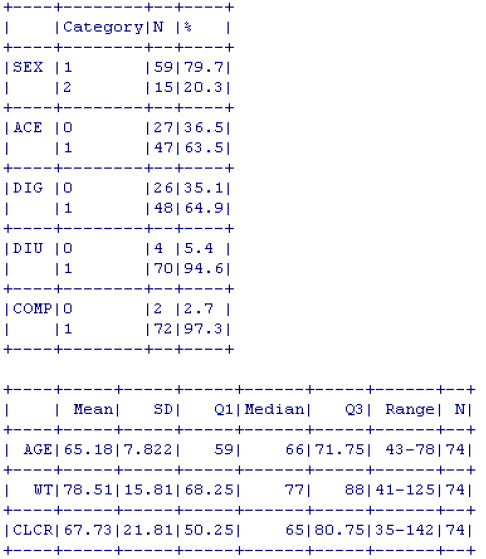
위 요건 중 어느 하나라도 간과한다면 적절한 공변량 분석이 이루어질 수 없으며, 이는 결과적으로 부적절한 공변량이 모델에 반영된다거나, 적절한 공변량이지만 잘못된 관계로 모델에 반영되는 결과를 초래한다. 이렇게 잘못 구축된 공변량 모델(covariate model, 공변량을 포함하는 약동-약력학 모델)은 비록 원래의 데이터를 잘 설명할 수 있다 하더라도, 시뮬레이션 과정에서 문제를 일으키므로 잘못된 의사 결정에 이르게 한다. (이후 절에서 구체적으로 설명) 따라서, 시뮬레이션에 활용하고자 하는 목적이 있는 모델링 과정에서는 특히 공변량 분석 과정에 주의를 기울어야 할 것이다.

* 1. **공변량 분석의 일반적 절차**
     1. **항목 평가 (variable evaluation)**

항목 평가란 약동-약력학 분석에 종속변수를 제외한 잠재적 공변량 항목 전체에 대하여 각 항목이 이후 공변량 분석 단계에 사용될 수 있을 것인가를 평가하는 과정이다. 모델링의 목적에 대한 고려 하에서 각 항목이 적절한 분포 특성을 보일 때 이를 적절한 항목이라 한다. 이 과정에서 적절하다고 판단된 데이터 항목만이 다음 단계인 *스크리닝*의 대상이 된다.

* + - 1. **기술통계분석과 부적절한 항목 또는 수준의 제거**

이 단계의 핵심적 절차는 각 항목이 어떠한 척도로 조사되었는가를 확인하고, 해당 척도에 적합한 방법으로 기술 통계 분석을 실시하는 것이다. 즉, 비척도로 조사된 항목은 평균, 표준편차, 분위 수 등의 값을 산출하여야 하고, 서열척도나 명목척도로 조사된 항목은 각 수준(level) 별로 빈도를 파악한다. 등간척도의 경우에는 상황에 따라 적절한 기술통계 방법을 선택한다. 실제 모델링 수행 시에는 완성된 기반 모델 (base model)의 제어구문 맨 뒤에 $TABLE로서 잠재적 공변량의 상세 사항을 출력한 후, R package인 xpose4 (ver. 4.5.X 기준)를 사용하면, 6: Covariate model > 2. Numerically summarize the covariates 기능을 이용하여 이 과정을 간단히 수행할 수 있다. (구체적인 방법은 xpose4 매뉴얼 참조) 이러한 분석의 결과로는 다음과 같은 표를 얻을 수 있다.



< 그림 2 > 항목 평가에 필요한 잠재적 공변량 요약표

각 개인의 특성 중 약동-약력학 파라미터에 유의미한 영향을 미치는 항목과 그 정량적 관계를 파악하는 것이 공변량 분석의 목적임을 고려할 때에, 기술통계분석 결과를 통해 1차적으로 수행하여야 하는 작업은 각 개인의 특성으로서 모델에 반영하기에 적절하지 않은 항목 또는 수준을 제외하는 것이다. 적절하지 않다 함은 다음의 두 가지 상황으로 요약할 수 있다.

1. 결측치가 너무 많음 (>25~30%)
2. 공변량 분석에 포함하기에 적절한 분포 특성을 가지지 않음

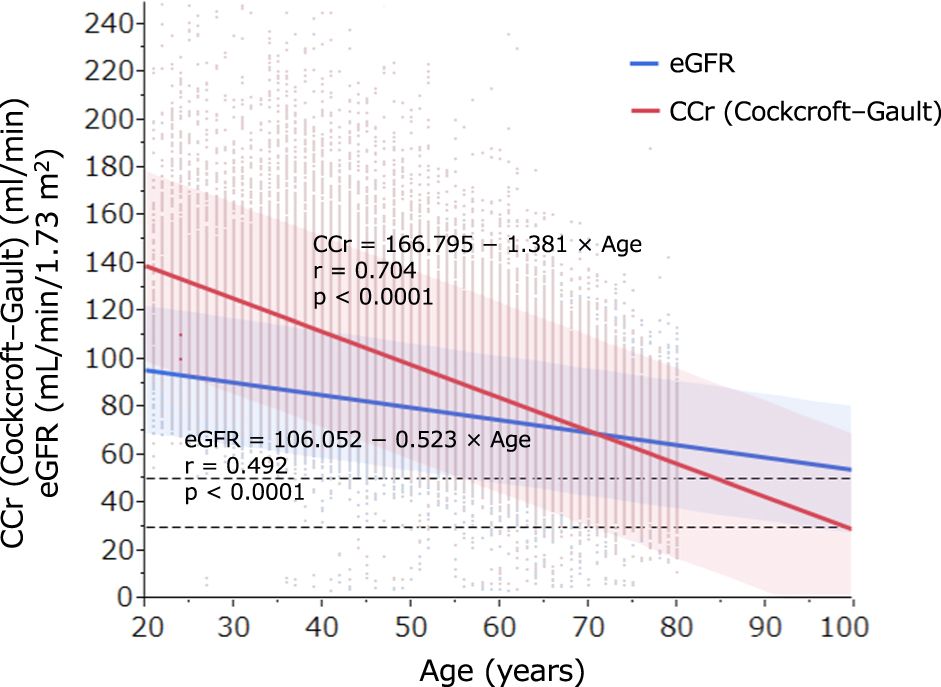
각각의 상황에서의 처리 방법은 이어지는 하위 절에서 다루도록 한다.

* + - * 1. **결측치가 많은 항목**

결측치는 많은 임상연구에서 필연적으로 발생한다. 그러기에 결측치를 대체(imputation)하기 위한 다양한 방법들이 존재한다. 그럼에도 불구하고, 지나치게 결측치가 많은 경우, 약동-약력학 파라미터에 대한 해당 항목의 영향을 정확하게 평가하는 것이 불가능하므로 해당 항목을 분석에서 제외하는 것이 좋다. 특정 항목을 분석에서 제외하여야 하는 결측치 비율이 정확하게 정해져 있는 것은 아니므로, 결측치가 발생함으로써 나타난 항목 분포와 모델링의 목적 등 다양한 요소를 고려하여 제외 여부를 결정한다. 그렇기 때문에 결측치의 비율이 상당히 높은 항목이라 할지라도, 해당 항목의 중요성이 높고, 충분히 정당화 가능한 근거가 있다면, 이를 분석에 포함하기도 한다. 예를 들어, 신장으로 배설되는 약물에 대해 혈중 크레아티닌 수치의 결측치가 발생하였을 때, 그러한 결측치가 특정 연령 대에서 집중적으로 발생하였다면, 해당 연령 대를 제외한 공변량 분석에 활용할 수 있을 것이다.

* + - * 1. **부적절한 분포를 보이는 항목**

공변량 분석은 물론, 모든 약동-약력학 모델링에서 비례적 관계를 다루는 경우에, 각 연구자는 그러한 비례적 관계가 독립변수의 어느 범위에서 유효한가를 먼저 고민해야 한다. 그러한 범위 이외에서 같은 관계를 외삽하는 경우, 실제 값과 다른 예측 결과를 보일 수 있기 때문이다. 공변량이라는 측면에서 이러한 문제를 보이는 대표적인 사례는 Cockcroft-Gault equation이다. 이 식은 토리여과율(Glomerular filtration rate)을 다른 신체적 변수 간의 관계를 통해 추정하는데, 해당 관계식에 포함된 체중 또는 연령의 값이 인구 평균에서 멀어질수록 예측 값과 실제 토리여과율 간의 차이가 확대되는 양상을 보인다.(그림 3)



Estimated glomerular filtration ratio is a better index than creatinine clearance (Cockcroft–Gault) for predicting the prevalence of atrial fibrillation in the general Japanese population

< 그림 3 > 독립 변수의 변화에 따른 토리여과율 예측값과 실제값의 차이: 연령의 예시

따라서, 이러한 식을 적용할 수 있는 독립변수의 범위가 어디인가를 설정하는 것은 매우 중요한 일이며, 약동-약력학 모델링에서도 청소율이나 효능 등 파라미터와 비례하는 공변량을 도입할 때 유사한 고려가 이루어져야 한다.

적절한 비례 관계 및 그 범위를 설정하는 첫 번째 작업은 기술통계분석 결과로부터 각 변수의 분포 특성을 확인하는 것이다. 많은 경우에 약동-약력학 데이터는 특정한 선정/제외기준을 만족하는 인구집단으로부터 얻어지는데, 이러한 이유로 인해 다음과 같은 문제가 발생할 수 있다.

1) 특정 독립변수 범위(주로 분포 범위의 양극단)에 대한 정보가 충분치 않음

2) 분포의 범위가 매우 좁거나, 편향됨

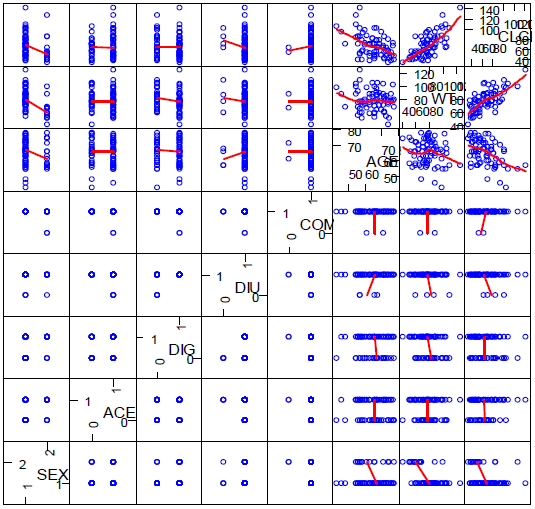
1)의 경우, 잠재적 공변량이 연속형 변수라면, 양극단에 있는 소수의 정보가 약동-약력학 파라미터와 잠재적 공변량 간의 비례 형태에 대한 잘못된 추론을 가능케 할 수 있으므로 이후 공변량 분석에서 이를 주의하여 수행하여야 한다. (그림 4-1) 범주형 변수라면, 어느 한 수준의 측정값이 다른 수준에 비해 절대적으로 부족하게 되며(그림 4-2), 이 경우에는 그러한 수준을 가장 가까운 수준과 병합하거나 그 수준을 제외하고 이후 공변량 분석을 수행하여야 한다. 만약 수준을 단 2개만 가진 항목(예> 성별, 환자 여부 등)이라면 아예 항목 자체를 제외하여야 할 수 있다. 그러나 특정 수준의 데이터 비율이 상대적으로 적은 경우라 할 지라도, 그 데이터 양이 충분(N > 30)하다고 판단된다면 이를 포함할 지 여부를 신중하게 고려하여 결정할 수 있다.

<그림 4> 부적절한 데이터 분포

2)의 경우는, 변수의 척도 특성과 상관 없이 해당 변수를 제외하는 것이 추천된다. 이는 첫번째로, 좁은 분포 범위로 인해 통계학적으로 유의한 상관관계를 얻기 힘들기 때문이며, 두번째로, 설혹 통계학적으로 유의한 상관관계를 보인다 할지라도, 임상적으로 유의한 범위의 상관성을 보여 주기 어렵기 때문이다. 대표적인 경우로는 건강인 대상자에서 얻은 인구학적 정보가 있다. 그러한 정보들 중 체중이나 연령 등이 분포용적 혹은 청소율과 선형적 비례 관계를 가지는 경우를 생각해 보자. 이 경우, 건강인 대상자의 특성으로 인해 체중이나 연령 등이 상당히 제한적인 범위에 머무를 가능성이 크며, 따라서 선형적 관계가 실제 관찰된 범위를 넘어 환자에게서도 동일하게 적용되리라는 가정을 하는 것은 상당히 위험하다. 따라서, 공변량 분석 자체가 의미 없는 상황이라고 할 수 있겠다.

* + - 1. **잠재적 공변량 간의 상관 관계 분석**

본 과정은 다음 2.3절에서 소개할 공선성 분석을 위한 준비 단계로서, 각 대상자 별로 확보된 잠재적 공변량의 값들 간에 나타나는 비례관계 유무를 파악하는 것이다. 우선적으로는 시각적 방법이 활용되며, 통계적 방법을 통해 이를 보강할 수 있다. 이 과정의 핵심적 절차는 각 대상자에서 관찰된 잠재적 공변량 항목들의 값을 그림 5와 같이 scatterplot matrix로 나타내는 것이다. 2.1.1.에서 언급한 바와 마찬가지로, xpose4 (ver. 4.5.X 기준)의 6: Covariate model > 3. Scatterplot matrix of covariates 기능을 통해 간단히 수행할 수 있다. 그 수행 결과는 아래 그림 5에 나타난 바와 같다.



< 그림 5 > 공변량 간의 상관 관계 탐색을 위한 scatterplot matrix

이러한 플롯을 통해 AGE, WT, CL 등 연속형 변수 간의 상관 관계, 연속형 변수와 AGE, WT, CL을 제외한 나머지 범주형 변수 간의 상관 관계에 대한 탐색적 정보를 확인할 수 있다. 다만, 범주형 변수 간의 상관 관계를 확인하기는 어렵다는 단점이 있으며, 범주형 변수 간의 관계 파악을 위한 통계학적인 방법(odds ratio 등)을 활용하면 도움이 된다. 연속형 변수의 경우에도 시각적인 정보가 충분치 않다면 잠재적 공변량 간의 상관 분석을 통해 상관계수 등을 구할 수 있다.

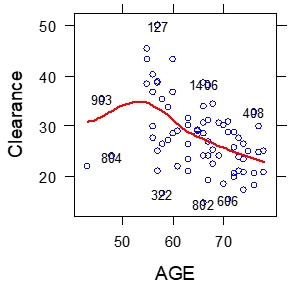
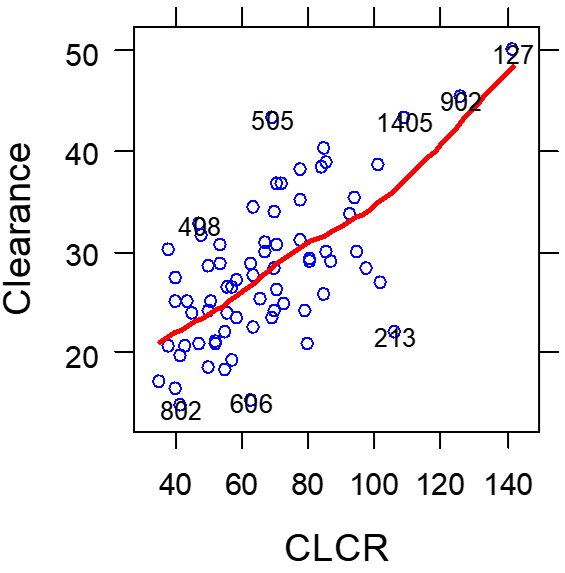
2.1.1.절에서는 수행 결과에 따라 일부 잠재적 공변량들을 제외 또는 수정하였지만, 이 절에서 수행한 사항의 결과에 따라 그러한 작업을 수행하지는 않는다. 그러나 궁극적으로 정확한 공변량 모델을 구축하기 위해 매우 핵심적인 작업이 이루어지는데, 이는 어떠한 잠재적 공변량들이 비례 관계를 가지고 있는가의 결과를 요약하여 기록해 두는 것이다. 그리고, 그러한 비례 관계가 나타난 원인에 대해 생리적, 의학적인 근거를 파악하여야 한다. 만일, 두 공변량 간의 비례 관계를 기존의 지식으로 설명할 수 없다면, 그것은 정보 부족 또는 우연에 의한 결과일 수 있다. 이러한 요약 및 부가 정보 자료들은 2.2.절의 수행 결과와 종합하여 실제 공변량 모델 구축을 위한 전략을 마련하는 데에 활용한다.

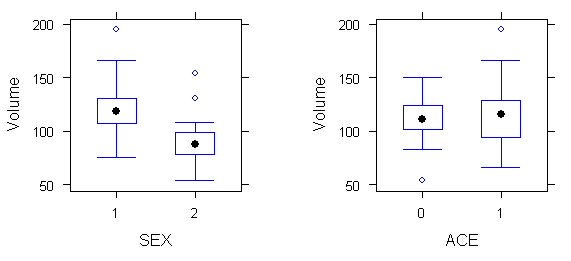
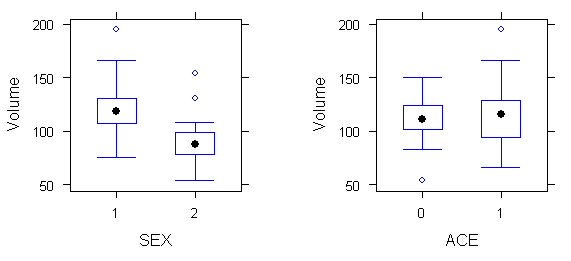
* + 1. **공변량 스크리닝 (covariate screening)**

개별 공변량을 약동-약력학 파라미터에 반영하여, 모델 개선에 대한 유의성 여부를 평가하는 작업이 공변량 분석의 핵심적인 단계이기는 하지만, 잠재적 공변량의 개수는 물론 약동-약력학 파라미터의 개수가 상당히 많은 경우가 대부분이므로 각 잠재적 공변량을 각 약동-약력학 파라미터에 대응시켜 하나하나 이 작업을 수행하는 것은 매우 비효율적일 수 있다. 예를 들어, 약동-약력학 파라미터(ETA가 추정된)의 개수가 8개이고, 잠재적 공변량의 개수가 10개라면 80개의 개별적 모델링 과정을 거쳐야만 모든 공변량 분석을 완료할 수 있으며, 그 결과에 따라 발생하는 다양한 조합을 함께 평가해 보아야 하는 문제도 생겨나게 된다. 이를 피하기 위해서는 대상자 별로 확보된 잠재적 공변량의 값과 각 파라미터의 ETA 값을 이용하여, 실제 모델링을 통해 평가해 볼 만한 가치가 있는 관계들을 사전에 파악하는 단계가 필요한데 이를 공변량 스크리닝이라 한다. 스크리닝은 시각적 방법과 수치적 방법으로 수행할 수 있으며, 두 방법 중 어느 하나에서라도 의미 있는 상관성을 보이는 잠재적 공변량과 파라미터 간의 관계가 있다면, 이를 실제 모델링하여 평가하는 것이 추천된다.

* + - 1. **시각적 스크리닝 (visual screening)**

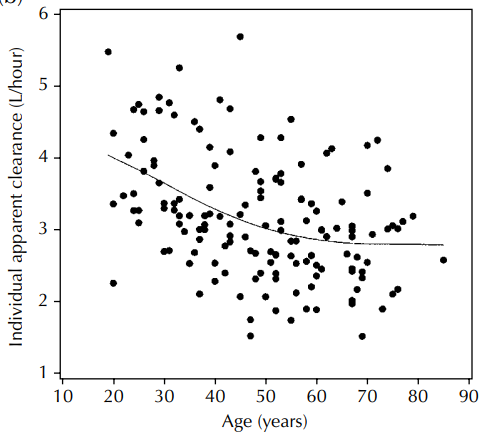
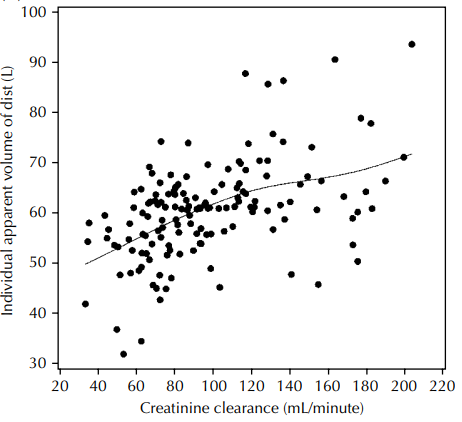
시각적 스크리닝의 핵심적 절차는 대상자 별로 확보된 잠재적 공변량의 값과 각 파라미터의 ETA 값을 각 조합 별로 하나의 그림에 나타내어 양의 혹은 음의 상관 관계가 보이는가를 시각적으로 판단해 보는 것이다. 이 때, 보통 공변량을 원인 변수로, 파라미터의 ETA 값을 종속 변수로 보기 때문에(약동-약력학적 파라미터는 각 개인의 특성 또는 특정 상황에 따라 발생한 결과값이라 생각함), 잠재적 공변량 값을 x축에, ETA 값을 y축으로 하여 나타내게 된다. 잠재적 공변량이 연속형 변수라면 이를 scatterplot으로 나타내고, 범주형 변수라면 각 수준 별 boxplot으로 나타낸다. 이 절차는 xpose4(ver. 4.5.X 기준)의 6: Covariate model > 4. Parameters vs covariates 기능을 이용하여 간단히 수행할 수 있다. 그림6은 그 예시로써, 패널(a), (b)는 scatterplot이며, (c), (d)는 boxplot이다.

(a) (b)

(c) (d)

< 그림 6 > 시각적 스크리닝을 위한 plot 예시

패널 (b)와 (c)에서는 명확한 상관관계를 확인할 수 있으며, (d)는 상관관계가 없는 것으로 보인다. 따라서, 이 경우, 청소율에 대해 CLCR을, 분포용적에 대해 SEX를 실제 모델링을 위한 잠재적 공변량으로 선정하고, ACE는 제외한다. 패널(a)의 추세선은 선형적 상관관계가 없는 것처럼 보이지만, 해당 패널을 좀 더 자세히 보면 불과 1-2개의 관측값이 추세선의 형태를 변형시키고 있다는 것을 알 수 있다. 이는 ‘2.1.1.2. 부적절한 분포를 보이는 항목’에서 언급한 “특정 독립변수 범위(주로 분포 범위의 양극단)에 대한 정보가 충분치 않음”에 해당하는 것이다. 관측값이 조밀한 영역에서는 분명한 음의 상관관계를 보이고 있음을 알 수 있으므로, AGE 역시 청소율에 대한 유효한 잠재적 공변량으로서 다음 단계의 평가에 포함하는 것이 바람직하며, 이를 제외하는 오류를 범해서는 안 된다. 이러한 오류는 시각적 스크리닝 과정에서 가장 빈번히 일어나는 것이기도 하다. 일부 경우에는 다음과 같이 반드시 선형적 상관관계를 가진다고 단정 짓지 못하는 사례도 있을 수 있다.



< 그림 7 > 비선형적 상관관계를 가질 가능성이 있는 경우 (Kelly 146 - 147 page)

이러한 잠재적 공변량들은 물론 다음 단계 평가에 포함하는 것이 바람직하며, 선형적 관계 이외의 구조로서 공변량-파라미터 관계를 설명하였을 때 더욱 좋은 결과를 얻을 수 있을 가능성을 반드시 염두에 두어 향후 과정에서 이에 대한 비교가 이루어질 수 있도록 해야 한다. 범주형 공변량에 대해서도 마찬가지 작업이 이루어져야 하며, 이는 이후 절에 공변량 모형의 구조와 관련하여 보다 구체적으로 설명하도록 한다. 이러한 절차 이외에 잠재적 공변량의 값에 따른 weighted residual 값의 분포 경향을 plot으로 보는 절차(xpose4의 6: Covariate model > 6. weighted residual vs covariates)도 있지만, 앞서 언급한 절차에 비해 중요성이 떨어지므로 여기에서는 구체적으로 다루지 않는다.

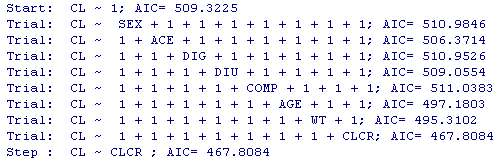
* + - 1. **수치적 스크리닝 (numerical screening)**

수치적 스크리닝의 대표적인 방법은 GAM(Generalized Additive Modeling)이다. GAM은 파라미터의 종류 별로 이루어진다. 또한, GAM은 우리가 직접 수행하는 것이 아니며, 컴퓨터가 파라미터 별로 최적의 모델을 찾을 때까지 자동으로 수행하는 절차이다. 따라서, 우리는 GAM을 실행할 수 있는 환경에서 파라미터를 선택하기만 하면 된다. 이는, Xpose4의 6: Covariate model > 7. GAM을 선택한 후, 파라미터 이름을 입력하는 것이다. 단, base model control stream에서 $TABLE에 지정한 파라미터명으로 입력해야만 실행이 가능하다.(개인 별 파라미터를 파라미터명(e.g. *CL*, *V*…)으로 출력했다면 파라미터명을, 각 파라미터에 대한 EBE 값(e.g. ETA1, ETA2…)으로 출력했다면 EBE명을 입력한다.

GAM을 특정 파라미터에 대해 실행하면 컴퓨터는 기본적으로 다음과 같은 작업을 수행한다.

1. dataset을 ETA 값 = 특정 값이라는 모델로 fitting하여 AIC 값을 얻는다. (AIC는 이전 장에서 설명함)
2. 모든 잠재적 공변량에 대해 ETA ~ 잠재적 공변량 값이라는 선형 모델로 fitting하여 AIC 값을 얻는다. 즉, 잠재적 공변량이 5개였다면, 총 5개의 AIC 값을 얻게 된다.
3. 2)에서 얻은 값 중, 가장 낮은 AIC 값을 보인 잠재적 공변량을 선택하고, 그 AIC 값이 1)에서 얻은 값보다 낮으면, 해당하는 잠재적 공변량을 유효한 것으로 판단한다.

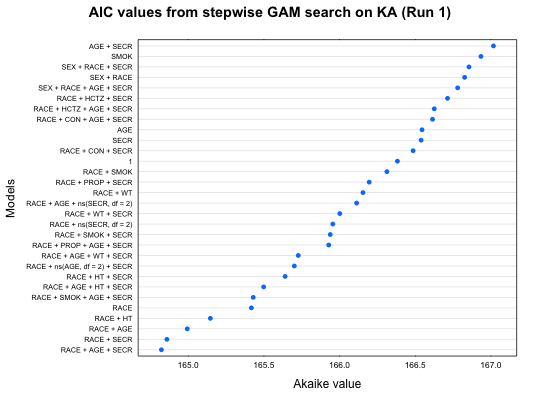
이러한 3가지의 절차를 1개의 ‘step’이라 한다. 하나의 step이 수행된 결과는 아래와 같이 표시되며, 이 사례에서는 CLCR이 최선의 결과로 선택되었다.



< 그림 8 > GAM에서 1개의 step을 실행한 결과

만일, 3)의 조건을 만족하는 잠재적 공변량이 없다면, GAM은 step 1에서 종료되고, 유의한 잠재적 공변량이 없다고 보고한다. 그러나 3)의 조건을 만족하는 공변량이 하나라도 있다면, GAM은 다른 step을 개시하며, 같은 작업을 반복한다. 다만, 다음 step에서는 1)의 기준 AIC가 이전 step에서 선택한 최적 모델의 AIC가 되며, 2)에서는 이전 step에서 선택된 잠재적 공변량을 모델에 유지한 채, 또 하나의 잠재적 공변량을 선형적 구조로서 가법적으로 이어 붙여 해당 모델의 AIC를 평가하게 된다.(예를 들어 CLCR이 이전에 step에서 포함된 상황이라면, ETA1 ~ CLCR + SEX와 같은 구조가 된다.) 이는 통계학적으로 다변량 선형회귀분석에서 사용하는 모델과 동일하다. 이 평가 과정 역시 모든 잠재적 공변량에 대해 이루어지며, 3)의 조건을 만족하는 잠재적 공변량이 없을 때까지 step이 반복된다. 중요한 것은 이미 모델에 포함되어 있는 잠재적 공변량일지라도, 각 step에서 한 번 더 평가가 이루어진다는 것이다.(e.g. ETA1 ~ CLCR + CLCR) 만일, 동일한 공변량이 두 번 이상 선택된다면, 이는 유의한 상관관계가 있지만, 잠재적 공변량과 파라미터 간의 상관 구조가 선형적이 아닐 수 있다는 것을 시사하는 것이므로, 앞서 그림 7에서 언급한 경우와 같이 기록해 두고 다음 과정을 진행하여야 한다.

이제 GAM이라는 의미를 다시 살펴보면, ‘모든 잠재적 공변량에 대해(generalized), 가법적으로(additive), 모델을 수행(modeling)하여, 각 결과를 비교하는 것이다.’라는 것을 알 수 있다. 파라미터 별 GAM 수행 결과는 그림 8의 결과를 반복적으로 보여 주는 형태로 확인하기도 하지만, 다음의 그림과 같이 모델 구조에 따른 AIC의 변화를 도식적으로 나타내어 선택된 잠재적 공변량들을 한 눈에 확인하는 경우가 많다.



< 그림 9 > GAM 실행 summary  
(https://uupharmacometrics.github.io/xpose4/reference/xpose.gam.html)

* + 1. **공선성(co-linearity)에 대한 고려**

앞선 2.1.2.절에서 잠재적 공변량 간의 상관 관계를 파악하는 작업에 대해 소개한 바 있다. 이는 이 절에서 논의하고자 하는 공선성에 대한 고려를 위한 준비 작업이라고 할 수 있다. 공선성이란 2개 이상의 잠재적 공변량이 서로 비례 관계를 가지는 상황에서, 그러한 잠재적 공변량들이 각각 하나의 약동-약력학 파라미터와 상관 관계를 가지는 경우를 나타내는 용어이다. 예를 들어, ‘공변량 스크리닝 결과 청소율에 체중과 신장이 동시에 가능성 있는 공변량으로 확인된 경우에 실제 공변량 평가 과정은 어떻게 이루어져야 하겠는가?’라는 질문이 발생할 수 있다. 이러한 질문에 대한 적절한 답을 얻기 위해서는 신장과 체중 간의 관계에 대한 파악이 우선되어야 하는 것이다.

어떠한 경우에는 약동-약력학 파라미터 간에도 서로 간에 상관 관계를 가지는 경우가 있어 2개 이상의 약동-약력학 파라미터와 2개 이상의 잠재적 공변량이 모두 상관관계를 보이는 상황도 발생할 수 있다. 예를 들어, 체중과 신장은 서로 비례 관계를 가지는 것이 일반적인 현상인데, 신장으로 제거되는 약물의 청소율 추정치(Cockcoroft-Gault 식 등)과 분포용적 등은 동시에 이러한 신체적 요인들과 비례 관계를 가질 수 있다. 이 경우에도 마찬가지로 2.1.2.절의 공변량 간의 상관성 탐색 결과가 필요하다. 다만, 약동-약력학 파라미터들 간의 상관성 분석은 엄밀하게는 공변량 분석 이전의 기반 모델 구축 단계에서 온전히 수행되어야 하는 작업이므로 이 장에서는 다루지 않았다(이미 상관성에 대한 정보가 갖추어져 있다고 전제함). 혼합효과모형 분석에서는 개별 약동-약력학 파라미터를 각 개인에서 분명하게 추정 가능한 수준의 정보가 데이터셋에 갖추어져 있지 않은 경우에 파라미터의 ETA 간의 상관성이 유도(induction)되는 특성이 있어, 이러한 문제가 심한 경우에는 특정 파라미터의 ETA를 추정하지 않거나, 파라미터 재설정(re-parametrization) 등의 해결책을 사용한다. 그러나 정보가 충분히 확보된 경우에도 생리학적으로 상관성을 가질 수밖에 없는 파라미터들을 모델링의 목적에 따라 각각 독립적으로 추정해야 하는 경우가 많아, 공선성에 대한 고려는 필수적이면서도 상당히 까다로운 작업이 된다.

위의 경우에는 공변량에 대한 본격적인 평가에 앞서 뚜렷한 기준을 설정하여 접근하여야 한다. 일반적으로 아래와 같은 기준으로 최선의 평가 전략을 구축한다.

1. 개별 파라미터-공변량 관계는 생리학적으로 타당한가?
2. 공변량 스크리닝 단계에서 특정 파라미터와 보다 큰 상관성 혹은 AIC 감소를 보여준 관계는 무엇인가?
3. 공선성 관계를 가지는 하나의 파라미터-공변량 관계를 모델에 반영한 후에 또다른 파라미터-공변량 관계를 추가하고자 한다면, 이를 지지할 수 있는 생리학적인 혹은 수학적인 관계가 있는가?

공선성은 매우 많은 상황에서 다양한 유형으로 발생하기 때문에 공변량 평가를 위한 표준적인 기준을 제시하기는 어렵다. 따라서, 모델링의 목적 등에 대한 고려와 관계된 다양한 전문가들의 의견을 종합하여 궁극적인 평가 계획 및 공변량 포함 여부 등을 결정해야 할 것이다.

* + 1. **공변량 평가: 전진선택(forward-selection)**

이 절의 논의에 앞서 다시 한번 주지하고자 하는 것은 본격적인 공변량 평가 방법은 시간이 많이 걸리는 작업이라는 것이다. 이는 개별 파라미터-공변량 관계를 실제 제어 구문에 반영한 후, 이를 실행하여 그러한 관계가 모델에 긍정적인 영향을 주는가를 평가하는 작업이 파라미터-공변량 관계 별로 일일이 이루어져야 하기 때문이다. 따라서, 효과적인 평가 계획을 수립하는 것이 무엇보다도 중요하며, 2.1~2.3의 절차가 올바르게 수행되었다면, 이 절에서 수행하고자 하는 공변량 평가 과정에 투입되는 시간과 노력을 절감할 수 있을 것이다. 이제부터는 공변량 평가의 첫 단계인 전진선택법에 대해 알아보도록 하겠다.

* + - 1. **전진선택이란?**

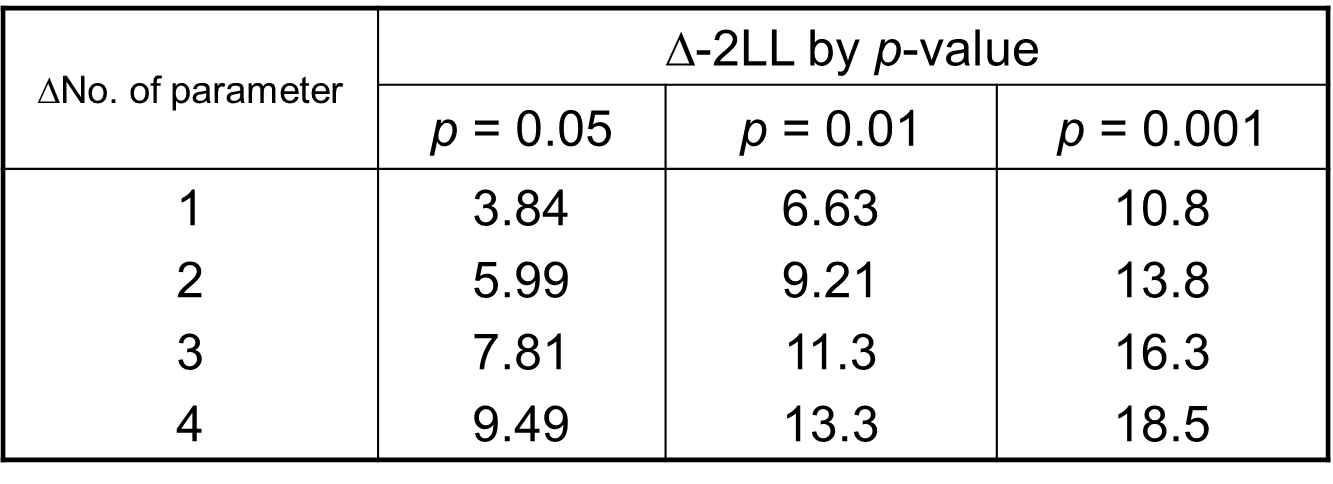
전진선택은 2.3절에서 약동-약력학 파라미터와 유의한 상관관계를 가질 가능성이 있다고 평가된 공변량들을 실제 모델의 제어구문에 반영하여 평가하는 것이다. 제어구문에 반영한다 함은 특정 약동-약력학 파라미터를 기반 모델에서와 같이 특정한 THETA에 대응시키는 것이 아니라, 공변량과 가장 가능성 있는 관계로서 표현하는 것을 의미한다.(공변량-파라미터 관계에 대해서는 다음 절에 설명함) 예를 들어, 분포용적과 체중이 선형적 상관관계를 가질 수 있다고 판단한다면, 기반 모델의 제어구문을 다음과 같이 변경하면 된다.

V = THETA(i) → V = WT\*THETA(i) + THETA(i+1)

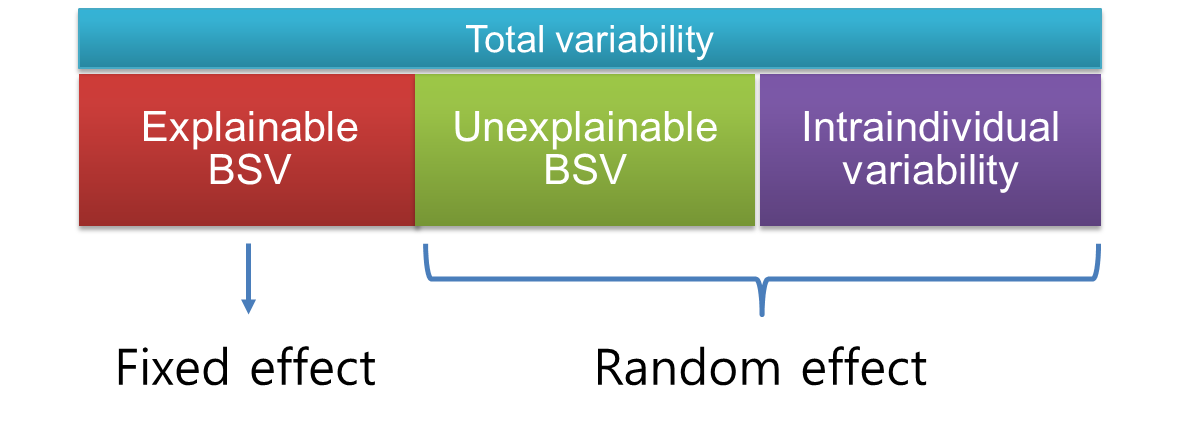
제어구문이 변경되었다면 이제 이를 실행하여, 그러한 변경(공변량의 반영)이 모델의 성능을 개선하는가를 평가하여야 한다. 평가의 요소는 다음의 세 가지이다.(각각에 대해서는 하위 절에서 구체적으로 설명한다.)

1. 목적함수값(objective function value, OFV)가 이전 모델에 비해 특정 cut-point 이상 감소하는가?
2. 공변량을 반영한 파라미터에 대응하는 개인 간 변이(between-subject variability, BSV)가 이전 모델에 비해 감소하는가?
3. 이전 모델에 비해 잔차의 뚜렷한 증가가 없는가?

1)은 공변량이 전체 모델의 성능을 개선하는가를 평가하는 기준이다. 제시한 특정 cut-point란 기반 모델 구축 시 적용한 바와 마찬가지로, 모델의 개선을 평가하기 위한 가능도비검정(Likelihood-Ration Test, LRT) 방법에 따라 정해진 p-value에 대응하는 OFV의 감소 값을 의미한다. 일반적으로 다음의 표를 참고한다.



2)는 매우 중요한 개념으로, *ωi*2의 감소로 판단한다. 이는 임의 효과(random effect)로 남겨져 있던 해당 파라미터의 개인 간 차가 고정 효과(fixed effect)로 적절히 설명되어, 시뮬레이션 과정 등에 활용이 가능한가를 평가하는 기준이다. 다음의 그림을 통해 그 의미를 파악할 수 있다.



특정 개인이 아닌 전체 인구집단에서 공변량과 파라미터가 일정한 관계를 갖는다면 이는 더 이상 설명 불가능한 개인차가 아니다. 따라서, 이는 설명 가능한 개인차로서 모델에 반영이 되며, 이렇게 적절히 반영된 공변량은 고정효과가 되는 것이다. 그렇다면 설명 불가능한 개인차의 크기는 줄어드는데 이것이 모델에서는 다음과 같이 표현된다.

V = THETA(i)\*exp(ETA(i)) → V = (WT\*THETA(i) + THETA(i+1))\*exp(ETA(i))

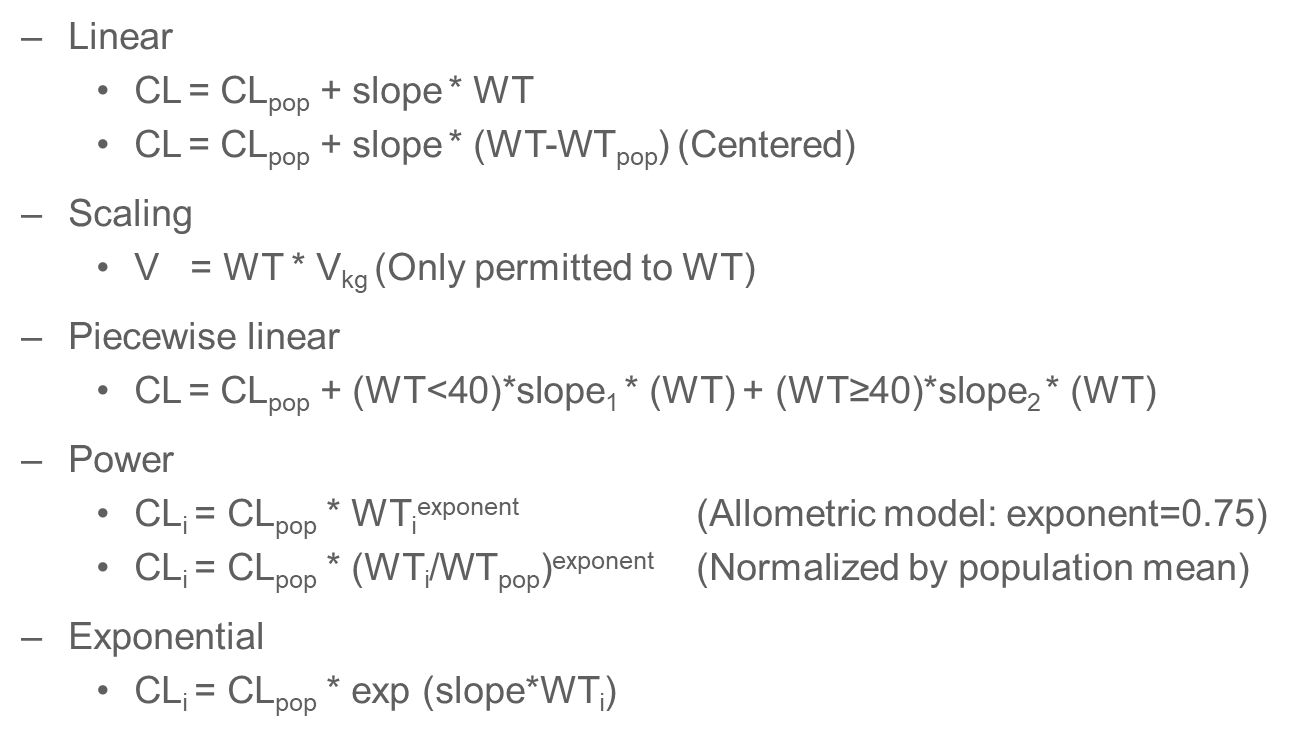
왼쪽의 경우와 오른쪽의 경우 ETA(i)의 의미는 어떻게 다르겠는가? 왼쪽의 경우, 분포용적에 대해 제공된 공정효과는 집단의 대표값인 THETA(i) 뿐이다. 따라서, 이 값과 각 개인의 분포용적 값 차이는 모두 ETA(i)로 설명된다. 이제 오른쪽을 보면, 고정효과로서 체중과의 선형적 관계가 제시되었다. 즉, 체중이 큰 사람은 분포용적이 크다는 일반적 경향은 고정효과로 설명이 되는 것이다. 결국 이 경우 ETA(i)는 체중의 차이가 무시된 집단의 대표값과 각 개인 값의 차이가 아닌, 같은 체중을 가진 사람들의 대표값과 각 개인 값의 차이가 된다. 따라서, 각 ETA(i)의 크기는 왼쪽에 비해 오른쪽에서 줄어들게 되고, 이는 궁극적으로 *ωi*2의 감소 효과를 일으키게 되는 것이다. 한 가지 추가할 사항은 시각적 스크리닝에서 본 바와 같은 공변량과 개인 별 파라미터 값 간의 scatterplot에 대한 것이다. 위에서 설명한 이유로 왼쪽의 ETA(i)는 체중과 일관된 경향을 가지는 값이 될 것이나, 오른쪽의 ETA(i)는 체중과 무관한 경향을 가지는 값이 되어야 한다.(체중의 영향은 고정효과로 설명이 되고, 남은 개인 간 차는 체중과는 무관한 차이일 것이므로) 따라서, 오른쪽 상황에서 체중과 ETA(i) 간의 scatterplot을 그리면 경향성이 없게 되는 것이 정상이며, 만일 경향성이 보인다면 공변량-파라미터 관계가 최적의 구조로 구현되지 않았다는 것을 의미하는 결과라 할 수 있다. 이 경우에는, 2.4.2를 참고하여 새로운 공변량-파라미터 구조 모델을 시도해 보아야 한다.

3)은 이상치(outlier)와도 관련이 있는 사항으로서, 대부분의 사람에서 설명력이 개선되는 경우 1), 2)의 조건을 만족하지만, 드물게는 새롭게 반영되는 공변량-파라미터 간의 고정효과와 완전히 다른 trend를 가지는 개인이 있어 잔차가 증가하는 상황이 발생하기도 한다. 따라서, 이 경우에는 추후 진행할 시뮬레이션 등의 목적 등을 고려하여 공변량을 반영할 것인가를 선택해야 할 수 있다.

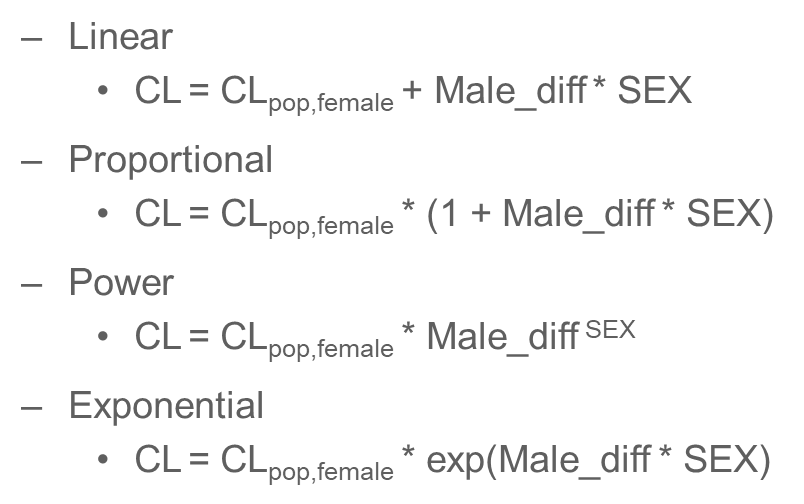
특정 공변량이 위의 세가지 기준을 모두 만족하여 유의미한 공변량으로 확인된다면, 이제 해당 공변량은 모델에 그대로 둔 채로 새로운 공변량을 같은 방법으로 모델에 반영 및 평가하게 된다. 공변량 스크리닝 단계에서 평가 필요성이 제시된 모든 공변량에 대하여 이 과정을 반복 수행한다. 어떤 순서로 공변량을 모델에 반영하여 평가하여야 하는가에 대해 정해진 이론은 없다. 하지만, 일반적으로 공변량 스크리닝 단계에서 시각적으로 좋은 상관관계가 보였거나 GAM 단계에서 AIC의 감소가 컸던 순으로 전진선택을 수행하게 된다. 추가 사항은 J.S.Owen과 J.Fiedler-Kelly가 저술한 Introduction to Population Pharmacokinetic / Pharmacodynamic Analysis with Nonlinear Mixed Effects Models의 5.7.6-7절을 참고할 수 있다.

* + - 1. **공변량-파라미터 간의 관계 모델**

2.4.1에서 선형적 관계로 공변량을 반영하는 예시를 제시하였다. 많은 경우에 이러한 선형적 모델은 공변량 평가의 기본 구조가 된다. 그러나, 시각적 스크리닝 단계에서 선형적 관계가 아닌 것으로 보이거나, GAM 과정에서 2번 이상 선택되는 잠재적 공변량 등은 선형적 관계 이외에 보다 적합한 공변량-파라미터 관계를 가질 수 있다. 심지어 이러한 잠재적 공변량들은 단순 선형 모델을 이용한 공변량 반영 및 평가 시 2.4.1.에 제시한 3가지 기준을 만족하지 못할 수 있으므로, 해당 잠재적 공변량들에 대해서는 선형적인 관계로 공변량의 유의성을 평가만을 수행해서는 안 되며, 이후 또다른 구조를 이용한 평가가 필요하다. 아래는 연속형 변수로 측정된 공변량에 대해 적용할 수 있는 다양한 구조들의 예시이다.



강조하고자 하는 바는 scaling에 관한 것으로, 이는 2.4.1에서 제시한 분포용적과 체중 간의 관계 중 절편 부분을 제거한 것이다. 일반적으로 이러한 관계 설정은 분포용적과 체중 간에만 정당화되는 것으로 알려져 있다. 체중 이외의 공변량에 대해서는 일반적으로 잘 사용하지 않는 구조이며, 다른 체중-파라미터 간의 관계에 대해 이 관계를 적용하기 위해서는 절편이 모델 개선에 도움이 되지 않는다는 객관적인 증거가 있어야만 한다. 다음은 범주형 변수로 측정된 공변량에 대한 구조들이다.



위의 예에서 SEX는 성별을 뜻하며, 여성은 0, 남성은 1의 방식으로 coding할 수 있다. 만일, 특정 공변량이 3개 이상의 범주를 가지는 경우에는 위의 예와 같이 일관된 수식으로서 표현하기 어려운 상황이 많아, 제어구문 중 IF문을 사용하여, 각 범주 별로 파라미터 값을 각각 추정하게 할 수도 있다. 이 때 범주 하나 당 하나의 THETA가 대응되므로, 2.4.1의 1) 기준 평가 시 이를 고려해야 한다. 또한, 범주 별로 분산을 같다고 가정할 지, 다르다고 가정할 지 역시 모두 OFV 등의 기준을 이용하여 평가해 보아야 한다는 어려움도 있다. 범주형 변수가 가진 수준의 수가 충분히 많아지면, 이를 연속형 변수와 유사하게 처리할 수 있다는 이론도 있으나, 일반적인 모델링에서는 용량(증량 시험 등에서)을 제외하고 이렇게 다양한 수준의 범주형 자료가 수집되는 경우는 많지 않아 여기에서는 다루지 않는다.

* + 1. **공변량 평가: 후진제거(backward-elimination)**

후진제거는 전진선택이 완료된 후 수행하는 절차이다. 전진선택은 스크리닝에서 확인된 잠재적 공변량 중, 2.4.1.의 세 가지 기준을 만족하여 유의한 공변량으로 선택되는 것이 더 이상 없을 때까지 진행한다고 하였으며, 이 절차가 끝났더라도 후진제거를 통해 선택된 공변량들을 재평가해야 한다는 것이 핵심이다. 후진제거 시에는 공변량이 모델에 포함된 순서의 역순으로 공변량을 하나씩 모델에서 빼 보면서, 그러한 공변량의 제거가 모델의 유의미한 악화를 유발하는가를 평가한다. 예를 들어, 전진선택 시 A, B, C, D 네 개의 공변량이 순서대로 유의한 공변량으로 선택되었다고 하자. 이 경우 후진 제거는 A,B,D를 가진 모델, A,C,D를 가진 모델, B,C,D를 가진 모델 순으로 수행한다. 마찬가지로 2.4.1.의 세 가지 기준을 이용하며, 특정한 공변량 제거 시에 유의미한 OFV의 증가가 있는가?, *ωi*2가 증가하는가?, 잔차에 변화가 없는가?의 기준을 통해 세 가지를 모두 충족하는 경우 유의미한 공변량으로 확정한다.

이 작업을 수행하는 이유는 아래와 같이 매우 간단하다.

1. 앞서 전진선택 수행 시, 일반적으로 공변량 스크리닝 단계에서 시각적으로 좋은 상관관계가 보였거나 GAM 단계에서 AIC의 감소가 컸던 순으로 수행한다고 하였으나, 이러한 순서가 반드시 지켜지는 것은 아니다.
2. 전진선택의 기본적 원리는 단변량 대 단변량의 관계이므로, 여러 가지의 공변량-파라미터 효과가 종합된 모델에서 개별 공변량-파라미터 관계가 여전히 유효한 것인지를 확인할 수 없다.

분포 용적에 대해 체중과 BMI 모두 가능성 있는 잠재적 공변량으로 선택된 상황에서, 체중-BMI 순으로 전진선택을 수행하였고, 둘 모두 공변량으로 선택되었다 하더라도, 후진제거 실시 과정에서 체중은 유의미한 공변량이 아닌 것으로 확인될 수 있다. 이것이 후진제거가 필요한 단적인 상황으로서, 보다 파라미터와 상관관계가 좋으며, 이전에 포함된 공변량의 정보를 일부 가진 공변량이 전진선택에서 이후에 평가되었을 때, 이전에 포함된 공변량이 더 이상 필요 없음에도 불구하고 모델에 남아 있게 되는 상황을 피하고자 하는 것이다. 이는 앞서 언급한 공선성과 관련이 있는 문제이며, 사전에 공선성을 파악하여 보다 좋은 공변량을 선별적으로 선택할 수도 있지만, 판단이 어려운 경우에 후진 제거를 실시하면 정확한 통계적 근거를 통해 불필요한 공변량을 제거할 수 있다. 후진 제거 시에는 OFV 변화에 대해 전진선택보다 더 엄격한 *p*-value를 적용하기도 한다. 추가 사항은 J.S.Owen과 J.Fiedler-Kelly가 저술한 Introduction to Population Pharmacokinetic / Pharmacodynamic Analysis with Nonlinear Mixed Effects Models의 5.7.8절을 참고할 수 있다.

* 1. **공변량 분석 관련 추가 고려 사항**

공변량 분석에서 무엇보다도 중요한 고려 사항은 바로 생리학적 타당성이다. 항목 평가에서 후진제거에 이르기까지 생리학적 타당성은 늘 고려되어야 한다. 실제 모델링 수행 시 접하게 되는 가장 흔한 상황은 전진선택과 후진제거의 모든 절차를 완료한 후 (통계학적 타당성을 확보하였다는 의미임) 특정 파라미터에 대해 생리학적으로 관련성이 적은 항목이 모델에 유의한 공변량으로 남아 있는 경우이다. 이 경우, 이러한 결과를 생리학적으로 해석해 낼 수 있는가가 그러한 공변량을 최종적으로 선택할 것인지 말 것인지를 결정하는 핵심 요소가 된다. 예를 들어, 신장으로 제거되는 약물의 청소율이 간의 기능을 나타내는 혈중 알부민 수치나 프로트롬빈 시간(prothrombin time) 등과 비례 관계를 가지는 결과가 도출되었다 한다면, 이것이 생리학적으로 타당하지 않은 것인가? 일반적인 관점에서 이는 우연의 산물이거나, 해석할 수 없는 외부 요인의 결과라고 볼 수도 있겠지만, 이러한 데이터가 진행된 간경변 환자에서 확보된 것이라면, 간 기능 저하에 의한 신 기능 저하를 충분히 의심할 수 있을 것이다. 따라서, 생리학적 타당성이란 고정된 개념이 아니라, 데이터가 확보된 대상자의 특성과 연구 대상 의약품의 특성 등이 종합적으로 해석되어야 확보할 수 있는 것이라 하겠다. 반대로 이미 잘 알려져 있는 공변량-파라미터 관계가 있고, 공변량에 대한 정보도 충분히 확보되었음에도 불구하고 그러한 공변량-파라미터가 최종 모델에서 통계적으로 유의하게 선택되지 않을 때, 그 이유에 대한 정당화 역시 필수적으로 요구된다. 이 경우에는 모델링 과정에서 세운 가정이나 모델 구조 등에 문제가 없었는지를 다시 한 번 확인해야 하며, 혼란 변인 등의 역할에 대해서도 생각해 보아야 한다. 해당 내용은 최종 모델링 보고서 등에 포함하는 것이 추천된다.

공변량이 포함된 모델에 대한 평가(evaluation) 시에는 주요 공변량의 수준 별로 Visual Predictive Check (VPC)를 수행하는 것이 필수적이다. 앞서 특정 개인에게는 일반화된 고정 효과로서의 공변량-파라미터 관계가 잘 적용되지 않을 수도 있다고 설명하였는데, 이는 해당 공변량의 특정 수준에 대해서도 마찬가지이다. 가장 흔하게 접하는 경우는 여러 개의 용량이 포함된 데이터셋을 모델링하는 경우에, 용량-생체이용률 간 관계를 도입하는 경우이다. 또는, 연령이 광범위하게 분포하는 데이터셋에서 연령-청소율 간의 관계도 유사한 경향성을 가진다. 모델링 과정에서 OFV의 감소가 가장 구조로서 공변량-파라미터 관계를 표현하게 되는데, 이러한 방법의 가장 큰 문제는 측정값의 개수가 충분치 않은 공변량의 특정 수준을 잘 반영하지 못할 수 있다는 것이다. 이에 따라, 공변량 수준(특정 용량 또는 연령 범위 등)에 따라 VPC를 수행하면, 모든 수준에서 측정값과 모델 예측값이 적절히 대응하고 있는지를 평가할 수 있음과 동시에 앞서 언급한 문제에 의해 잘 예측되지 않은 수준이 어느 것인지도 확인할 수 있다. 심지어 공변량 모델링 시 간과한 부적절한 공변량-파라미터 관계까지도 확인할 수 있으므로, 이러한 평가 과정은 반드시 수행되어야 할 것이다.