

## Regression model for the preparation of calibration curve in the quantitative LC-MS/MS analysis of urinary methamphetamine, amphetamine and 11-nor- $\Delta^9$ -tetrahydrocannabinol-9-carboxylic acid using R

Jin Young Kim<sup>★</sup> and Dong Won Shin

*Forensic Genetics and Chemistry Division, Supreme Prosecutors' Office, Seoul 06590, Korea*

(Received May 20, 2021; Revised June 21, 2021; Accepted June 22, 2021)

### 소변 중 메트암페타민, 암페타민 및 대마 대사체 LC-MS/MS 정량분석에서 검량선 작성을 위한 R을 활용한 회귀모델 선택

김진영<sup>★</sup> · 신동원

대검찰청 디엔에이 · 화학분석과

(2021. 5. 20. 접수, 2021. 6. 21. 수정, 2021. 6. 22. 승인)

**Abstract:** Calibration curves are essential in quantitative methods and for improving the accuracy of analyte measurements in biological samples. In this study, a statistical analysis model built in the R language (The R Foundation for Statistical Computing) was used to identify a set of weighting factors and regression models based on a stepwise selection criteria. An LC-MS/MS method was used to detect the presence of urinary methamphetamine, amphetamine, and 11-nor-9-carboxy- $\Delta^9$ -tetrahydrocannabinol in a sample set. Weighting factors for the calibration curves were derived by calculating the heteroscedasticity of the measurements, where the presence of heteroscedasticity was determined via variance tests. The optimal regression model and weighting factor were chosen according to the sum of the absolute percentage relative error. Subsequently, the order of the regression model was calculated using a partial variance test. The proposed statistical analysis tool facilitated selection of the optimal calibration model and detection of methamphetamine, amphetamine, and 11-nor-9-carboxy- $\Delta^9$ -tetrahydrocannabinol in urine. Thus, this study for the selection of weighting and the use of a complex regression equation may provide insights for linear and quadratic regressions in analytical and bioanalytical measurements.

**요약:** 검량선 작성은 기기분석을 통해 생체시료에서 분석물질의 농도를 측정하는 정량분석법 개발과 측정값의 정확도 향상에 있어서 필수적인 요소이다. 본 연구에서는 R 기반 통계분석기법을 이용하여 개별 분석물질 정량에 적합한 가중계수와 회귀모델 선정하기 위한 단계별 선택 기준을 적용하여 분석 프로그램을 설정하였다. 국내에서 남용빈도가 가장 높은 필로폰과 대마 복용여부 확인을 위해 액체크로마토그래피-질량분석법(LC-MS/MS)이 적용되었으며, 분석물질로 마약의 복용 여부를 확인에 일반적으로 사

<sup>★</sup> Corresponding author

Phone : +82-(0)2-3480-3730 Fax : +82-(0)2-3480-3606

E-mail : paxus@spo.go.kr

This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

용되는 대상 마약의 모체와 대사체를 소변 시료에서 분석하였다. 검량선 작성에 있어서 가중계수 적용 여부는 원본데이터의 이분산성 검정을 통해 확인하였고, 가중계수가 필요하다고 판단된 경우 분산분석을 통해 적정 가중계수를 선정하였다. 다음으로 편분산분석을 이용하여 회귀모델에 적합한 차수를 결정하였다. 분석물질인 메트암페타민, 암페타민, 대마 대사체에 대해 R 기반 프로그램에 적용한 결과, 단계별 결과 및 최종 모델식을 직관적으로 이해하기 쉽고 신속하게 얻을 수 있었다. 이러한 결과는 문서 파일로 저장이 가능하여 보관의 편의성을 제고하였으며, 본 연구를 통해 제작된 R 기반 프로그램을 활용하여 검량선 작성을 필요로 하는 다양한 약물분석 분야에 확대 적용할 수 있을 것으로 판단된다.

**Key words:** quantification, calibration curve, regression model, weighting factor, bioanalysis

## 1. 서 론

마약류 중 보통 ‘필로폰’이라 불리는 메트암페타민(methamphetamine, MA)은 국내에서 가장 많이 남용되는 마약이다.<sup>1</sup> 그 다음으로 남용 빈도가 높은 마약은 대마이다. 국내와 달리 전 세계적으로는 대마의 남용 빈도가 가장 높다. 이는 미국 등 여러 나라에서 대마초(cannabis) 합법화 과정에서 대마 성분을 의약품으로 사용을 허가하거나 기호용으로 판매가 가능하도록 허용하는 등에 따른 과급효과로 볼 수 있다.<sup>2</sup>

마약 복용여부는 일반적으로 소변, 모발 등 생체시료를 이용하여 확인한다. 마약 복용 후 일정시간이 경과하면 모약물과 대사체가 체외로 모두 배설되므로 이들 시료를 분석하여 복용여부에 대한 정보를 얻을 수 있다. 이러한 이유로 생체시료를 마약분석에 활용하게 되었고 약물 특성에 따라 모약물과 대사체를 동시에 분석하거나 대사체만을 분석하기도 한다.

MA와 대마초 내 활성 성분인 테트라하이드로칸나비놀( $\Delta^9$ -tetrahydrocannabinol)은 체내에서 대사과정을 거쳐 각각의 주요 대사체로서 amphetamine(AP)과 11-nor-9-carboxy- $\Delta^9$ -tetrahydrocannabinol(THCCOOH) 이 생성된다.<sup>3</sup> 필로폰 복용여부 확인을 위해서 MA와 AP를 분석대상 물질(target compound)로 선정하여 분석한다. 대마의 경우 대체로 THCCOOH만을 분석한다. MA와 AP는 모두 작용기로 아민(amine)이 포함되어 있으며, MA의 이성질체 중 하나인 *d*-MA의 pKa 값은 9.87, AP의 이성질체 중 하나인 *d*-AP의 pKa 값은 10.1로 모두 약염기성을 나타내고 있다. 반면에 대마초 복용여부 확인에 사용되는 분석대상 물질인 THCCOOH는 작용기로 카르복실산(carboxylic acid)을 포함하고 있고 pKa 값이 4.2로 약산성을 띠고 있다.<sup>4,6</sup>

마약류 분석은 효율성 측면에서 국내 남용빈도가 가장 높은 필로폰과 대마를 동시에 분석하는 것이 유

리하다. 또한 신속하고 정확한 분석결과를 얻는 것도 중요한 요소이다. 이러한 점을 고려하여 이전의 연구에서는 필로폰과 대마 복용 여부를 동시에 확인하기 위해 기체크로마토그래피-질량분석기(GC-MS)로 MA, AP 및 THCCOOH를 한 번에 분석하여 효율성을 높이고자 하였다.<sup>7</sup> 여기에 사용된 GC-MS는 마약류 분석에 가장 널리 사용되는 분석 장비 중 하나로 질량스펙트럼 라이브러리(library) 검색을 통해 미지 물질 확인에 유용할 뿐 아니라 액체크로마토그래피-질량분석기(LC-MS/MS)에 비해 상대적으로 기질(matrix)의 영향을 덜 받는다는 장점이 있다.<sup>8,9</sup> 그러나 분석대상 물질에 극성기가 포함되어 있는 경우 추가 전처리 단계로 유도체화 과정이 포함되어야 하므로 이러한 과정을 필요로 하지 않는 LC-MS/MS를 이용하여 분석하는 경우를 쉽게 찾아 볼 수 있다. 이는 GC-MS 보다 생체시료에서 상대적으로 높은 극성을 띤 물질 분석에 유용하고, 유도체화 과정 없이 분석이 가능하므로 상대적으로 시료 전처리 시간을 단축할 수 있다는 점에서 LC-MS/MS가 유리하기 때문이다.<sup>10</sup>

검량선(calibration curve) 작성을 위한 적정 가중계수(weighting factor)와 회귀모델(regression model)의 선택은 기기분석을 통해 생체시료에서 분석물질(analyte)의 농도를 측정하는 정량분석법 개발에 있어서 중요한 요소 중 하나이지만 이러한 회귀모델 선정을 위해 직관적이면서도 체계적인 관별 프로세스를 제시하는 연구 사례는 흔치 않다. 따라서 분석에 적합한 가중계수와 회귀모델을 찾아내기 위해 어떤 요인을 검정해야 할 것인가에 대한 면밀한 검토가 수반되어야 한다. 앞서 정량분석에서 측정의 편차를 줄이고 분석결과의 정확도를 높이기 위한 검량선 회귀모델 선정의 중요성에 대해서는 여러 연구를 통해 언급되었다.<sup>11-15</sup>

검량선 작성에 사용되는 통계적 방법 중 하나가 회귀분석(regression analysis)이다. 회귀분석은 독립변수

(입력량)의 종속변수(산출량)에 대한 영향을 추정할 수 있는 기법으로 독립변수에 대한 종속변수 사이의 관계를 나타내고 있다. 이러한 회귀모델은 크게 선형(linear)과 비선형으로 나눌 수 있으며, 비선형에는 이차함수(quadratic equation), 지수함수, 로그함수 등 다양한 응용이 가능하다. 보통 회귀분석을 수행하는 경우, 비선형 보다 직관성이 뛰어나고 쉽게 구현이 가능한 가중계수가 적용되지 않은 선형 회귀모델이 많이 사용되고 있다.<sup>16</sup> 선형 회귀모델을 적용했을 때 정량 한계에 해당하는 QC 측정값의 정확도가 20%를 넘어서는 경우, 분석법 유효화 가이드라인에서 제시한 규격을 벗어나게 되므로 QC 측정값의 정확도가 20% 이내가 되도록 가중계수를 적용하게 된다.<sup>17</sup> 가중계수가 없는 선형 회귀모델(linear regression model)에서 조금 더 복잡한 형태인 가중계수  $1/x$ ,  $1/y$ ,  $1/x^2$  및  $1/y^2$ 인 선형 회귀모델, 가중계수가 없는 이차다항회귀모델(quadratic regression model), 가중계수  $1/x$ ,  $1/y$ ,  $1/x^2$  및  $1/y^2$ 인 이차다항회귀모델의 가중계수와 회귀모델의 조합을 순차적으로 적용하여 최적의 모델을 찾아나가게 된다.

이론적으로 검량선은 선형 회귀모델로 작성되어야 하지만, 물리·화학적 영향에 의해 이차다항회귀모델로 작성되는 경우가 있다. LC-MS 이온화 과정에서 dimer-multimer-cluster의 생성, 고농도에서 검출기(detector) 포화현상, 넓은 정량범위의 설정, 기질효과, 낮은 질량값에서 검출기의 선택성 상실, 부적당한 내부표준물질의 사용 등을 들 수 있다.<sup>18-21</sup> 이차다항회귀모델을 적용한 LC-MS/MS 정량분석법으로 생체시료 분석(bioanalysis)을 수행한 여러 연구 사례가 있다.<sup>19,20,22,23</sup>

다양한 회귀모델 중에서 어떤 가중계수를 적용하고 회귀모델의 차수를 얼마로 결정할지 여부는 표준화된 잔차분석(standardized residual analysis)과 분산분석(F-test) 등을 통해 결정하게 된다.<sup>24</sup> 우선 이분산성(heteroscedasticity) 검정 수행을 통해 가중계수 도입이 필요한가를 먼저 확인한 후, 회귀모델에 가중계수 도입 필요성이 있는 경우 적합한 가중계수 선정을 위해 분산분석을 수행하고, 이어서 회귀모델의 적정 차수를 결정하기 위한 분산분석-적합도 검정을 순차적으로 수행하게 된다.

본 연구에서는 LC-MS/MS를 이용한 MA, AP 및 THCCOOH 분석의 검량선 작성에 있어서 R을 이용한 통계분석기법을 적용하여 ① 독립변수(농도)와 종속변수(면적비)로 구성된 원본데이터의 이분산성 검정을 통해 가중계수 적용 여부를 판단하였고, ② 가중계수

도입이 필요한 경우 분산분석을 통해  $1/x$ ,  $1/y$ ,  $1/x^2$ 과  $1/y^2$  중 하나의 가중계수를 선정하였으며, ③ 선형과 이차다항회귀모델의 적합성을 편분산분석(partial F-test)을 통해 평가하여 적합한 차수를 결정하였고, ④ 가중계수, 회귀모델의 차수 등의 결과를 정리해 최종 모델식을 산출하였으며, ⑤ 최종 결과를 문서 파일로 만들어 확인 및 저장할 수 있도록 하였다. 본 연구의 분석물질 MA, AP, THCCOOH의 정량분석에 사용될 최종 회귀모델 산출을 위해 5 단계로 구분하여 R 코드(R script)로 작성된 프로그램을 사용하여 계산하였다.

## 2. 실험

### 2.1. 시약 및 기구

표준물질 MA, AP, THCCOOH 및 내부표준물질 AP-d<sub>8</sub>, MA-d<sub>11</sub>, THCCOOH-d<sub>9</sub> 모두 Cerilliant사(Austin, TX, USA)에서 구입하였다. 이동상에 첨가된 formic acid (LC-MS LiChropur grade)는 Sigma-Aldrich사에서 구입하였다. 이동상으로 사용된 HPLC 등급의 메탄올(methanol)은 J.T. Baker/Avantor사(Center Valley, PA, USA)에서 구입하였고, 증류수(LiChrosolv grade)는 Merck (Darmstadt, Germany)에서 구입하였다. 이외의 시약은 ACS 등급 이상의 시약을 사용하였고, polypropylene microcentrifuge tube (1.5 mL)는 Eppendorf 사(Hamburg, Germany)에서 구입하였다.

표준물질 MA, AP, THCCOOH 및 내부표준물질 AP-d<sub>8</sub>, MA-d<sub>11</sub>, THCCOOH-d<sub>9</sub>은 각각 메탄올로 희석하여 혼합용액으로 제조하였으며 -20 °C에서 보관하였다.

### 2.2. 소변 시료

바탕시료는 약물을 복용하지 않은 사람의 소변을 사용하였으며, 검량선 작성에 사용하였다. 검량선을 작성하기 위해 바탕시료에 MA는 6, 10, 25, 50, 125, 250, 500, 750, 1500 ng/mL가 되도록 제조하였고, AP는 2, 6, 10, 25, 50, 125, 250, 500, 750 ng/mL로 제조하였으며, THCCOOH는 0.4, 1.2, 2, 5, 10, 25, 50, 100, 150 ng/mL의 농도로 제조하였다.

### 2.3. 분석기기

본 연구에 사용한 액체크로마토그래프는 진공탈기 장치(vacuum degasser), 이차 펌프(binary pump), 자동 시료주입기 및 컬럼오븐으로 구성된 Nanospace NASCA HPLC (Shiseido, Tokyo, Japan)를 사용하였다. 컬럼은

Hypersil Gold C18 column (150 × 2.1 mm I.d., 5 μm, Thermo Scientific, Waltham, MA, USA)를 사용하였다. 이동상 A는 0.4% formic acid가 포함된 증류수를 사용하였고, 이동상 B 용매로는 메탄올을 사용하였다. 이동상의 유속은 400 μL/min으로 이동상 B는 10%로 설정하여 1 min 동안 흘려준 뒤에 3.5 min이 될 때까지 20%로 증가시킨 후, 7 min이 되었을 때 90%가 되도록 설정하였고, 이후 1 min 동안 90%로 유지하였고 이동상 B를 10%로 설정한 후 8.2 min 까지 용매를 흘려주었다. 이후 이동상 B가 10%인 조건으로 2.8 min 간 등용매 용리하였다.

액체크로마토그래프에 연결된 질량분석기는 전자분무이온화 (electrospray ionization) 장치와 장착된 SCIEX QTrap 6500 triple-quadrupole mass spectrometer (AB SCIEX, Foster city, CA, USA)를 사용하였다. 두 가지 방식 (polarity switching mode)으로 전자분무이온화 장치를 사용하였다. MA와 AP 분석을 위한 양이온 방식으로, nebulizing gas 55, curtain gas 60, turbo ion spray heater gas 80 (제조사 설정 단위), 이온화 전압은 5000 V로 0~6 min 동안 설정하였고, THCCOOH 분석을 위한 음이온 방식은 6~11 min, nebulizing gas 35, curtain gas 75, turbo ion spray heater gas 50, 이온화 전압은 -4500 V로 설정하였다. 이때 Turbo-gas의 온도는 600 °C였고, gas 유량의 단위는 제조사에서 설정한 임의의 단위이다. 질량분석을 위해 MRM (multiple reaction monitoring) 방법을 사용하였고, 모 이온 (precursor ion)의 조각화 (fragmentation)를 위한 충돌기체 (collision gas)로는 질소 가스를 사용하였다. 분석에 사용된 MRM 이온쌍은 분석물질은 1개의 정량이온쌍과 정성이온쌍을 각각 선정하였고 내부표준물질의 경우 1개의 정량이온쌍을 선정하였다. MA는  $m/z$  150→91/ $m/z$  150→119, AP은  $m/z$  136→119/ $m/z$  136→91, THCCOOH는  $m/z$  343→245/ $m/z$  343→219이었으며, 내부표준물질 MA-d<sub>11</sub>은  $m/z$  161→127, AP-d<sub>8</sub>은  $m/z$  144→97, THCCOOH-d<sub>9</sub>는  $m/z$  352→254로 설정하였다.

## 2.4. 시료 전처리

소변 시료 100 μL를 1.5 mL polypropylene tube에 넣고, 이동상 50 μL와 AP-d<sub>8</sub>, MA-d<sub>11</sub>, THCCOOH-d<sub>9</sub>가 포함된 내부표준물질 50 μL를 첨가하여 혼합하였다. Sigma 3-30K 원심분리기 (Sigma, Munich, Germany)를 이용하여 50000 g에서 3 min 동안 초고속원심분리한 후, 상층액 8 μL를 LC-MS/MS에 주입하였다.

## 2.5. 회귀분석 모델

본 연구에서 회귀모델의  $x$ 축은 독립변수로 농도이고,  $y$ 축은 종속변수로 내부표준물질에 대한 분석물질의 면적비이다. 검량선 작성에 적용된 모델은 선형과 이차다항회귀모델 2가지로 선형 회귀모델은 다음의 식으로부터 구할 수 있다.<sup>25,27</sup>

$$y = a + b \cdot x$$

$$\left( \begin{aligned} a &= \frac{\sum x_i^2 \cdot \sum y_i - \sum x_i \cdot \sum (x_i \cdot y_i)}{n \sum x_i^2 - (\sum x_i)^2}, \\ b &= \frac{n \sum (x_i \cdot y_i) - \sum x_i \cdot \sum y_i}{n \sum x_i - (\sum x_i)^2} \end{aligned} \right)$$

여기서  $(x_i, y_i)$ 는  $n$  개의 데이터쌍 중에서  $i$  번째를 의미하며,  $x_i$ 는 농도,  $y_i$ 는 내부표준물질에 대한 분석물질의 면적비를 나타낸다.

다음으로 검토된 모델은 비선형 형태인 이차다항회귀모델로 다음의 식에서 계산되었다.

$$y = a + b \cdot x + c \cdot x^2$$

$$\left( \begin{aligned} a &= \frac{\sum y_i}{n} - b \cdot \frac{\sum x_i}{n} - c \cdot \frac{\sum x_i^2}{n}, \\ b &= \frac{\left( \left( \sum (x_i \cdot y_i) - \frac{\sum x_i \cdot \sum y_i}{n} \right) \cdot \left( \sum x_i^4 - \frac{(\sum x_i^2)^2}{n} \right) - \left( \sum (x_i^2 \cdot y_i) - \frac{\sum x_i^2 \cdot \sum y_i}{n} \right) \cdot \left( \sum x_i^3 - \frac{\sum x_i^2 \cdot \sum x_i}{n} \right) \right)}{\left( \sum x_i^2 - \frac{(\sum x_i)^2}{n} \right) \cdot \left( \sum x_i^4 - \frac{(\sum x_i^2)^2}{n} \right) - \left( \sum x_i^3 - \frac{\sum x_i^2 \cdot \sum x_i}{n} \right)^2}, \\ c &= \frac{\left( \left( \sum (x_i^3 \cdot y_i) - \frac{\sum x_i^2 \cdot \sum y_i}{n} \right) \cdot \left( \sum x_i^2 - \frac{(\sum x_i)^2}{n} \right) - \left( \sum (x_i \cdot y_i) - \frac{\sum x_i \cdot \sum y_i}{n} \right) \cdot \left( \sum x_i^3 - \frac{\sum x_i^2 \cdot \sum x_i}{n} \right) \right)}{\left( \sum x_i^2 - \frac{(\sum x_i)^2}{n} \right) \cdot \left( \sum x_i^4 - \frac{(\sum x_i^2)^2}{n} \right) - \left( \sum x_i^3 - \frac{\sum x_i^2 \cdot \sum x_i}{n} \right)^2} \end{aligned} \right)$$

회귀모델의 차수 선택을 위해 선형과 이차다항회귀모델을 비교하여 잔차값에 대한 잔차평균오차 (residual standard error)가 작은 것을 선택하여 결정하였다. 회귀모델의  $y$ -절편과 기울기, 분산분석표의 유의확률인

Pr(>F) 등의 값은 R(www.r-project.org)을 사용하여 구하였다.

## 2.6. 가중 계수

검량선 주위의 분포한 측정값들의 분산은 동일하게 분포하는 것이 이상적이다. 그러나 특정 범위에서 표준편차가 크게 발생하는 경우, 오차가 무작위로 고르게 분포하지 않을 뿐 아니라 등분산 조건을 만족하지 못하게 된다. 이를 개선하기 위한 방법으로 회귀모델에 일반적으로 가중계수  $1/x$ ,  $1/y$ ,  $1/x^2$  및  $1/y^2$ 를 적용하여 적합도를 개선하는 방법이 있다.

우선 가중계수 도입이 필요한가를 확인하기 위해 독립변수(농도)와 종속변수(면적비)로 구성된 원본데이터를 F-test를 통해 이분산성을 구하고 이분산성이 확인되면 가중계수를 적용하였다. 검량선 작성에 적합한 가중계수 선정을 위해서 가중계수가 적용된 회귀모델식을 통해 계산된 농도( $C_{calculated}$ )와 명목 농도( $C_{nominal}$ )의 상대오차(relative error, RE)를 농도에 대한 잔차도(residual plot)를 작성하여 비교하였다. 여기서 상대오차를 구하는 식은 다음과 같다.

$$RE = \frac{C_{calculated} - C_{nominal}}{C_{nominal}} \times 100$$

가중계수의 적합성을 평가하고 선정하기 위해  $1/x$ ,  $1/x^2$ 의 가중계수가 적용된 회귀모델의 F값(F-value)을 계산하여 비교한 후 적정 가중계수를 선정하였다. 가중계수의 적합성 평가는 R 프로그램을 사용하여 수행하였다.

## 3. 결과 및 고찰

### 3.1. 기질 효과

본 연구에서는 원심분리된 시료를 희석 주입법(dilute-and-shoot)을 사용하여 분석하였다. 이러한 전처리 방식은 신속한 분석에 이점이 있는 반면, 추출 및 정제 과정이 포함된 고체상 추출법이나 액체상 추출법과 달리 기기분석 시 소변 시료 자체에서 기인하는 기질 효과(matrix effect)에는 상대적으로 영향을 많이 받게 된다. 이러한 영향을 줄이기 위해 50000 g의 초고속원심분리 과정을 도입하여 소변 시료로부터 유래하는 분석 방해물질의 영향을 일정 부분 제거하였다. 분석물질의 선택성을 높이기 위해 크로마토그램 바탕선(baseline)의 수준과 분석물질의 머무름 시간을 고려하여 이동상 선정하였고, 그 결과 이동상 A는 0.4%

formic acid가 포함된 탈이온수 및 이동상 B 용매로는 메탄올로 선정하였다. THCCOOH의 경우 테이세이라 등의 연구를 참고하여 음이온 전자분무화 방식을 적용하여 기질 효과를 최소화하였다.<sup>28</sup>

### 3.2. 이분산성 검정

회귀모델의 x축은 독립변수로 농도이고, y축은 종속변수로 내부표준물질에 대한 분석물질의 면적비를 나타내는 두 변수 간의 연관성을 직관적으로 파악할 수 있도록 분석물질에 대한 산점도(scatter plot)를 Fig. 1에 나타내었다. Fig. 1(A)와 Fig. 1(B)에서 데이터의 추정 회귀선이 직선의 형태를 보이는 것과 달리, Fig. 1(C)에서 x축의 농도와 y축의 면적비는 2차식의 패턴을 보일 뿐 아니라 농도가 증가하면 면적비의 증가 정도가 점점 커짐을 알 수 있다. 이러한 경우 선형 회귀모형에 2차항을 추가하면 독립변수의 수가 하나 더해져 다중 회귀분석의 2차식 관계임을 산점도를 통해 추정해 볼 수 있었다.

선형 회귀모델에서 검량선이 직선인 경우 상관계수는 1에 근접한다고 볼 수 있으나, 반대로 상관관계가 1에 근접한 값을 갖는다고 해서 직선성을 갖는다고 볼 수는 없다.<sup>13</sup> 따라서 상관계수가 갖는 한계점을 보완하고 직선성 여부를 보다 명확히 설명하기 위해서 잔차분석과 분산분석을 수행한다. 잔차는 회귀분석에서 종속변수인 측정값과 회귀모델의 예측값 사이의 차이를 나타내는 값으로 데이터가 회귀선 주위로 편향된 형태가 아니라 동일 범위 내에 분포하는 경우 오차의 등분산성이 있다고 본다.<sup>29</sup> 한 예로 낮은 농도에서의 오차 분포와 높은 농도에서 오차 분포가 상이한 경우, F-test를 통해서도 등분산성 여부를 명확히 확인할 수 있다. F-test로부터 신뢰구간 95%에서 얻은 p-value가 0.05 보다 작은 경우 이분산성이 있으며 가중계수 도입이 필요하다고 볼 수 있다. 가중계수의 도입 여부 확인을 위해 이분산성 검정을 F-test를 사용하여 수행하였다. 분산분석 결과, 분석물질 MA, AP, THCCOOH 모두에서 가중계수 도입이 필요한 것으로 확인되었다.

분석물질에 대해 가중계수가 적용된 회귀모델식을 통해 계산된 농도와 명목 농도 사이의 상대오차를 농도에 대한 잔차도를 Fig. 2에 나타내었다. 분석물질의 잔차도는 가중계수와 회귀모델을 조합하여 분석물질 각각에 대해 3개의 잔차도를 작성하였다. 잔차도를 확인한 결과, MA와 AP는 각각 Fig. 2(A)와 Fig. 2(B)의 네 번째 위치한 잔차도, THCCOOH는 Fig. 2(C)의 다

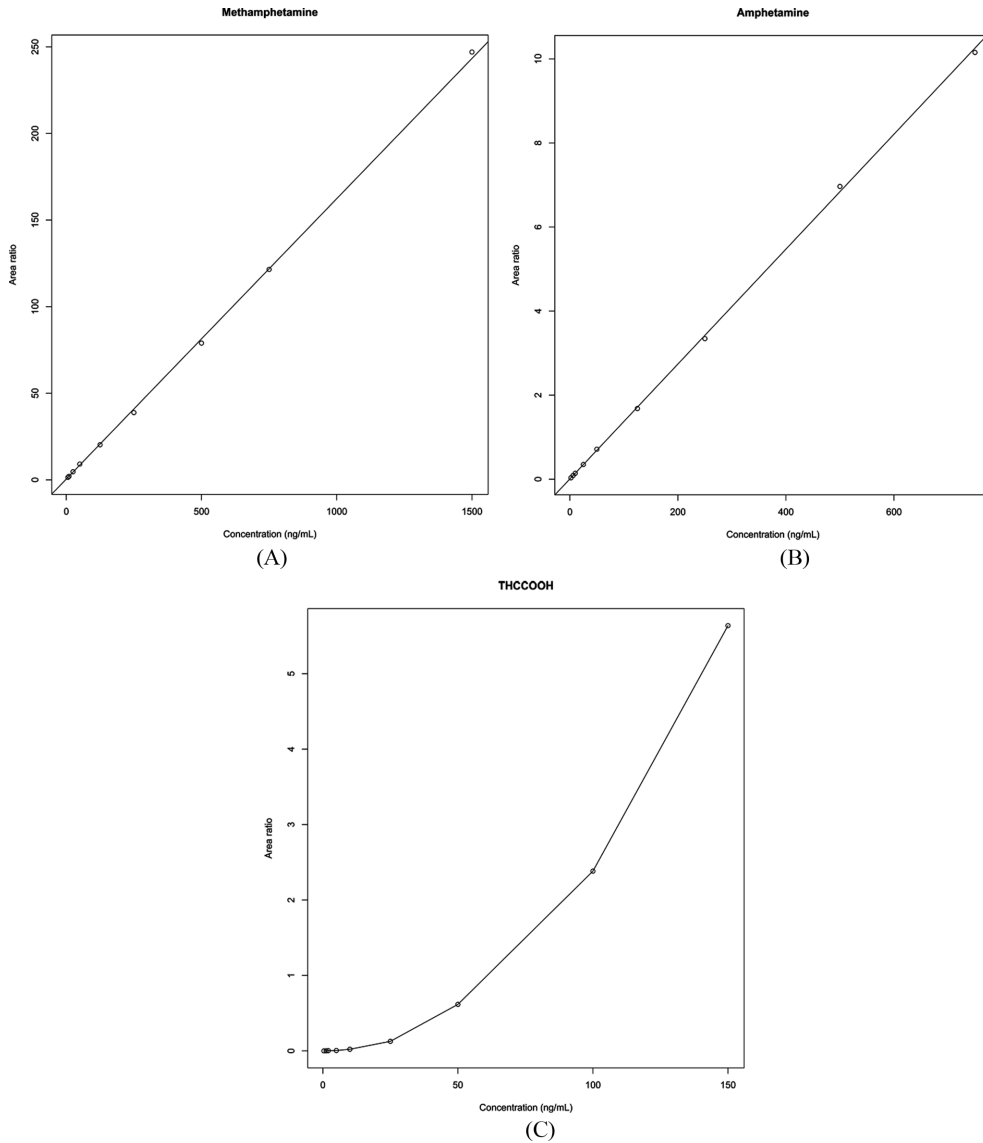


Fig. 1. The scatter plots generated for each correlation between concentration (x-axis) and area (y-axis) ratio for linear regression models of (A) methamphetamine, (B) amphetamine and quadratic regression model of (C) THCCOOH.

섯 번째 잔차도에서 오차 분포의 무질서도가 감소하였음을 확인할 수 있었다.

### 3.3. 가중계수 선정

검량선 작성에 사용되는 측정값에 이분산성을 확인하여 가중계수 도입이 필요하다는 결론이 내려지면, 가중계수  $1/x$ ,  $1/y$ ,  $1/x^2$  및  $1/y^2$ 을 적용한 각각의 회귀모형을 구하고 얻어진 4종의 회귀모형에 대해 분산분석 결과로부터 얻어진 F-value를 상호 비교하여 낮은 값을

가진 가중계수가 반영된 회귀모형을 선정하였다. 분석물질 MA, AP 및 THCCOOH에 대해 가중계수가 적용된 회귀모형식으로부터 얻어진 F-value를 비교한 결과, MA, AP는 가중계수  $1/x^2$ , THCCOOH는  $1/y^2$ 을 적용하였을 때 가장 좋은 결과를 얻을 수 있었다. 이러한 결과는 Fig. 2의 잔차도 비교를 통해서도 확인이 가능하였다.

### 3.4. 회귀모형 선정

검량선은 생체 시료 내 분석물질의 농도를 계산하

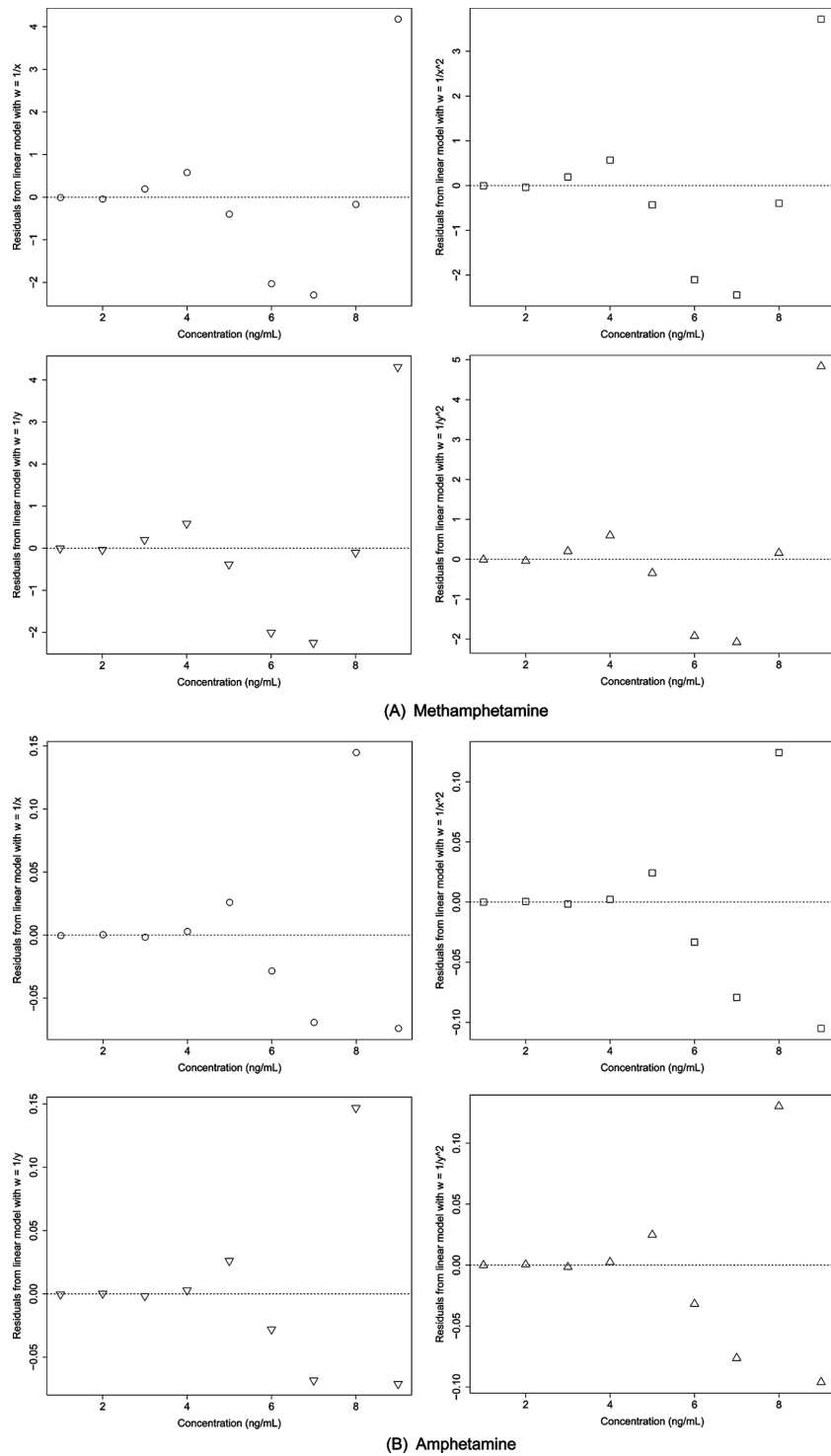


Fig. 2. Comparison of residual plots against  $1/x$ ,  $1/x^2$ ,  $1/y$  and  $1/y^2$  for selecting the correct weighting factors for (A) methamphetamine, (B) amphetamine and (C) THCCOOH.

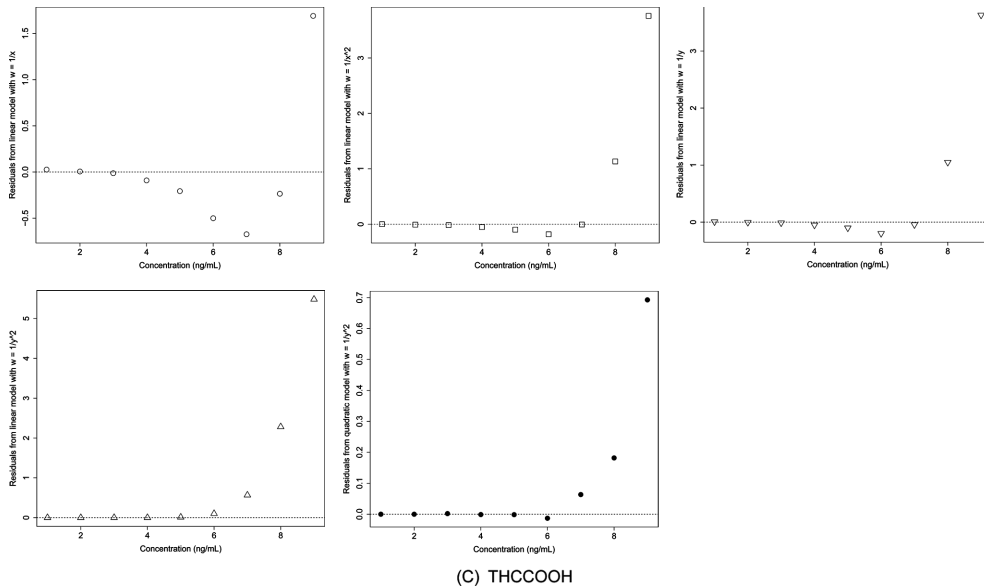


Fig. 2. Continued

Table 1. The results of R stepwise regression model selection

Compound	Heteroscedasticity test	Weighting factor	Model order	R <sup>2</sup>	Regression equation*
Methamphetamine	Weighting needed	$1/x^2$	Linear	0.998	$y=0.467+0.162 \cdot x$
Amphetamine	Weighting needed	$1/x^2$	Linear	0.999	$y=0.0614+0.0137 \cdot x$
11-Nor- $\Delta^9$ -tetrahydrocannabinol-9-carboxylic acid	Weighting needed	$1/y^2$	Quadratic	0.933	$y=0.000231+0.000107 \cdot x+0.000219 \cdot x^2$

\*Regression equations were exported into text file from R.

는 회귀모델식으로 분석의 정확도 뿐 아니라 정밀도에 영향을 미치며, 분석법의 정량범위 설정에 매우 중요한 요소이다. 따라서 검량선 작성 시 최저 농도(LLOQ, lower limit of quantification)와 최고 농도(ULOQ, upper limit of quantification)를 결정하는 정량 범위 설정이 적정한지 확인할 필요성이 크다. 검량선의 낮은 농도에서의 측정값의 오차 분포와 높은 농도에서 오차 분포가 일정 범위 내에 고르게 유지되는 것이 바람직하다. 분포가 고르지 않은 경우 가중계수를 도입하게 되고 선형 회귀모델이 적합하지 않은 경우 이어서 회귀모델의 적정 차수를 결정하게 된다.

본 연구에서는 비선형 회귀모델에 비해 직관성이 뛰어나고 쉽게 구현이 가능하다고 알려진 선형 회귀모델과 이차다항회귀모델을 대상으로 하였다. 회귀모델식을 결정하기 위해 오차의 분포 보다 모델식으로부터 유래하는 오차를 감소하는데 비중은 둔 편분산분석

결과를 사용하였다. 분석물질 MA, AP 및 THCCOOH에 대해 선형과 이차다항회귀모델을 적용한 결과, MA, AP와 달리 THCCOOH의 경우에서만 이차다항회귀모델식이 적합한 것으로 확인되었다.

최종 회귀모델 산출을 위해 전 과정에서 R을 사용하였고, 적정 가중계수, 회귀모델의 차수 등의 결과를 정리해 최종 모델식을 산출하였다. 얻어진 최종 결과는 Rstudio 콘솔을 통해서도 확인이 가능할 뿐만 아니라 별도의 TXT 형태의 문서 파일로 출력할 수 있도록 하였다. 단계별 과정을 거쳐 구한 회귀모델식의 결과는 Table 1에 나타내었다.

#### 4. 결 론

설정된 가중계수와 회귀모델의 적정성 여부에 따라 검량선 작성 시 농도에 대한 면적비의 측정값의 정확



도에 직접적 영향을 미치게 되고 결국 분석법 개발 및 유효화 과정에도 큰 영향을 미친다. 본 연구에서는 액체크로마토그래피-질량분석법을 이용한 소변 중 메트암페타민, 암페타민 및 대마 대사체 정량 분석과정에 선형 및 비선형 회귀모델의 적합성, 잔차분석 및 분산분석 결과 및 가중계수 도입·선정을 위해 이분산성 검정과 분산분석을 수행함으로써 가중계수를 선정하였다. 또한 회귀모델의 차수 선정을 위해 편분산분석을 수행하여 최적의 회귀모델을 선정할 수 있었다. 정량분석에서 검량선의 작성은 분석법의 유효화뿐만 아니라 측정의 오차에 있어서 가장 큰 영향을 주는 요소 중 하나임에 틀림없다. 따라서 본 연구에서와 같이 단계별 최적의 회귀모델을 결정하고 그 결과를 문서 파일로 확인과 저장하는 일련의 과정은 새로운 분석법 개발 시 검량선의 농도 범위, 최저 정량한계, 최고 정량한계 설정 시 결과의 정확도를 향상시킬 뿐만 아니라 일상적인 정량 분석 결과 값에 포함된 측정 오차를 줄일 수 있는 유용한 도구가 될 것으로 판단된다.

### 감사의 글

This study was supported in part by a grant from the National R&D program of Supreme Prosecutors' Office (SPO) and Ministry of Science and ICT (MSIT) of Korea.

### Supplemental Material

The R script for the selection of regression model for the preparation of calibration curve in the quantitative LC-MS/MS measurements is available at <https://github.com/paxus11>.

### References

1. N. J. Kwon and E. Han, *Forensic Sci. Int.*, **286**, 81-85 (2018).
2. K. Huh, *Korean Police Stud. Rev.*, **17**, 291-316 (2018).
3. J. Caldwell, L. G. Dring and R. T. Williams, *Biochem. J.*, **129**, 11-22 (1972).
4. C. B. Salocks, X. Hui, S. Lamel, P. Qiao, J. R. Sanborn and H. I. Maibach, *Food Chem. Toxicol.*, **50**, 4436-4440 (2012).
5. M. Carvalho, H. Carmo, V. M. Costa, J. P. Capela, H. Pontes, F. Remiao and M. L. Bastos, *Arch. Toxicol.*, **86**, 1167-1231 (2012).
6. K. Purschke, S. Heintl, O. Lerch, F. Erdmann and F. Veit, *Anal. Bioanal. Chem.*, **408**, 4379-4388 (2016).
7. S. Y. Kim, J. Y. Kim, W. Kwon, M. K. In, Y. E. Kim and K. J. Paeng, *Microchem. J.*, **110**, 326-333 (2013).
8. D. K. Lee, M. H. Yoon, Y. P. Kang, J. Yu, J. H. Park, J. Lee and S. W. Kwon, *Food Chem.*, **141**, 3931-3937 (2013).
9. J. M. Halket, D. Waterman, A. M. Przyborowska, R. K. P. Patel, P. D. Fraser and P. M. Bramley, *J. Exp. Bot.*, **56**, 219-243 (2005).
10. E. R. Perez, J. A. Knapp, C. K. Horn, S. L. Stillman, J. E. Evans and D. P. Arfsten, *J. Anal. Toxicol.*, **40**, 201-207 (2016).
11. J. C. Miller and J. N. Miller, *Analyst*, **116**, 3-14 (1991).
12. H. Gu, G. Liu, J. Wang, A. F. Aubry and M. E. Arnold, *Anal. Chem.*, **86**, 8959-8966 (2014).
13. Analytical Methods Committee, *Analyst*, **119**, 2363-2366 (1994).
14. P. Araujo, *J. Chromatogr. B*, **877**, 2224-2234 (2009).
15. V. Barwick, Preparation of calibration curves: a guide to best practice, Teddington, UK (2003).
16. D. A. Ratkowsky and D. E. Giles, Handbook of non-linear regression models (No. 04; QA278. 2, R3.), M. Dekker, New York (1990).
17. F. T. Peters, O. H. Drummer and F. Musshoff, *Forensic Sci. Int.*, **165**, 216-224 (2007).
18. B. Desharnais, F. Camirand-Lemyre, P. Mireault and C. D. Skinner, *J. Anal. Toxicol.*, **41**, 261-268 (2017).
19. G. Liu, Q. C. Ji and M. E. Arnold, *Anal. Chem.*, **82**, 9671-9677 (2010).
20. L. Yuan, N. Zheng, M. Jemal and M. E. Arnold, *Rapid Commun. Mass Spectrom.*, **26**, 1465-1474 (2012).
21. L. B. Nilsson and P. Skansen, *Rapid Commun. Mass Spectrom.*, **26**, 1399-1406 (2012).
22. B. Chou, R. S. Adler, M. Meng, S. Percey, B. Dean, C. E. Hop and Y. G. Shin, *J. Pharm. Biomed. Anal.*, **70**, 354-361 (2012).
23. X. Ding, B. Chou, R. A. Graham, S. Cheeti, S. Percey, L. C. Matassa and C. E. C. A. Hop, *J. Chromatogr. B*, **878**, 785-790 (2010).
24. B. Desharnais, F. Camirand-Lemyre, P. Mireault and C. D. Skinner, *J. Anal. Toxicol.*, **41**, 269-276 (2017).
25. D. C. Montgomery, E. A. Peck and G. G. Vining, Intro-

- duction to linear regression analysis, John Wiley & Sons (2021).
26. I. Lavagnini and F. Magno, *Mass Spectrom. Rev.*, **26**, 1-18 (2007).
27. B. R. Heo, N. Kwon and J. Y. Kim, *Anal. Sci. Technol.*, **31**, 161-170 (2018).
28. H. Teixeira, A. Verstraete, P. Proenca, F. Corte-Real, P. Monsanto and D. N. Vieira, *Forensic Sci. Int.*, **170**, 148-155 (2007).
29. A. M. Almeida, M. M. Castel-Branco and A. C. Falcao, *J. Chromatogr. B*, **774**, 215-222 (2002).

---

**Authors' Position**

Jin Young Kim : Forensic Chemist  
Dong Won Shin : Researcher