Original Article

Excel의 추세선을 이용한 표준곡선 검증

동아대학교 병원 핵의학과
이경화 · 박형기 · 신영만

Standard Curve Validation using Trendlines in Excel

Kyung-Hwa Lee, Hyung-Ki Park and Young-Man Shin
Dong-A University Hospital Nuclear Medicine

Purpose Using a regression formula of the trendline near the coefficient of determination (R2) "1" by substituting the dependent variable of the standard curve to calculate the values of the independent variable. To determine the suitability of a regression equation by comparing the difference between the independent variables of the standard curve and the predicted independent variables .

Materials and Methods Perkin Elmer Gamma-Counter machine was used for Standard curve of regression methods. TSH. TG-Ag (Thyroglobulin Antigen), Insulin that used materials and method test to compare the result from the Excel trendline of the regression formula.

Results Each of the value of coefficient of determination (R^2) and Trendline R^2=1, Polynomial Trendline for TSH , R^2=1, Polynomial Trendline for TG-Ag, R^2=0.9994, Polynomial Trendline for Insulin.

Conclusion We confirmed that IRMA immune method is found to the nearest trends elected a standard curve using polynomial trendline . The independent variables to predict the trend by using a polynomial trendline formula containing the error was a limitation .

Key Words Regression, Standard curve, Trendline

서 론

DREAM G-10 (Shin-jin, KOREA)의 Graph Algorithm에는 Point to Point, Linear Regression, 2'nd Order Polynomial, Cubic Spline, Smoothing Cubic Spline, Lagrange, Four Parameter Logistics으로 설정되어 있으며, 각각의 Graph Algorithm의 종류에 따라 오차를 포함하고 있다. Graph Algorithm에 따라 데이터에 오차를 포함한다는 점은 검사실내에서 설정된 표준곡선을 기준으로 하여 장비 Graph Algorithm이외에도 신뢰성이 높은 Graph Algorithm을 이용할 수 있다는 점도 시사한다. 그런 의미에서 간편하게 사용하는 Excel의 추세선(=회귀분석)을 이용하여 검사실내에서 설정된 표준곡선과 비교분석하여 결정계수가 "1"에 가까운 추세선 식을 먼저 유도하고, 추세선에서 유도된 예측값과의 적합성을 회귀직선 모형을 통하여 그 값을 비교분석하였다. 또한 결정된 추세선식을 이용하여 시약(Kit)내의 각각의 표준농도별로 추세선에 미치는 영향을 비교분석하였다.

따라서 본 연구의 목적은 검사실의 표준곡선을 기준으로 Excel의 추세선을 유도하고, 시약의 농도별로 추세선에 미치는 영향을 비교분석하여 Excel의 추세선이 예측값에 대한 용이성을 제고하는 것이다.

• Received: September 27, 2016 Accepted: October 6, 2016
• Correspoding author : **Kyung-Hwa Lee**
Address for correspondence : Dong-A University Hospital, Dongdaesin-dong 3(sam)-ga, Seo-gu, Busan-si, Korea
Tel : +82-51-240-5632
Fax : +82-51-242-7237
E-mail : khlee6710@hanmail.net

방 법

1. 대상 및 장비

저, 중, 고농도의 값이 고루 분포하고 방사면역측정법(radioimmunoassay: IRMA)법인 종목- Insulin 83건을 검사

대상으로 하고, 장비는 WIZARD (PerkinElmer, USA)와 DREAM G-10 (Shinjin, KOREA)을 이용하여 장비의 Graph Algorithm과 Excel의 추세선 프로그램을 이용하였다.

2. 방법

1) 모형설정 방법

(순서 1) Excel 데이터의 입력

Table 1. Input of Insulin data in Excel

No	A:CPM	A:Dose	Pat/ID
1	156.1	0	BLANK
2	401.5	5.3	STD 1
3	1142.1	12.1	STD 2
4	6054.6	37.5	STD 3
5	16464.8	88	STD 4
6	42503.8	315	STD 5
76	32,248	225.593	SAMPLE 70
77	34,689	246.873	SAMPLE 71
78	34,689	246.873	SAMPLE 72
79	39,679	290.374	SAMPLE 73
80	40,482	297.374	SAMPLE 74
81	40,482	297.374	SAMPLE 75
82	40,678	299.083	SAMPLE 76
83	41,065	302.457	SAMPLE 77
84	41,065	302.457	SAMPLE 78
85	41,678	307.801	SAMPLE 79
86	41,899	309.727	SAMPLE 80
87	41,899	309.727	SAMPLE 81
88	42,308	313.293	SAMPLE 82
89	42,676	316.501	SAMPLE 83

(순서 2) 분석 도구의 호출

- 메뉴에서 [도구(T)]-[데이터 분석(D)]을 선택

(순서 3) 회귀분석의 선택

- [통계 데이터 분석] 대화상자에서, [분석 도구(A) : 회귀 분석]을 선택하고 [확인] 버튼을 클릭

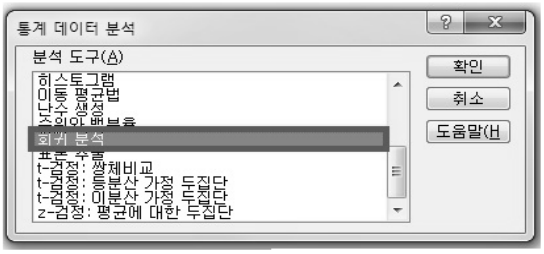


Fig. 1. Regression analysis selection.

- [회귀분석] 대화상자가 나타나면, Table 1.의 사항과 신뢰수준, 잔차를 지정하고 [확인] 버튼을 클릭

Fig. 2. Regression analysis term input.

(순서 4) 실행결과

- [그림을 통한 회귀분석] 데이터 하나를 선택하고 추세선을 추가하고 추세선 옵션에서 추세 유형별로 선택

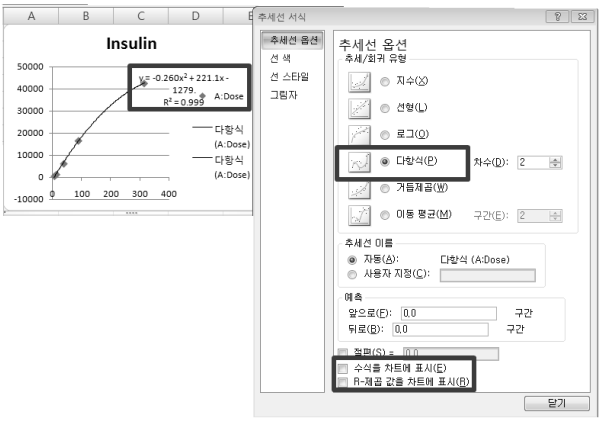


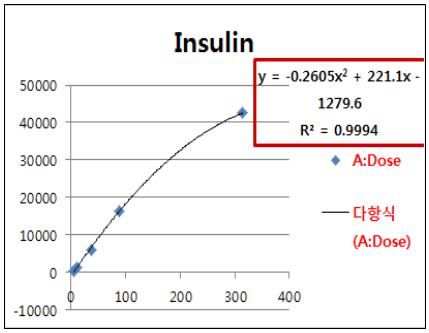
Fig. 3. Select the type of trendline.

2) 추세선을 이용한 농도별 변화 비교 방법

(순서 1) Excel의 다항식 추세선을 이용한 Insulin 표준 농도의 다항식의 계산

- Insulin 표준 농도 0, 5.3, 12.1, 37.5, 88, 315 μ IU/㎖ (6개)
- 표준농도 회귀식 $y = -0.260x^2 + 221.1x - 1279.$ $R^2 = 0.999$

($a : x^2$(이차항 상수), $b : x$(일차항 상수), c : 절편)



2차 방정식과 근의 공식 이용

$$y = ax^2 + bx + c$$

$$\bar{x} = \frac{-b \pm \sqrt{b^2 - 4a(c-y)}}{2a}$$

Fig. 4. Polynomial trendline in Excel– Standard curve.

(순서 2) 표준농도 한 개씩을 제외시켜 농도별로 추세선의 예측 값에 미치는 영향 비교

- 다항식의 추세선을 이용하여 표준농도 하나씩 제외시켜 기울기와 절편을 비교

3) Algorithm종류와 추세선 비교 방법

(순서 1) 평균값의 비교

- WIZARD (PerkinElmer, USA)의 표준 평균값과 DREAM G-10 (Shinjin, KOREA)의 Graph Algorithm의 종류별로 평균값 비교
- WIZARD (PerkinElmer, USA)의 표준 평균값과 Excel의 다항 회귀 추세선의 평균값 비교
- Excel의 다항 회귀 추세선의 평균값과 DREAM G-10 (Shinjin, KOREA)의 Graph Algorithm의 종류별로 평균값 비교

(순서 2) 평균값을 기준으로 적합성이 높은 Graph Algorithm의 비교

4) 결과 해석방법

(해석 1) 결정계수

□ 결정계수는 목적변수의 분산 중 설명변수에 의한 회귀 예측변량이 설명하는 분산의 비율을 나타내는 것으로 목적변수 y의 변동 중에서 x를 사용한 회귀식에 의해서 설명되는 변동의 비율을 나타낸다.

$$\text{결정계수}(R) = \frac{\text{회귀방정식의 편차}}{\text{전체편차}} = \frac{\sum(y_i - \bar{y})^2 - \sum(y_i - \bar{y}_{cal})^2}{\sum(y_i - \bar{y})}$$

□ 결정계수가 1에 가까울수록 회귀식이 데이터에 가장 적합하다.

(해석 2) 회귀식

□ 분석결과 얻어진 회귀식 중 다항식이 결정계수 "1"에 가장 가깝다.

$y = -0.260x^2 + 221.1x - 1279$

$R^2 = 0.999$

□ 회귀직선의 결정에 기울기와 절편값의 결정에는 최소제곱법으로 계산한다.

최적의 직선= 회귀선 y = ax + b (기울기: a, 절편값: b)

a, b를 계산한 최종결과

$$a = \frac{\left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right) - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} x_i y_i\right)}{N\left(\sum_{i=1}^{N} x_i^2\right) - \left(\sum_{i=1}^{N} x_i\right)^2} \qquad b = \frac{N\left(\sum_{i=1}^{N} x_i y_i\right) - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right)}{N\left(\sum_{i=1}^{N} x_i^2\right) - \left(\sum_{i=1}^{N} x_i\right)^2}$$

결 과

1) 모형설정 결과

□ WIZARD (PerkinElmer, USA)의 Graph Algorithm-Point to Point를 표준곡선으로 하여 Excel의 다항식 추세선으로 결정

- 다항식 추세선 결정계수= 0.999로 결정계수= 1에 가장 근접

2) 추세선을 이용한 표준 농도별 변화 비교 결과

□ 6개의 표준농도를 1개씩 표준농도 값을 제외시켜 각각의 표준 농도가 다항식 추세선에 미치는 영향을 다항 회귀식의 기울기와 결정계수를 비교한 결과 표준농도 값에서 가장 높은 농도인 315 μIU/mℓ를 제외한 결과가 표준 다항회귀식과 차이가 가장 큼.

- 표준 다항식 추세선

  $y = -0.2605x2 + 221.1x - 1279.6$, $R^2 = 0.9994$

  315 μIU/mℓ를 제외한 추세선

  $y = 0.353x2 + 163.4x - 647.9$, $R^2 = 0.999$

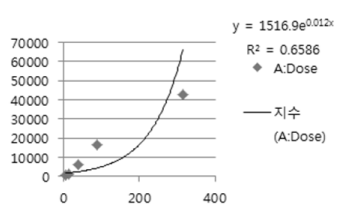
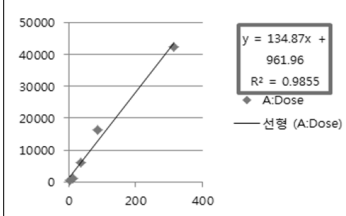
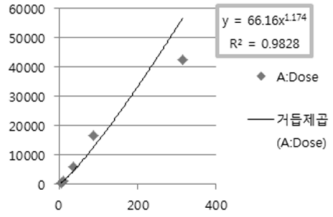
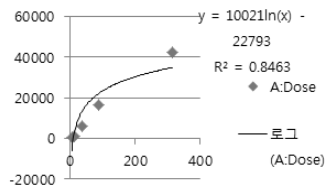
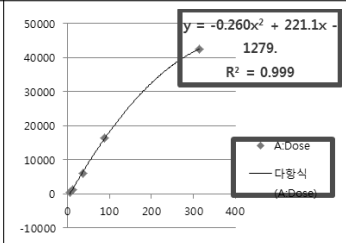
유형	지수	선형	거듭제곱
추세선			
회귀식	$y = 1516.9e^{0.012x}$ $R^2 = 0.6586$	$y = 134.87x + 961.96$ $R^2 = 0.9855$	$y = 66.16x^{1.174}$ $R^2 = 0.9828$
유형	로그	다항식	
추세선			
회귀식	$y = 10021\ln(x) - 22793$ $R^2 = 0.8463$	$y = -0.260x^2 + 221.1x - 1279.$ $R^2 = 0.999$	

Fig. 5. Depending on the trendline types compare with linear regression formula and coefficients of determination.

(Insulin 단위: μ IU/㎖)

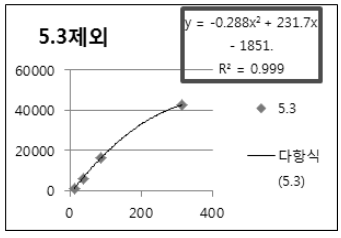
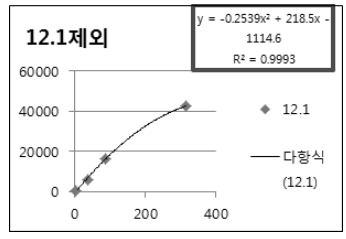
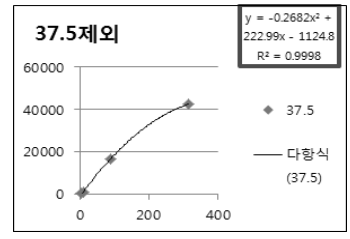
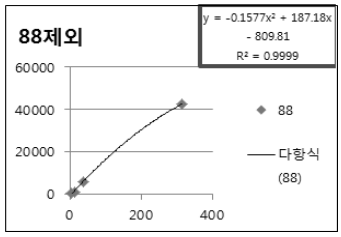
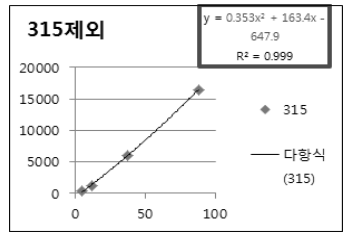
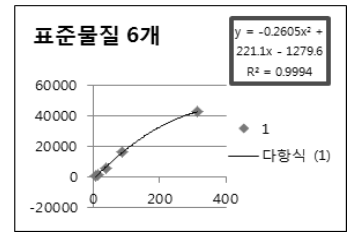
Std.	5.3 제외	12.1 제외	37.5 제외
추세선			
회귀식	y = -0.288x2 + 231.7x - 1851. $R^2 = 0.999$	y = -0.2539x2 + 218.5x - 1114.6 $R^2 = 0.9993$	y = -0.2682x2 + 222.99x - 1124.8 $R^2 = 0.9998$
Std.	88 제외	315 제외	표준물질 유지
추세선			
회귀식	y = -0.1577x2 + 187.18x - 809.81 $R^2 = 0.9999$	y = 0.353x2 + 163.4x - 647.9 $R^2 = 0.999$	y = -0.2605x2 + 221.1x - 1279.6 $R^2 = 0.9994$

Fig. 6. Using regression trendline compare with except standard concentration by changing the regression formula and coefficient of determination.

□ 표준농도 88, 315 μIU/mℓ를 제외한 회귀식으로 예측된 평균값과 표준농도 회귀식의 평균값과의 차이 비교 결과

- 88 μIU/mℓ는 102.87로 7.05의 차이, 315 μIU/mℓ는 74.42로 34.45만큼 감소된 평균값이 나옴.

Table 2-1. Compared with the average value of the each concentration of exclusion

단위(μIU/mℓ)	평균	평균차이
표준 농도	96.75	0
5.3	96.7	0.05
12.1	96.54	0.16
37.5	95.79	0.75
88	102.84	7.05
315	74.42	-34.45

Table 2-2. Compared with the data, average data of standard and except concentration in Excel

Pat/ID	다항식	5.3	12.1	37.5	88	315
BLANK	-0.26	-0.288	-0.253	-0.268	-0.157	0.353
STD 1	221.1	231.7	218.5	222.9	187.1	163.4
STD 2	-1279	-1851	-1114	-1124	-809.8	-647.9
STD 3	y=ax+bx+c					
STD 4						
STD 5	0	5.3	12.1	37.5	88	315
SAMPLE 58	150.17	147.51	150.65	148.46	167.03	128.04
SAMPLE 59	155.23	152.45	155.74	153.51	172.39	130.85
SAMPLE 60	155.23	152.45	155.74	153.51	172.39	130.85
SAMPLE 61	174.95	171.72	175.53	173.21	192.93	141.13
SAMPLE 62	178.93	175.62	179.52	177.19	197.02	143.08
SAMPLE 63	178.93	175.62	179.52	177.19	197.02	143.08
SAMPLE 64	187.84	184.38	188.46	186.12	206.07	147.30
SAMPLE 65	187.84	184.38	188.46	186.12	206.07	147.30
SAMPLE 66	188.20	184.73	188.82	186.48	206.43	147.46
SAMPLE 67	188.20	184.73	188.82	186.48	206.43	147.46
SAMPLE 68	192.73	189.19	193.36	191.02	210.98	149.52
SAMPLE 69	192.73	189.19	193.36	191.02	210.98	149.52
SAMPLE 70	197.51	193.90	198.15	195.82	215.74	151.64
SAMPLE 71	219.16	215.35	219.80	217.60	236.77	160.56
SAMPLE 72	219.16	215.35	219.80	217.60	236.77	160.56
SAMPLE 73	272.69	269.56	272.98	272.02	284.16	178.20
SAMPLE 74	283.17	280.52	283.31	282.85	292.47	180.97
SAMPLE 75	283.17	280.52	283.31	282.85	292.47	180.97
SAMPLE 76	285.85	283.34	285.94	285.63	294.54	181.64
SAMPLE 77	291.30	289.13	291.28	291.29	298.66	182.96
SAMPLE 78	291.30	289.13	291.28	291.29	298.66	182.96
SAMPLE 79	300.42	298.97	300.18	300.84	305.30	185.05
SAMPLE 80	303.87	302.75	303.54	304.48	307.73	185.80
SAMPLE 81	303.87	302.75	303.54	304.48	307.73	185.80
SAMPLE 82	310.54	310.17	310.00	311.56	312.29	187.19
SAMPLE 83	316.88	317.39	316.13	318.36	316.45	188.43
평균	96.75	96.70	96.54	95.79	102.87	74.42

3) Algorithm종류와 추세선 비교 결과

□ Excel의 다항 회귀식의 WIZARD (PerkinElmer, USA)과 DREAM G-10 (Shinjin, KOREA)의 Graph Algorithm 이용 가능성 여부

- WIZARD (PerkinElmer, USA)- Point to Point의 표준곡선은 DREAM G-10 (Shinjin, KOREA)-Point to Point곡선과 적합성이 높음.
- Excel 다항식과 DREAM G-10 (Shinjin, KOREA)-2'nd order Polynomial과 평균값의 적합성이 높음

		WIZARD	DREAM10				Excel 다항식
No	A:CPM	A:Dose	Point-to-Point	Lineaar regression	2'nd order polynomial	Cubic Spline	Lab
1	156						-0.2605
2	401	5.3	5.3	5.30	5.30	5.30	221.09
3	1142	12.1	12.1	12.10	12.10	12.10	-1279.9
4	6054	37.5	37.5	37.50	37.50	37.50	
5	16464	88	88	88.00	88.00	88.00	
6	42503	315	315	315.00	315.00	315.00	
7	2082	21.612	18.31	19.8	15.87	18.23	15.49
41	29444	227.28	192.5	173.36	189.45	180.69	175.08
42	42308	313.898	313.06	232.24	289.82	312.66	311.42
43	18115	110.93	100.12	117.07	110.8	97.72	99.36
44	19608	129.928	111.41	124.82	120.61	107.01	108.30
45	41678	310.315	306.83	229.45	284.64	305.18	301.18
46	9817	62	56.72	71.21	59.42	54.19	53.57
47	39679	298.577	287.27	220.53	268.39	282.17	273.20
48	41899	311.578	309.01	230.43	286.46	307.79	304.67
49	26778	204.576	169.46	160.58	170.09	158.52	155.34
50	41065	306.777	300.8	226.72	279.63	298.01	291.97
51	5551	36.233	35.4	44.7	35.07	35.26	32.11
52	10794	66.783	61.49	76.93	65.19	58.71	58.67
53	34689	266.47	239.89	197.88	229	229.36	219.41
54	5841	36.926	36.62	46.6	36.69	36.56	33.53
55	9621	60.982	55.75	70.05	58.28	53.3	52.56
56	17695	105.3	97	114.87	108.07	95.19	96.88
57	8721	56.021	51.27	64.67	53.04	49.25	47.94
58	16429	87.891	87.84	108.17	99.92	87.8	89.55
59	31676	244.752	212.35	183.9	206.04	200.57	192.91
60	19120	123.883	107.68	122.3	117.38	103.92	105.35
평균		76.3	70.5	64.5	69.0	68.8	67.9

Fig. 7. Compared with the trendline and Algorithm

고 찰

본 연구의 근거는 WIZARD (PerkinElmer, USA)와 DREAM G-10 (Shinjin, KOREA)과 같은 장비가 Graph Algorithm이 다양하게 표현이 가능하다는 것이다. 이점은 간편한 Excel 프로그램인 추세선(Trendline; 회귀분석)도 Graph Algorithm으로써 이용이 가능하다는 것을 시사한다고 보았다. Excel의 회귀분석을 이용한 추세선을 비교분석하게 된 근거는 통계학에서 회귀분석은 연속형 변수들에 대한 독립변수와 종속변수 사이의 상관관계를 나타내는 선형관계식을 구하는 기법이며, 모형의 적합도를 측정하는 분석방법이라고 정의하므로 이용 가능하다고 보았다.

본 연구에서 Graph Algorithm 상관성의 정도는 결정계수(R^2)로 해석을 하였으며 결정계수가 "1"에 가까울수록 Y값

에 대한 X 값의 변화를 충분히 설명한다고 보았다. 그러나 여기서 '충분히 설명한다.'는 것이 'Y값으로 X값을 100%를 나타내는 것이 아니다.' 라는 문제를 가진다.[1)]

장비회사에서 제공되는 Graph Algorithm을 이용한 방법에서부터 이미 Errors를 포함한 99%의 농도를 계산하는 점을 보아도 이미 Error를 포함한 경우임을 알 수 있다. 이러한 오차를 포함한 표준곡선을 기준으로 하여 다시 오차를 포함한 식을 적용한 예측값을 얻는다는 것에 대한 문제점은 「수치해석」적인 면에서 더 연구가 되어야 할 부분이므로 본 연구에는 오차를 포함한 Excel의 추세선이 과연 어느 정도 장비의 Graph Algorithm과 상관성을 가지는지 알아본 것이다.

상관성을 측정하는 궁극적인 목적은 Excel의 추세선을 이용한 예측값이 임상에 적용 가능성의 정도를 알고자 한 것이다. 또 시약회사에서 제공된 6개의 표준 농도값에서 1개씩 농도를 제외한 추세선 회귀식을 이용하여 표준농도가 고농도일수록 Graph Algorithm에 대해 미치는 영향이 크다는 것을 알 수 있다.

이것은 검사실내 표준농도 조절이 요구될 경우 고농도의 문제를 먼저 개선해야한다는 것을 의미한다. 여기서 표준농도 조절이라는 의미는 결정계수(R^2)를 조정한다는 의미이기도 하다. 그리고 Graph Algorithm과 Excel 추세선과의 적합성 정도를 결정계수(R^2)뿐만 아니라 다른 기준의 적용에 대한 과제도 전제한다.

요 약

Insulin 83건을 검사 대상으로, 장비 WIZARD(PerkinElmer, USA)와 DREAM- G-10(Shinjin, KOREA)의 Graph Algorithm 중에 표준곡선과 신뢰성이 가장 높은 Excel의 추세선은 다항식 추세선이다. 다음으로 다항식 추세선식을 이용하여 표준농도 1개씩을 제외하여 표준농도의 회귀식에 미치는 영향을 비교한 결과 최고농도(315 μ IU/㎖)의 평균값이 표준물질 6개로 실시한 표준 회귀식을 이용한 평균값과 비교하여 49%나 평균값이 저하되었다. 단 낮은 농도에서는 영향이 미비하였다.

마지막으로 WIZARD의 Point to Point형식과 DREAM G-10의 Point to Point형식이 적합성이 높고, DREAM G-10(Point to Point)와 Excel 다항식 추세선이 적합성이 높으며, Excel 다항식 추세선과 DREAM G-10(2'nd order Polynomial)의 적합성이 높다.

REFERENCES

1. PerkinElmer User guide to MultiCalc functions, Part 2 Standard curve theory, 5:24-5:47.
2. 진성태, Excel로 배우는 통계분석, 대경북스, 2011.
3. 이일현, EasyFlow 회귀분석, 한나래, 2014:41-119.
4. 노형진, Excel을 활용한 통계적 품질관리 제10부 회귀분석, 중소기업청. 대한 상공회의소 품질혁신 교육 제28호. 2005.
5. 황창선, 김승범, 유.무인 회전익기 주요 설계변수의 추세선 및 비교분석 연구, 한국우주항공학회지, 2005.