บทที่ 3

ความถูกต้องของผลการตรวจวิเคราะห์

ในการปฏิบัติงานทางด้านเคมีคลินิคความถูกต้องของผลการตรวจวิเคราะห์เป็นสิ่งแรกที่จะต้องคำนึงถึง ทั้งนี้ โดยการตั้งมาตรฐานการตรวจให้อยู่ในระดับที่น่าเชื่อถือได้ และเปรียบเทียบกรรมวิธีการตรวจวิเคราะห์ต่าง ๆ แล้วจึงเลือกวิธีที่ให้ผลดีที่สุดมาใช้ในการปฏิบัติงานต่อไป

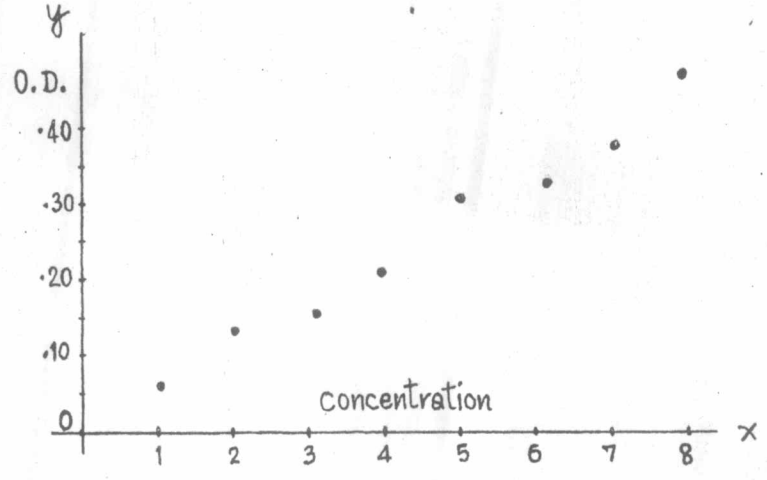
3.1 การตั้งมาตรฐานการตรวจวิเคราะห์ (Standardization)

นำ Primary standard มาละลายด้วยตัวทำละลายเพื่อให้มีค่าความเข้มข้น (Concentration) ต่าง ๆ กัน แล้วนำไปผ่านกรรมวิธีการตรวจวิเคราะห์และวัดปริมาณความหนาแน่นของแสง (Optical density) ได้ ดังนั้นถ้าให้

ความเข้มข้นของ Standard (concentration) = x_i

ความหนาแน่นของแสงที่วัดได้ (O.D.) = y_i

จะเห็นได้ว่า x_i และ y_i เป็น bivariate population คือทุก ๆ จุดที่เราหาบค่าความเข้มข้น (x_i) เราจะหาบค่า Optical density (y_i) ด้วยเสมอไป ดังนั้นใน x, y coordinate system สามารถเขียนแผนผังการกระจาย (Scatter diagram) ได้ดังภาพที่ 6



ภาพที่ 6 แผนผังการกระจายสำหรับ bivariate population (x_i, y_i)

เมื่อลากเส้นตรงผ่านจุดต่าง ๆ บนแผนผังการกระจาย ก็จะได้เส้นมาตรฐาน (Standard curve) สำหรับไว้ใช้ในการปฏิบัติงานกับคนไข้ต่อไป ในทางปฏิบัติเราทำกลับกันกับการตั้งมาตรฐานการตรวจวิเคราะห์ โดยนำเลือดจากผู้ป่วยมาวิเคราะห์แล้วอ่านค่า O.D. (y) แล้วจึงเปรียบเทียบค่าเป็น Concentration (x)

จากรูปที่ 6 ถ้าเราไม่สามารถลากเส้นผ่านจุดต่าง ๆ ได้ควรตรวจสอบดูความสัมพันธ์ระหว่าง x_i กับ y_i โดย Linear regression analysis แล้วจึงสร้าง Linear regression equation ขึ้นใหม่

สูตรคำนวณ Linear regression

$$b_x = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

$$b_y = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$r = \sqrt{b_x \cdot b_y}$$

$$t_{(n-2)} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

เมื่อ r = correlation coefficient

t = t-value

b_x = regression of x on y

b_y = regression of y on x = slope

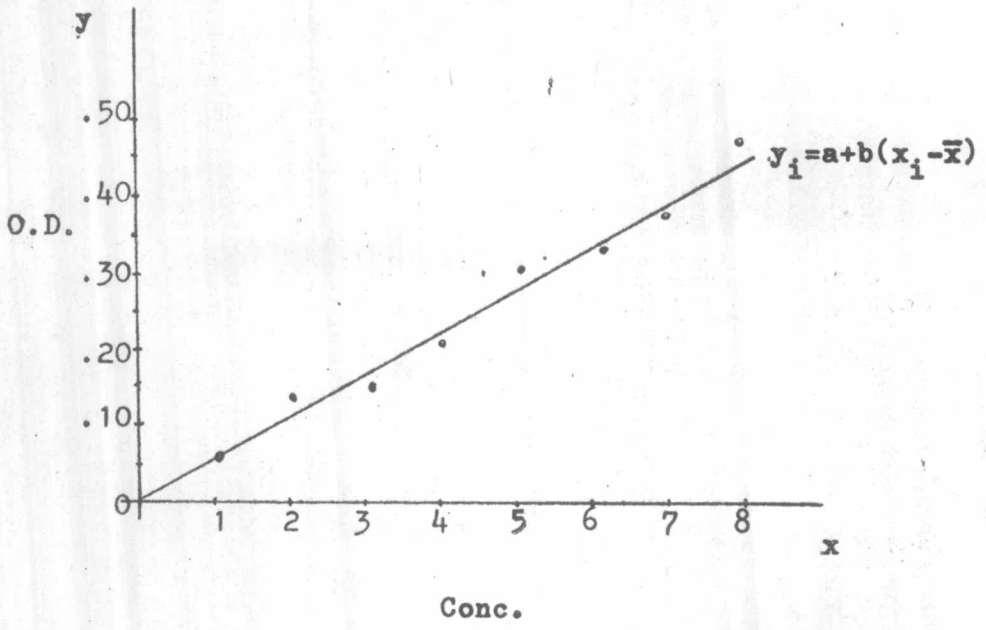
n = pairs of x and y

สูตรคำนวณ Linear regression equation

$$y_i = a + b\,(x_i - \bar{x})$$

เมื่อ $$b = \frac{\sum_{i=1}^{n} \left[y_i (x_i - \bar{x}) \right]}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

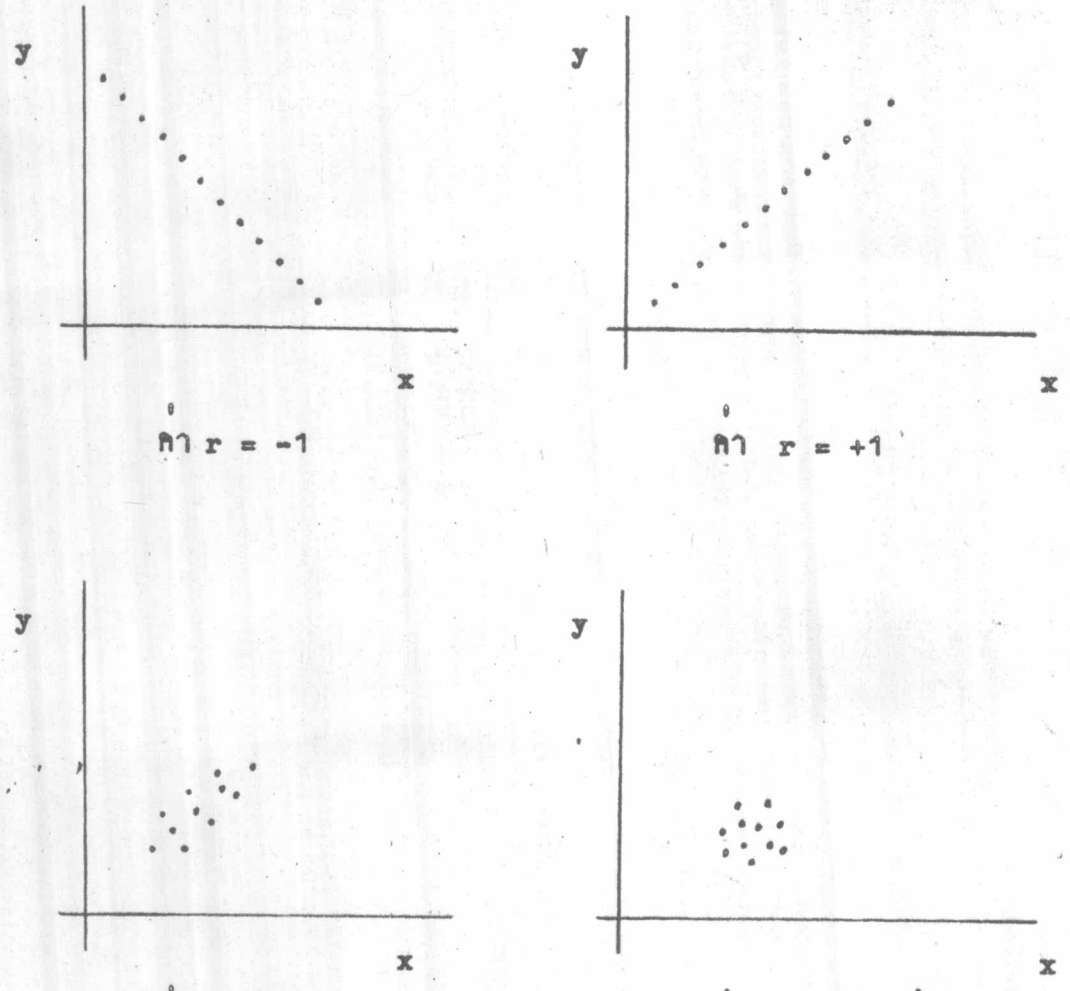
และ $$a = \bar{y}$$



ภาพที่ 7 เส้นมาตรฐานที่ได้จาก Linear regression equation

ความสำคัญของ Correlation coefficient

ถ้าค่า Correlation coefficient (r) มีค่าเข้าใกล้ ± 1 แสดงว่ามีความสัมพันธ์กันเป็นเส้นตรง (Linear relationship) แต่ถ้า r เข้าใกล้ 0 แสดงว่าไม่มีความสัมพันธ์กันและไม่เป็นเส้นตรง



ภาพที่ 8 ลักษณะของ regression line ของค่า correlation coefficient ต่าง ๆ

วิธีทดสอบสมมติฐาน

ใน Normal bivariate population correlation coefficient $\rho = 0$ ตัวอย่างที่มีจำนวนคู่ = n ที่สุ่มออกมาจาก population ดังกล่าวมีค่า correlation coefficient = r ดังนั้น

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

ค่า t ที่คำนวณได้จะใช้ในการตัดสินว่า ตัวอย่างที่มีจำนวนคู่ = n และมี correlation coefficient = r นั้น มีความสัมพันธ์กันเป็นเส้นตรงหรือไม่ ถ้าความน่าจะเป็น (probability) ที่ได้มีค่าน้อยกว่า 5 % ก็ปฏิเสธสมมติฐานดั้งเดิมที่ว่าตัวอย่างนั้นมาจาก bivariate population ที่มี $\rho = 0$

ซึ่งหมายความว่า birariate population นั้น มี $\rho \neq 0$ แสดงว่ามี linear relationship

สัญญลักษณ์ที่ใช้ในโปรแกรมคอมพิวเตอร์ (โปรแกรมที่ 1 ผนวก ง.)

correlation coefficient	= R
t-value	= T
regression of x on y	= BX
x-intercept	= XINT
y_i	= OD (I)
x_i	= CONC (I)
number of pairs of x_i and y_i	= N

การจัดลำดับข้อมูล

บัตรที่	สดมภ์ที่	เขตพิกัด	ตัวแปร
1	1-2	I2	N = 10
2-11	1-8	F8.3	CONC (I)
	9-16	F8.3	OD (I)

มิติ (Dimension)

โปรแกรมนี้สามารถรับข้อมูลที่ได้จากความเข้มข้นของ Standard (CONC(I)) กับความหนาแน่นของแสงที่วัดได้ (OD(I)) รวมทั้งหมดไม่เกิน 20 คู่ ซึ่งมากพอที่จะใช้ในงานวิจัย หรืองานประจำวันเป็นอย่างดี ในการทดลองนี้ได้นำข้อมูลเข้าเครื่องคอมพิวเตอร์เพียง 10 คู่ (N = 10) การทำงานของคำสั่งต่าง ๆ ในโปรแกรมก็ยังถูกต้องเรียบร้อย โดยไม่ต้องมีการแก้ไขหรือเพิ่มเติมคำสั่งต่าง ๆ เลย

3.2 การเลือกวิธีตรวจวิเคราะห์

ในการตรวจวิเคราะห์สารเคมีแต่ละชนิดอาจมีวิธีตรวจได้หลายวิธีด้วยกัน จึงจำเป็นต้องเลือกเอาวิธีที่ดีที่สุดไว้ใช้ในการปฏิบัติงานต่อไป สมมติว่าจะเลือกวิธีที่ดีที่สุดจาก 2 วิธี การสร้างแบบการทดลอง (Experimental design) จะแบ่งออกได้เป็น

3.2.1 เปรียบเทียบผลของวิธีตรวจวิเคราะห์ทั้ง 2 วิธี

วิธีทดลองที่ 1 (Pair t-test)

นำเลือดของผู้ป่วย (unknown sample) ประมาณ 30 ราย มาผ่านกรรมวิธีการตรวจวิเคราะห์ทั้ง 2 วิธี ผลที่วิเคราะห์ได้เป็น x_i และ y_i ตามลำดับ ถ้าให้ x_i และ y_i เป็นตัวแปร และมีลักษณะการแจกแจงความถี่เป็นแบบปกติ และ n เป็นจำนวน population (คู่ของ x_i และ y_i) จะสามารถตรวจสอบได้ว่า ค่าเฉลี่ยที่วัดได้จากการตรวจวิเคราะห์ทั้ง 2 วิธี นั้น มีค่าแตกต่างกันหรือไม่

สูตร

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (d_i)^2 - (\sum_{i=1}^{n} d_i)^2 / n}{n - 1}}$$

$$t = \sqrt{n}\,\frac{\bar{d}}{SD}$$

จำนวน d.f. = n - 1

เมื่อ n = number of pairs of observation

d_i = difference $(x_i - y_i)$

$\bar{d}$ = mean of the differences

SD = standard deviation of the differences

สมมติฐาน $\bar{x} = \bar{y}$ ที่ระดับความเชื่อมั่น 95 %

หมายเหตุ วิธีการคำนวณนี้ใช้ได้ในกรณีที่มีค่าเป็น pair variates เท่านั้น (ทั้ง x และ y มี equal replication)

สัญญลักษณ์ที่ใช้ในโปรแกรมคอมพิวเตอร์ (โปรแกรมที่ 2 ผนวก ง.)

ข้อมูลที่ได้จากวิธี A = A

ข้อมูลที่ได้จากวิธี B = B

Mean of the differences (d) = DBAR

S.D. of the differences = SD

Number of pairs of observation = N

Sum of differences ($\sum_{i=1}^{n} d_i$) = SUMD

Sum of differences square $(\sum_{i=1}^{n} d_i)^2$ = SDSQ

t-value = T

ข้อมูลที่ให้เครื่องอ่านครั้งละ 10 ค่า = DATAA (I)

การจัดลำดับข้อมูล

บัตรที่	สดมภ์ที่	เซตพิกัด	ตัวแปร
1-3	1-80	10F8.2	DATAA (I) จากวิธี A
4	1-8	F8.2	77777777
5-7	1-80	10F8.2	DATAA (I) จากวิธี B
8	1-8	F8.2	88888888

มิติ (Dimension)

โปรแกรมนี้สามารถรับข้อมูลที่ได้จากวิธีตรวจวิเคราะห์ทั้ง 2 วิธี ไม่เกิน 30 คู่ คือ A(30) และ B(30) แต่ถ้าในการทดลองเกิดอุปสรรคขึ้นทำให้ข้อมูลที่ได้ไม่ครบ 30 คู่ โปรแกรมนี้ก็สามารถทำงานได้เป็นอย่างดีโดยไม่ต้องมีการแก้ไข

ในกรณีที่ข้อมูลทั้ง 2 ชุด มีจำนวนการทดลองไม่เท่ากัน (Unequal replication) ถ้าให้ x และ y เป็นตัวแปรซึ่งมีลักษณะการแจกแจงความถี่เป็นแบบปกติ โดยเป็นตัวแปรซึ่งมีอิสระต่อกัน มีค่าเฉลี่ยเป็น m_x, m_y และมี population เป็น n_1, n_2 ตามลำดับ จะสามารถคำนวณค่า t ได้จากสูตร

$$t = \frac{(\bar{x} - \bar{y}) - m_{\bar{x}-\bar{y}}}{S_{\bar{x}-\bar{y}}}$$

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + \sum_{i=1}^{n}(y_i - \bar{y})^2}{n_1 + n_2 - 2}}$$

$$S_{\bar{x}-\bar{y}} = \sqrt{(S_{\bar{x}})^2 + (S_{\bar{y}})^2}$$

$$S_{\bar{x}} = \frac{S}{\sqrt{n_1}}$$

$$S_{\bar{y}} = \frac{S}{\sqrt{n_2}}$$

จำนวนชั้นแห่งความเป็นอิสระ $= n_1 + n_2 - 2$

สมมติฐาน $m_x = m_y$

หรือ $m_x - m_y = 0$ ที่ระดับความเชื่อมั่น 95 %

วิธีทดลองที่ 2 (Bartletts test)

นำ Pool serum ที่เก็บรวบรวมไว้ออกมาวิเคราะห์ โดยวิธีทั้งสองที่ต้องการจะเปรียบเทียบกัน ทำการตรวจซ้ำ ๆ กันประมาณ 30 ครั้ง

ให้ x_k = ผลที่ได้จากการตรวจวิธีที่ 1

y_j = ผลที่ได้จากการตรวจวิธีที่ 2

s_i = SD ของแต่ละวิธี $i = 1, 2$

a = จำนวนของ s_i^2

$\bar{s}^2 = \sum_{i=1}^{n} f_i s_i^2 / \sum_{i=1}^{n} f_i$

f_i = จำนวนชั้นแห่งความเป็นอิสระ (degree of freedom) ของแต่ละวิธี

สูตร $SD = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}$

$M = 2.3026 \left\{ (\sum_{i=1}^{n} f_i) \log \bar{s}^2 - \sum_{i=1}^{n} f_i \log s_i^2 \right\}$

$C = 1 + \frac{1}{3(a - 1)} \left\{ \sum_{i=1}^{n} \frac{1}{f_i} - \frac{1}{\sum_{i=1}^{n} f_i} \right\}$

$X^2 = M / C$

สมมติฐาน $s_1^2 = s_2^2$ ที่ระดับความเชื่อมั่น 95 %

ในกรณีนี้สามารถคำนวณเปรียบเทียบกันได้ แม้ $k \neq j$ (จำนวน replicate ไม่เท่ากัน) ถ้าความแปรปรวนของวิธีตรวจทั้ง 2 นั้น มีค่าเท่ากัน(นัยสำคัญ 5%) ก็สรุปได้ว่าวิธีการตรวจทั้ง 2 วิธี นั้น ให้ผลไม่แตกต่างกัน

สัญญลักษณ์ที่ใช้ในโปรแกรมคอมพิวเตอร์ (โปรแกรมที่ 3 ผนวก ง.)

จำนวนข้อมูลของวิธี A	= N
จำนวนข้อมูลของวิธี B	= M
SD. ของวิธี A	= SDA
SD. ของวิธี B	= SDB
Chi-square	= CHISQ
ข้อมูลที่ให้เครื่องหมายอ่านครั้งละ 10 ค่า	= DATA (I)

การจัดลำดับข้อมูล

บัตรที่	สดมภ์ที่	เขตพิกัด	ตัวแปร
1-3	1-80	10F8.2	DATA (I) จากวิธี A
4	1-80	F8.2	77777777
5-7	1-80	10F8.2	DATA (I) จากวิธี B
8	1-80	F8.2	88888888

มิติ (Dimension)

โปรแกรมนี้สามารถรับข้อมูลจากวิธีตรวจวิเคราะห์ทั้ง 2 วิธี ไม่เกินวิธีละ 30 ค่า (A(30) และ B(30)) ซึ่งเพียงพอที่จะใช้ในการสรุปผลการคำนวณ และโปรแกรมนี้สามารถใช้งานได้โดยไม่ต้องแก้ไข ถึงแม้ข้อมูลที่ได้จากวิธีตรวจวิเคราะห์ทั้ง 2 วิธี มีจำนวนไม่เท่ากัน ทั้งนี้เพราะมีการเรียกใช้โปรแกรมย่อย (Subroutine) สองครั้ง ทำให้มีการคำนวณในลักษณะเดียวกันสองครั้ง และมีอิสระต่อกัน (ค่า N จึงไม่จำเป็นต้องเท่ากับ M)

3.2.2 การทดสอบ Recovery ของวิธีการตรวจวิเคราะห์

ถ้ามีวิธีตรวจหาปริมาณน้ำตาลในเลือดอยู่ 2 วิธี และทั้ง 2 วิธีนี้ ตรวจผู้ป่วยรายเดียวกันได้ 100 mg % เท่ากัน เมื่อเติม Primary standard คือ glucose ที่มีความเข้มข้น 100 mg % ลงไปอีก แล้วนำไปตรวจด้วยวิธีทั้ง 2 อีก ในครั้งนี้ถ้าวิธีทั้ง 2 ดีที่สุดเหมือน ๆ กัน ก็จะวัดได้ 200 mg % ซึ่งถือว่ามี recovery เป็น 100 % แต่ในทางปฏิบัติค่า % recovery ของแต่ละวิธีจะไม่เท่ากัน จึงจำเป็นต้องมีการเปรียบเทียบกันเพื่อเลือกเอาวิธีที่ดีที่สุดไว้ใช้ปฏิบัติต่อไป

วิธีทดลอง

นำเลือดของผู้ป่วย (unknown sample) ที่ได้ทำการตรวจเรียบร้อยแล้วว่า มีสาร A อยู่เป็นจำนวนเท่าใดมาประมาณ 30 ราย (ทำต่อจากหัวข้อ 3.2.1) เติม primary standard ของสาร A ที่ทราบค่าความเข้มข้นแน่นอนลงไป แล้วจึงนำไปตรวจวิเคราะห์ด้วยวิธีทั้ง 2 นั้น อีกครั้ง แล้วจึงนำค่าที่ได้มาเปรียบเทียบกัน

สูตร $\%\ \text{Recovery} = \dfrac{y_i}{x_i + c} \times 100$

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}$$

$$F = \frac{SD_1^2}{SD_2^2}$$

สมมติฐาน $SD_1 = SD_2$

เมื่อ C = ความเข้มข้นของ primary standard ที่เติมลงไป

x_i = ค่าที่วัดได้ก่อนเติม primary standard

y_i = ค่าที่วัดได้หลังเติม primary standard

S_1 = S.D. ของวิธีตรวจที่ 1

S_2 = S.D. ของวิธีตรวจที่ 2

สัญญลักษณ์ที่ใช้ในโปรแกรมคอมพิวเตอร์ (โปรแกรมที่ 4 ผนวก ง.)

ความเข้มข้นของ primary standard	=	CONC
ค่าที่วัดได้ก่อนเติม primary standard	=	AMED
ค่าที่วัดได้หลังเติม primary standard	=	AADD
F - value	=	F
ชื่อวิธี	=	NAME (I)
จำนวนข้อมูลที่จะให้เครื่องอ่าน	=	ITEST

การจัดลำดับข้อมูล

บัตรที่	สดมภ์ที่	เขตพิกัด	ตัวแปร		
1	1-2	I2	IIEST	=	30
	3-10	F8.2	CONC	=	100
2-4	1-8	F8.2	AMED	จากวิธี	A
	9-16	F8.2	AADD		
5-7	1-8	F8.2	AMED	จากวิธี	B
	9-16	F8.2	AADD		

มิติ (Dimension)

ในโปรแกรมนี้มีข้อจำกัดอยู่เล็กน้อยคือ ข้อมูลที่ได้จากวิธีตรวจวิเคราะห์ทั้งสองจะต้องมีค่าเท่ากัน (ITEST = 30) ทั้งนี้เพราะค่าดัชนีในกระทงความ DO ทุกอันมีพิสัยเท่ากับ ITEST ทั้งสิ้น ซึ่งในทางปฏิบัติก็ไม่น่ามีปัญหาเกิดขึ้น เมื่อป้อนข้อมูลจากวิธีวิเคราะห์ A และ B ชุดละ 30 ค่า แล้ว ผลการคำนวณ % recovery ก็จะอยู่ในตัวแปรหมวดคือ RECOA(30) และ RECOB(30) ทำให้มิติของโปรแกรมนี้ต้องบังคับใช้เท่ากับ 30 เพียงค่าเดียวเท่านั้น