

SO SÁNH SỰ TƯƠNG ĐỒNG MỘT SỐ XÉT NGHIỆM HÓA SINH TRÊN HAI MÁY ALINITY VÀ ATELLICA SOLUTIONS

Phan Thị Thanh Hải¹, Đặng Thị Ngọc Dung²✉.

¹Bệnh viện TWQĐ 108

²Trường Đại học Y Hà Nội

Đánh giá sự tương đồng kết quả xét nghiệm từ hai hay nhiều thiết bị thực hiện cùng xét nghiệm là công việc phải thực hiện để đảm bảo chất lượng xét nghiệm, giúp cung cấp các kết quả xét nghiệm chính xác, tin cậy cho chẩn đoán và điều trị. Khi một phòng xét nghiệm có 2 hay nhiều hơn các thiết bị xét nghiệm khác nhau, cần phải đánh giá sự tương đồng kết quả thu được giữa các thiết bị này giúp cho các bác sỹ lâm sàng chẩn đoán và theo dõi kết quả điều trị một cách chính xác và hiệu quả. Đề tài được tiến hành với mục tiêu so sánh kết quả một số xét nghiệm hóa sinh trên 2 máy Alinity và Atellica Solutions. Nghiên cứu áp dụng hướng dẫn EP09-A3 của CLSI, 40 mẫu bệnh phẩm có nồng độ trải khắp khoảng tuyến tính của phương pháp được sử dụng để so sánh phương pháp của các xét nghiệm glucose, creatinin, ALT, canxi và HDL giữa 2 máy. Nghiên cứu sử dụng vật liệu là mẫu bệnh nhân để tiến hành thực nghiệm so sánh phương pháp theo hướng dẫn EP09-A3 của CLSI. Theo phân tích hồi quy tuyến tính Passing-Bablok, các xét nghiệm glucose, creatinin, ALT, canxi và HDL trên 2 máy có tương quan khá chặt chẽ với $R > 0,98$ và $p < 0,001$. Kiểm định tuyến tính theo test cusum của tất cả xét nghiệm đều được chấp nhận với $p > 0,05$. Độ dốc của các xét nghiệm là từ 0,919 đến 1,164 và khoảng tin cậy của độ dốc của tất cả xét nghiệm đều không bao gồm 1. Khoảng tin cậy của giao điểm của tất cả xét nghiệm đều không bao gồm 0, trừ xét nghiệm glucose. Trung bình sự khác biệt (%) và 95%CI của sự khác biệt (%) của các xét nghiệm glucose, HDL, creatinin, ALT, canxi lần lượt là -5,5% (-6,0% - (-4,9%)); -3,9% (-7,2% - (-0,6%)), 5,8% (3,5%-8,2%), -10,7% (-12,3%-(-8,3%)) và 1,3% (0,1%-2,6%). Các xét nghiệm creatinin, canxi, ALT và HDL có sự khác biệt giữa hai máy với độ lệch được chấp nhận. Xét nghiệm glucose có sự khác biệt giữa hai máy với độ lệch không được chấp nhận.

Từ khóa: So sánh phương pháp, EP09-A3

I. ĐẶT VẤN ĐỀ

Xét nghiệm đóng vai trò quan trọng trong chẩn đoán, điều trị, tiên lượng và dự phòng bệnh, do vậy đòi hỏi kết quả xét nghiệm phải chính xác, tin cậy.¹

Kết quả xét nghiệm cận lâm sàng đóng một vai trò rất quan trọng, chúng đưa ra những kết quả phân tích có ích cho chẩn đoán, theo dõi quá trình điều trị, tiên lượng và dự phòng bệnh tật. Do đó, đòi hỏi các kết quả xét nghiệm phải tin cậy, chính xác.¹

Để đáp ứng với yêu cầu ngày càng cao của bác sỹ lâm sàng, các phòng xét nghiệm phải không ngừng đổi mới cải tiến để, nâng cao chất lượng. Thực tế các phòng xét nghiệm ở Việt Nam đang phải đối mặt với khó khăn là sử dụng nhiều hệ thống xét nghiệm từ các nhà sản xuất khác nhau để thực hiện cùng loại xét nghiệm. Việc triển khai thêm các hệ thống xét nghiệm hiện đại với phương pháp xét nghiệm tiên tiến là xu hướng tất yếu trong các phòng xét nghiệm y học. Thực tế cho thấy, một phòng xét nghiệm thường sử dụng nhiều hơn một hệ thống thiết bị chẩn đoán để xét nghiệm cùng một chỉ số cận lâm sàng, gây khó khăn cho việc kiểm soát chất lượng xét nghiệm và truy xuất nguồn gốc

Tác giả liên hệ: Đặng Thị Ngọc Dung

Trường Đại học Y Hà Nội

Email: dzunghmu@gmail.com

Ngày nhận: 22/10/2020

Ngày được chấp nhận: 01/11/2020

(traceability).

So sánh phương pháp được dùng để đánh giá tính tương quan, ước lượng độ chệch lệch giữa hai phương pháp. Nếu độ chệch lệch (bias) giữa hai phương pháp đạt yêu cầu nhỏ hơn tiêu chuẩn cho phép thì hai phương pháp này có thể hoán đổi cho nhau mà không gây ảnh hưởng đến việc phiên giải kết quả xét nghiệm của bệnh nhân, do đó không làm ảnh hưởng đến quá trình điều trị và theo dõi bệnh nhân.^{2,3,4} Thực nghiệm So sánh phương pháp được tiến hành khi áp dụng một phương pháp, hay một thiết bị mới hoặc một lô hóa chất mới, khi phòng xét nghiệm sử dụng hệ thống đã phân tích, hoặc khi sử dụng một dịch vụ của phòng xét nghiệm khác nhằm đảm bảo rằng kết quả xét nghiệm từ những sự thay đổi đó không làm sai lệch việc phiên giải kết quả xét nghiệm của bệnh nhân.⁵

Abbott và Siemens là hai hệ thống máy xét nghiệm đã được sử dụng rộng rãi trên toàn thế giới và đã chứng minh được tính ưu việt của nó. Hệ thống Alinity đã được đưa vào sử dụng tại phòng xét nghiệm A5, khoa xét nghiệm, Bệnh viện Đại học Y Hà Nội. Gần đây, hệ thống Atellica Solution đang được triển khai để đưa vào sử dụng tại phòng xét nghiệm. Để đảm bảo kết quả được tin cậy, chính xác, đồng thời để đưa ra các khuyến cáo trong việc chỉ định cùng một chỉ số xét nghiệm trên hai hệ thống máy khác nhau mà vẫn đảm bảo được việc sử dụng kết quả xét nghiệm trong theo dõi điều trị bệnh nhân một cách chính xác, có hiệu quả, chúng tôi tiến hành đề tài với mục tiêu: So sánh phương pháp xét nghiệm một số xét nghiệm trên 2 hệ thống máy hóa sinh miễn dịch Alinity và Atellica Solutions.

II. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP

1. Chất liệu nghiên cứu

Chất liệu nghiên cứu: Huyết thanh thừa của bệnh nhân, thu thập từ các mẫu máu sau khi đã được phân tích tại phòng xét nghiệm A5, khoa xét nghiệm, Bệnh viện Đại học Y Hà Nội.

Thiết bị và hóa chất sử dụng: Hệ thống hóa sinh miễn dịch Alinity của hãng Abbott và Atellica của hãng Siemens tương ứng là thiết bị tham chiếu và thiết bị được so sánh, là “thiết bị đóng” hoàn toàn, sử dụng chất chuẩn và hóa chất đúng nhà sản xuất, kiểm soát chất lượng bằng thang Sigma và các luật của Westgard, tham gia ngoại kiểm định kỳ 1 tháng/1 lần, sử dụng giá trị tham chiếu của nhà sản xuất. Cả 2 thiết bị đã được xác nhận giá trị sử dụng về độ chụm, độ đúng, khoảng tuyến tính, giới hạn đo mẫu trắng, giới hạn phát hiện, giới hạn định lượng phù hợp với công bố của nhà sản xuất trước khi đưa vào sử dụng

Địa điểm, thời gian: Phòng xét nghiệm A5, khoa xét nghiệm, Bệnh viện Đại học Y Hà Nội từ tháng 07/2019-11/2020.

2. Phương pháp

Nghiên cứu thực nghiệm trong phòng xét nghiệm

Thực nghiệm so sánh phương pháp tiến hành theo hướng dẫn EP09-A3 của CLSI. Các xét nghiệm được tiến hành so sánh bao gồm glucose, creatinin, HDL, ALT và canxi. Thiết kế nghiên cứu chạy 40 mẫu bệnh phẩm trong vòng 5 ngày, mỗi ngày chạy 8 mẫu, chạy lặp lại 2 lần đối với mỗi mẫu trên từng thiết bị. Phân tích 40 mẫu bệnh nhân, được lựa chọn các mức nồng độ theo các mức quyết định lâm sàng, phủ hết khoảng đo. Tiến hành phân tích mỗi mẫu 2 lần trên cùng một thiết bị, chạy trong khoảng thời gian 5 ngày. Các mẫu được phân tích trong vòng 2 giờ trên 2 máy. Dữ liệu phương pháp của các xét nghiệm nghiên cứu được trình bày ở bảng 1.

Bảng 1. Dữ liệu phương pháp của các xét nghiệm nghiên cứu

Chỉ số	Nội dung	Alinity	Atellica
Glucose (mmol/l)	Nguyên lý	Hexokinase	Hexokinase
	Tham chiếu chuẩn	- SRM 965/ NIST - ID-LC/MS	- SRM 965a/NIST
	Khoảng đo	0,28-44,4	0,2-38,9
	Giá trị tham chiếu	4,1-5,6	4,1-5,9
Creatinin (μ mol/l)	Nguyên lý	Jaffe	Jaffe
	Tham chiếu chuẩn	- SRM 967/ NIST - ID-LC/MS	- SRM 967/NIST - IDMS
	Khoảng đo	17,7-3270,8	9-2652
	Giá trị tham chiếu	Nam: 63,6-110,5; Nữ: 50,4-98,1	Nam: 53-97; Nữ: 44-71
HDL (mmol/l)	Nguyên lý	So màu	So màu
	Tham chiếu chuẩn	- CDC - HDL-C	- NCEP
	Khoảng đo	0,13-4,66	0,52-3,34
	Giá trị tham chiếu	Có nguy cơ tim mạch: < 1,04 Không có nguy cơ tim mạch: > 1,55	Có nguy cơ tim mạch: < 1,04 Không có nguy cơ tim mạch: > 1,55
ALT (U/l)	Nguyên lý	Động học enzyme (NADH không pyridoxal phosphate)	Động học enzyme (NADH không pyridoxal phosphate)
	Tham chiếu chuẩn	-Molar extinction factor -NADH Molar extinction factor	- SRM 454 - IFCC
	Khoảng đo	5-3899	7-1100
	Giá trị tham chiếu	0-55	10-49
Canxi (mmol/l)	Nguyên lý	Asenazo III	Tạo phức với sắt
	Tham chiếu chuẩn	SRM 956/ NIST ID-ICP-MS	SRM915 và SRM 909b/NIST. AAS/ NIST
	Khoảng đo	0,5-6,0	0,25-7,5
	Giá trị tham chiếu	2,1-2,55	2,08-2,65

CDC = Centers for Disease Control and Prevention, IDMS - Isotope Dilution Mass Spectrometry

ID-ICP-MS - Isotope Dilution Inductively Coupled Plasma Mass Spectrometry

ID-LC/MS - Isotope Dilution-Liquid Chromatography Mass Spectrometry

NIST - National Institute of Standards and Technology, SRM - Standard Reference Materials

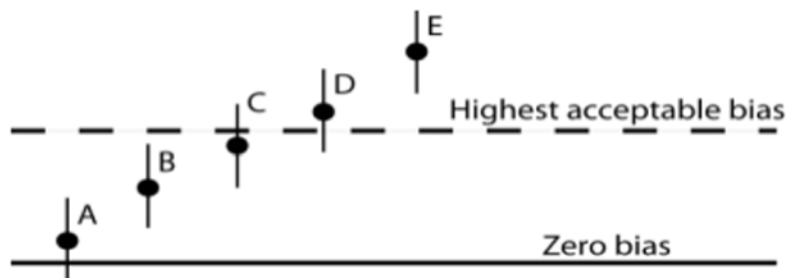
NCEP - The National Energy Customer Framework?

IFCC - The International Federation of Clinical Chemistry and Laboratory Medicine

AAS - Atomic Absorption Spectrometric

Sử dụng phần mềm Medcalc 19.2 để xử lý số liệu. Kiểm tra phân phối chuẩn bằng giả thuyết Kolmogorov-Smirnov và Shapiro-Wilk, chấp nhận giả thuyết phân phối chuẩn với $p > 0,05$. Phân tích hồi quy Passing - bablok xác định độ dốc, giao điểm và khoảng tin cậy của chúng, đưa ra phương trình hồi quy $y = a + bx$ giữa 2 phương pháp. Vẽ biểu đồ khác biệt Bland - Altman. Phiên giải kết quả

độ lệch ước tính với 5 trường hợp A, B, C, D và E theo EP09-A3, với tiêu chuẩn chấp nhận là độ lệch cho phép của Ricos.⁶ Trong đó trường hợp A, B, C, D là bias giữa 2 phương pháp được chấp nhận. Trường hợp E không được chấp nhận.²



Hình 1. Các khả năng xảy ra với độ lệch ước tính

Bảng 2. Phiên giải kết quả độ lệch

Trường hợp	Kết quả	Kết luận
A	95% CI của khác biệt trung bình có chứa giá trị 0.	Không có sự khác biệt đáng kể giữa hai phương pháp đo.
B	Khác biệt cho phép có chứa 95% CI của khác biệt trung bình.	Sự khác biệt giữa hai phương pháp và đạt tiêu chuẩn chấp nhận độ lệch với 95% CI
C	Khác biệt cho phép có chứa khác biệt trung bình nhưng không chứa 95%CI. Khác biệt trung bình nằm trong giới hạn độ khác biệt cho phép.	Có thể chấp nhận độ khác biệt trung bình nhưng không đảm bảo độ tin cậy 95% và sẽ có một số kết quả khác biệt đáng kể giữa hai phương pháp.
D	Khác biệt cho phép không chứa khác biệt trung bình nhưng có chứa 95% CI. Khác biệt trung bình nằm ngoài giới hạn tiêu chí khác biệt cho phép.	Khác biệt nằm ngoài tiêu chí chấp nhận nhưng vì 95% CI có bao gồm khác biệt cho phép nên vẫn có thể kết luận sự khác biệt giữa hai phương pháp là chấp nhận được nhưng mức độ tin cậy sẽ nhỏ hơn tình huống C.
E	Khác biệt trung bình và 95% CI của nó nằm ngoài khoảng chấp nhận của khác biệt cho phép.	Khác biệt không được chấp nhận. Phương pháp đo mới không thể thay thế phương pháp cũ. Nếu áp dụng, phương pháp mới cần xây dựng lại khoảng tham chiếu và xem xét nồng độ quyết định lâm sàng.

95% CI: 95% khoảng tin cậy

3. Đạo đức nghiên cứu

Việc tiến hành nghiên cứu được thực hiện theo đúng mục tiêu nghiên cứu. Các thông tin thu thập chỉ phục vụ cho mục đích nghiên cứu.

III. KẾT QUẢ

Kết quả so sánh tương quan bằng phương pháp hồi quy Passing- Pablok thể hiện ở bảng 3. Hệ số tương quan của các xét nghiệm từ 0,980-0,999 đều lớn hơn 0,9 với $p < 0,0001$. Kết quả kiểm tra giả thuyết tuyến tính theo Cusum test đều được chấp nhận với $p > 0,05$. 95%CI của độ dốc của tất cả xét nghiệm đều không bao gồm 1. 95% CI của giao điểm của tất cả xét nghiệm đều không bao gồm 0 trừ xét nghiệm Glucose. Nồng độ mẫu glucose, creatinine, canxi, HDL, ALT được lựa chọn phân bố đều trên khoảng đo (Hình 2).

Kết quả so sánh sự khác biệt của mẫu bệnh nhân trên 2 máy Atellica và Alinity bằng phương

pháp đồ thị khác biệt Bland-Altman được trình bày ở bảng 4 và hình 3. Khác biệt cho phép có chứa 95% CI của khác biệt trung bình (trường hợp B) gồm có các xét nghiệm: ALT và HDL. Khác biệt cho phép có chứa khác biệt trung bình nhưng không chứa 95%CI, khác biệt trung bình nằm trong giới hạn độ khác biệt cho phép (trường hợp C) có xét nghiệm creatinin. Khác biệt cho phép không chứa khác biệt trung bình nhưng có chứa 95%CI của khác biệt trung bình (trường hợp D) có xét nghiệm canxi. Khác biệt trung bình và 95%CI của nó nằm ngoài khoảng chấp nhận của khác biệt cho phép (trường hợp E) có xét nghiệm Glucose.

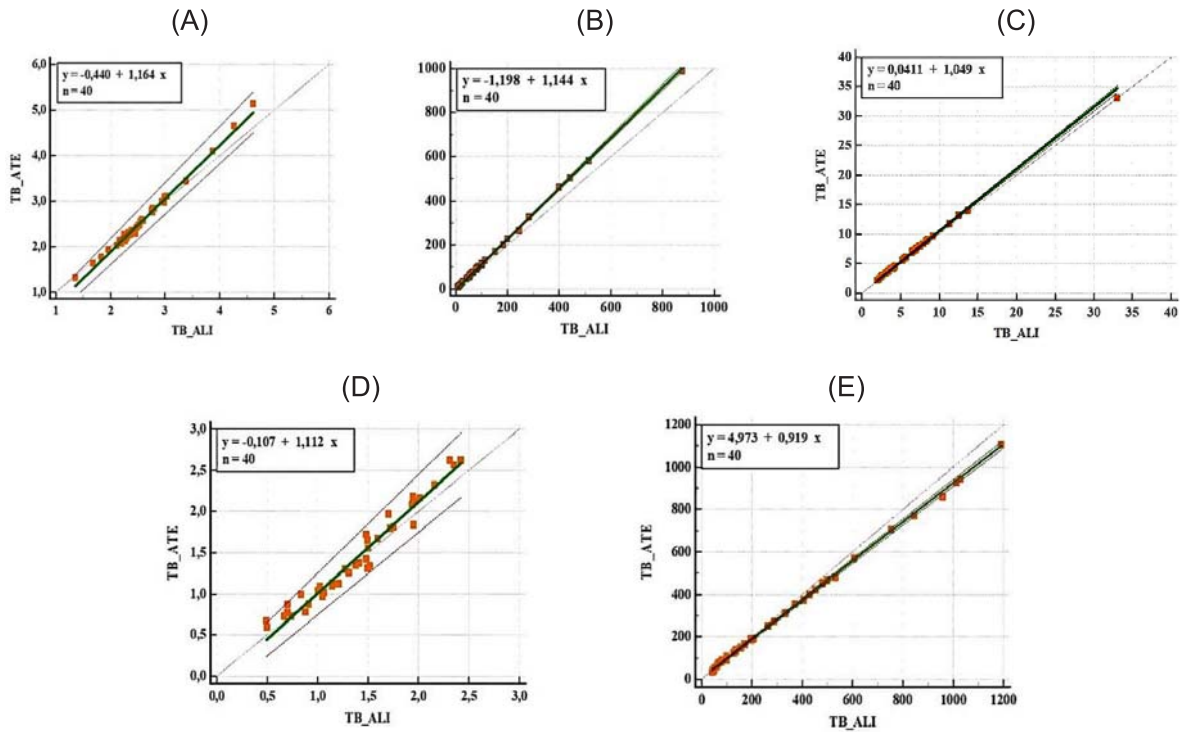
Bảng 3. Kết quả phân tích tương quan bằng phương pháp hồi quy Passing- Pablok

Xét nghiệm	Độ dốc	95% CI của độ dốc	Giao điểm	95% CI của giao điểm	Tương quan		Cusum test
					R	p	
Glu	1,049	1,031-1,062	0,041	-0,043-0,117	0,999	< 0,0001	P = 0,3
Cre	0,919	0,907-0,929	4,973	0,760-8,774	0,999	< 0,0001	P = 0,53
ALT	1,144	1,131-1,161	-1,198	-1,938-(-0,048)	0,998	< 0,0001	P = 0,3
Canxi	1,164	1,100-1,228	-0,440	-0,616-(-0,275)	0,986	< 0,0001	P = 0,53
HDL	1,112	1,002-1,201	-0,107	-0,262-0,046	0,980	< 0,0001	P = 0,07

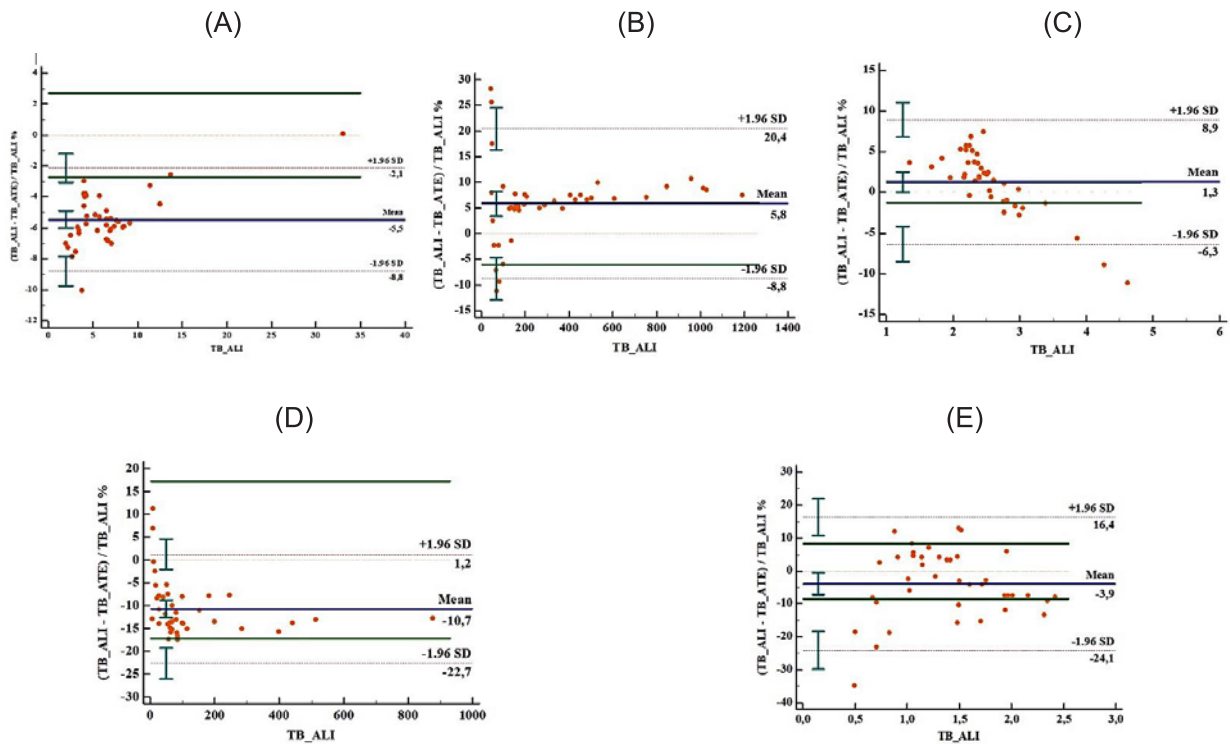
Bảng 4. Kết quả so sánh sự khác biệt của mẫu bệnh nhân trên 2 máy Atellica và Alinity bằng phương pháp đồ thị khác biệt Bland-Altman

Xét nghiệm	Trung bình khác biệt (%)	95% CI của khác biệt (%)	* Bias cho phép (%)	Đánh giá (Trường hợp)
Glucose	-5,5	-4,9 - (-6,0)	2,7	E
Creatinin	5,8	3,5 - 8,2	6,0	B
ALT	-10,7	-12,3 - (-8,3)	17,2	B
Canxi	1,3	-0,1 - (-2,6)	1,2	D
HDL	-3,9	- 7,2 - (-0,6)	8,4	B

*Lấy theo nguồn của RICOS 6



Hình 2. Biểu đồ Passing- Pablok so sánh kết quả canxi (A), ALT (B), glucose (C), HDL (D) và creatinin (E) giữa 2 máy Alinity và Atellica.



Hình 3. Biểu đồ khác biệt trong so sánh kết quả glucose (A), creatinin (B), canxi (C), ALT (D) và HDL (E) của 40 mẫu bệnh phẩm trên 2 máy Alinity và Atellica bằng đồ thị Bland- Altman.

IV. BÀN LUẬN

So sánh phương pháp được tiến hành bởi phòng xét nghiệm để đánh giá sự tương đồng giữa hai phương pháp. Sự tương đồng này sẽ làm giảm độ lệch (Bias) phân tích và cho phép sử dụng so sánh các kết quả xét nghiệm trên hai phương pháp khác nhau mà vẫn đảm bảo tính chính xác giúp cho các bác sỹ lâm sàng chẩn đoán và theo dõi kết quả điều trị một cách chính xác và hiệu quả hơn.

Mục tiêu của nghiên cứu này của chúng tôi được thực hiện nhằm mục đích so sánh kết quả một số chỉ số xét nghiệm hóa sinh được thực hiện trên 2 hệ thống máy Alinity (của hãng Abbott) và Atellica (của hãng Siemens) tại phòng xét nghiệm A5, khoa xét nghiệm Bệnh viện Đại học Y Hà Nội để đánh giá mức độ tương đồng của hai hệ thống và có giải pháp cho các xét nghiệm có độ lệch vượt giới hạn cho phép.

Về mặt phương pháp, Nghiên cứu được thực hiện theo hướng dẫn EP09-A3 của CLSI, dựa trên việc phân tích kết quả mẫu bệnh nhân được lựa chọn và bệnh phẩm được tiến hành phân tích trên 2 hệ thống máy. Các xét nghiệm được tiến hành so sánh bao gồm glucose, creatinin, HDL, ALT và canxi. Thiết kế nghiên cứu chạy 40 mẫu bệnh phẩm trong vòng 5 ngày, mỗi ngày chạy 8 mẫu, chạy lặp lại 2 lần đối với mỗi mẫu trên từng thiết bị.

Trên thực tế có nhiều đề xuất cho thực nghiệm so sánh giữa hai phương pháp, trong một số trường hợp mẫu QC có thể được sử dụng xác minh khả năng so sánh kết quả giữa các thiết bị và phương pháp khác nhau. Khi các thiết bị giống hệt nhau và sử dụng cùng một lô thuốc thử, thì mối quan hệ về kết quả QC nhiều khả năng sẽ rất giống với mối quan hệ đối với các mẫu bệnh phẩm. Tuy nhiên, khi nghiên cứu so sánh phương pháp được thực hiện trên hai hệ thống khác nhau, thuốc thử được sử dụng

khác nhau thì khả năng phản ánh mối quan hệ về kết quả khi phân tích mẫu bệnh nhân thông qua kết quả của mẫu QC là hạn chế và có thể dẫn đến những kết luận sai lầm về khả năng so sánh của mẫu bệnh nhân. Các mẫu bệnh nhân thu thập từ chính các mẫu bệnh phẩm được sử dụng cho phân tích tại phòng xét nghiệm là nguồn vật liệu lý tưởng cho thực nghiệm so sánh phương pháp, vì chúng là các mẫu dự kiến được phân tích bằng hệ thống đo lường.⁷ trong đó có phương pháp đề xuất so sánh độ tương đồng của kết quả xét nghiệm hai thiết bị sử dụng các mẫu QC đơn giản, tiện lợi, dễ sử dụng. Khi hai thiết bị xét nghiệm tương tự sử dụng cùng lô thuốc thử, khả năng mối quan hệ giữa các kết quả QC của hai máy là tương tự như mối quan hệ giữa các mẫu bệnh phẩm giữa 2 máy. Tuy nhiên, khi các thiết bị xét nghiệm khác nhau, đặc biệt là từ các nhà sản xuất khác nhau, rất có khả năng kết quả phân tích các mẫu QC bởi thiết bị khác nhau không có cùng mối quan hệ như kết quả của các mẫu bệnh phẩm và vì vậy mà kết luận đưa ra về độ tương đồng của kết quả xét nghiệm của các mẫu bệnh phẩm trên 2 máy có thể là không chính xác.⁴ Hơn nữa, việc sử dụng mẫu QC sẽ không đánh giá sự tương đồng giữa 2 thiết bị trên toàn hệ thống đo, không cho biết sai số hệ thống giữa 2 máy nếu có là sai số tỷ lệ hay sai số hằng định. Vì vậy, các mẫu bệnh phẩm là lựa chọn tối ưu khi sử dụng để so sánh độ lệch giữa hai phương pháp trên các hệ thống máy xét nghiệm khác nhau.⁷

Thông thường, các thực nghiệm so sánh phương pháp chỉ thực hiện phép đo 1 lần đối với mỗi mẫu trên từng thiết bị. Tuy nhiên, việc chạy lặp lại nhiều lần các mẫu trên từng hệ thống, sẽ cung cấp nhiều điểm dữ liệu so sánh hơn, sẽ giúp kiểm tra tính hợp lệ của phép đo trên từng thiết bị và giúp tăng khả năng phát hiện các lỗi do trộn mẫu (mix-ups), lỗi dịch

chuyển (transposition) và các lỗi khác. Một hoặc hai lỗi như thế cũng gây nên ảnh hưởng lớn đến kết quả cuối cùng của nghiên cứu. Những phát hiện lỗi này giúp phòng xét nghiệm giải thích sự khác biệt trong kết quả nghiên cứu có phải là giá trị ngoại lai hay không.⁸ Ít nhất, các lỗi này cũng giúp phát hiện được các kết quả trái ngược nhau có phải là giá trị ngoại lai hay không.⁹

Các hướng dẫn hiện thời về việc phân tích kết quả của thực nghiệm so sánh phương pháp nói chung và so sánh thiết bị nói riêng khuyến cáo việc sử dụng kết hợp phân tích thống kê và đồ thị bao gồm biểu đồ phân tán (scatter plot) và biểu đồ khác biệt (difference plot) kết hợp với tính toán độ khác biệt trung bình và khoảng tin cậy của nó.²

Biểu đồ phân tán mô tả sự thay đổi theo cặp giá trị đo lường thông qua các giá trị đo. Mỗi một cặp của phép đo lường được biểu diễn bằng 1 điểm ứng với 1 giá trị trên trục hoành x (thường là phương pháp tham chiếu) và một giá trị trên trục tung y (thường là giá trị của phương pháp so sánh). Trong nghiên cứu này, trục x và trục y tương ứng là giá trị trung bình của phép đo lặp lại của máy Alinity và máy Atellica.

Thu thập và Quan sát trực quan các giá trị được biểu thị dưới dạng biểu đồ ở hình 3 cho thấy các nồng độ mẫu của được sử dụng trong thực nghiệm so sánh phương pháp đối với xét nghiệm glucose, creatinin, ALT, canxi và HDL được phân phối tương đối đều trên khoảng đo. Không có giá trị ngoại lai nào được quan sát trên các biểu đồ phân tán.

Biểu đồ khác biệt biểu diễn kết quả của thực nghiệm so sánh phương pháp với trục hoành x là nồng độ đo lường và sự khác biệt giữa 2 phương pháp được biểu diễn trên trục tung y. Bland-Altman là một ví dụ của biểu đồ khác biệt. Những biểu đồ như thế này có thể kiểm tra một cách trực quan để xác định mối liên hệ cơ bản

của các biến. Người sử dụng có thể lựa chọn 4 dạng khác nhau của biểu đồ khác biệt dựa vào 2 yếu tố. Yếu tố đầu tiên là xác định liệu mong muốn của phòng xét nghiệm là xem phương pháp so sánh như là một phương pháp tham chiếu chuẩn hay là xác định giá trị của trung bình kết quả của 2 phương pháp là cách tốt nhất để ước tính giá trị thực của 1 mẫu. Trong trường hợp đầu, trục hoành của đồ thị sẽ là kết quả của phương pháp so sánh. Trường hợp thứ 2, trục hoành của đồ thị là kết quả trung bình của 2 phương pháp.⁶

Trong phòng xét nghiệm lâm sàng có thể xem phương pháp đang sử dụng là phương pháp so sánh và nó được xem là một phương pháp tham chiếu để so sánh kết quả của phương pháp hiện tại với 1 phương pháp mới. Trong trường hợp này, trục hoành nên là kết quả xét nghiệm của phương pháp so sánh.²

Phân tích định lượng trong so sánh phương pháp gồm có: phân tích hồi quy tuyến tính Deming và Passing-Bablok; ước tính sự khác biệt trung bình hoặc sự khác biệt tại điểm quyết định lâm sàng. Kỹ thuật Deming và Passing-Bablok với giả thuyết 2 phương pháp có mối quan hệ tuyến tính để tìm ra mối tương quan tuyến tính giữa 2 phương pháp thông qua phương trình $y = ax + b$. Trong đó a được gọi là độ dốc, b gọi là giao điểm. Hai phương pháp sẽ được xem là không có sự khác biệt khi khoảng tin cậy của a bao gồm 1 và khoảng tin cậy của b bao gồm 0. Khi a gần tiến đến 1 và b gần tiến đến 0 thì phương trình $y = ax + b$ dần tiến đến $y = x$.^{7,8}

Phân tích kết quả nghiên cứu bằng hồi quy tuyến tính Passing-Bablok (được trình bày ở bảng 3) cho thấy các xét nghiệm glucose, creatinin, ALT, canxi và HDL trên 2 máy có tương quan khá chặt chẽ với $R > 0,98$ và $p < 0,001$. Kiểm định tuyến tính theo test cusum của tất cả xét nghiệm đều được chấp nhận với

$p > 0,05$. Giá trị độ dốc của các xét nghiệm là từ 0,919 đến 1,164 và khoảng tin cậy của độ dốc của tất cả xét nghiệm đều không bao gồm 1, cho thấy giữa 2 máy có sai số tỷ lệ. Khoảng tin cậy của giao điểm của xét nghiệm creatinin, ALT, canxi và HDL đều không bao gồm 0 chứng tỏ có sai số hằng định giữa 2 thiết bị. Đối với xét nghiệm glucose có khoảng tin cậy của giao điểm chứa 0, chứng tỏ không có sai số hằng định.

Theo CLSI, EP 09-A3, chỉ đưa ra duy nhất tiêu chí so sánh sự khác biệt trung bình giữa 2 phương pháp với độ lệch được chấp nhận và hướng dẫn phân giải kết quả so sánh độ lệch theo 5 trường hợp là A, B, C, D, E. Trong đó trường hợp E có độ lệch là không chấp nhận được.² Điều quan trọng là các phòng xét nghiệm phải tự thiết lập tiêu chuẩn chấp nhận là ở mức độ lệch tối ưu, độ lệch mong muốn hay độ lệch cho phép dựa vào đặc điểm, điều kiện của phòng xét nghiệm. Thông thường, các tiêu chí này là theo nhà sản xuất công bố hoặc theo biến thiên sinh học từ các nguồn tin cậy. Nghiên cứu này lấy bảng biến thiên sinh học của Ricos làm tiêu chí để đánh giá, với tiêu chí chấp nhận là độ lệch cho phép. Tiêu chí này có sẵn trên Westgard và được chấp nhận rộng rãi.⁶

Chúng tôi cũng chưa tìm được nghiên cứu nào so sánh phương pháp trên 2 thiết bị Alinity và Atellica. Tuy nhiên, độ lệch của các xét nghiệm giữa 2 thiết bị trong nghiên cứu của chúng tôi nhìn chung là tương đối lớn, trong đó có xét nghiệm glucose có độ lệch 5,47% là không được chấp nhận nếu so với độ lệch cho phép của Ricos.

Theo nghiên cứu của Sten Wesgard và cộng sự,¹⁰ đánh giá bias của 53 hệ thống Alinity cho thấy giá trị bias của các xét nghiệm glucose, creatinin, ALT, canxi và HDL tương ứng là 0,22%, 1,94%, 3,38%, 1,35% và 0,93% thấp

hơn nhiều so với bias giữa 2 hệ thống trong nghiên cứu của chúng tôi. Điều này được giải thích là do nghiên cứu của Sten Wesgard và cộng sự là nghiên cứu bias giữa các thiết bị của cùng một hãng, nên không bao gồm sai số do hóa chất, quá trình hiệu chuẩn, chất chuẩn và các thông số máy móc khác.

Sự khác biệt về kết quả của mẫu bệnh nhân giữa các thiết bị có thể là do khác nhau về nguyên lý xét nghiệm, quá trình hiệu chuẩn (calibration), lô chất chuẩn khác nhau, khác nhau trong “khả năng hoán đổi” của các chất chuẩn với phép đo khác nhau từ các nhà sản xuất, khác nhau về hóa chất, các thông số phân tích khác nhau - như là tỉ lệ pha loãng, thời gian ủ ấm giữa các thiết bị khi sử dụng cùng hóa chất hoặc các ảnh hưởng trước phân tích mẫu gồm cách xử lý mẫu giữa các loại hệ thống thiết bị khác nhau.⁷

Cụ thể, trong nghiên cứu của chúng tôi hai thiết bị được so sánh của 2 hãng khác nhau, có các đặc tính kỹ thuật khác nhau. Cùng với đó là sự khác nhau về nguyên lý xét nghiệm được quan sát thấy ở chỉ số canxi (nguyên lý của xét nghiệm canxi trên máy Alinity là sự tạo thành phức hợp màu với asenazo III, trong khi đó trên máy Atellica là tạo phức với sắt). Bên cạnh đó cũng có sự khác nhau về các giá trị mức nồng độ calibrator được sử dụng trên hai hệ thống. Những sự khác biệt này có thể dẫn đến những sự khác biệt về kết quả phân tích trên hai hệ thống được xem xét. Tuy nhiên, để đưa hai hệ thống này vào áp dụng thực tế và kết quả xét nghiệm từ hai hệ thống tạo giá trị tin tưởng như nhau, đảm bảo được tính tin cậy cho mục tiêu theo dõi và điều trị bệnh nhân thì những sự khác biệt này phải nằm trong giới hạn chấp nhận. Và thực tế, kết quả nghiên cứu này chỉ ra rằng có sự khác biệt về kết quả xét nghiệm creatinin, canxi, ALT và HDL khi thực hiện trên hai hệ thống, tuy nhiên sự khác biệt

này là chấp nhận được. Duy chỉ có xét nghiệm Glucose có sự khác biệt không được chấp nhận. Theo hướng dẫn của EP09-A3 thay vì kết luận phương pháp không được chấp nhận, thay vào đó, tập hợp các kết quả trên có thể thúc đẩy phòng thí nghiệm lâm sàng điều chỉnh các khoảng tham chiếu bằng cách sử dụng kết quả của nghiên cứu so sánh theo tài liệu CLSI EP28. Dữ liệu từ nghiên cứu so sánh phương pháp đã có có thể được sử dụng để đánh giá để xem liệu chúng có thể được sử dụng để cập nhật khoảng tham chiếu hiện có hay thực tế là cần một nghiên cứu mới.¹¹

V. KẾT LUẬN

Các xét nghiệm creatinin, canxi, ALT và HDL có sự khác biệt về kết quả mẫu bệnh nhân trên hai máy với độ lệch được chấp nhận. sự tương đồng về kết quả của mẫu bệnh nhân trên 2 hệ thống máy. Xét nghiệm glucose có sự khác biệt về kết quả mẫu bệnh nhân trên hai máy với độ lệch không được chấp nhận. cần có sự chuyển đổi giá trị tham khảo trên máy Atellica để đạt được sự tương đồng trong áp dụng lâm sàng.

Lời cảm ơn

Các tác giả xin cảm ơn Khoa xét nghiệm, Bệnh viện Đại học Y Hà Nội đã hỗ trợ kỹ thuật cho triển khai nghiên cứu này.

TÀI LIỆU THAM KHẢO

1. Nerenz RD, Pittman ME, Scott MG. Impact of Errors and Variability on Clinical Laboratory Test Interpretation. In: McManus LM, Mitchell RN, eds. *Pathobiology of Human Disease*. Academic Press; 2014:3222-3236. doi:10.1016/B978-0-12-386456-7.06303-6
2. Budd JR, Durham AP, Gwise TE, et al. *Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline*. Clinical Laboratory Standards Institute; 2013.
3. Panteghini M, Forest JC. Standardization in laboratory medicine: New challenges. *Clin Chim Acta Int J Clin Chem*. 2005;355:1-12. doi:10.1016/j.cccn.2004.12.003
4. Guzel O, Guner EI. ISO 15189 accreditation: Requirements for quality and competence of medical laboratories, experience of a laboratory I. *Clin Biochem*. 2009;42(4-5):274-278. doi:10.1016/j.clinbiochem.2008.09.011
5. Statistical analysis in method comparison studies part one. Accessed November 1, 2020. <https://acutecaretesting.org/en/articles/statistical-analysis-in-method-comparison-studies-part-one>
6. Desirable Biological Variation Database specifications - Westgard. Accessed October 9, 2020. <https://www.westgard.com/biodatabase1.htm>
7. EP31AIRE: Patient Results Within One Health Care System. Accessed October 28, 2020. <https://clsi.org/standards/products/method-evaluation/documents/ep31/>
8. The Comparison of Methods Experiment - Westgard. Accessed October 13, 2020. <https://www.westgard.com/lesson23.htm>
9. Assessing precision, bias and sigma-metrics of 53 measurands of the Alinity ci system - ScienceDirect. Accessed October 27, 2020. <https://www.sciencedirect.com/science/article/pii/S0009912017306616>
10. Westgard S, Petrides V, Schneider S, Berman M, Herzogenrath J, Orzechowski A. Assessing precision, bias and sigma-metrics of 53 measurands of the Alinity ci system. *Clin Biochem*. 2017;50(18):1216-1221. doi:10.1016/j.clinbiochem.2017.09.005
11. EP28A3C: Define and Verify Reference Intervals in Lab. Accessed November 1, 2020. <https://clsi.org/standards/products/method-evaluation/documents/ep28/>

Summary

ASSESSMENT OF THE SIMILARITY OF SOME BIOCHEMISTRY ASSAY RESULTS ON ALINITY AND ATELLICA SOLUTION

Evaluating similar test results from two or more devices performing the same test is a must-do to ensure test quality and help to providing accurate and reliable test results for diagnosis and treatment. Using CLSI's EP09-A3 guideline, this study analyzed 40 serum samples with concentration across the linear range of the method used in this study and compare the test results of glucose, creatinine, ALT, calcium, and HDL between two test machines, Alinity from Abbot and Atellica from Siemens. Passing-Bablok linear regression analysis showed that test results of glucose, creatinine, ALT, calcium and HDL on 2 machines were highly correlated ($R > 0.98$ and $p < 0.001$). Linearity tests according to cusum test for all assays were accepted with $p > 0.05$. The slope of the assays was between 0,919 and 1,164, and the confidence intervals of the slope of all assays excluded the value 1. Except for the glucose test, the confidence intervals of the intersection of all assays did not include 0. The mean difference (%) and 95% CI of the difference (%) of the glucose, HDL, creatinine, ALT, and calcium tests were -5.5% (-6.0% - (-4.9%)); -3.9% (-7.2% - (-0.6%)), 5.8% (3.5% - 8.2%), -10.7% (-12,3% - (-8.3%)) and 1.3% (0.1%-2.6%), respectively. Creatinine, calcium, ALT and HDL differed from the two machines with an accepted bias. Glucose assay result was different between the two machines and was not acceptable.

Keywords: method comparison, EP09-A3