

EVALUATION OF THE CONCORDANCE OF FT3, FT4, AND TSH ASSAYS ACROSS DIFFERENT IMMUNOASSAY ANALYZER SYSTEMS

Le Ha Phuong¹, Nguyen Thi Thanh Hai^{1,2*}

¹Hanoi Medical University - 1 Ton That Tung, Kim Lien Ward, Hanoi City, Vietnam

²National Hospital For Tropical Diseases - 78 Giai Phong, Kim Lien Ward, Hanoi City, Vietnam

Received: 08/09/2025

Revised: 25/09/2025; Accepted: 21/10/2025

ABSTRACT

Objective: To evaluate the comparability of FT3, FT4, and TSH assays using external quality assessment (RIQAS) data over 40 months (2020–2025) across analytical platforms Abbott, Beckman Coulter (BC), Siemens and Roche.

Methods: A cross-sectional study was conducted to assess comparability based on Biasgroup, CV%group, TEa_TT, TEa_BV, and HI.

Results: TSH assays showed comparability across all systems with HI ≤ 1 and biases falling within the TEa compared with Roche, Siemens demonstrating the highest level of agreement. FT4 and FT3 exhibited substantial negative bias (-10% to -30%) across systems.

Conclusion: TSH assays achieved inter-system comparability, whereas FT4 showed considerable negative bias of 15–30% and displayed consistent bias across all concentration ranges when comparing Abbott, BC, and Siemens with Roche. For FT3, bias decreased with increasing concentrations in Abbott, but increased in BC; Siemens was comparable to Roche at concentrations < 25 pmol/L.

Keywords: Harmonization, thyroid function tests, external quality assessment (EQA).

*Corresponding author

Email: nguyenthanhhai@hmu.edu.vn **Phone:** (+84) 702810229 **Https://doi.org/10.52163/yhc.v66iCD19.3559**

ĐÁNH GIÁ MỨC ĐỘ TƯƠNG ĐỒNG CỦA CÁC XÉT NGHIỆM FT3, FT4 VÀ TSH TRÊN CÁC HỆ THỐNG MÁY XÉT NGHIỆM MIỄN DỊCH

Lê Hà Phương¹, Nguyễn Thị Thanh Hải^{1,2*}

¹Trường Đại học Y Hà Nội - 1 Tôn Thất Tùng, P. Kim Liên, Tp. Hà Nội, Việt Nam

²Bệnh viện Bệnh Nhiệt đới Trung ương - 78 Giải Phóng, P. Kim Liên, Tp. Hà Nội, Việt Nam

Ngày nhận: 08/09/2025

Ngày sửa: 25/09/2025; Ngày đăng: 21/10/2025

TÓM TẮT

Mục tiêu: Đánh giá mức độ tương đồng của các xét nghiệm FT3, FT4 và TSH dựa trên dữ liệu ngoại kiểm (RIQAS) trong 40 tháng (2020–2025) giữa các nhóm máy xét nghiệm: Abbott, Beckman Coulter (BC), Siemens và Roche.

Phương pháp: Nghiên cứu mô tả cắt ngang đánh giá tính tương đồng dựa trên các chỉ số Bias/nhóm, CV%nhóm, TEa_TT, TEa_BV và HI.

Kết quả: TSH có sự tương đồng ở tất cả hệ thống máy với HI ≤ 1 và độ lệch nằm trong khoảng tổng sai số cho phép khi so sánh với hệ thống Roche, trong đó Siemens cho thấy mức độ tương đồng cao nhất. Xét nghiệm FT4 và FT3 có độ lệch lớn (10-30%) giữa các hệ thống máy.

Kết luận: Xét nghiệm TSH đạt tương đồng giữa các hệ thống; xét nghiệm FT4 có độ lệch âm khá lớn từ 15-30% và đồng nhất ở mọi nồng độ khi so sánh các hệ thống Abbott, BC, Siemens với Roche; xét nghiệm FT3 có xu hướng độ lệch giảm dần ở hệ thống Abbott và tăng dần ở hệ thống BC khi nồng độ tăng lên từ thấp đến cao; Siemens tương đồng với Roche ở dải nồng độ < 25 pmol/L.

Từ khoá: Độ tương đồng, xét nghiệm tuyến giáp, ngoại kiểm.

1. ĐẶT VẤN ĐỀ

Xét nghiệm sinh hoá nói chung, đặc biệt là xét nghiệm miễn dịch thường xuất hiện sự sai khác kết quả giữa các hệ thống máy do ảnh hưởng bởi vị trí kháng nguyên, ái lực kháng thể, nồng độ và yếu tố nền của mẫu, nguyên lý phương pháp...[1] Các phòng xét nghiệm sử dụng các phương pháp phân tích khác nhau, kết quả không thống nhất, gây khó khăn cho việc theo dõi điều trị, liên thông kết quả xét nghiệm, đặc biệt đối với các bệnh lý mãn tính như bệnh lý tuyến giáp.[2] Xét nghiệm đánh giá chức năng tuyến giáp đóng vai trò thiết yếu trong thực hành lâm sàng, bao gồm TSH, T3 và T4 tự do (FT3, FT4), T3 và T4 toàn phần (T3,T4), trong đó các chỉ số FT3, FT4 và TSH được ưu tiên sử dụng để chẩn đoán và theo dõi các bệnh lý tuyến giáp nhờ khả năng phản ánh chính xác tình trạng chức năng tuyến giáp sinh lý; xét nghiệm T3 và T4 ít được chỉ định hơn do nồng độ của chúng bị ảnh hưởng mạnh bởi các protein gắn như TBG hay albumin, dẫn đến biến thiên không đại diện cho tình trạng hormone tự do thực tế.[1] Các xét nghiệm này cũng gặp vấn đề khó khăn tương tự

như xét nghiệm miễn dịch nói chung do sự đa dạng về phương pháp phân tích, thiết bị và hoá chất giữa các phòng xét nghiệm có thể gây sai lệch kết quả, dẫn đến khó khăn trong diễn giải lâm sàng và ảnh hưởng đến quyết định điều trị. Do vậy, phân tích độ tương đồng của các xét nghiệm hormone tuyến giáp FT3, FT4 và TSH là ưu tiên cần thiết phục vụ nhu cầu chẩn đoán và theo dõi điều trị các bệnh lý tuyến giáp trên lâm sàng.

Trong những năm gần đây, các hướng dẫn của CLSI và IFCC đã nhấn mạnh tầm quan trọng của việc nghiên cứu và đánh giá độ tương đồng giữa các hệ thống xét nghiệm, nhằm hướng tới khả năng liên thông kết quả – đặc biệt đối với xét nghiệm hormon tuyến giáp, vốn là nhóm chỉ số cần được theo dõi lâu dài trên bệnh nhân và so sánh giữa nhiều cơ sở y tế khác nhau. [3,4] Các hướng dẫn này đều khuyến cáo sử dụng đa dạng loại mẫu để đánh giá chính xác sự tương đồng và khả năng liên thông. Tuy nhiên, mỗi loại mẫu đều có hạn chế nhất định: mẫu nội kiểm nội bộ chỉ phản ánh hiệu năng của một hệ thống riêng lẻ và không

*Tác giả liên hệ

Email: nguyenthanhhai@hmu.edu.vn Điện thoại: (+84) 702810229 <https://doi.org/10.52163/yhc.v66iCD19.3559>

đại diện cho các hệ thống khác; mẫu bệnh nhân là tiêu chuẩn sinh học lý tưởng nhưng khó triển khai trên diện rộng do hạn chế về số lượng, biến thiên sinh học cá thể và khó chuẩn hóa quy trình thu thập, bảo quản, phân phối. Trong bối cảnh đó, mẫu ngoại kiểm (EQA) có ưu thế nổi bật vì vừa phản ánh thực hành thường quy, vừa cung cấp dữ liệu khách quan từ nhiều hệ thống, do đó trở thành công cụ phù hợp để đánh giá mức độ tương đồng của các xét nghiệm miễn dịch.[5] Trong số các chương trình EQA hiện nay, Randox International Quality Assessment Scheme (RIQAS) là chương trình được áp dụng rộng rãi tại Việt Nam đồng thời đáp ứng được các yêu cầu để đánh giá sự tương đồng bao gồm mẫu EQA có tính tương đồng với mẫu bệnh nhân và bao phủ tất cả các khoảng nồng độ. Theo dữ liệu được cung cấp bởi RIQAS và thực tế các phòng xét nghiệm tại Việt Nam, Roche chiếm ưu thế về số lượng thiết bị tham gia ngoại kiểm so với các nhóm máy lớn khác bao gồm Abbott, Beckman Coulter (BC), Siemens. Do đó, Roche được lựa chọn làm nhóm thiết bị tham chiếu trong nghiên cứu này để tạo thuận lợi cho việc áp dụng kết quả nghiên cứu tại các cơ sở y tế.

Vì vậy, nghiên cứu này khai thác dữ liệu RIQAS giai đoạn 2020–2025 tại Việt Nam nhằm đánh giá mức độ tương đồng của FT3, FT4 và TSH giữa các hệ thống xét nghiệm, từ đó cung cấp bằng chứng khoa học hỗ trợ cho việc liên thông kết quả xét nghiệm tuyến giáp, tiết kiệm chi phí và góp phần nâng cao chất lượng chăm sóc sức khỏe người dân.

2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Thiết kế nghiên cứu: Nghiên cứu mô tả cắt ngang - hồi cứu.

2.2. Đối tượng nghiên cứu: Dữ liệu EQA của 3 xét nghiệm hormon tuyến giáp: FT3, FT4 và TSH trên 4 hệ thống máy xét nghiệm thông dụng Roche, Siemens, BC và Abbott thu thập từ chương trình EQA của RIQAS trong giai đoạn 40 tháng từ 2020 đến tháng 2025.

2.3. Cỡ mẫu và chọn mẫu: 40 chu kỳ EQA của RIQAS của các xét nghiệm FT3, FT4 và TSH của các hệ thống máy Roche, Abbott, BC và Siemens, ngoại trừ các máy có số lượng mẫu dưới 30 và các nhóm máy có tổng dưới 100 trong một tháng.

2.4. Các biến số trong nghiên cứu

N nhóm thiết bị ($N_{nhóm}$): Là tổng số lượng kết quả từ tất cả các máy trong một nhóm thiết bị cùng nhà sản xuất, với n_i là số kết quả của từng loại máy i trong nhóm.

$$N_{nhóm} = \sum_i n_i$$

Trung bình nhóm: Là giá trị trung bình có trọng số của nhóm thiết bị cùng nhà sản xuất được tính trên

số lượng kết quả (n) của từng máy trong nhóm ở một chu kỳ ngoại kiểm.

$$\bar{x} = \frac{\sum_i n_i \cdot \bar{x}_i}{\sum_i n_i}$$

với $\bar{x}_i$ là giá trị trung bình của máy i .

SD_{gộp} và CV%_{gộp} (Hệ số biến thiên gộp)

SD_{gộp} là độ lệch chuẩn của nhóm thiết bị cùng nhà sản xuất trong 1 chu kỳ ngoại kiểm.

$$SD_{gộp} = \sqrt{\frac{\sum [(n_i - 1)SD_i^2 + n_i(\bar{x}_i - \bar{x})^2]}{\sum (n_i - 1)}}$$

với SD _{i} là độ lệch chuẩn của máy i .

CV%_{gộp} là độ biến thiên tương đối của các máy trong cùng một nhóm thiết bị trong một chu kỳ ngoại kiểm. Được tính thông qua chỉ số độ lệch chuẩn gộp SD_{gộp}[6].

$$CV_{gộp} = \frac{SD_{gộp}}{\bar{x}} \times 100$$

Bias_{nhóm}: Là mức độ sai lệch có hệ thống giữa kết quả trung bình của một nhóm thiết bị cùng nhà sản xuất so với nhóm tham chiếu – (hệ thống Roche), được biểu thị dưới dạng phần trăm (%) [7]. Với $\bar{x}_{nhóm}$ là giá trị trung bình của nhóm thiết bị đang được so sánh.

$$Bias = \frac{\bar{x}_{nhóm} - \bar{x}_{RC}}{\bar{x}_{RC}} \times 100$$

CV%_{nhóm}: Là hệ số biến thiên của các kết quả trong cùng nhóm thiết bị ở tất cả chu kỳ ngoại kiểm (m), dùng để đánh giá độ ổn định của nhóm và tính toán chỉ số TEa_T [7].

$$CV_{nhóm} = \frac{1}{N_{months}} \sum_{m=1}^{N_{months}} CV_{gộp,m}$$

với m là số tháng tham gia ngoại kiểm của nhóm thiết bị (40 tháng)

TEa_T: Là mức sai số tổng cho phép được xác định dựa trên hiệu năng thực tế của các nhóm thiết bị tham gia ngoại kiểm [7].

$$TEa_T = |Bias_{nhóm}| + 1.96 \times CV_{nhóm}$$

TEa_{BV}: Là mức sai số tổng cho phép dựa trên biến thiên sinh học, được tính từ hai thành phần biến thiên trong cá thể (CV _{i}) và giữa các cá thể (CV _{g}) theo dữ liệu tham khảo từ EFLM Biological Variation Database [6, 7].

$$TEa_{BV} = k \cdot \sqrt{CV_I^2 + CV_G^2 + |Bias_{cho phép}|}$$

Khoảng tin cậy (Confidence Interval - CI): Là khoảng giá trị ước tính dùng để đánh giá độ chính xác của các tham số thống kê như bias, CV% và TEa. Trong nghiên cứu này sử dụng khoảng tin cậy 95% (CI

95%). $CI = Estimate \pm (t \text{ or } z) \cdot SE$ với Estimate là giá trị ước lượng của đại lượng ($Bias_{nhóm}$, $CV\%_{nhóm}$, hoặc TEa_{TT}). Hệ số nhân, z-score = 1.96 cho 95% CI mẫu lớn ($n > 30$). SE: Sai số chuẩn của ước lượng.

Chỉ số tương đồng - HI (Harmonization Index): Là chỉ số đánh giá mức độ tương đồng giữa các nhóm thiết bị, được tính bằng tỷ số giữa TEa_{TT} và TEa_{BV} [1].

$$HI = \frac{TEa_{TT}}{TEa_{BV}}$$

2.4. Quy trình nghiên cứu

- Thu thập dữ liệu EQA về xét nghiệm FT3, FT4, TSH từ 40 chu kỳ của các hệ thống máy Roche, Abbott, BC và Siemens. Lọc dữ liệu bằng cách loại bỏ các máy có số lượng mẫu dưới 30 và các nhóm máy có tổng dưới 100 trong một tháng.

- Tính toán các chỉ số Nhóm, Trung bình có trọng số và $CV\%$ gộp cho từng nhóm máy theo từng chu kỳ EQA để đánh giá đặc điểm kết quả xét nghiệm; Ước lượng $Bias_{nhóm}$, $CV\%_{nhóm}$ và TEa_{TT} kèm khoảng tin cậy (CI 95%) để phân tích độ lệch và biến thiên giữa các hệ thống.

- Tính TEa_{BV} dựa trên CVi và CVg từ cơ sở dữ liệu

EFLM Biological Variation để làm chuẩn so sánh, 3 mức TEa_{BV} Tối thiểu, Mong muốn và Tối ưu của TSH lần lượt là 41,4; 27,5 và 13,8 (%); FT4: 10,6; 7,0 và 3,6 (%); FT3: 11,0; 7,3 và 3,7 (%). [8] So sánh TEa_{TT} với TEa_{BV} đồng thời đánh giá $Bias_{nhóm}$ và $CV\%_{nhóm}$, HI để xác định mức độ tương đồng. Tiêu chuẩn chấp nhận độ tương đồng giữa các dòng máy khi độ lệch $Bias_{nhóm} < TEa_{BV}$ tối thiểu và $HI < 1$.

-Phân tích xu hướng sai khác theo dải nồng độ của các hệ thống thông qua biểu đồ Bland-Altman, đề xuất hướng liên thông phù hợp.

2.5. Xử lý và phân tích số liệu: Dữ liệu được xử lý và phân tích bằng phần mềm Microsoft Excel.

2.6. Đạo đức trong nghiên cứu: Nghiên cứu được thực hiện độc lập, sử dụng hoàn toàn dữ liệu thứ cấp xét nghiệm FT3, FT4 và TSH từ các hệ thống máy xét nghiệm miễn dịch từ chương trình ngoại kiểm chất lượng của RIQAS trong giai đoạn 2020–2025, không chứa bất kỳ thông tin nhận dạng cá nhân hay dữ liệu bệnh nhân. Nghiên cứu không can thiệp trên người hoặc động vật, không thu thập mẫu sinh học mới.

3. KẾT QUẢ NGHIÊN CỨU

3.1. Nhóm các máy tham gia đánh giá

Nhóm máy	Loại máy	FT3		FT4		TSH	
		n	N _{nhóm}	n	N _{nhóm}	n	N _{nhóm}
Roche	Roche Cobas 4000/e411	308,05	966,48	391,55	1211,2	411,27	1245,1
	Roche Cobas 601/602	377		484,98		487,60	
	Roche Cobas e402/801	281,42		334,68		346,23	
Abbott	Abbott Architect	283,43	522,27	343,85	627,83	358,93	656,15
	Abbott Alinity	238,85		283,98		297,23	
BC	Beckman Dxl	182,20	315,02	216,85	381,32	279,25	353,95
	Beckman Access	132,82		164,47		74,70	
Siemens	Siemens Centaur	101,75	221,53	117,25	248,43	113,10	246,4
	Siemens Atellica IM	119,78		131,18		133,30	

Nhận xét: Hệ thống máy Roche chiếm tỷ lệ lớn nhất, đó cũng là lý do được lựa chọn để sử dụng làm nhóm tham chiếu trong nghiên cứu. Abbott, BC, Siemens có số lượng máy từ 200-500 thiết bị, đủ để đánh giá khách quan độ tương đồng giữa các hệ thống phân tích.

3.2. Bias_{nhóm}, CV%_{nhóm} và TEa_TT của từng nhóm thiết bị trong 40 chu kỳ ngoại kiểm

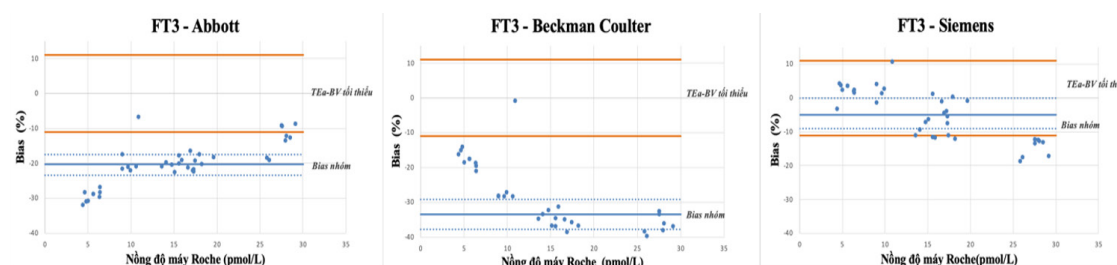
Loại xét nghiệm	Chỉ số	Abbott	BC	Siemens
TSH	Bias _{nhóm}	-20,37 [17,39 - 23,45]	-33,47 [29,18 - 37,76]	-4,94 [0,07 - 9,01]
	CV% _{nhóm}	5,95 [5,5 - 6,4]	8,285 [8,02-8,56]	5,285 [4,95-5,63]
	TEa_TT	32,03 [28,72- 35,35]	49,71 [45,27 - 54,19]	15,3 [11,02%- 19,56]
FT4	Bias _{nhóm}	-28,14 [24,06-32,22]	-15,18 [12,05 -18,31]	-22,99 [15,06 -30,94]
	CV% _{nhóm}	6,03 [5,65- 6,41]	7,23 [7,04 - 7,42]	5,34 [5,07 - 5,61]
	TEa_TT	39,94 [35,82-44,10]	29,35 [26.20 - 32.50]	33,46 [25,51 - 41,43]
FT3	Bias _{nhóm}	-18,08 [15,83-20,33]	-12,65 [8,37-16,93]	-0,99 [-4,03-2,01]
	CV% _{nhóm}	4,94 [4,74-5,14]	5,14 [5,01-5,27]	5,36 [4,94-5,78]
	TEa_TT	27,76 [25,48-30,06]	22,73 [18,45-27,01]	11,50 [7,40-15,60]

Nhận xét: Ở xét nghiệm TSH, Abbott và BC có sự khác biệt tương đối lớn khi so với Roche (-20,37% và -33,47%), Siemens thấp nhất (-4,94%). Với FT4, Abbott có độ lệch cao nhất (-28,14%), tiếp theo là Siemens (-22,99%) và BC (-15,18%). FT3 có độ lệch thấp, đặc biệt Siemens gần như không chênh lệch so với Roche (-0,99%). CV% dao động từ -5 đến -8% với BC và Abbott cao hơn Roche, cho thấy độ phân tán kết quả lớn hơn.

3.3. Chỉ số tương đồng HI

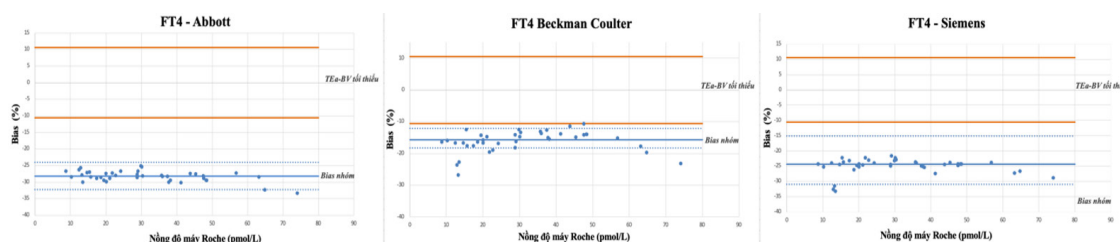
HI	FT3			FT4			TSH		
	Abbott	BC	Siemens	Abbott	BC	Siemens	Abbott	BC	Siemens
Tối thiểu	2,91	4,51	1,39	3,76	2,76	3,15	0,67	0,54	0,27
Mong muốn	4,38	6,80	2,09	5,70	4,19	4,78	1,00	0,82	0,41
Tối ưu	8,65	13,43	4,13	11,09	8,15	9,29	2,01	1,64	0,83

Nhận xét: Xét nghiệm TSH có HI < 1 ở mức mong muốn với tất cả các nhóm máy, thể hiện sự tương đồng tốt. Ngược lại, FT3 và FT4 có HI > 1, đặc biệt Abbott và BC cao nhất, phản ánh sự sai khác lớn. Siemens có HI thấp nhất trong cả ba xét nghiệm, gợi ý rằng hệ thống này cho kết quả tương đồng với Roche hơn so với các nhóm khác.



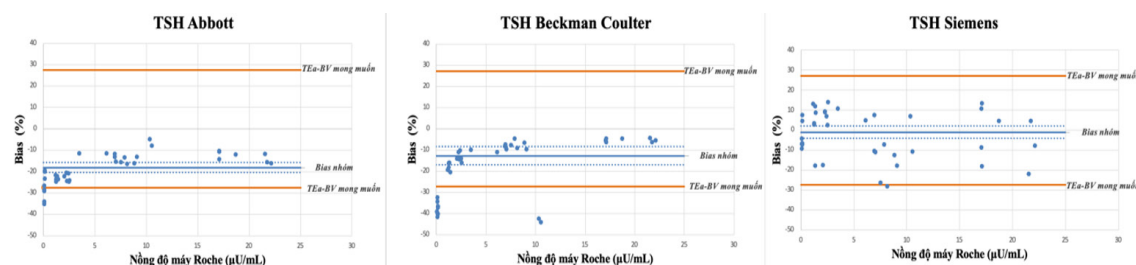
Biểu đồ 1. Độ tương đồng giữa các nhóm máy theo mức nồng độ trong xét nghiệm FT3

Nhận xét: Abbott và BC không có sự tương đồng với Roche ở mọi nồng độ, kết quả đều thấp hơn từ 5-42%, trong đó, độ lệch của Abbott có xu hướng giảm khi nồng độ càng cao và độ lệch của BC có xu hướng giảm ở nồng độ thấp. Siemens tương đồng tốt với Roche ở vùng nồng độ thấp và trung bình, nhưng ở nồng độ cao > 25 pmol/L độ lệch vượt ra ngoài ngưỡng sai số cho phép khiến tỷ lệ các điểm nằm trong TEa_BV chỉ còn 60-70%.



Biểu đồ 2. Độ tương đồng giữa các nhóm máy theo mức nồng độ trong xét nghiệm FT4

Nhận xét: Cả 3 hệ thống Abbott, BC và Siemens đều không tương đồng với Roche với độ lệch âm lớn và đồng nhất trên toàn dải nồng độ. Abbott và Siemens có độ lệch lớn, -25% đến -30%, trong khi BC có độ lệch thấp hơn, khoảng -10% đến -15%. Hầu hết các điểm độ lệch của ba hệ thống này đều không đạt TEa_BV tối thiểu nhưng phân bố song song với trục hoành, cho thấy có thể hiệu chỉnh nhờ hệ số chuyển đổi đơn giản.



Biểu đồ 3. Độ tương đồng giữa các nhóm máy theo mức nồng độ trong xét nghiệm TSH

Nhận xét: Cả 3 hệ thống Abbott, BC và Siemens đều tương đồng kết quả xét nghiệm TSH với Roche ở toàn dải nồng độ, độ lệch đều nằm trong giới hạn TEa_BV mong muốn, cho phép có thể liên thông kết quả xét nghiệm. Abbott và BC chỉ không đạt độ tương đồng ở các điểm với mức nồng độ rất thấp ($<1 \mu\text{U/mL}$), không có ý nghĩa quyết định lâm sàng.

4. BÀN LUẬN

Nghiên cứu của chúng tôi dựa trên dữ liệu EQA từ nhiều phòng xét nghiệm tại Việt Nam trong 40 tháng cho thấy sự khác biệt về kết quả xét nghiệm hormon tuyến giáp FT4 và FT3 giữa các hệ thống máy tương đối rõ rệt, trong khi TSH thể hiện mức độ tương đồng cao hơn. Kết quả của chúng tôi tương đồng với nhiều báo cáo khác trên thế giới về lĩnh vực này. Kong và cộng sự khi phân tích dữ liệu EQA từ nhiều phòng xét nghiệm đã nhận định rằng FT4 và FT3 là các xét nghiệm có sự khác biệt giữa các hệ thống lớn nhất, trong khi TSH dễ đạt được sự tương đồng hơn[7]. Zhang S. và cộng sự trong một nghiên cứu khác đã ghi nhận sự khác biệt giữa các hệ thống miễn dịch FT4, FT3 và đề xuất các chiến lược chuẩn hóa để giảm thiểu sai lệch[5]. Roche là hệ thống phổ biến nhất tại Việt Nam, chiếm tỷ lệ lớn trong chương trình EQA RIQAS, do đó được lựa chọn làm tham chiếu trong nghiên cứu này nhằm thuận tiện để so sánh và ứng dụng kết quả vào thực tiễn.

Sự khác biệt trong giữa các mức độ tương đồng của các hormon này có thể giải thích do 2 loại yếu tố chính là sinh học và kỹ thuật xét nghiệm. Về mặt sinh học, TSH chịu ảnh hưởng lớn từ nhịp sinh học, stress,

tuổi, giới, ... nên CVi và CVg đều cao, dẫn đến TEa_BV cao (41%). Vì vậy, kết quả TSH có thể dao động đáng kể giữa các lần đo hoặc giữa các cá nhân mà vẫn trong giới hạn chấp nhận lâm sàng. Ngược lại, FT4 và FT3 chỉ chiếm 0,02–0,05% tổng hormone, luôn cân bằng động với nồng độ ổn định, nên CVi và CVg rất thấp, dẫn đến TEa_BV của FT4/FT3 thường $<10\%$. Về kỹ thuật, TSH được đo bằng phương pháp miễn dịch sandwich với hai kháng thể đặc hiệu nhắm vào các epitope khác nhau trên phân tử TSH- β , ít bị ảnh hưởng bởi protein gán hoặc chất gây nhiễu. Đồng thời, các chương trình quốc tế của IFCC đã triển khai các giải pháp tương đồng hóa TSH bằng phương pháp thống kê như All-Procedure Trimmed Mean (APTM), giúp giảm đáng kể sự khác biệt giữa các hệ thống máy. Nhờ đó, kết quả TSH giữa các hãng thường khá tương đồng và dễ so sánh. Ngược lại, FT4 và FT3 sử dụng phương pháp miễn dịch cạnh tranh phức tạp hơn và dễ bị ảnh hưởng bởi các yếu tố nhiễu. Mặc dù FT4 đã có phương pháp tham chiếu tiêu chuẩn (RMP), nhưng chuỗi truy nguyên sang các hệ thống miễn dịch vẫn còn nhiều bất cập do khác biệt kháng thể, hiệu ứng protein gán và các yếu tố nền. Đặc biệt, FT3 hiện chưa có RMP khiến việc chuẩn hóa toàn cầu chưa hoàn thiện, dẫn đến sai số khó khắc phục giữa các hệ thống. Bên cạnh đó, sự khác biệt giữa các hệ thống còn xuất phát từ thiết kế phản ứng và công nghệ phát tín hiệu. Roche và Siemens áp dụng kháng thể đơn dòng với quy trình 2 bước, sử dụng biotin/streptavidin hoặc acridinium; BC dùng kháng thể từ chuột kết hợp alkaline phosphatase; Abbott sử dụng CMIA với hạt từ và acridinium. Những khác biệt này ảnh hưởng đến ái lực gắn và khả năng loại bỏ nhiễu, gây sai lệch kết quả, đặc biệt ở nồng độ

thấp hoặc rối loạn protein huyết tương. Miller và cộng sự nhấn mạnh rằng sự khác biệt giữa hệ thống trong xét nghiệm miễn dịch không thể loại bỏ hoàn toàn, nên cần định lượng và kiểm soát mức sai khác thay vì theo đuổi sự trùng khớp tuyệt đối [2].

Trong nghiên cứu của chúng tôi, TSH cho thấy mức độ tương đồng cao giữa các hệ thống với HI < 1 trên toàn bộ dải nồng độ. Trên lâm sàng, kết quả TSH từ Roche và 3 hệ thống còn lại có thể dùng liên thông trong theo dõi bệnh nhân, tuy nhiên phòng xét nghiệm vẫn cần đánh giá độ tương đồng trên từng dòng máy cụ thể. Đối với FT4, sự khác biệt giữa các hệ thống khá rõ rệt và có thể ảnh hưởng đến quyết định lâm sàng. Theo kết quả phân tích của chúng tôi, cả 3 hệ thống máy Siemens, Abbott và BC đều có sai lệch lớn với xu hướng thấp hơn khi so sánh với hệ thống Roche lên tới 25–30%. Thực tế, đây là ngưỡng FT4 đủ để thay đổi quyết định điều trị của bác sĩ lâm sàng. Vì vậy, các đơn vị xét nghiệm cần thận trọng trong việc phiên giải kết quả FT4 được phân tích trên các hệ thống máy khác nhau, nên sử dụng thiết bị duy nhất trong theo dõi điều trị và sử dụng dải tham chiếu cho từng thiết bị nhất định. Mặt khác, chúng tôi nhận thấy độ lệch của xét nghiệm FT4 trên các hệ thống máy đều tương đối đồng nhất trên các mức nồng độ của mẫu, vì vậy có thể xây dựng hệ số chuyển đổi đơn giản giữa các hệ thống. FT3 cũng thể hiện sai khác giữa các hệ thống, mặc dù mức độ thấp hơn FT4 nhưng độ lệch khá phân tán trên các dải nồng độ khác nhau. Cụ thể trong nghiên cứu này, Abbott và BC khi so sánh với Roche đều có kết quả thấp hơn từ 10-20%, trong đó, độ lệch của Abbott có xu hướng giảm khi nồng độ càng cao; ngược lại, độ lệch của BC có xu hướng giảm ở nồng độ thấp. Tuy nhiên, xét nghiệm FT3 trên Siemens có sự tương đồng tốt với Roche ở vùng nồng độ thấp < 25 pmol/L. Do đó, đối với xét nghiệm FT3 cũng tương tự FT4 không khuyến khích sử dụng trên nhiều hệ thống khác nhau, đặc biệt lưu ý khi liên thông xét nghiệm, nên thống nhất sử dụng cùng thiết bị với dải tham chiếu tương ứng.

Tóm lại, xét nghiệm TSH đạt tương đồng giữa các hệ thống có thể cân nhắc sử dụng liên thông các hệ thống xét nghiệm, xét nghiệm FT4 và FT3 có độ lệch âm khá lớn từ 15-25% khi so sánh các hệ thống Abbott, Beckman Coulter, Siemens với Roche, cần lưu ý trong diễn giải xét nghiệm trong thực hành lâm sàng và liên thông xét nghiệm.

5. KẾT LUẬN

Xét nghiệm TSH đạt tương đồng giữa các hệ thống; xét nghiệm FT4 có độ lệch âm khá lớn từ 15-30% và đồng nhất ở mọi nồng độ khi so sánh các hệ thống Abbott, BC, Siemens với Roche; xét nghiệm FT3 có xu hướng độ lệch giảm dần ở hệ thống Abbott và tăng dần ở hệ thống BC khi nồng độ tăng lên từ thấp đến cao; Siemens tương đồng với Roche ở dải nồng độ < 25 pmol/L.

TÀI LIỆU THAM KHẢO

- [1] Welsh KJ, Soldin SJ. How reliable are free thyroid and total T3 hormone assays? Clin Biochem. 2016;49(18):1362–1367.
- [2] Miller WG, Myers GL, Lou G, et al. Harmonization: the sample, the measurement, and the report. ClinChem.2015;61(7):911–917.
- [3] Clinical and Laboratory Standards Institute (CLSI). Measurement Procedure Comparison and Bias Estimation Using Patient Samples. CLSI guideline EP09-A3. Wayne (PA): CLSI; 2013.
- [4] Clinical and Laboratory Standards Institute (CLSI). Characterization and Qualification of Commutable Reference Materials for Laboratory Medicine. CLSI guideline EP30-A. Wayne (PA): CLSI; 2010.
- [5] Zhang S, Cheng F, Wang H, Wen J, et al. Comparability of thyroid-stimulating hormone immunoassays using fresh frozen human sera and external quality assessment data. PLoS One. 2021;16(6):e0253422.
- [6] Clinical and Laboratory Standards Institute (CLSI). Evaluation of Linearity of Quantitative Measurement Procedures.2nd ed. CLSI guideline EP06-A2. Wayne (PA): CLSI; 2010.
- [7] Kong L, Liu Y, Zhou C, et al. Evaluation of harmonization among thyroid hormone testing systems: a comparative study based on external quality assessment data. Pract Lab Med. 2025;46:e00491.
- [8] Ricós C, Alvarez V, Cava F, et al. Biological Variation Database. EFLM Biological Variation Working Group, European Federation of Clinical Chemistry and Laboratory Medicine (EFLM); 2019.