Research Article

For reprint orders, please contact: info@becaris.com

Journal of Patient Centricity

Analytical comparability of thyroid-stimulating hormone and free thyroxine across capillary and venous collection matrices for remote thyroid diagnostics

Natalie Getreu*[1,2], Zoe Grant[1], Tharni Vasavan[1], Adrian Timpson[3], Timothy Woolley[4] & Helen C O'Neill[1,2]

[1]Research & Development, Hertility Health Limited, London, W1W 5PF, UK
[2]EGA Institute of Women's Health, University College London, London, WC1E 6HU, UK
[3]Department of Genetics, Evolution & Environment, University College London, London, WC1E 6BT, UK
[4]Inuvi Diagnostics, Gloucester, GL2 8AX, UK
*Author for correspondence: natalie@hertilityhealth.com

Aim: To evaluate the analytical comparability of thyroid-stimulating hormone (TSH) and free thyroxine (FT4) in capillary blood samples collected using serum separator tubes compared with venous reference standards. **Background:** Advances in at-home blood collection have enabled individuals to monitor key reproductive hormones with clinical accuracy. However, the hydrophobic nature of serum separating gels in small-volume capillary tubes can cause matrix-related interference with certain lipophilic analytes, such as sex steroids. This study investigates whether TSH and FT4 maintain analytical integrity within this specific collection matrix, a critical requirement for ensuring the accuracy of remote thyroid diagnostics. **Materials & methods:** This secondary analysis, drew data from two collection events of a validated primary study: an at-home collection setting (Trial 1, capillary samples only) and a supervised clinic-based setting (Trial 2, concurrent capillary and venous samples). Samples from 39 participants were analyzed for TSH and FT4 using the Cobas® e801 immunoassay platform (Roche Diagnostics). Statistical significance was assessed using paired t-tests on log-transformed values, with a Šidák-adjusted threshold (α = 0.000814) to account for multiple comparisons. **Results:** TSH and FT4 demonstrated robust analytical comparability across all seven tube-pair comparisons. Mean percentage differences for TSH ranged from -2.96% to +2.43%, and for FT4 from -1.49% to +0.50%. All differences fell substantially below the desirable total allowable error (TEa) thresholds derived from the European Federation of Clinical Chemistry and Laboratory Medicine Biological Variation Database (24.7% for TSH; 6.3% for FT4). In contrast to the pronounced gel-related suppression previously documented for estradiol (-13.3%) and testosterone (-9.0%), thyroid analytes exhibited no clinically significant matrix-related bias. **Conclusion:** TSH and FT4 demonstrate consistent analytical comparability across capillary serum separator tube collection matrices. These findings provide preliminary analytical evidence supporting capillary-based sampling as an alternative to venipuncture for thyroid monitoring within the pre-analytical conditions studied, and support the integration of these markers into decentralised, at-home diagnostic frameworks.

Plain language summary: What does capillary blood collection mean for thyroid hormone testing at home?
What is this article about? At-home blood testing services allow people to collect small samples by finger-prick rather than visiting a clinic for a standard blood draw. These samples are collected in small tubes that often contain a gel which separates the liquid part of the blood. Previous research shows that this gel can interfere with the sample, and artificially lower the levels of certain hormones, this was seen in estrogen and testosterone. This study investigated whether the same problem occurs for two thyroid hormones, thyroid stimulating hormone (TSH) and free thyroxine (FT4), which are important for diagnosing and monitoring thyroid conditions.

What were the results? Samples from 39 participants were collected in four types of blood collection tubes (with and without gel, using both finger-prick and standard blood draw). TSH and FT4 measurements were essentially identical across tube types. The small differences observed were well within the internationally accepted limits for clinical accuracy. Unlike estrogen and testosterone, TSH and FT4 were not affected by the gel.

What do these results suggest? TSH and FT4 can reliably be measured from finger-prick samples collected in standard gel containing tubes used in at-home testing kits. The gel does not interfere with these thyroid hormones. This means that people with thyroid conditions may be able to monitor their health from home without needing to visit a clinic for a blood draw, making monitoring thyroid treatment more accessible.

First draft submitted: 21 January 2026; Accepted for publication: 23 June 2026; Published online: 8 July 2026

Keywords: capillary-based sampling • free thyroxine • patient-centric sampling • thyroid monitoring • thyroid-stimulating hormone

Capillary blood sampling whether via finger-prick, heel lance or upper arm microsampling devices, has become an increasingly adopted method for remote hormone testing, driven by advantages including minimal discomfort, reduction of blood volume, improved cost-efficiency and enabling self-sampling, especially in decentralised diagnostic settings [1,2]. By reducing reliance on trained phlebotomists for routine sample collection, this approach can lower geographic and financial barriers to testing, though venipuncture may still be required in some clinical scenarios. While these approaches have successfully expanded access to care, their accuracy is highly dependent on pre-analytical factors, including collection technique, hemolysis, and in particular the type of blood collection tube used, all of which can influence analyte concentrations. Capillary tubes often contain a serum separator gel (SST) designed to improve sample processing and stability during transport [3]. However, the hydrophobic nature of these gels is known to cause matrix-related interference with certain lipophilic analytes, potentially via adsorption, absorption or other hydrophobic partitioning mechanisms, effects that are likely analyte dependent [4]. Previous validation research has demonstrated clinically significant reductions in estradiol (E2) and total testosterone (T) measurements (up to 13% and 8%, respectively) when collected in capillary SSTs [5]. This underscores the critical need for analyte-specific validation across all components of an at-home testing platform.

Despite this vulnerability shown by sex steroids, similar methodologies hold significant potential for enhancing accessibility for other endocrine markers, most notably thyroid hormones. Thyroid disorders are highly prevalent, especially in women, impacting metabolic health, reproductive function, and pregnancy outcomes [6,7]. Accurate and frequent monitoring of key thyroid biomarkers, such as thyroid-stimulating hormone (TSH) and free thyroxine (FT4), is essential for timely diagnosis and therapeutic management. Remote, at-home collection of TSH and FT4 could significantly increase access to care for this large patient population.

The potential for pre-analytical interference is not limited to steroid hormones, yet the analytical comparability of TSH and FT4 measurements across capillary SST and venous collection matrices remains underexplored. Existing literature on gel-related analyte loss has focused primarily on small lipophilic molecules and therapeutic drugs [8]. To address this, the present study evaluates three critical pillars of remote diagnostic accuracy: the analytical comparability of thyroid analyte measurements across the transport and storage conditions of at-home diagnostics, the comparability of measurements across capillary and venous matrices and the potential for matrix-related interactions with serum separator gels in small-volume collection devices. Crucially, despite the widespread use of capillary SSTs in the diagnostic industry, no prior study has rigorously examined the analytical comparability of TSH and FT4 in this specific at-home collection matrix. This lack of formal validation means that the assumption of equivalence in commercial capillary SST kits is often based on the stability observed in large-volume venous samples, overlooking potential matrix-analyte-volume interactions, such as the increased surface-area-to-volume ratio in small-bore capillary tubes, which may magnify contact between analytes and gel components relative to standard venous tubes [9].

In contrast to the clear vulnerability demonstrated by lipophilic steroid hormones, TSH and FT4 may be less affected by serum separator gels due to their distinct physicochemical properties [10]. Preliminary results from the original validation trial, or parent study, suggested comparable measurements for both TSH and FT4 across capillary tube types, in contrast to the suppression observed for E2 and T, providing the scientific justification for this predefined secondary analysis [5]. Here, we specifically examine the analytical comparability of TSH and FT4

Table 1. Blood collection tube specifications and terminology.

Tube acronym	Description	Brand	Volume	Method
cS-CAT	Capillary serum clot activator	MiniCollect®	1.0 ml	Finger-prick
cSS-CAT	Capillary serum separator clot activator	MiniCollect®	0.8 ml	Finger-prick
vS-CAT	Venous serum clot activator	Vacuette®	4.0 ml	Phlebotomy venipuncture
vSS-CAT	Venous serum separator clot activator	Vacuette®	5.0 ml	Phlebotomy venipuncture

Technical specifications of the capillary and venous tubes utilised across the at-home (Trial 1) and clinic-based (Trial 2) datasets. All tubes were provided by Grenier Bio-One (Kremsmunster, Austria). All four tube types contain a silicon dioxide (SiO_2) particle-based clot activator.
cSS-CAT: Capillary serum separator clot activator tube; cS-CAT: Capillary serum clot activator tube; vS-CAT: Venous serum clot activator tube; vSS-CAT: Venous serum separator clot activator tube.

across multiple tube types (serum separator and non-gel tubes), aiming to determine whether capillary sampling produces analytically equivalent results compared with the venous reference method.

Materials & methods

Ethical compliance & study cohort

All procedures for the primary data collection were approved by the London Research Ethics Committee (20/LO/0265) and conducted in accordance with the Declaration of Helsinki. This secondary analysis utilised de-identified data from a cohort of thirty-nine participants assigned female at birth, aged ≥18 years, recruited through a digital health platform, including employees and volunteers. Participants were not financially compensated for their participation. Each participant provided informed consent for the use of their samples in research validation studies. As this study constitutes a predefined secondary analysis of pseudonymised data collected under original ethical approval, no additional ethical clearance was required. The original approval encompasses secondary analyses of this dataset. A participant flow diagram detailing enrolment, exclusions, and final paired sample numbers for each of the seven tube-pair comparisons is provided in Supplementary Figure 1.

Study design

This study is a predefined secondary analysis of data derived from a previously published paired venous-capillary blood collection validation trial [5]. Data were drawn from two distinct collection events reported in the parent study: Trial 1 (February to March 2023), in which 34 paired capillary samples were collected from 13 participants into cSS-CAT and cS-CAT tubes in a domestic setting; and Trial 2 (September to October 2023), in which 26 participants each provided four concurrent samples (cSS-CAT, cS-CAT, vSS-CAT and vS-CAT) in a supervised clinic setting. Hereafter these are referred to as the at-home dataset (Trial 1) and the clinic-based dataset (Trial 2). Across both collection events, a total of 39 unique participants were enrolled. To evaluate the impact of the collection matrix on thyroid analytes, four tube types were compared (Table 1).

Sample collection & logistics

Two distinct settings contributed data to this analysis. The at-home dataset (Trial 1) includes capillary samples (cS-CAT and cSS-CAT) self-collected by participants in a domestic setting using standardised finger-prick test kits. In the clinic-based dataset (Trial 2), venous samples (vS-CAT and vSS-CAT) were collected by a trained phlebotomist and capillary samples were concurrently self-collected by the participant, all within a 30-minute window. To capture a diverse distribution of analyte concentrations, samples were collected across all phases of the menstrual cycle; menstrual cycle phase was self-reported by participants at the time of collection. All samples were posted on the day of collection to the accredited laboratory via Royal Mail Tracked 24® service, under ambient conditions. Paired measurements were excluded where samples were flagged as haemolysed, where analyte concentrations fell outside the assay's limit of quantification, or where a sample collection failed entirely. In the clinic-based dataset, one venipuncture collection failed, reducing the maximum available pairs for venous comparisons. The resulting paired n for each tube-type comparison is reported in Table 2.

Analytical & laboratory methods

TSH and free thyroxine (FT4) concentrations were measured using the Cobas® e801 immunoassay module (Roche Diagnostics, Basel, Switzerland) at a UKAS-accredited diagnostic laboratory (ISO 15189:2022), with TSH and FT4 within the accredited scope. The Cobas e801 module was used to measure estradiol and testosterone, as

Table 2. Analytical comparability of thyroid-stimulating hormone and free thyroxine across all seven tube-type comparisons.

Metric	At-home (Trial 1)	Clinic-based (Trial 2)					
	cSS-CAT vs cS-CAT	cSS-CAT vs cS-CAT	cSS-CAT vs vSS-CAT	cSS-CAT vs vS-CAT†	cS-CAT vs vS-CAT	vSS-CAT vs vS-CAT	vSS-CAT vs cS-CAT
Thyroid-stimulating hormone (TSH) mIU/l							
Paired n	58	—	26	25	26	26	27
Difference, %‡	-0.72%	-0.50%	-2.92%	-2.87%	-2.03%	+0.53%	+2.38%
Log difference	-0.009	-0.005	-0.030	-0.029	-0.021	+0.005	+0.024
Lin's CCC	0.996	—	0.987	0.988	0.993	1.000	0.993
Pearson r	0.996	—	0.993	0.994	0.996	1.000	0.995
p-value	0.123	0.409	0.021	0.023	0.057	0.051	0.026
Free thyroxine (FT4) pmol/l							
Paired n	58	—	26	25	25	26	26
Difference, %‡	+0.44%	+0.50%	+0.38%	-1.12%	-1.51%	-1.42%	0.00%
Log difference	+0.004	+0.005	+0.004	-0.011	-0.015	-0.014	0.000
Lin's CCC	0.981	—	0.992	0.989	0.973	0.989	0.982
Pearson r	0.982	—	0.993	0.993	0.978	0.994	0.982
p-value	0.109	0.147	0.289	0.003	0.012	<0.001§	0.992

† Primary comparison (cSS-CAT vs vS-CAT): the clinically relevant comparison between the at-home capillary SST collection format and the venous reference standard, further compared in Figures 1–3.
‡ % Difference = back-transformed percentage difference, calculated as (exp[mean log difference] - 1) × 100. This is the geometric mean ratio expressed as a percentage and is preferred over arithmetic percent difference for log-transformed paired data.
§ Statistically significant after Šidák correction for multiple testing ($\alpha = 0.000814$). All other comparisons are non-significant after correction.
*Statistical significance after Šidák correction ($p < 0.000814$). p-values reported to three decimal places; values below 0.001 reported as <0.001.
Paired comparisons of TSH and FT4 across all seven tube-type combinations. Values reported are back-transformed percentage differences calculated as (exp[mean log difference] -1) × 100. Paired n reflects the number of participants contributing valid measurements to each comparison after exclusion.
TSH concentrations (median, IQR) by tube type: cSS-CAT 1.57 (1.26–2.13) mIU/l; cS-CAT 1.56 (1.29–2.17) mIU/l; vS-CAT 1.42 (1.13–1.87) mIU/l; vSS-CAT 1.45 (1.15–1.86) mIU/l.
FT4 concentrations (median, IQR) by tube type: cSS-CAT 16.3 (15.3–17.4) pmol/l; cS-CAT 16.3 (15.2–17.4) pmol/l; vS-CAT 17.2 (15.4–18.3) pmol/l; vSS-CAT 16.9 (15.3–18.4) pmol/l.
CCC = Lin's Concordance Correlation Coefficient. Values ≥ 0.99 indicate near-perfect concordance. — = not calculated (insufficient venous data in Trial 1).
EFLM thresholds: TSH desirable total allowable error (TEa) = 24.7%; optimal allowable bias = 6.3%. FT4 desirable TEa = 6.3%; optimal allowable bias = 1.9%. All differences for both analytes fell within optimal bias thresholds.
AMR: TSH 0.005–100 mIU/l; FT4 0.3–100 pmol/l (Cobas® e801, Roche Diagnostics, Basel, Switzerland).
cSS-CAT: Capillary serum separator clot activator tube; cS-CAT: Capillary serum clot activator tube; FT4: Free thyroxine; TEa: Total allowable error; TSH: Thyroid-stimulating hormone; vS-CAT: Venous serum clot activator tube; vSS-CAT: Venous serum separator clot activator tube. Paired n: Number of participants contributing valid paired measurements after exclusion.

well as all other hormones included in the parent study except for anti-Müllerian hormone which was measured using the Access 2 Immunoassay System (Beckman Coulter Diagnostics). All collection tubes were supplied by Greiner Bio-One (Kremsmünster, Austria). Samples were centrifuged in accordance with the blood collection tube manufacturer's recommendations prior to analysis. Hemolysis was assessed using the hemolysis rejection index specified in the relevant kit insert; samples exceeding this threshold were excluded from analysis. At-home samples were collected using a registered Class A *in vitro* diagnostic medical device (IVD) intended for capillary blood collection and transport to a clinical laboratory. Further details of the specimen collection device, including technical specifications, are provided in Supplementary Data 1. All analyses were performed within 24-72 h of sample collection, consistent with the operational logistics of decentralised, at-home diagnostic services. Where collection timestamps were available in the dataset, the interval from sample collection to laboratory receipt ranged from approximately 10 to 18 h (median 13 h), consistent with Royal Mail Tracked 24® delivery. According to the manufacturer, the assays demonstrate high analytical precision in venous serum, with total coefficients of variation of 2.1% for TSH and 2.0% for FT4.

Statistical analysis

Paired comparisons were conducted across the seven possible tube-pair combinations, and all calculations were performed using R (v4.3.1). Raw TSH and FT4 concentrations exhibit right-skewed distributions typical of endocrine analytes measured across a population; natural log-transformation normalizes these distributions, satisfying the assumptions of the paired *t*-test. Log-transformed TSH and FT4 values were therefore used, and paired t-tests

were applied to assess the significance of the differences between tube types. Back-transformed percent differences were calculated as (exp[mean log difference] -1) × 100, which is mathematically equivalent to the geometric mean ratio expressed as a percentage and is preferred over the arithmetic percent difference formula when working with log-transformed paired data to preserve distributional symmetry. The parent study performed 63 unique paired t-tests (9 analytes × 7 tube-pair combinations). To maintain statistical rigor and minimize the risk of Type I errors (false positives), the significance threshold was adjusted using the Šidák method. This yielded a corrected significance level of $\alpha = 1 - (1–0.05)^{(1/63)} = 0.000814$. This threshold is retained in the present secondary analysis to maintain consistency with the original statistical framework, as TSH and FT4 represent a pre-specified subset of those 63 comparisons. Full methodological details of the parent study are available in the original publication [5].

Results

Thyroid analyte comparability

TSH and FT4 displayed consistently low variation and robust analytical comparability across all seven paired tube comparisons. As summarised in Table 2, mean percentage differences for TSH remained within the narrow range of -2.96% to +2.43%. For FT4, variations were even more constrained, ranging from -1.49% to +0.50%.

Across all comparisons, the magnitude of variation was well within the desirable TEa limits derived from the EFLM Biological Variation Database [11]. For TSH, all differences were substantially below the desirable TEa of 24.7%, and no statistically significant differences were detected after Šidák correction. For FT4, a single comparison (vSS-CAT vs vS-CAT) reached a statistical significance ($p < 0.000814$); however, the magnitude of the difference (-1.39%) was well below the desirable EFLM TEa threshold of 6.3%, and is unlikely to affect clinical interpretation (Table 2).

The lack of systematic bias across collection matrices is further illustrated in Figures 1 & 2.

The strong concordance between cSS-CAT and vS-CAT measurements is further confirmed by Passing–Bablok regression analysis, which demonstrated no proportional or systematic bias for either analyte (Figure 3).

In contrast to the stability observed for TSH and FT4, the sex steroid hormones estradiol (E2) and total testosterone (T) exhibited pronounced and statistically significant suppression when collected in tubes containing serum separator gel (cSS-CAT). Comparing capillary tubes with and without gel (cSS-CAT vs cS-CAT) revealed significant reductions for E2 (-13.3% at-home, $p < 0.001$) and T (-9.0% at-home, $p < 0.001$). The magnitude of this effect was most pronounced in the clinic-based dataset when comparing capillary gel samples to the venous non-gel reference standard (cSS-CAT vs vS-CAT), resulting in mean reductions of -18.8% for E2 and -14.0% for T. In the same comparison, TSH and FT4 differences were negligible at -2.86% and -1.09%, respectively. These findings highlight that while steroid hormones are highly vulnerable to matrix-related interference in small volume tubes, thyroid analytes maintain robust analytical integrity.

Discussion

Analytical comparability of thyroid analytes in at-home sampling

To our knowledge, this is the first study to directly evaluate the analytical comparability of TSH and FT4 in capillary serum separator micro-collection tubes under real-world at-home transport conditions. While previous studies have examined gel-related effects on thyroid analytes in standard venous SSTs [12], the capillary format presents distinct pre-analytical challenges, including a higher surface-area-to-volume ratio and real-world postal transport, that have not been formally evaluated for TSH and FT4 prior to this study. Previous research of capillary-venous concordance for these analytes includes work by Suzuki *et al.* and Rondeel *et al.*, which demonstrated good agreement between sampling sites, but did not investigate the impact of serum separator gel on analyte measurements [13,14]. These studies used alternative collection matrices (lithium heparin anticoagulant in the case of Suzuki *et al.*), and did not assess analytical performance under real-world ambient transport conditions. The present study therefore extends the existing evidence base by specifically evaluating SST-based capillary collection under the pre-analytical conditions of at-home diagnostics. Across all paired comparisons, mean percentage difference for TSH (range: -2.96% to +2.43%) fell within the EFLM-derived optimal allowable bias of 6.3% for TSH, and differences for FT4 (range: -1.49% to +0.50%) were close to or within the optimal allowable bias of 1.9% for FT4. Crucially, all differences, including the single statistically significant FT4 comparison, remained well below the respective desirable TEa thresholds defined by the EFLM Biological Variation Database (24.7% for TSH and 6.3% for FT4), indicating that the matrix-related bias introduced by capillary SST collection is unlikely to affect clinical decision-making or therapeutic monitoring of thyroid disorders [11].

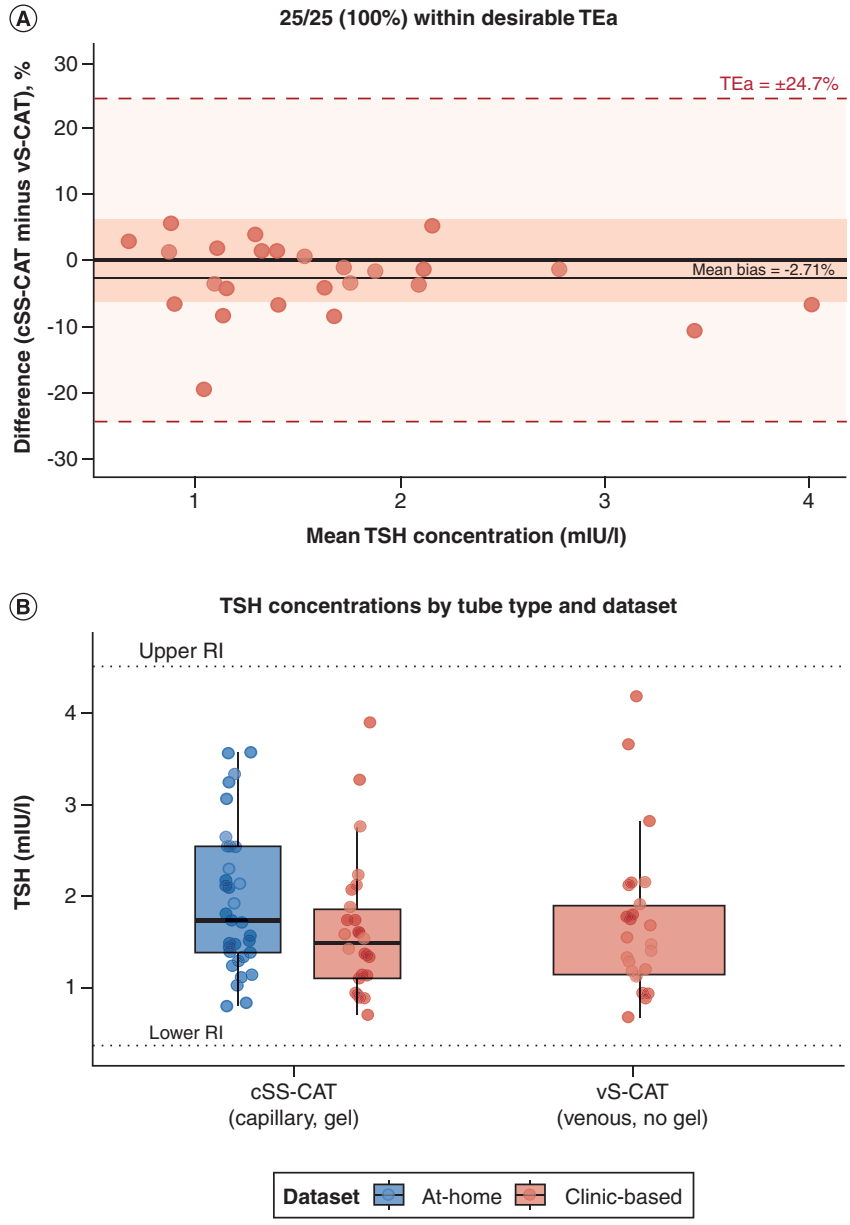

Figure 1. Analytical comparability of thyroid-stimulating hormone – capillary serum separator clot activator tubes versus venous serum clot activator tubes (primary comparison). (A) Bland–Altman plot of percentage difference (cSS-CAT minus vS-CAT) against mean TSH concentration (mIU/l). The solid horizontal line indicates the mean bias (-2.71%). Dashed lines indicate the EFLM desirable TEa (±24.7%); shaded regions represent the desirable TEa (light) and optimal allowable bias (±6.3%, dark). All 25/25 (100%) paired measurements fell within the desirable TEa; no statistically significant difference was detected after Šidák correction. **(B)** Distribution of TSH concentrations (mIU/l) by tube type (cSS-CAT and vS-CAT) and dataset. Box plots display the median (centre line), interquartile range (box limits), and 1.5× IQR (whiskers); individual data points are overlaid. Dotted lines indicate the upper and lower reference interval boundaries (0.35–4.50 mIU/l). Panel B is presented for descriptive purposes only; no statistical comparisons between tube types are shown.
Dashed lines = EFLM desirable TEa (±24.7%); shaded zones = desirable TEa (light) and optimal allowable bias (±6.3%) (dark). Blue = at-home dataset (Trial 1); Red = clinic-based dataset (Trial 2).
cSS-CAT: Capillary serum separator clot activator tube; FT4: Free thyroxine; RI: reference interval; TEa: Total allowable error; TSH: Thyroid-stimulating hormone; vS-CAT: Venous serum clot activator tube.

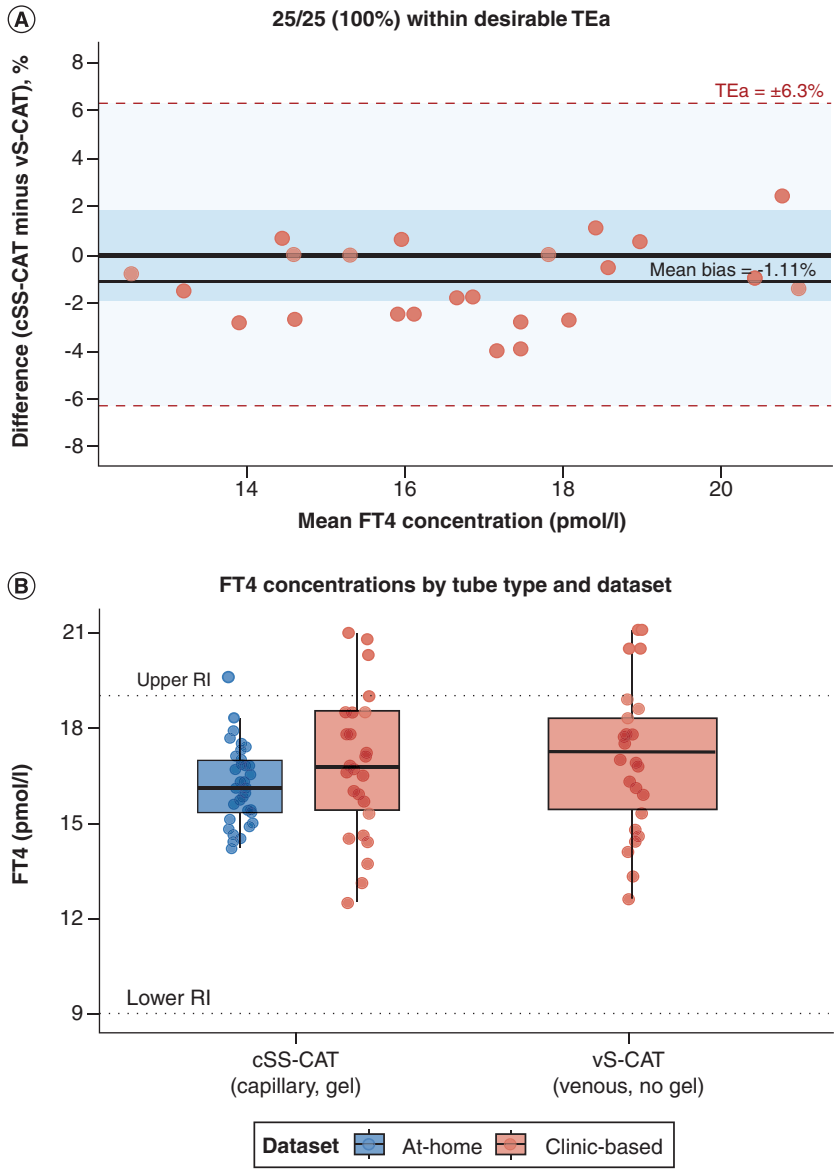

Figure 2. Analytical comparability of free thyroxine – capillary serum separator clot activator tubes versus venous serum clot activator tubes (primary comparison). (A) Bland–Altman plot of percentage difference (cSS-CAT minus vS-CAT) against mean FT4 concentration (pmol/l). The solid horizontal line indicates the mean bias (-1.11%). Dashed lines indicate the EFLM desirable TEa (±6.3%); shaded regions represent the desirable TEa (light) and optimal allowable bias (±1.9%, dark). All 25/25 (100%) paired measurements fell within the desirable TEa; no statistically significant difference was detected after Šidák correction. **(B)** Distribution of FT4 concentrations (pmol/l) by tube type (cSS-CAT and vS-CAT) and dataset. Box plots display the median (centre line), interquartile range (box limits), and 1.5× IQR (whiskers); individual data points are overlaid. Dotted lines indicate the upper and lower reference interval boundaries (9.0–19.0 pmol/l). Panel B is presented for descriptive purposes only; no statistical comparisons between tube types are shown.
Dashed lines = EFLM desirable total allowable error (TEa = ± 6.3%); shaded zones = desirable TEa (light) and optimal allowable bias (± 1.9%) (dark). Blue = at-home dataset (Trial 1); Red = clinic-based dataset (Trial 2).
cSS-CAT: Capillary serum separator clot activator tube; FT4: Free thyroxine; RI: reference interval; TEa: Total allowable error; TSH: Thyroid-stimulating hormone; vS-CAT: Venous serum clot activator tube.

This comparability is likely due to the distinct physicochemical properties of the analytes. TSH is a high-molecular-weight glycoprotein, and while FT4 is a smaller molecule, it is heavily protein-bound in circulation. These characteristics contrast sharply with lipophilic, low-molecular weight steroid hormones, which are more prone to absorptive interactions. Our findings align with literature confirming the biochemical stability of TSH

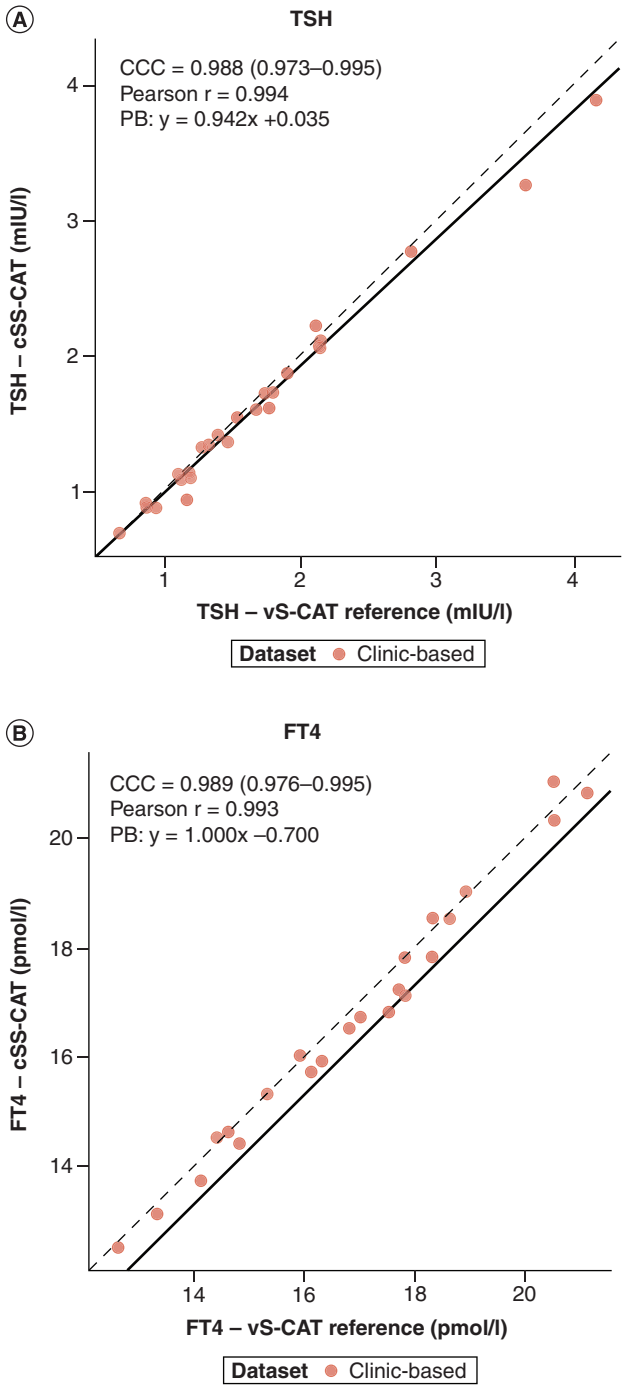

Figure 3. Method comparison scatter plots for thyroid-stimulating hormone and free thyroxine between capillary serum separator clot activator tubes (y-axis) and venous serum clot activator tube (x-axis). (A) Scatter plot of TSH measurements in cSS-CAT (y-axis) against vS-CAT reference (x-axis), both in mIU/L. CCC = 0.988 (95% CI: 0.973–0.995); Pearson r = 0.994; Passing–Bablok regression: y = 0.942x + 0.035, indicating no statistically significant proportional or systematic bias. **(B)** Scatter plot of FT4 measurements in cSS-CAT (y-axis) against vS-CAT reference (x-axis), both in pmol/l. CCC = 0.989 (95% CI: 0.976–0.995); Pearson r = 0.993; Passing–Bablok regression: y = 1.000x - 0.700, indicating no statistically significant proportional or systematic bias.
Dashed line = line of identity (y = x). Solid line = Passing–Bablok regression. Clinic-based dataset (Trial 2) only; n = 25 paired observations per panel.
CCC: Lin's Concordance Correlation Coefficient (95% CI); cSS-CAT: Capillary serum separator clot activator tube; FT4: Free thyroxine; TEa: Total allowable error; TSH: Thyroid-stimulating hormone; vS-CAT: Venous serum clot activator tube.

and FT4 across various pre-analytical conditions, including stability in different collection tubes (serum clot, plasma-gel) and over prolonged storage periods (up to 72 h) at ambient and refrigerated temperatures [12,15]. This inherent resilience makes TSH and FT4 ideal candidates for decentralised testing, where sample transport and processing delays are an inherent variable.

Thyroid analytes versus steroid hormones

The robustness of TSH and FT4 contrasts the systematic suppression documented for the sex steroid hormones estradiol and testosterone in the parent study [5]. For E2 and T, significant reductions of up to -18.8% and -14.0%, respectively, were observed when comparing capillary gel tubes (cSS-CAT) to the venous reference standards (vS-CAT) in the clinic-based dataset.

This disparity reinforces the hypothesis that the hydrophobic nature of the serum separator gel leads to matrix-related interference with specific lipophilic analytes via hydrophobic partitioning mechanisms [4]. The enhanced suppression linked to serum separator gels in capillary collection may be driven by several variables. Notably, the increased surface-area-to-volume ratio in small-bore tubes means a limited serum volume (0.8 ml) is in contact with a proportionally larger gel surface area than in standard 5.0 ml venous tubes. This configuration, combined with storage temperatures and the duration of ambient postal transit, extends the interaction time between the gel and the analyte. While the specific influence of each of these factors remains to be determined through formal mechanistic research, they likely contribute to the observed matrix effects. The fact that TSH and FT4 remained analytically comparable despite this high-contact environment emphasises that thyroid analytes are not susceptible to this specific matrix interference. This underscores a critical methodological takeaway, that analytical validation is non-transferable across endocrine markers. While E2 and T require non-gel capillary tubes or the application of conversion formulae [5], TSH and FT4 can be accurately measured using standard SST platforms.

Clinical implications

The demonstrated analytical comparability of TSH and FT4 in SST-based capillary collection systems has significant implications for expanding accessible thyroid care. Thyroid disorders are highly prevalent, and frequent monitoring is required for therapeutic management (e.g. levothyroxine dose titration) and continuous symptom management. Our data provides analytical evidence to support the use of validated at-home collection in innovative telehealth models for screening and continuous monitoring [16,17].

Capillary SSTs offer a practical advantage for at-home diagnostics by reducing reliance on trained venipuncture, therefore lowering barriers to testing for patients limited by geographical or financial factors or those requiring frequent monitoring [18–20]. It should be noted, however, that analytical comparability does not confer clinical equivalence. Capillary self-collection carries non-trivial pre-analytical failure rates, including insufficient sample volume, hemolysis, clotting or delayed processing, each of which may limit utility in clinical practice and are not captured by method comparison studies conducted under controlled research conditions. Gel-free capillary tubes may also be viable for TSH and FT4 measurement, with SSTs primarily improving sample processing efficiency and serum yield during transport. This facilitates a shift toward patient-centric care where timing of blood collection can be optimised based on clinical need rather than clinic availability.

The foundation of remote testing rests on the premise that capillary results are analytically comparable to venous results. While previous published data on TSH and FT4 concordance between venous and capillary matrices has shown variable results [12,15], our findings specifically extend this evidence base by demonstrating their analytical comparability within the SST matrix under real-world transport conditions, which is required for efficient, decentralised collection. These findings are also relevant to ISO 15189 verification for decentralised thyroid testing workflows and align with regulatory expectations for remote diagnostics [21].

Study limitations & future work

We acknowledge several limitations of this study. The remote nature of the at-home dataset collection resulted in a variable interval between sample collection and assaying, consistent with the typical logistics of at-home diagnostic services (up to 72 h). Importantly, sample collection was performed using a registered IVD medical device, ensuring a standardised and regulatory-compliant collection method across all at-home participants. Available collection timestamps indicate that laboratory receipt occurred within approximately 10–18 h of collection for the study samples, likely reflecting prioritised handling and direct laboratory receipt during the study period. However, as the precise collection-to-analysis interval was not systematically recorded for all participants, granular assessment of analyte behavior across the full potential 24–72 h window was not possible. While every effort was made to process paired samples from the same participant within the same analytical run, this could not be guaranteed for all comparisons, which may introduce a small degree of between-run variability. Additionally, the study cohort was recruited from a general population and the observed TSH (0.46–4.19 mIU/l) and FT4 (12.1–21.7 pmol/l) concentrations fell largely within or close to reference intervals, indicating that the majority of participants were euthyroid. As such, these findings cannot be readily extrapolated to clinical populations with hypothyroidism or hyperthyroidism, where analyte concentrations may span a broader portion of each assay's analytical measurement range. It should also be noted that this study demonstrates analytical comparability across collection matrices under real-world conditions, rather than intrinsic analyte stability in the traditional sense; true stability studies require

serial time-point measurements against a defined baseline under controlled conditions. Furthermore, the precision of effect estimates could be improved with a larger sample size.

Future work should investigate the stability of these analytes over longer durations (e.g., up to 7 days) and across various temperature fluctuations to model real-world transport failure scenarios. Additional mechanistic experiments would be beneficial to formally confirm the lack of interaction between TSH/FT4 and the serum separator gel. Despite these limitations, key strengths include the investigation of capillary performance across multiple tube types and the use of a cohort with a diverse distribution of analyte concentrations, lending strong external validity to our results.

Conclusion

TSH and FT4 demonstrate consistent analytical comparability across all tested collection matrices, including SST-containing capillary tubes commonly employed in at-home diagnostics. This resilience to matrix-related effects, particularly the absence of the gel-related interference observed for steroid hormones, provides preliminary analytical evidence for the comparability of capillary SST measurements with venous reference standards under the tested pre-analytical conditions. Confirmation in larger cohorts spanning a broader range of thyroid function, including patients with known thyroid dysfunction, will be needed before broader clinical implementation can be recommended. By reducing reliance on phlebotomists and centralised clinical infrastructure, these findings nonetheless provide a meaningful scientific foundation for the integration of thyroid function testing into remote diagnostic frameworks, supporting a more equitable and patient-centric approach to endocrine care.

Summary points

- At-home blood collection using finger-prick capillary samples is increasingly used for hormone testing, but the serum separator gel in small-volume capillary tubes can suppress certain analytes making analyte-specific validation essential.
- Previous research demonstrated clinically significant suppression of estradiol (up to -18.8%) and total testosterone (up to -14.0%) in capillary serum separator tubes, highlighting the need to validate each analyte independently.
- This predefined secondary analysis evaluated the analytical comparability of thyroid-stimulating hormone (TSH) and free thyroxine (FT4) across seven tube-type comparisons using paired capillary and venous samples from 39 participants collected under real-world at-home and clinic-based conditions.
- Mean percentage differences for TSH ranged from -2.96% to + 2.43% and for FT4 from -1.49% to +0.50% across all seven comparisons, all substantially below the EFLM desirable total allowable error (TEa) thresholds (24.7% for TSH; 6.3% for FT4).
- All TSH comparisons were non-significant after Šidák correction; a single FT4 comparison (vSS-CAT vs vS-CAT) reached statistical significance, but the difference (-1.39%) remained well within the desirable TEa and is clinically negligible.
- Lin's Concordance Correlation Coefficient for the primary comparison (cSS-CAT vs vS-CAT) was 0.988 for TSH and 0.989 for FT4, confirming near-perfect concordance between capillary SST and venous reference measurements.
- Unlike lipophilic sex steroids, TSH and FT4 are not susceptible to serum separator gel interference, consistent with their distinct physicochemical properties as a high-molecular-weight glycoprotein and a protein-bound hormone respectively.
- These findings provide preliminary analytical evidence supporting the use of serum separator tube-based capillary blood collection for TSH and FT4 measurement in at-home diagnostic settings, without requirement for gel-specific correction factors.
- Analytical validation is analyte-specific and cannot be extrapolated across endocrine markers; assay-by-assay validation is required before assuming equivalence between capillary and venous collection matrices.
- Confirmation in larger, clinically diverse cohorts including people with hypothyroidism and hyperthyroidism is needed before broader clinical implementation of at-home thyroid monitoring can be recommended.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at https://becarispublishing.com/doi/epdf/10.57264/jpc-2026-0005

Author contributions

The corresponding author takes full responsibility that all authors on this publication have met the following required criteria of eligibility for authorship: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpreta-

tion of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved. Nobody who qualifies for authorship has been omitted from the list. Z Grant was responsible for writing the original draft and investigation. T Vasavan was responsible for conceptualisation, investigation, methodology and project administration. A Timpson was responsible for data curation and formal analysis. T Woolley was responsible for conceptualisation, investigation and resources. HC O'Neill was responsible for conceptualisation, funding acquisition, resources and supervision. N Getreu was responsible for conceptualisation, funding acquisition, methodology, resources, supervision and writing review and editing.

Acknowledgments

The authors thank the participants who consented to take part in this study as well as the two anonymous reviewers for their valuable feedback.

Financial disclosure

This work was funded by Hertility Health Limited (London, UK). The authors have received no other financial and/or material support for this research or the creation of this work apart from that disclosed.

Competing interests disclosure

Z Grant, T Vasavan, HC O'Neill and N Getreu are employees or affiliates of Hertility Health Limited. T Woolley is an employee of Inuvi Diagnostics. The authors have no other competing interests or relevant affiliations with any organization or entity with the subject matter or materials discussed in the manuscript apart from those disclosed.

Ethical conduct of research

All procedures for primary data collection were reviewed and approved by the London Research Ethics Committee (20/LO/0265) and conducted in strict accordance with the Declaration of Helsinki. All participants provided written informed consent for the use of their pseudonymized samples in research validation studies. Because this study constitutes a predefined secondary analysis of fully pseudonymised data collected under the scope of the original ethical approval, no additional formal ethical review or clearance was required.

AI assistance disclosure

Claude (Anthropic) was used to assist in addressing peer-review comments and suggestions, and to improve the readability and clarity of the manuscript. All AI-assisted content was fully reviewed and edited by the authors as needed. No funding was received for writing assistance.

Data sharing statement

The datasets analyzed during the current study are not publicly available due to commercial confidentiality and data privacy constraints regarding participant health information. De-identified analytical data may be made available from the corresponding author upon reasonable request for the purposes of scientific verification or regulatory compliance.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit https://creativecommons.org/licenses/by-nc-nd/4.0/

References

1. Hoffman MSF, McKeage JW, Xu J, Ruddy BP, Nielsen PMF, Taberner AJ. Minimally invasive capillary blood sampling methods. *Expert Rev. Med. Devices* 20(1), 5–16 (2023).
2. Blicharz TM, Gong P, Bunner BM *et al.* Microneedle-based device for the one-step painless collection of capillary blood samples. *Nat. Biomed. Eng.* 2, 151–157 (2018).
3. Bowen RAR, Remaley AT. Interferences from blood collection tube components on clinical chemistry assays. *Biochem. Med. (Zagreb.)* 24(1), 31–44 (2014).
4. Hepburn S, Wright MJ, Boyder C *et al.* Sex steroid hormone stability in serum tubes with and without separator gels. *Clin. Chem. Lab. Med.* 54(9), 1451–1459 (2016).
5. Vasavan T, Timpson A, Woolley T, O'Neill HC, Getreu N. Systematic reduction in estradiol and testosterone measurements due to serum separator gel in blood collection tubes: implications for at-home fertility testing. *J. Appl. Lab. Med.* 10(4), 779–792 (2025).

6. Dosiou C. Thyroid and fertility: recent advances. *Thyroid* 30(4), 479–486 (2020).
7. Busnelli A, Beltratti C, Cirillo F, Bulfoni A, Lania A, Levi-Setti PE. Impact of thyroid autoimmunity on assisted reproductive technology outcomes and ovarian reserve markers: an updated systematic review and meta-analysis. *Thyroid* 32(9), 1010–1028 (2022).
8. Garza KY, Carter J, Mercer A *et al.* Evaluation of serum and rapid serum separator collection tubes for therapeutic drug assays. *Clin. Biochem.* 115, 81–85 (2023).
9. Parikh M, Wimmer C, DiPasquale C, Barr RL, Jacobson JW. Evaluation of a novel capillary blood collection system for blood sampling in nontraditional settings as compared with currently marketed capillary and venous blood collection systems for selected general chemistry analytes. *J. Appl. Lab. Med.* 10(3), 639–652 (2025).
10. Hedayati M, Razavi SA, Boroomand S, Kheradmand Kia S. The impact of pre-analytical variations on biochemical analytes stability: a systematic review. *J. Clin. Lab. Anal.* 34(12), e23551 (2020).
11. Aarsand AK, Webster C, Fernandez-Calle P *et al.* The EFLM Biological Variation Database. (Accessed 30 April 2026). https://biologicalvariation.eu/
12. Ercan M, Fırat Oğuz E, Akbulut ED, Yilmaz M, Turhan T. Comparison of the effect of gel used in two different serum separator tubes for thyroid function tests. *J. Clin. Lab. Anal.* 32(6), e22427 (2018).
13. Suzuki N, Takeuchi M, Miyazaki N *et al.* Determination of capillary blood TSH and free thyroxine levels using digital immunoassay. *J. Endocrine Soc.* 8(4), bvae030 (2024).
14. Rondeel JM, Albers JS, Schroor EJ, Schindhelm RK. Does sampling site influence levels of free thyroxine and thyrotropin determined by a current immunoassay. *Clin. Biochem.* 43(9), 771–772 (2010).
15. Schakelaar MY, Lentjes EGWM, Visser TSQ, Schoneveld AH, Hoefer IE, Tiel Groenestege WM. A preliminary analysis of thyrotropin measurement from finger stick dried blood spot with an automated high-throughput immunoassay analyzer. *Thyroid* 33(4), 449–455 (2023).
16. Gavin KM, Kreitzberg D, Gaudreau Y, Cruz M, Bauer TA. Identification and management of thyroid dysfunction using at-home sample collection and telehealth services: retrospective analysis of real-world data. *J. Med. Internet Res.* 25, e43707 (2023).
17. Syed ST, Gerber BS, Sharp LK. Traveling towards disease: transportation barriers to health care access. *J. Commun. Health* 38(5), 976–993 (2013).
18. Collier BB, Brandon WC, Chappell MR *et al.* Comparing capillary blood collection technologies: assessing patient experience, device performance, & clinical accuracy. *Bioanalysis* 17(21), 1329–1336 (2025).
19. Zhou Y, Abel G, Warren F, Roland M, Campbell J, Lyratzopoulos G. Do difficulties in accessing in-hours primary care predict higher use of out-of-hours GP services? Evidence from an English National Patient Survey. *EMJ* 32(5), 373–378 (2015).
20. Yazdi-Feyzabadi V, Sabermahani M, Borhaninejad V, Iranmanesh M. Breaking barriers: challenges faced by the older adults in accessing primary healthcare in a developing country. *BMC Geriatr.* 25(1), 946 (2025).
21. Woolley T, Hepburn S, Perkins K. UK recommendations for the validation and adoption of capillary blood testing within the routine clinical laboratory: consensus opinion from the LabMed patient-centred testing and sampling (PaCTS) group. *Ann. Clin. Biochem.* https://doi.org/10.1177/00045632261454300 (2026).