Data-Driven Optimization of Diagnostic Thresholds in Clinical Laboratories

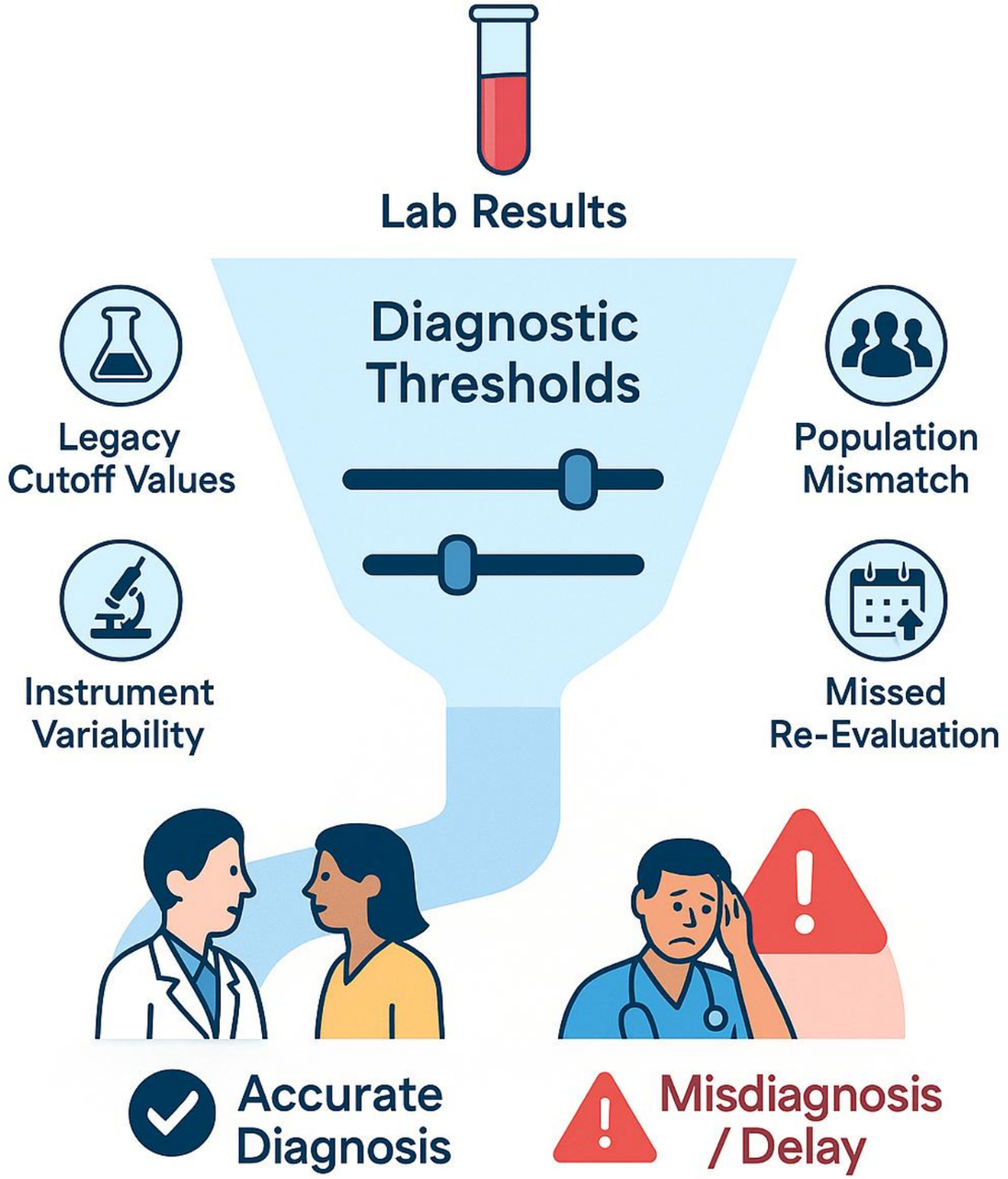
Jieli Shirley Li, MD, PhD, DABCC, FCACB
Associate Professor/Director of Clinical Pathology Services
The Ohio State University Wexner Medical Center

September 23, 2025

Learning Objectives

- Describe current challenges in setting and applying diagnostic thresholds in laboratory testing
- Recognize the clinical implications of outdated or assay-misaligned thresholds
- Apply outcome-based, data-driven approaches to evaluate and refine thresholds across different disciplines
- Translate case-based strategies into practice to enhance diagnostic accuracy, reduce unnecessary test rejection, and improve patient care

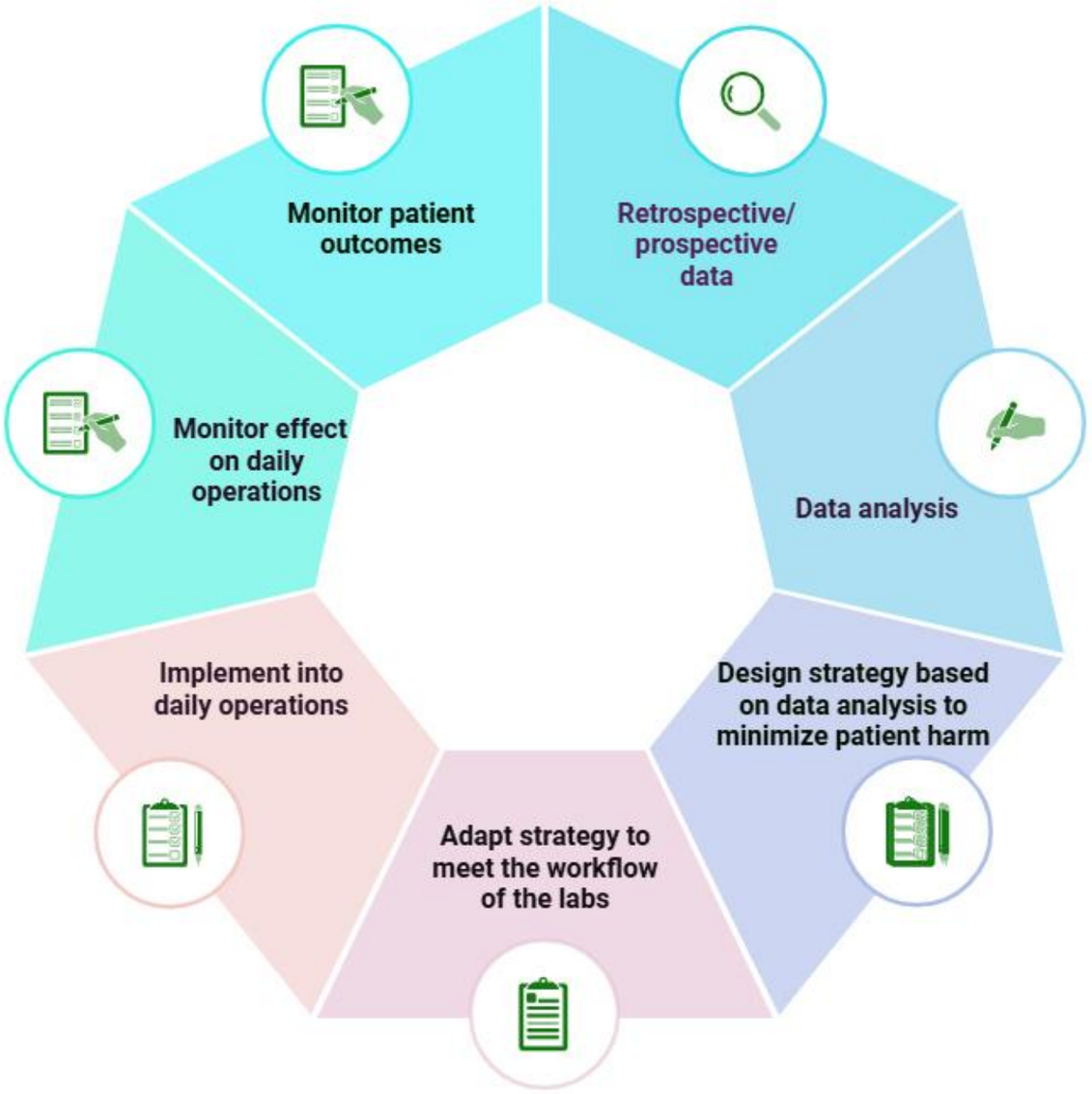
Importance of Accurate Diagnosis



Limitations of Existing Thresholds

- Many thresholds in use today have not been re-evaluated.
- Manufacturer-provided cutoffs
- Inappropriate reference intervals
- Instrument-specific variability and differences in calibration or analytical performance
- Lack of harmonization

Closing the Loop: Data, Strategy, and Outcomes



Modified from Crit Rev Clin Lab Sci. 2019 Nov 2;57(3):146-160.

Case 1- Assay-Specific Cortisol Cutoffs for Adrenal Insufficiency Diagnosis

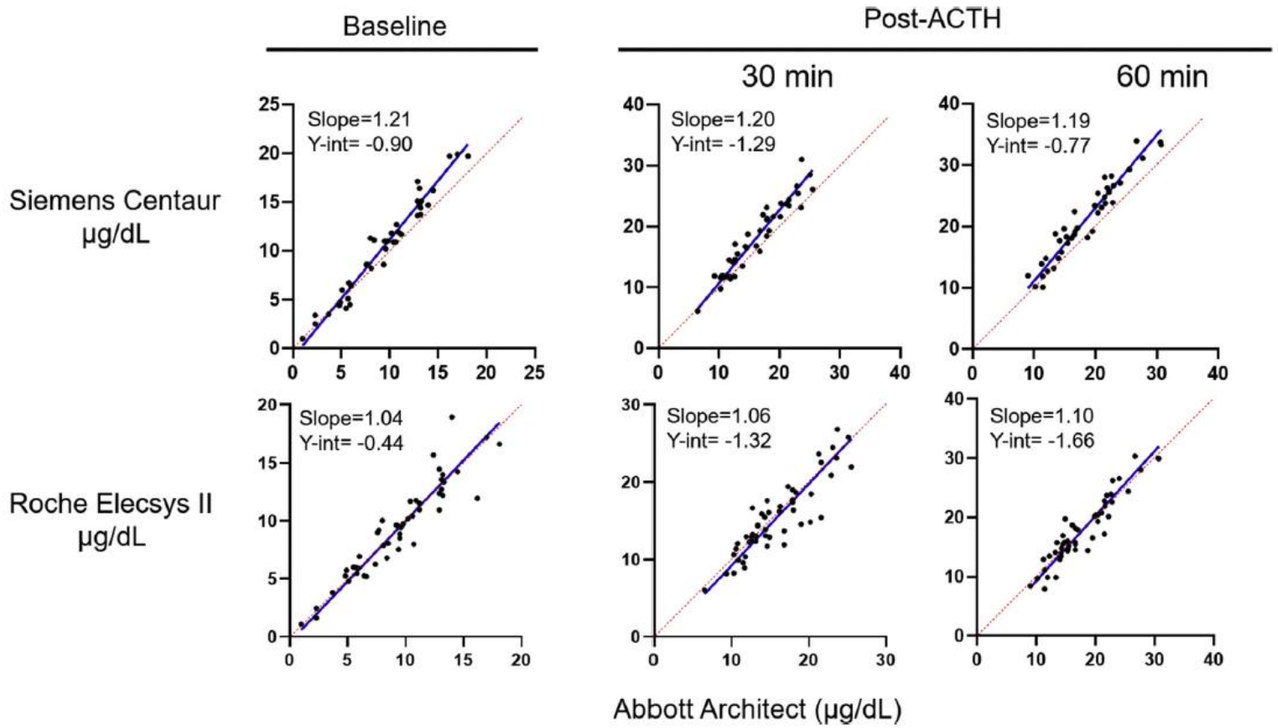
- According to the Endocrine Society, a cortisol level above 18 µg/dL at either 30 or 60 minutes is used as a cutoff to evaluate adrenal insufficiency after an ACTH stimulation test.
- However, the original source for this 18 µg/dL cutoff lacks detailed methodology information. This cutoff was likely based on immunoassay technology using polyclonal antibodies.
- More recently, one commercial immunoassay, which uses monoclonal antibodies, has recommended a lower cutoff value.
- Is the 18 µg/dL cutoff accurate for our analyzer?

Case 1- Assay-Specific Cortisol Cutoffs for Adrenal Insufficiency Diagnosis

- **Challenge**: Historical cortisol cutoffs used in ACTH stimulation tests (e.g., 18 µg/dL) were derived from polyclonal immunoassays and do not account for newer monoclonal antibody-based assays. This discrepancy can lead to overdiagnosis and unnecessary treatment of adrenal insufficiency.

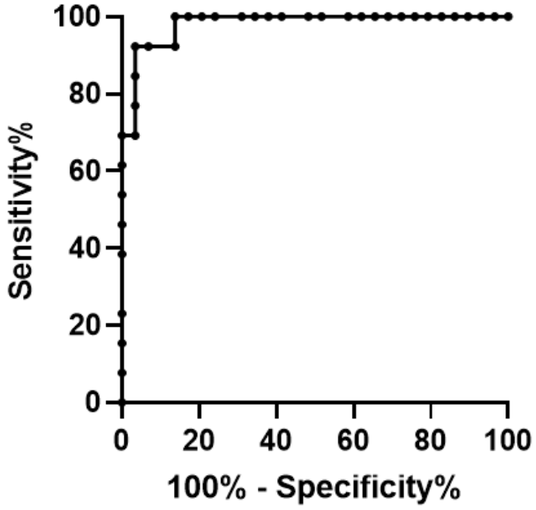
- **Action**: Compared cortisol levels from the three commercial immunoassays in patients undergoing ACTH stimulation. Conducted ROC analysis to establish new diagnostic thresholds, using the LC-MS/MS as a reference.

- *Endocr Pract. 2022 Jul;28(7):684-689.*



New Diagnostic Cutoffs for Adrenal Insufficiency After Cosyntropin Stimulation Using Abbott Architect Cortisol Immunoassay

Li Zha, PhD [1, 2], Jieli Li, MD, PhD [3], Subhashree Mallika Krishnan, MD [4], Michael R. Brennan, MD [5, 6], Yan Victoria Zhang, PhD, MBA [1], Patrick Povse, BS [3], Rae Kerlin, BS [3], Kevin Shively, BS, MHA [3], Felicia Oleksik, BS [4], JoAnna Williams, MD [3], Elizabeth Sykes, MD [4, 6], Qian Sun, PhD [4, 6, *]



Results

- A cortisol level of 18 µg/dL on our analyzers is accurate and practical for diagnosing adrenal insufficiency, as it uses polyclonal antibodies.

Case 2- Addressing Clinical Concerns in the Switch to High-Sensitivity Troponin

- The lab needs to switch from contemporary to high-sensitivity troponin due to its higher analytical sensitivity, which allows for more accurate detection of low levels of troponin, improving early diagnosis and patient care.

- Physicians are seeking clarification on the differences between these two assays, as both measure troponin. They are concerned about whether the high-sensitivity troponin assay might cause discrepancies in diagnosing or managing patient cases.

Case 2- Addressing Clinical Concerns in the Switch to High-Sensitivity Troponin

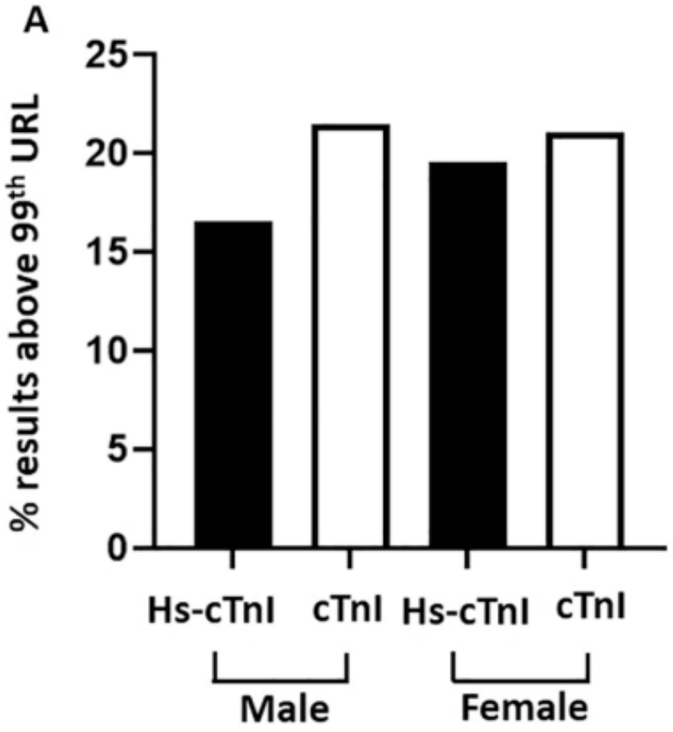
- **Challenge**: Both assays measure the same biomarker, but differences in analytical sensitivity, reporting units, and thresholds, especially with sex-specific 99th percentiles, raised questions about potential overdiagnosis, age or sex disparities, and clinical discordance.

- **Action**: Two retrospective studies were conducted to evaluate concordance between contemporary and high-sensitivity troponin assays. One focused on incidence of elevations above the 99th percentile across ED and inpatient populations. The other assessed diagnostic patterns across age, sex, and different high-sensitivity cutoff definitions. Concordance and survival analyses were also performed.

Clinical Biochemistry 89 (2021) 77–80

No increase in the incidence of cardiac troponin I concentration above the 99th percentile by Siemens Centaur high-sensitivity compared to the contemporary assay

He S. Yang [a,b,1], Amos Shemesh [b,c,1], Jieli Li [d,1], Tingting Xie [a], Fred S. Apple [e], JoAnna Williams [d], Zhen Zhao [a,b,*,2], Peter A.D. Steel [b,c,*,2]

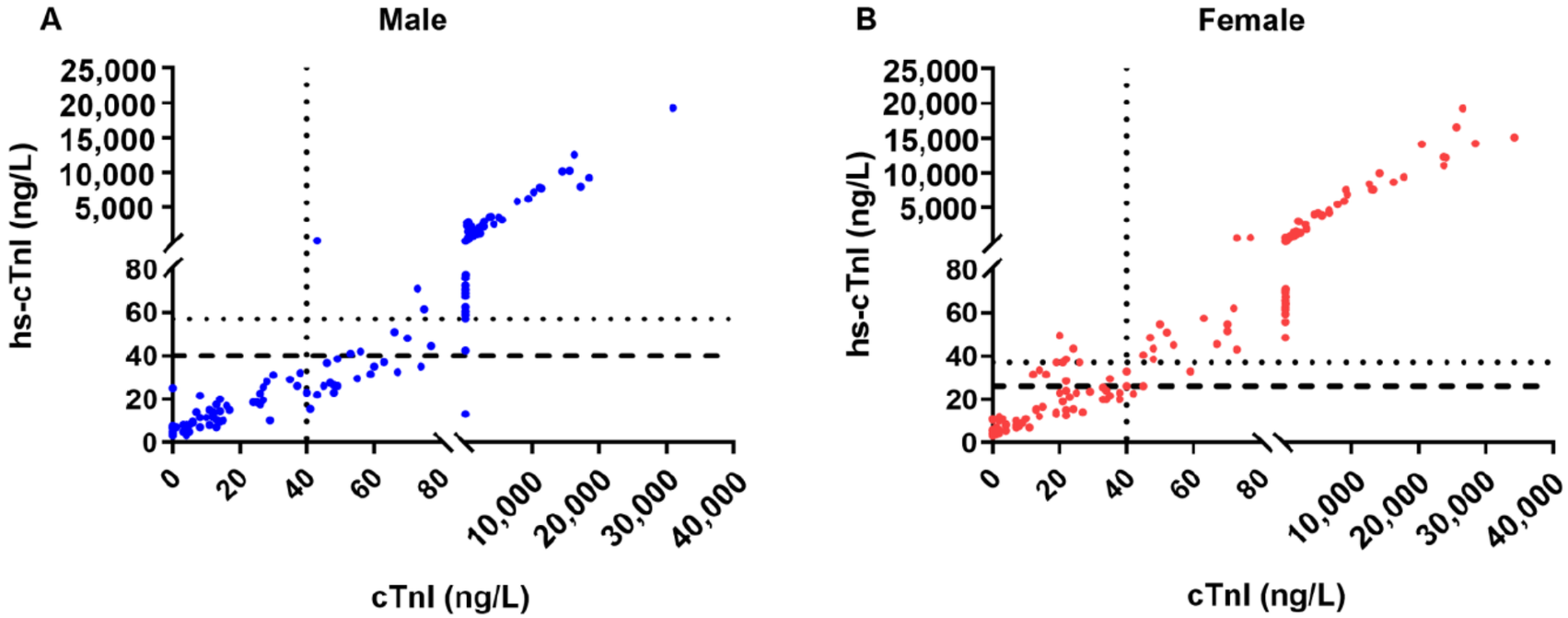


Clin Biochem. 2021 Mar:89:77-80.

Article

Evaluation of Age and Sex Differences in Contemporary versus High-Sensitivity Troponin I Measurement in Hospitalized Patients

Hussam Alkhalaileh [1], Ruhan Wei [2], Ashly Cordero Rivera [3], Mustafa Goksel [3], Jason K. Y. Lee [4], Ernest Mazzaferri, Jr. [5], JoAnna Jones [3] and Jieli Li [3,*]

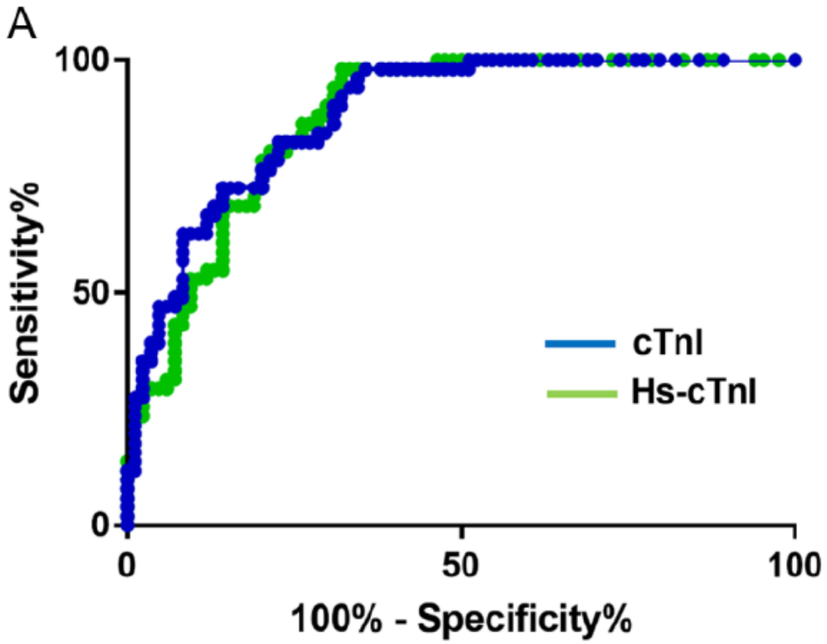


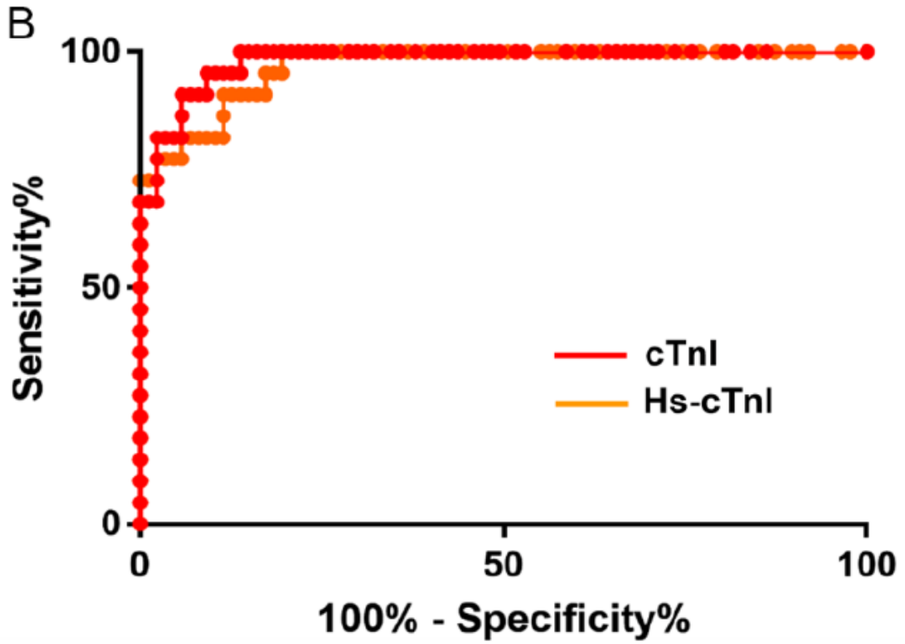
J Clin Med. 2024 Apr 21;13(8):2428.

Article

Evaluation of Age and Sex Differences in Contemporary versus High-Sensitivity Troponin I Measurement in Hospitalized Patients

Hussam Alkhalaileh [1], Ruhan Wei [2], Ashly Cordero Rivera [3], Mustafa Goksel [3], Jason K. Y. Lee [4], Ernest Mazzaferri, Jr. [5], JoAnna Jones [3] and Jieli Li [3,*]





J Clin Med. 2024 Apr 21;13(8):2428.

Case 3-Harmonizing Testosterone Reference Intervals with Clinical Guidelines

- Hi Shirley –

➢ In all the clinical reading that I have seen, normal testosterone levels are between 300-800. Ours has a much lower range, which makes some physicians think that a reading in the 'normal' level is not actually normal. Most of what I have read indicates that free testosterone should be considered if ran on an equilibrium dialysis assay.

➢ The clinical literature is fairly consistent that the former is much more clinically useful than the latter.

Verification of Reference Interval

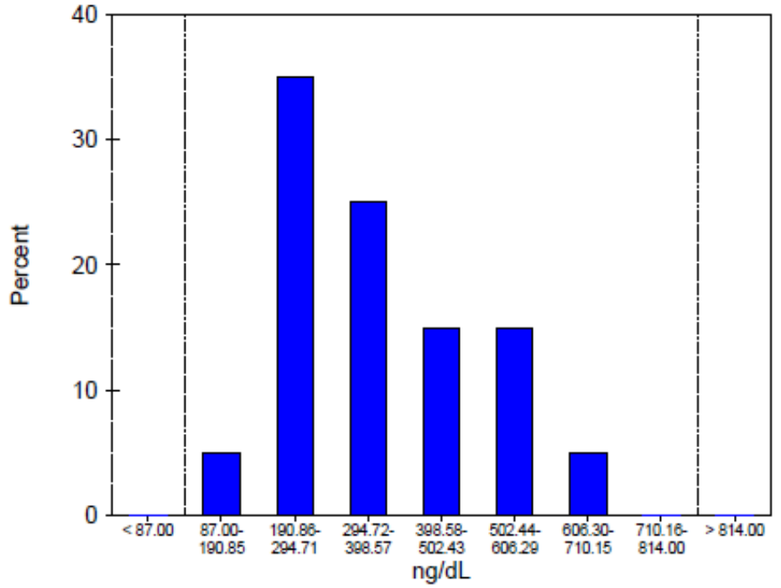
Analyst Patrick P
Specimen Criteria

Date 06 Apr 2022

Reference Interval	
Proposed	87 to 814 ng/dL
Results (Total/Excl)	20 / 0
Max/Obs outside	10.0% / 0.0%
Passes	Yes

Statistical Analysis	
Mean	355.296 ng/dL
SD	138.448
Median	325.085
Range	144.04 to 686.01
Central 95% Interval	--
Central 95% Index	--

Reference Interval Histogram



Results Distribution

Interval	Percent	Count
< 87.00	0	0
87.00-190.85	5	1
190.86-294.71	35	7
294.72-398.57	25	5
398.58-502.43	15	3
502.44-606.29	15	3
606.30-710.15	5	1
710.16-814.00	0	0
> 814.00	0	0

ENDOCRINE SOCIETY

Landmark Study Defines Normal Ranges for Testosterone Levels

Washington, DC | January 10, 2017

New defined reference range can help limit misdiagnoses and unnecessary treatments

264-916 ng/dL

American Urological Association

Evaluation and Management of Testosterone Deficiency (2024)

Clinicians should use a total testosterone level below **300 ng/dL** as a reasonable cut-off in support of the diagnosis of low testosterone.

C28-A3c
Vol. 28 No. 30
Replaces C28-A2
Vol. 20 No. 13

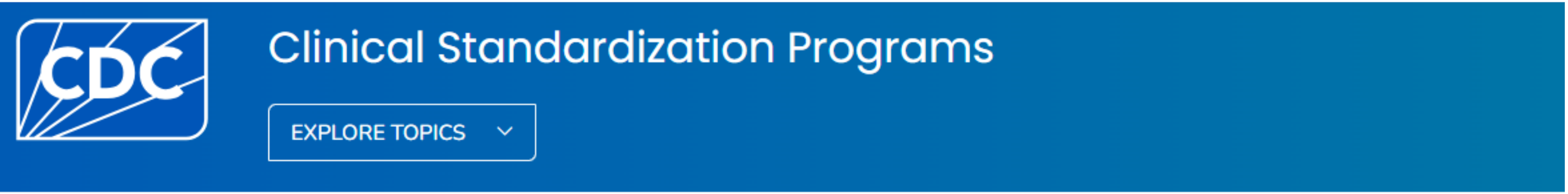
Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition

CLSI Recommendations

- **Establishment**: Minimal 120 healthy individuals for each age and set partition
  - **Challenges**: Difficult to achieve these numbers when using traditional direct sampling methods

- **Verification**: Minimal 20 samples to verify manufacturer-suggested reference intervals
  - **Challenges**: These reference intervals may not accurately reflect the lab's patient population.

10.1 Transference: Comparability of the Analytical System



Steroid Hormones Standardization Programs

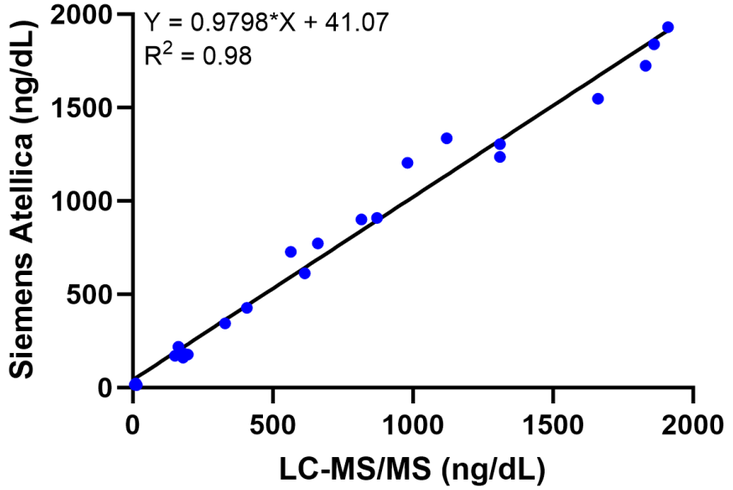
CDC's hormone standardization programs work to:

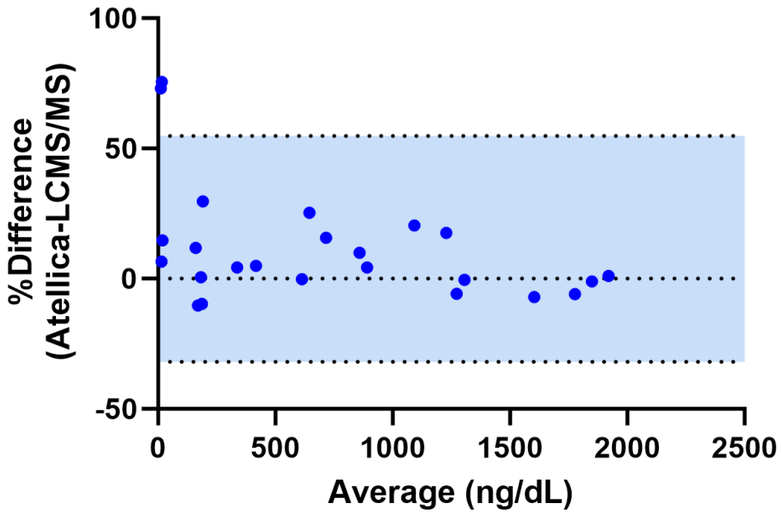
- Provide metrological reference measurements procedures. These procedures are internationally recognized and are used to assist with calibration (also called establishing metrological traceability). [Hormone Reference Laboratories]
- Assess and certify the analytical performance of testosterone and estradiol tests used in patient care, research, and public health. This includes two ***independent phases***, known respectively as, "**HoSt Phase 1"** and "**HoSt Phase 2"** (this process is also called verification of metrological traceability) [Improving Performance]
- Monitor the accuracy of measurements performed in routine laboratories over time. In this program samples are analyzed along with regular patient or study samples. [Accuracy-based Monitoring Program (AMP)]. In addition, customized blinded quality control samples to monitor the accuracy of measurements conducted as part of research studies can be provided.

A-078 Harmonization of Immunoassay Reference Intervals with LC-MS/MS for Total and Calculated Free Testosterone Measurement

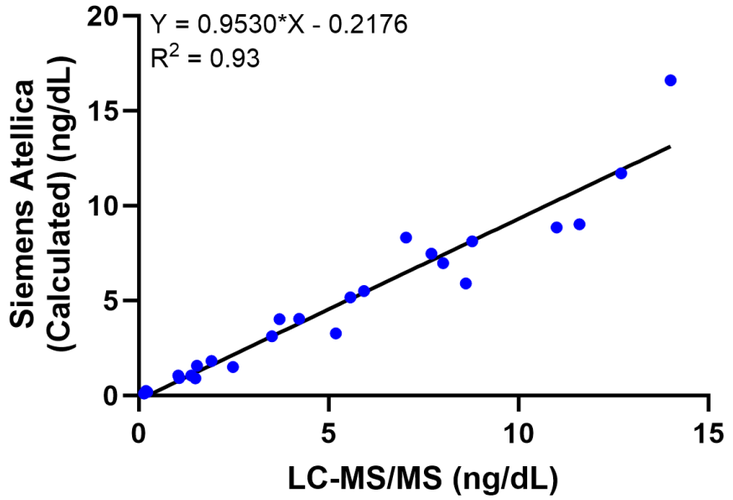
S Li, P Povse, F Lattanzio, K Patel, R Kerlin, J Jones

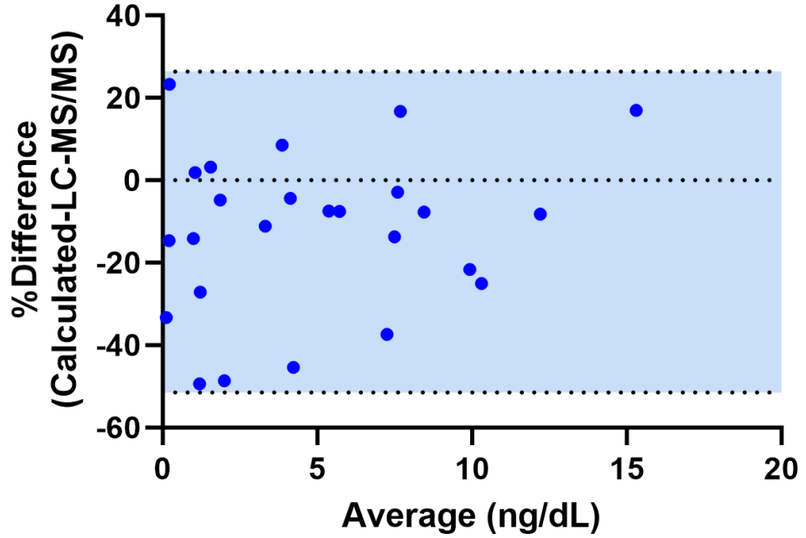
Clinical Chemistry, Volume 70, Issue Supplement_1, October 2024, hvae106.077,
https://doi.org/10.1093/clinchem/hvae106.077
Published: 02 October 2024





Total testosterone





Free testosterone

Updated Reference Intervals

> or = 19 years	Male (ng/dL)	Female (ng/dL)
Total testosterone	240-950	8-60
Free testosterone	2.29-20.70	0.00-1.08

ENDOCRINE SOCIETY

Landmark Study Defines Normal Ranges for Testosterone Levels

Washington, DC | January 10, 2017

New defined reference range can help limit misdiagnoses and unnecessary treatments

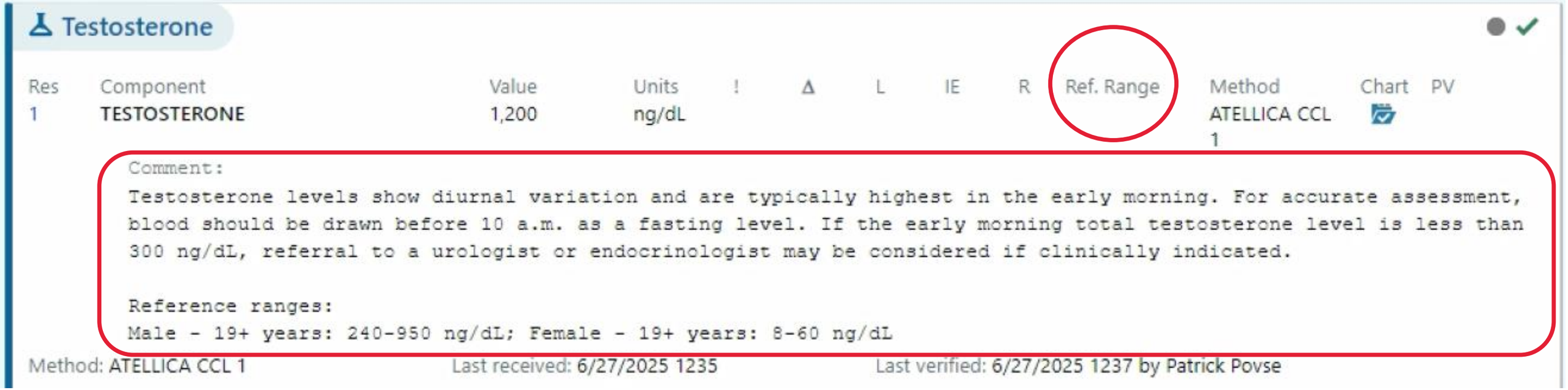
264-916 ng/dL

American Urological Association

Evaluation and Management of Testosterone Deficiency (2024)

Clinicians should use a total testosterone level below **300 ng/dL** as a reasonable cut-off in support of the diagnosis of low testosterone.

Updated Reference Intervals



Testosterone

Res	Component	Value	Units	!	Δ	L	IE	R	Ref. Range	Method	Chart	PV
1	TESTOSTERONE	1,200	ng/dL							ATELLICA CCL 1		

Comment:
Testosterone levels show diurnal variation and are typically highest in the early morning. For accurate assessment, blood should be drawn before 10 a.m. as a fasting level. If the early morning total testosterone level is less than 300 ng/dL, referral to a urologist or endocrinologist may be considered if clinically indicated.

Reference ranges:
Male - 19+ years: 240-950 ng/dL; Female - 19+ years: 8-60 ng/dL

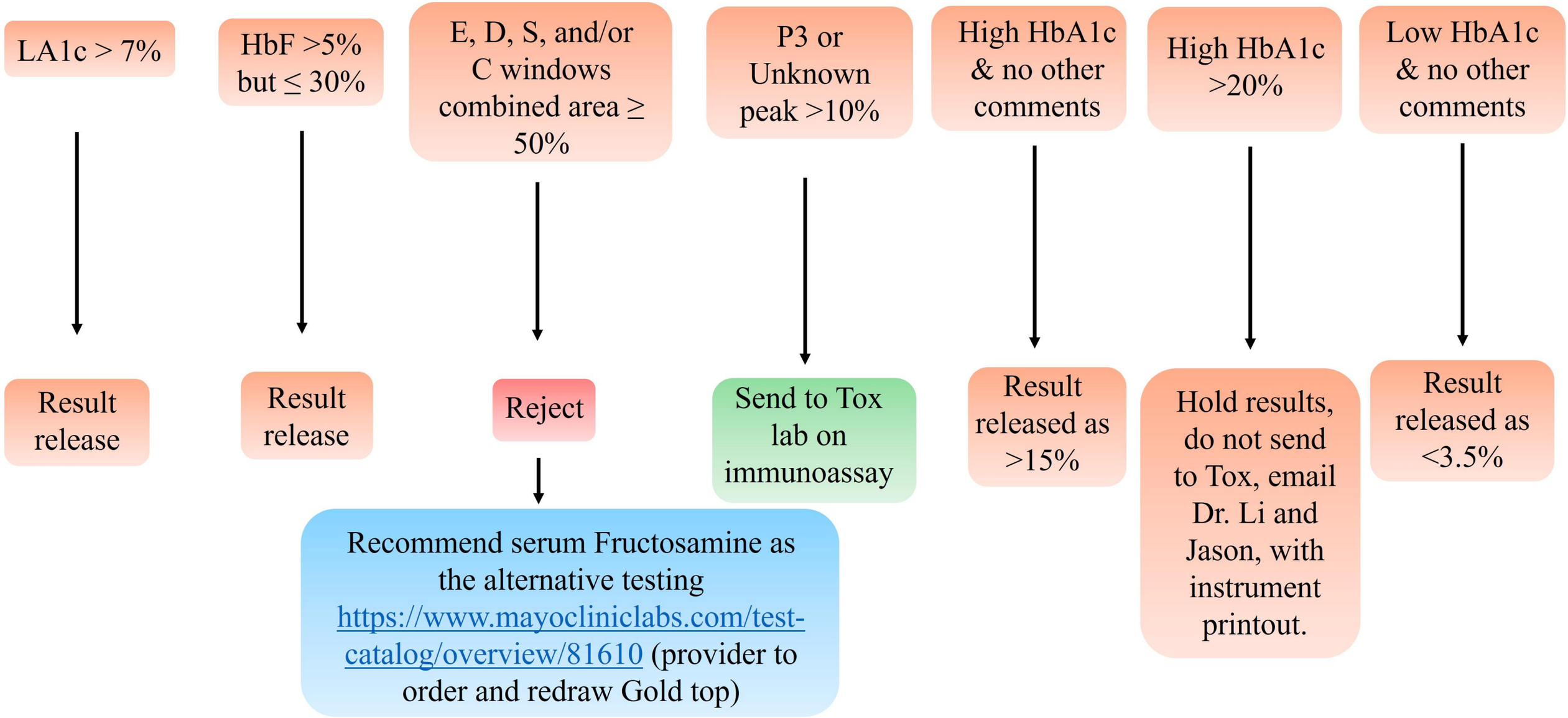
Method: ATELLICA CCL 1 Last received: 6/27/2025 1235 Last verified: 6/27/2025 1237 by Patrick Povse

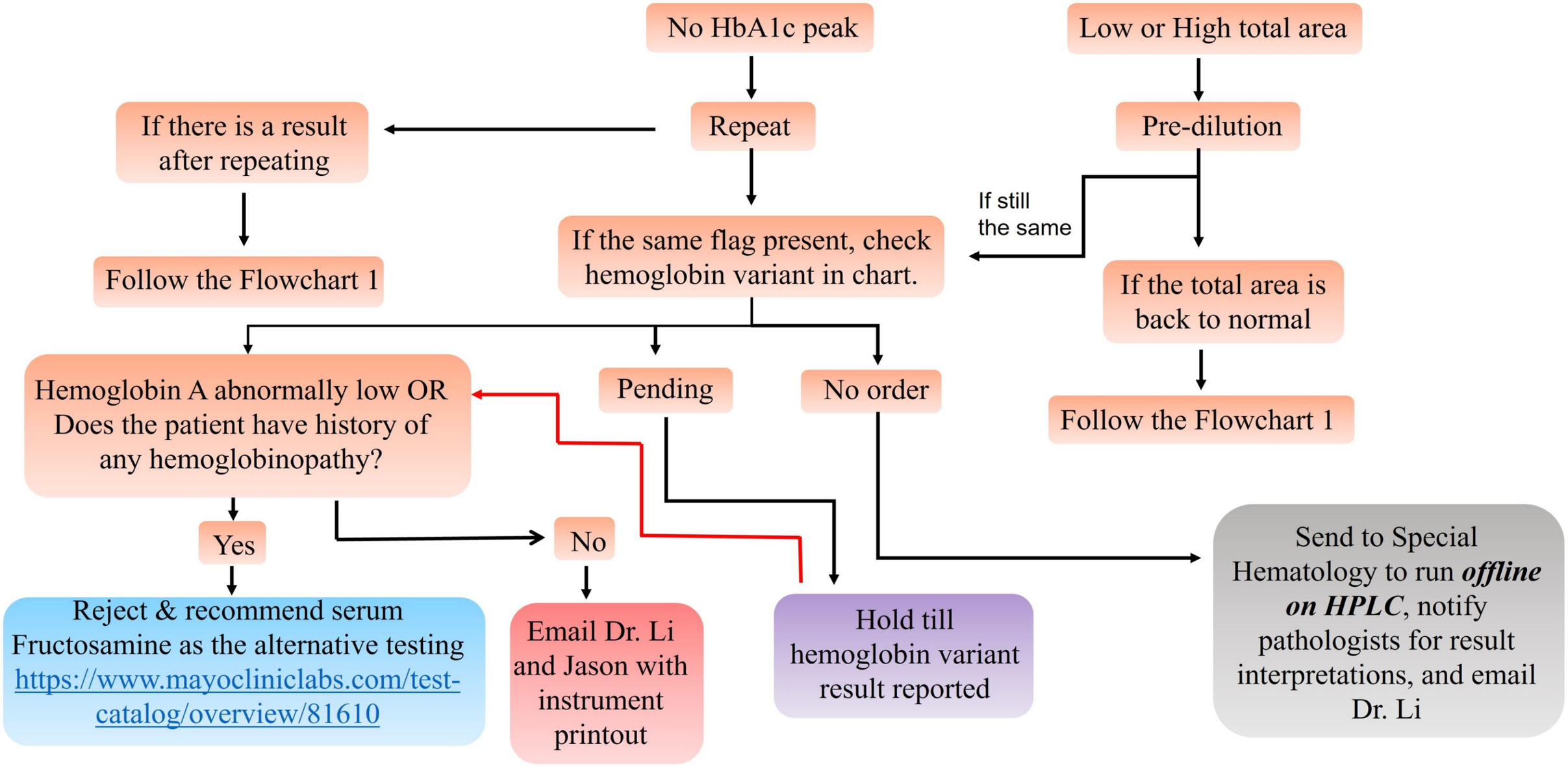
Case 4-Reducing HbA1c Rejection Rates Through Algorithm-Driven Decision Support

- **Challenge**: The HbA1c assay on the HPLC platform exhibited a high rejection rate (30%), leading to clinician complaints, repeat testing, and diagnostic delays. Flags on the instrument—some due to actual hemoglobin variant interference, others due to instrument limitations—make it difficult for staff to determine which results were reliable.
- **Action**: Reviewed 3 months of flagged results and categorized rejection causes. Developed and implemented a laboratory workflow algorithm to guide staff on when to release, hold, or investigate flagged results. This also included clear communication protocols with clinicians and recommendations for alternative testing (e.g., fructosamine) when variant interference was suspected.

Hb A1c Rejection Rate – 30%

Flags on the instrument	% (3 months before)
No HbA1c results	2
Possible variant interference	32
High HbA1c	11
Low HbA1c	6





Hb A1c Rejection Rate – 21%

Flags on the instrument	% (3 months before)	% (3 months after)
No HbA1c results	2	2
Possible variant interference	32	18
High HbA1c	11	7
Low HbA1c	6	3

Case 5 –Not All Insulin Tests Are Equal: Implications for Factitious Hypoglycemia Workups

- Dr. Li – I have a 62-year-old patient with type 2 diabetes who has been treated with Lispro. Recently, he has been experiencing recurrent episodes of symptomatic hypoglycemia, primarily occurring in the late afternoon, despite reduced insulin doses. His laboratory results show abnormally high insulin levels that seem inconsistent with his dosing schedule. I am wondering if your method can detect Lispro specifically, as I need to confirm whether the elevated insulin is due to exogenous Lispro or if there is unexpected endogenous insulin production. Assessing his endogenous insulin status would be very helpful in guiding further management.

Case 5 –Not All Insulin Tests Are Equal: Implications for Factitious Hypoglycemia Workups

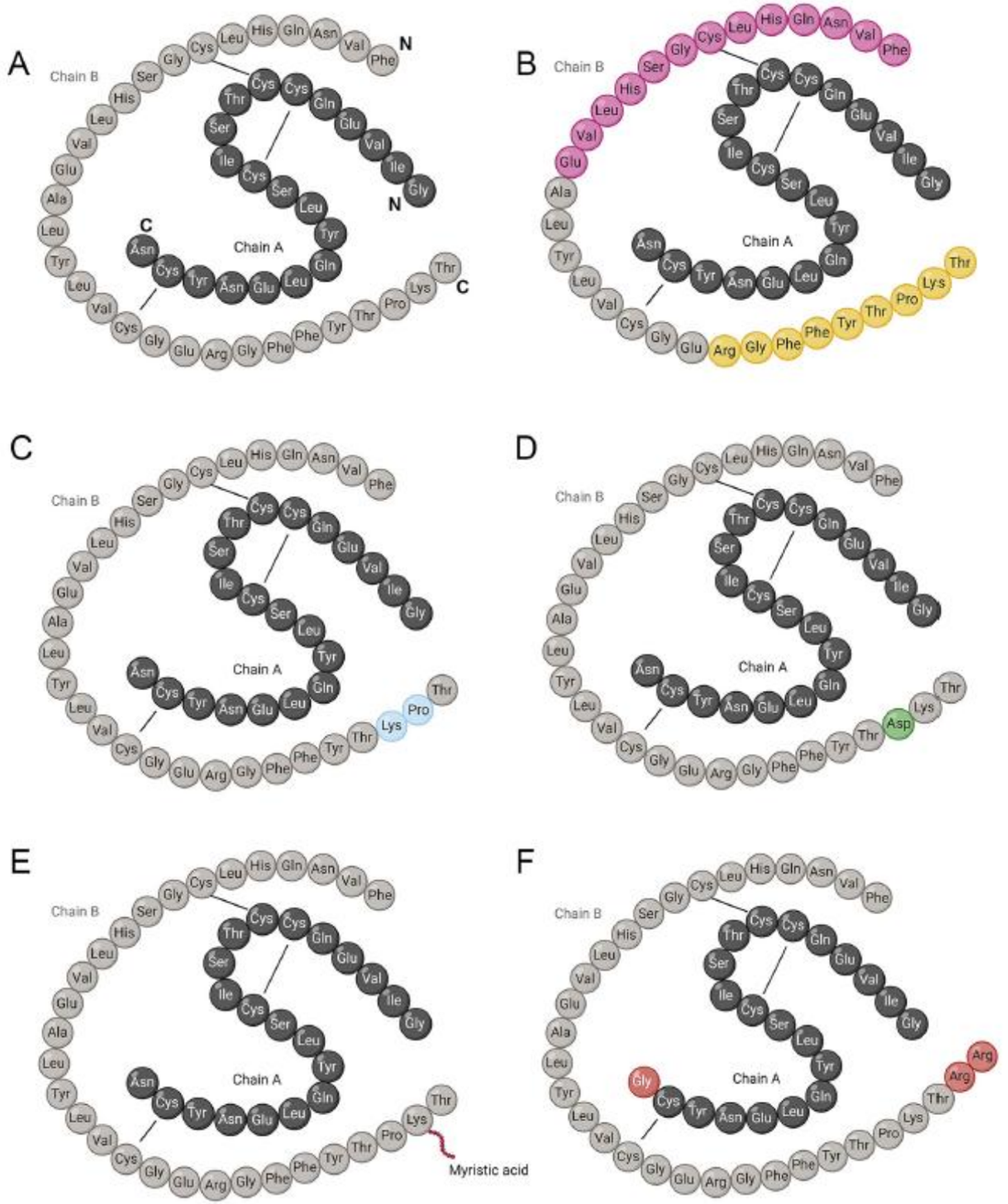
- **Challenge**: Inconsistent detection of insulin analogs across different immunoassay platforms has contributed to diagnostic uncertainty in suspected cases of factitious hypoglycemia, potentially leading to misdiagnosis or delayed care.

- **Action**: Conducted a comparative study evaluating the performance of our immunoassay versus LC-MS/MS. Five insulin analogs were spiked into serum at clinically relevant concentrations and tested to assess cross-reactivity and recovery. Results were analyzed against recombinant human insulin controls.

• *Endocrine. 2024 Jul 19. doi: 10.1007/s12020-024-03970-6.*

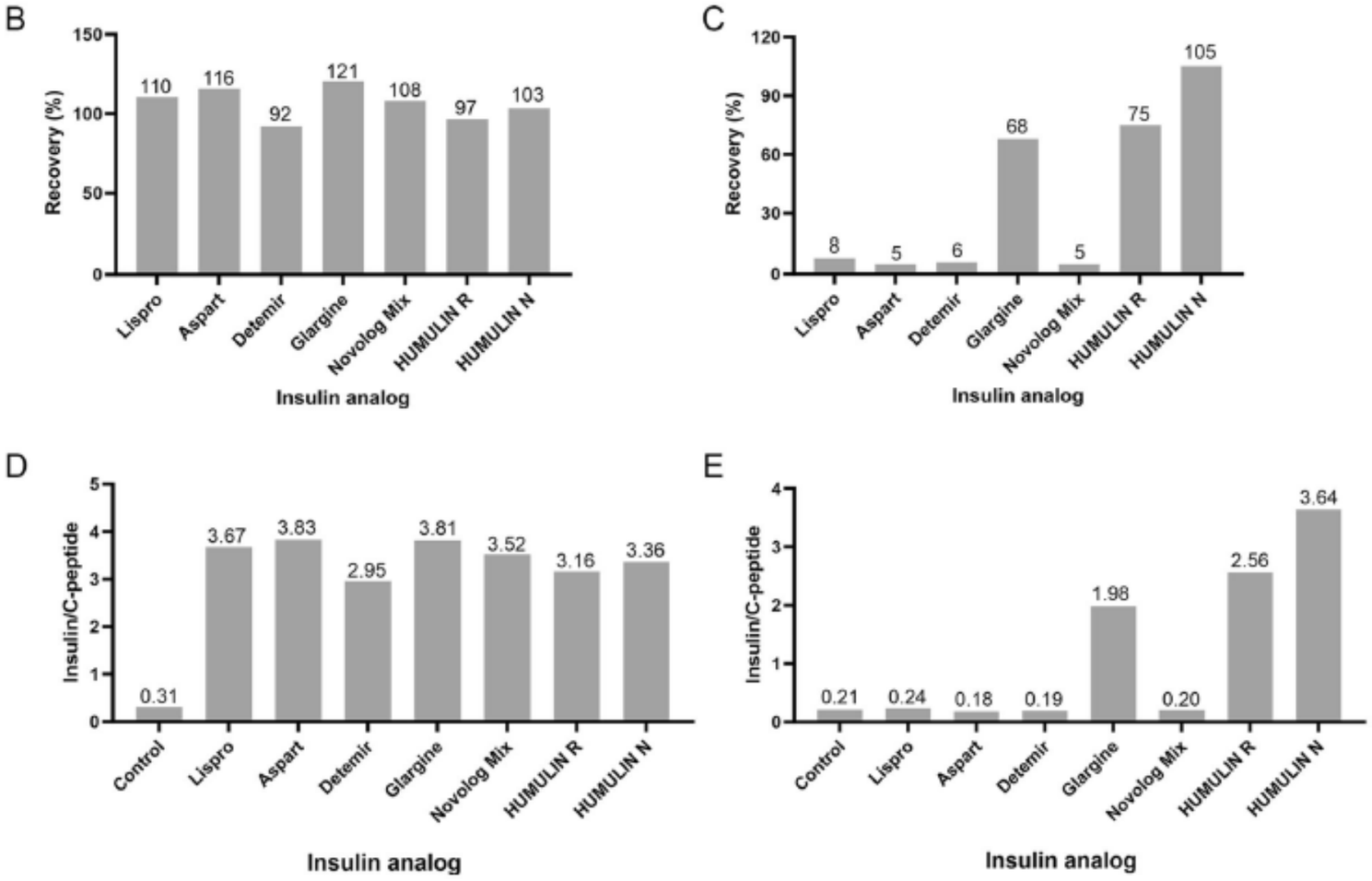
Insights into insulin analog cross-reactivity: a comparative study of Siemens Atellica and LC-MS/MS

Jieli Li[1] · Maya Hatten-Beck[2,3] · Jason K. Y. Lee[4] · Andrew N. Hoofnagle[2,3]



Insights into insulin analog cross-reactivity: a comparative study of Siemens Atellica and LC-MS/MS

Jieli Li[1] · Maya Hatten-Beck[2,3] · Jason K. Y. Lee[4] · Andrew N. Hoofnagle[2,3]



Endocrine. 2024 Jul 19. doi: 10.1007/s12020-024-03970-6.

Case 6 - Establishing Clinically Meaningful Calcitonin Thresholds Post-Thyroidectomy

- Hi Shirley,

➢ I wanted to discuss the calcitonin assay. I have several patients with medullary thyroid cancer (MTC) and MEN2.

➢ According to the ATA guidelines, imaging should be followed up if calcitonin is detectable in MTC patients after thyroidectomy. However, I now have more patients with detectable calcitonin but no evidence of recurrence after imaging.

➢ Should we reconsider the cutoff for detectable calcitonin, or is further investigation into these cases warranted?

Case 6 - Establishing Clinically Meaningful Calcitonin Thresholds Post-Thyroidectomy

- **Challenge**: Following implementation of the new immunoassay platform with increased analytical sensitivity for calcitonin, more patients with medullary thyroid carcinoma (MTC) had detectable calcitonin levels post-thyroidectomy. According to ATA guidelines, detectable calcitonin warrants imaging, but many of these patients showed no evidence of recurrence. This discrepancy raised concerns about unnecessary follow-ups and clinical confusion.

- **Action**: Conducted a retrospective review of 56 samples from 40 post-thyroidectomy patients. Evaluated calcitonin levels alongside imaging studies and carcinoembryonic antigen (CEA) results to determine a clinically meaningful threshold using the new assay.

Original Article

Defining the Functional Sensitivity for the Siemens Atellica Calcitonin Assay: Insight From a Single-Center Study

Jieli Li, MD, PhD [1,*], Ashley Patton, MD [1], Jason K.Y. Lee, MLSMS [2], Matt Scheidegger, MT [2], Irina Azaryan, MD [3], Jennifer A. Sipos, MD [3], Fadi Nabhan, MD [3], JoAnna Jones, MD [1], Alicia Algeciras-Schimnich, PhD [4], Matthew D. Ringel, MD [3,5,**]

Calcitonin and Imaging Results

Patient #	Age (y.o.)	Sex	Date of calcitonin results	Calcitonin result (pg/mL)	CEA result (ng/mL)	Neck US	Neck CT w/wo contrast	Chest CT w/wo contrast	Abdomen/Pelvis CT w/wo contrast	Neck MRI	Abdomen/ Pelvis MRI	Evidence of disease
1	45	Male	1/20/2023	4.31	NA	NA	neg	NA	NA	NA	NA	No
2	41	Female	6/22/2023	2.24	NA	neg	NA	NA	NA	NA	NA	No
3	60	Female	7/18/2023	3.25	<2.0	neg	NA	NA	NA	NA	NA	No
4	40	Female	4/5/2023	2.69	<2.0	neg	NA	NA	NA	NA	NA	No
5	72	Female	1/10/2023	<1.89	**32.1**	NA	neg	stable T12 metastasis	neg	NA	NA	Yes
5	72	Female	4/4/2023	<1.89	**18.5**	NA	neg	stable T12 metastasis	neg	NA	NA	Yes
5	72	Female	6/28/2023	2.12	**20.20**	NA	neg	stable T12 metastasis	neg	NA	NA	Yes
6	66	Female	2/7/2023	<1.89	2.2	NA	neg	neg	neg	NA	NA	No
6	66	Female	5/2/2023	2.77	<2.0	NA	neg	neg	neg	NA	NA	No
6	66	Female	7/25/2023	4.4	<2.0	NA	neg	neg	neg	NA	NA	No
7	70	Male	7/26/2023	3.82	**5.4**	neg	NA	NA	NA	NA	NA	No
8	40	Female	11/9/2022	<1.89	<2.0	neg	NA	NA	NA	NA	NA	No
8	40	Female	4/26/2023	3.21	<2.0	neg	NA	NA	NA	NA	NA	No
9	26	Female	3/21/2023	3.05	<2.0	neg	NA	NA	NA	NA	NA	No
10	83	Male	11/8/2022	<1.89	3.5	NA	Stable osseous	neg	NA	NA	NA	Yes

Endocr Pract. 2024 Nov;30(11):1044-1050.

Original Article

Defining the Functional Sensitivity for the Siemens Atellica Calcitonin Assay: Insight From a Single-Center Study

Jieli Li, MD, PhD [1,*], Ashley Patton, MD [1], Jason K.Y. Lee, MLSMS [2], Matt Scheidegger, MT [2], Irina Azaryan, MD [3], Jennifer A. Sipos, MD [3], Fadi Nabhan, MD [3], JoAnna Jones, MD [1], Alicia Algeciras-Schimnich, PhD [4], Matthew D. Ringel, MD [3,5,**]

Highlights

- Calcitonin 1.89 pg/mL: 43% sensitivity, 67% specificity for post-thyroidectomy medullary thyroid carcinoma (MTC).
- Calcitonin at 5 pg/mL cutoff: 0% sensitivity, 100% specificity in same patients.
- Laboratories should establish their own clinical cutoffs.
- Accurate cutoffs crucial for monitoring MTC persistence/recurrence, enabling timely management.

Clinical Relevance

Establishing accurate calcitonin cutoffs is important for monitoring persistent or recurrent medullary thyroid carcinoma post-thyroidectomy, supporting timely, and precise clinical decisions.

Endocr Pract. 2024 Nov;30(11):1044-1050.

Recommendations for Practice

- Periodically re-evaluate diagnostic thresholds
- Engage clinicians early and align diagnostic criteria
- Use published guidelines and peer benchmarking
- Build infrastructure to document, track, and audit threshold changes

Thank you