Лабораторная служба
2024, Т. 13, №3, с.50–60
https://doi.org/10.17116/labs20241303150

Laboratory Service
2024, vol. 13, no. 3, pp. 50–60
https://doi.org/10.17116/labs20241303150

Должны ли мы зависеть от референтных интервалов, указанных в инструкции производителя? Сравнение референтных интервалов тиреотропного гормона и свободного тироксина от четырех производителей с результатами современных непрямых методов и прямого метода

© НИК Ф. ДИРКС, ВЕНДИ П.ДЖ. ДЕН ЭЛЬЗЕН, ЖАКЛИН ДЖ. ХИЛЛЕБРАНД, ХЕЛЕЕН И. ЯНСЕН, ЭДВИН ТЕН БОЕКЕЛЬ, ЖАКОЛИН БРИНКМАН, МАДЕЛОН М. БУЙС, АЙШЕ Ю. ДЕМИР, ИНЕКЕ М. ДЕЙКСТРА, СИЛЬВИЯ К. ЭНДЕНБУРГ, ПАУЛА ЭНГБЕРС, ЖАННЕТТ ГУТДЖЕС, МАРСЕЛЬ ДЖ.В. ЯНССЕН, ВИЛЬГЕЛЬМИНА Х.А. КНИСТ-ДЕ ЙОНГ, МААРТЕН Б. КОК, СТЕФАН КАМФУИС, АДРИАН КРУИТ, ЭТЬЕН МИШЕЛЬСЕН, АЛЬБЕРТ ВОЛТУИС, АНИТА БОЕЛЕН, АННЕМИЕКЕ К. ХЕЙБОЕР

РЕЗЮМЕ

Цель исследования. Правильная интерпретация лабораторных тестов щитовидной железы зависит от верности референтных интервалов (РИ) для тиреотропного гормона (ТТГ) и свободного тироксина (FT4). ISO15189 требует периодического пересмотра РИ, но в лабораторном сообществе наблюдаются трудности с экономически эффективными подходами. В статье исследована возможность заменить прямой подход (с использованием здоровой референсной популяции) непрямыми методами (с использованием исторических лабораторных данных) и проведено сравнение результатов с предоставленными производителями РИ для ТТГ и FT4.

Материал и методы. Мы собрали исторические данные (2008—2022 гг.) из 13 лабораторий Нидерландов, чтобы установить РИ с использованием непрямых методов, TMC (для ТТГ) и RefineR (для FT4). В лабораториях использовались тест-системы распространенных производителей (Roche, Abbott, Beckman и Siemens). РИ с помощью непрямого метода определялись для каждой лаборатории и для каждого года и группировались по каждому производителю (>1 000 000 точек данных на каждого производителя). Прямые РИ были установлены у 125 здоровых лиц на каждой платформе.

Результаты. Уровень РИ ТТГ с расчетом непрямым методом на протяжении многих лет оставался стабильным для всех производителей. Для FT4 непрямой расчет оказался надежным для трех производителей (Roche, Beckman и Siemens), но верхний референтный предел (ВРП) производителя Abbott показал снижение на 2 пмоль/л по сравнению с 2015 г. Сравнение РИ, определенных прямым и непрямым методом для ТТГ и FT4, показало близкое совпадение при адекватно использованной возрастной стратификации. РИ от производителя, в частности Abbott, Roche и Beckman, показали менее адекватные значения ВРП (общая разница 0,5—1,0 мкМЕ/мл) для ТТГ. Для FT4 ВРП, предоставленные Roche, Abbott и Siemens, были завышены на 1,5—3,5 пмоль/л.

Выводы. Результаты проведенного исследования подчеркивают важность верификации РИ, поскольку предоставленные производителем РИ часто не адекватны популяции, исследуемой в лаборатории, и могут быть ненадежными. Непрямые методы предлагают экономически эффективную альтернативу для верификации РИ в конкретной лаборатории или на конкретной тест-системе.

Ключевые слова: *прямой метод, непрямой метод, референтные интервалы, заболевание щитовидной железы.*

Автор, ответственный за переписку: Аннемике К. Хейбоер — e-mail: a.heijboer@amsterdamumc.nl
Перевод: Ольга Плеханова

КАК ЦИТИРОВАТЬ:
Ник Ф. Диркс, Венди П.Дж. ден Эльзен, Жаклин Дж. Хиллебранд, Хелеен И. Янсен, Эдвин тен Боекель, Жаколин Бринкман, Маделон М. Буйс, Айше Ю. Демир, Инеке М. Дейкстра, Сильвия К. Энденбург, Паула Энгберс, Жаннетт Гутджес, Марсель Дж.В. Янссен, Вильгельмина Х.А. Книст-де Йонг, Маартен Б. Кок, Стефан Камфуис, Адриан Круит, Этьен Мишельсен, Альберт Волтуис, Анита Боелен, Аннемиеке К. Хейбоер. Должны ли мы зависеть от референтных интервалов, указанных в инструкции производителя? Сравнение референтных интервалов тиреотропного гормона и свободного тироксина от четырех производителей с результатами современных непрямых методов и прямого метода. *Лабораторная служба*. 2024;13(3):50–60. https://doi.org/10.17116/labs20241303150

Should we depend on reference intervals from manufacturer package inserts? Comparing TSH and FT4 reference intervals from four manufacturers with results from modern indirect methods and the direct method

© NIEK F. DIRKS, WENDY P.J. DEN ELZEN, JACQUELIEN J. HILLEBRAND, HELEEN I. JANSEN, EDWIN TEN BOEKEL, JACOLINE BRINKMAN, MADELON M. BUIJS, AYSE Y. DEMIR, INEKE M. DIJKSTRA, SILVIA C. ENDENBURG, PAULA ENGBERS, JEANNETTE GOOTJES, MARCEL J.W. JANSSEN, WILHELMINA H.A. KNIEST-DE JONG, MAARTEN B. KOK, STEPHAN KAMPHUIS, ADRIAN KRUIT, ETIENNE MICHIELSEN, ALBERT WOLTHUIS, ANITA BOELEN, ANNEMIEKE C. HEIJBOER

ABSTRACT

Objective. Correct interpretation of thyroid function tests relies on correct reference intervals (RIs) for thyroid-stimulating hormone (TSH) and free thyroxine (FT4). ISO15189 mandates periodic verification of RIs, but laboratories struggle with cost-effective approaches. We investigated whether indirect methods (utilizing historical laboratory data) could replace the direct approach (utilizing healthy reference individuals) and compared results with manufacturer-provided RIs for TSH and FT4.

Material and methods. We collected historical data (2008—2022) from 13 Dutch laboratories to re-establish RIs by employing indirect methods, TMC (for TSH) and refineR (for FT4). Laboratories used common automated platforms (Roche, Abbott, Beckman or Siemens). Indirect RIs (IRIs) were determined per laboratory per year and clustered per manufacturer (>1,000.000 data points per manufacturer). Direct RIs (DRIs) were established in 125 healthy individuals per platform.

Results. TSH IRIs remained robust over the years for all manufacturers. FT4 IRIs proved robust for three manufacturers (Roche, Beckman and Siemens), but the IRI upper reference limit (URL) of Abbott showed a decrease of 2 pmol/L from 2015. Comparison of the IRIs and DRIs for TSH and FT4 showed close agreement using adequate age-stratification. Manufacturer-provided RIs, notably Abbott, Roche and Beckman exhibited inappropriate URLs (overall difference of 0.5—1.0 μIU/mL) for TSH. For FT4, the URLs provided by Roche, Abbott and Siemens were overestimated by 1.5—3.5 pmol/L.

Conclusions. These results underscore the importance of RI verification as manufacturer-provided RIs are often incorrect and RIs may not be robust. Indirect methods offer cost-effective alternatives for laboratory-specific or platform-specific verification of RIs.

Keywords: *direct method, indirect method, reference intervals, thyroid disease.*

Corresponding author: Annemieke C. Heijboer — e-mail: a.heijboer@amsterdamumc.nl

TO CITE THIS ARTICLE:
Niek F Dirks, Wendy PJ den Elzen, Jacquelien J Hillebrand, Heleen I Jansen, Edwin ten Boekel, Jacoline Brinkman, Madelon M Buijs, Ayse Y Demir, Ineke M Dijkstra, Silvia C Endenburg, Paula Engbers, Jeannette Gootjes, Marcel JW Janssen, Wilhelmina HA Kniest-de Jong, Maarten B Kok, Stephan Kamphuis, Adrian Kruit, Etienne Michielsen, Albert Wolthuis, Anita Boelen, Annemieke C Heijboer. Should we depend on reference intervals from manufacturer package inserts? Comparing TSH and FT4 reference intervals from four manufacturers with results from modern indirect methods and the direct method. *Laboratory Service*. 2024;13(3):50–60. (In Russ.). https://doi.org/10.17116/labs20241303150

Введение

Эутиреоз определяется как нормальная функция щитовидной железы, при которой концентрации тиреотропного гормона (ТТГ) и свободного тироксина (FT4) находятся в пределах референтного интервала (РИ). Другие комбинации ТТГ и FT4 могут обозначать первичный гипотиреоз (повышение ТТГ и снижение FT4) или гипертиреоз (снижение ТТГ и повышение FT4).

Очевидно, что правильный диагноз во многом зависит от правильного определения РИ, предоставленного лабораторией.

Согласно принципам хорошей лабораторной практики, лаборатории устанавливают свои собственные нижние и верхние референтные пределы (НРП и ВРП) для РИ, адаптированные к конкретному используемому методу и популяции пациентов. Институт клинических лабораторных стандартов (CLSI) предоставляет руководство (CLSI C28-A3c) по определению РИ [1].

Однако определение РИ в соответствии с данным руководством, с участием как минимум 120 подходящих референтных лиц, требует времени и средств. Вместо этого лаборатории часто полагаются на РИ из инструкций производителя или на данные другой лаборатории. После этого при переходе на новую методологию существующие РИ часто корректируются после сравнительного исследования методов.

Вышеупомянутое руководство содержит рекомендации по верификации РИ, установленных производителем или другой лабораторией.

К сожалению, РИ, установленные извне, не всегда подлежат переносу из-за различий в преаналитической, аналитической и биологической вариациях [2, 3, 4]. Обеспечение пригодности РИ требует их соответствия с внешней тест-системой и референсной популяции с текущей тест-системой и популяцией пациентов.

Сопоставимость тест-систем можно оценить посредством сравнительного исследования методов в соответствии с CLSI EP09 [5]. Сравнимость популяции испытуемых более субъективна и может быть объективно проверена только путем верификации установленного извне РИ в популяции испытуемых. CLSI C28-A3c рекомендует верифицировать новый РИ с помощью референсной популяции, состоящей как минимум из 20 образцов, и принимать устанавливаемый РИ в случае, если не более 2 из 20 значений выходят за пределы предложенного РИ. Следует отметить, что, когда все 20 эталонных значений попадают в предложенный РИ, он может оказаться слишком широким для популяции исследуемых. При этом риск принятия слишком широкого РИ не минимизируется.

Альтернативный подход с использованием современной методологии заключается в установлении или верификации РИ с использованием существующих исторических лабораторных данных [3]. Использование непрямых методов на основе исторических данных дает возможность получить расчетные РИ, которые отражают используемую аналитическую лабораторную тест-систему и актуальную исследуемую популяцию. Кроме того, при условии наличия достаточных исторических данных, оно позволяет проводить более обширное разделение (например, по полу и возрасту), что зачастую является слишком дорогостоящим для лабораторий при использовании прямого метода. В последние годы было разработано некоторое количество современных методов непрямого определения РИ, преодолевших ограничения известных методов Бхаттачарая и Хоффманна [6].

В недавнем исследовании T. Ammer и соавт. сравнили метод Хоффманна с четырьмя современными непрямыми

методами: усеченным методом максимального правдоподобия (TML), kosmic, усеченным минимальным хи-квадратом (TMC) и refineR, а также непараметрическим прямым методом с использованием 120 референсных образцов [7]. TMC показал наилучшие результаты для ТТГ, для которого характерно асимметричное распределение, тогда как refineR показал лучшие результаты для FT4 с нормальным распределением результатов исследований. При доле результатов, не входящих в «нормальный» диапазон (патологическая фракция) 10% или менее для ТТГ и FT4, TMC и refineR даже превзошли прямой метод (с использованием 120 референсных индивидуумов).

После установления РИ, лаборатории в соответствии со стандартом ISO 15189 обязаны периодически проводитть их пересмотр, частота которого фактически сводится к пересмотру во время перехода на новый метод. Поэтому была исследована надежность РИ ТТГ и FT4 за последние 15 лет, чтобы определить, необходима ли более частая верификация РИ. Использовались два непрямых метода оценки РИ: TMC (для ТТГ) и RefinR (для FT4) на исторических данных (2008—2022 гг.). Кроме того, была оценена пригодность РИ ТТГ и FT4 с помощью непрямых методов относительно РИ с использованием прямого метода.

Материал и методы

Сбор проб и данных

Прямое определение

ТТГ и FT4 измеряли в аликвотах сыворотки крови, полученных от 125 референсных индивидуумов (здоровых сотрудников Амстердамского научно-исследовательского медицинского центра, в том числе 54 мужчин и 71 женщин, в возрасте 19—68 лет). Эти образцы были собраны в период с сентября 2022 г. по февраль 2023 г. У 69 из 125 контрольных лиц также были доступны аликвоты гепаринизированной плазмы. От каждого участника было получено письменное согласие. Образцы сыворотки и гепаринизированной плазмы собирали, центрифугировали при 3000×g в течение 10 мин при 18°C, аликвотировали и хранили при –80°C. Срок хранения не превышал 10 мес.

Непрямое определение РИ

Все 13 участвующих лабораторий расположены в Нидерландах, территории с достаточным содержанием йода. Эти лаборатории обслуживают в общей сложности 28 больниц по всей территории Нидерландов и предоставляют дополнительные лабораторные услуги, включая флеботомию, более чем в 500 различных кабинетах по всей стране.

Результаты непрямого определения РИ ТТГ и FT4 за период с 2008 по 2022 г. были получены из лабораторной информационной системы (ЛИС) участвующими лабораториями и перенесены в таблицы Excel вместе с анонимным и уникальным идентификатором пациента, уникальным идентификатором образца, датой проведения анализа, полом и возрастом. Нечисловые результаты и результаты пациентов в возрасте до 18 лет и старше 100 лет были исключены. Чтобы уменьшить процент патологических результатов в наборах данных, все результаты пациента исключались из анализа, если для этого пациента в течение одного года было получено несколько результатов (повторные измерения). Если лаборатории применяли факторы к результатам ТТГ или FT4, эти коэффициенты отменялись перед расчетом.

В **табл. 1** представлены обзор аналитических систем, примерный процент от общего количества результатов, количество определений ТТГ и FT4, а также распределение по возрасту и полу результатов в 2022 г. для каждой из участвующих лабораторий. Следует отметить, что некоторые лаборатории сменили производителя и могут быть также включены в анализ по производителю, отличному от того, который указан в **табл. 1**.

Аналитические процедуры

Все участвующие лаборатории были аккредитованы в соответствии с ISO 15189 (или CCKL (Национальный координационный комитет по обеспечению качества лабораторий здравоохранения в Нидерландах) в прошлом) и проводили внутреннюю и внешнюю оценку качества. Концентрации ТТГ и FT4 определялись на их тест-системах с использованием соответствующих реагентов и калибраторов в соответствии с инструкциями производителя. Образцы крови были взяты путем венепункции в соответствии с местными инструкциями. Во всех лабораториях использовались пробирки с гепаринизированной плазмой для определения ТТГ и FT4.

Референтные интервалы, определенные прямым способом

Концентрации ТТГ и FT4 анализировали с использованием Roche Cobas e601, Abbott Alinity I и Siemens Atellica IM 1300 в лаборатории эндокринологии Амстердамского научно-исследовательского центра. В больнице Ниджа Смеллингхе анализировали концентрации ТТГ и FT4, полученные на тест-системе Beckman Unicel DxI. Расчеты прямым способом были основаны на исследовании 125 образцов. РИ рассчитывались с помощью параметрических или непараметрических методов в зависимости от распределения результатов. 2,5-й и 97,5-й процентили и их 90% доверительные интервалы были определены в соответствии с CLSI 28-A3c с использованием статистического программного обеспечения MedCalc v. 18.5 [1]. Учитывая, что во всем мире ТТГ и FT4 обычно измеряются в сыворотке, а все голландские лаборатории, участвовавшие в этом исследовании, использовали образцы гепаринизированной плазмы, было проведено сравнение методов (регрессия Пассинга—Баблока с использованием Analyse-it для Microsoft Excel v. 2.30) между 69 парными сыворотками и гепаринизированными образцами плазмы для определения потенциальных существенных различий между двумя матрицами.

Референтные интервалы, определенные непрямым способом

Методология

Непрямые методы TMC и RefineR основаны на предположении, что большинство включенных данных не яв-

Таблица 1. **Характеристики лабораторий-участников в 2022 г.**
Table 1. **Characteristics in 2022 of participating laboratories**

Лаборатория	Аналитическая система	Количество выполненных тестов (взрослые) ТТГ/FT4	Доля результатов от общего	Возраст испытуемых (взрослые), Me (IQR)	Доля женщин
1	Abbott Architect	>200,000/>100,000	~50%	59,8 (42,3—75,1)	67,9%
2	Abbott Architect	>50,0007/>20,000	~35%	59,9 (43,2—74,2)	67,0%
3	Roche Cobas	>20,0007/>20,000	~10%	58,9 (42,1—72,4)	65,4%
4	Roche Cobas	>200,0007/>50,000	~45%	60,7 (44,4—73,8)	68,2%
5	Roche Cobas	>50,0007/>20,000	~35%	61,1 (45,6—73,5)	64,8%
6	Siemens Atellica	>50,0007/>50,000	~50%	61,8 (45,8—74,9)	65,8%
7	Beckman Unicel Dxl	>200,0007/>50,000	100%	54,5 (37,6—70,0)	69,3%
8	Beckman Unicel Dxl	>20,0007/>5,000	~20%	59,1 (44,0—72,4)	66,0%
9	Abbott Architect	>20,0007/>20,000	~55%	63,0 (49,0—75,0)	73,0%
10	Siemens Atellica	>50,0007/>20,000	~40%	59,3 (42,9—73,1)	64,4%
11	Roche Cobas	>50,0007/>20,000	~30%	63,1 (48,5—75,3)	62,4%
12	Siemens Atellica	>20,0007/>10,000	–30%	58,6 (40,3—72,7)	67,8%
13	Siemens Atellica	>100,0007/>50,000	100%	58,2 (41,0—71,9)	68,4%

Примечание. Здесь и в табл. 2, 3: Me (IQR) — медиана (интерквартильный размах).

ляются патологическими [8, 9]. Оба алгоритма предполагают наличие диапазона результатов исследований, в котором вклад патологических результатов незначителен, а распределение непатологической фракции можно эффективно смоделировать с помощью преобразования нормального распределения Бокса—Кокса. Модели определяются параметром мощности λ, средним значением μ и стандартным отклонением *σ*. Для TMC оценка РИ проводилась согласно руководству по программному обеспечению TMC (версия 13, редакция от 15 ноября 2022 г.). Для RefineR использовался пакет программной среды R RefineR v. 1.6.0. Как для TMC, так и для RefineR для выполнения этапов исключения и непосредственного анализа данных использовалась комбинация статистического программного обеспечения Excel и R (версия 4.2.1, Foundation for Statistical Computing, Вена, Австрия). Для оценки РИ FT4 с помощью RefineR было использовано 200 бутстрепов. Полученные гистограммы с предложенными референтными пределами (РП) от TMC и RefineR были просмотрены визуально и проверены на наличие ошибочно смоделированных распределений. Годы с ошибочным смоделированным распределением были исключены из последующего анализа.

Стратификация

Поскольку популяции для оценки прямым (18—79 лет) и непрямым (18—100 лет) способами были разными, была проведена оценка, требуется ли для ТТГ или FT4 стратификация по возрасту для сравнения прямого и непрямого определения РИ.

Смещение при рефлекс-тестировании

В ряде лабораторий для оценки функции щитовидной железы используют рефлекс-тестирование. Это означает, что измерение FT4 автоматически запускается, когда результат ТТГ выходит за пределы РИ. Выполнение измерения только ТТГ, FT4 или их комбинации независимо от результатов измерения ТТГ оставалось возможным. Чтобы оценить, был ли набор данных FT4 из какой-либо лаборатории смещен в сторону неэутиреоидных состояний из-за более высокого процента лиц с гипотиреозом и гипертиреозом, мы рассчитали РИ для FT4 непрямым способом в 2022 г. на основе данных, содержащих только результаты FT4 от пациентов с нормальным ТТГ (на основании непрямого определения РИ ТТГ в 2022 г.).

Долгосрочная устойчивость референтных интервалов

Пределы эквивалентности (ПЭ), как описано у R. Haeckel и соавт., использовались для сравнения рассчитанных непрямым способом РИ [10]. ПЭ определяются допустимой неопределенностью в два референтных предела РИ и основан на допустимом аналитическом стандартном отклонении, полученном на основе эмпирической биологической вариации (рассчитанной на основе РИ). Годы, в которых РП превышали ПЭ, соответствующие общим медианным референтным пределам, считались ненадежными и были исключены из последующего анализа. В этом случае была определена новая медиана без ненадежных РП. Эту процедуру повторяли до тех пор, пока ни один предел не превысил ПЭ.

Сравнение РИ, полученных непрямым способом и предоставленных производителем, с РИ, определенными прямым способом

Определенные непрямым способом РИ на одного производителя (медиана всех лабораторий) с 2022 г. и предоставленные производителем РИ сравнивались с РИ, определенными прямым способом, с использованием коэффициентов ошибки смещения (bias ratios — BR), как описано у Y. Ozarda и соавт. [11]. BR основаны на допустимом смещении на минимальном уровне и не должны превышать 0,375 (минимальный коэффициент смещения). Также определены желательный BR (<0,250) и оптимальный BR (<0,125).

Таблица 2. **Сравнение непрямых и прямых РИ ТТГ с рекомендованными производителем РИ ТТГ**

Table 2. **Comparison of the indirect and direct TSH RIs with the manufacturer-suggested TSH RIs**

	Прямой метод (2003)				Непрямой метод (2022)						Рекомендованные производителем			
	19—68 лет		18—100 лет				18—60 лет							
	НРП (90% ДИ)	ВРП (90% ДИ)	НРП, Me (IQR)	ВРП, Me (IQR)	BR НРП	BR ВРП	НРП, Me (IQR)	ВРП, Me (IQR)	BR НРП	BR ВРП	НРП	ВРП	BR НРП	BR ВРП
Roche Cobas (n=4)	0,79 (0,60—0,95)	4,79 (4,41—6,46)	0,66 (0,65—0,67)	5,02 (4,86—5,32)	0,127	0,225	0,65 (0,63—0,71)	4,73 (4,64—5,04)	0,137	0,059	0,27	4,20	0,510	0,578
Abbott Alitity (n=3)	0,67 (0,43—0,77)	3,97 (3,82—5,71)	0,55 (0,47—0,56)	4,39 (4,34—5,10)	0,151	0,528	0,56 (0,49—0,57)	4,06 (3,94—4,32)	0,138	0,113	0,35[a]	4,94[a]	0,402	1,219
Siemens Atllica (n=4)	0,76 (0,60—0,89)	4,76 (4,40—6,94)	0,66 (0,65—0,67)	5,09 (5,06—5,58)	0,098	0,323	0,66 (0,62—0,67)	4,71 (4,62—5,05)	0,098	0,049	0,55	4,78	0,206	0,020
Beckman Dxl (n=2)	0,71 (0,60—0,83)	4,67 (4,27—6,71)	0,64 (0,61—0,66)	5,33 (5,16—5,49)	0,069	0,653	0,64 (0,61—0,67)	4,84 (4,61—5,06)	0,069	0,168	0,38	5,33	0,327	0,653

Примечание. РИ — референтный интервал. а — Abbott предоставляет РИ для ТТГ на основе центрального 99% интервала вместо обычного центрального 95% интервала. РИ Abbott ТТГ, основанный на центральном 95% интервале, составит 0,47—3,67 мкМЕ/мл. Смещения для этих референтных пределов составляют 0,251 (НРП) и 0,377 (ВРП). $BR_{НРП} = \frac{НРП - НРП_0}{SD_{РИ}}$, $BR_{ВРП} = \frac{ВРП - ВРП_0}{SD_{РИ}}$, $SD_{РИ} = \frac{ВРП_0 - НРП_0}{3{,}92}$, минимальный BR<0,375, желательный BR<0,250, оптимальный BR<0,125. НРП — нижний референтный предел; ВРП — верхний референтный предел; BR — коэффициент смещения; ДИ — доверительный интервал.

Встречи экспертной группы

Была создана группа экспертов, в которую вошли специалисты по клинической химии, обладающие опытом в эндокринологии и статистике, из каждой из участвующих лабораторий. Эта группа неоднократно собиралась для обсуждения методологии и полученных результатов.

Результаты

Определение РИ прямым способом

Результаты исследования концентрации ТТГ, измеренной с помощью тест-системы Roche Cobas, показали очень небольшую относительную разницу (4%) между сывороткой и гепаринизированными образцами, которая считалась клинически незначимой. Результаты прямого определения РИ показаны в **табл. 2**. Как и ожидалось, распределения ТТГ были смещены, и для определения РИ ТТГ прямым способом использовался непараметрический метод. Результаты Roche, Abbott и Beckman показали нормальное распределение для FT4 (критерий Д'Агостино-Пирсона, D'Agostino-Pearson test), поэтому для определения FT4 прямым способом использовался параметрический метод. Результаты FT4, полученные с помощью Siemens Atellica, были ненормально распределены, что потребовало использования непараметрического метода.

Расчетные референтные интервалы

Стабильность данных

Ежемесячные данные скользящих медиан ТТГ указывали на стабильность результатов ТТГ в изучаемые периоды времени почти во всех лабораториях. Данные по FT4 показывают, что результаты FT4 в целом были менее стабильными.

Стратификация

Детальное изучение данных ТТГ и FT4 для непрямого определения РИ каждого возрастного интервала в 10 лет показало необходимость стратификации ТТГ и FT4 с более старшим возрастом. Начиная с 60 лет РИ ТТГ с расчетом непрямым методом превосходили 95% ДИ. Для FT4 расчетные РИ превзошли 95% ДИ от 70 лет и старше. Для сравнения возрастные группы 18—60 лет и 18—100 лет использовались для ТТГ и 18—70 лет и 18—100 лет — для FT4.

Долгосрочная устойчивость референтных интервалов

В период с 2008 по 2022 г. из ЛИС было получено более 7 млн результатов по ТТГ у взрослых, из которых около 55% — от пациентов в возрасте 18—60 лет. Было получено более 2 млн результатов по FT4 у взрослых, из которых около 70% — от пациентов в возрасте 18—70 лет. На **рис. 1, 2** показаны расчетные РИ в возрастной группе 18—60 лет (ТТГ) или 18—70 лет (FT4), сгруппированные по тест-системам исследования (Roche, Abbott, Beckman или Siemens). В некоторых случаях в одном и том же году на графиках скользящих медиан наблюдалось неоднократное превышение РИ, эти годы были исключены из анализа. Все медианные расчетные РИ ТТГ и FT4 Roche, Abbott, Beckman и Siemens

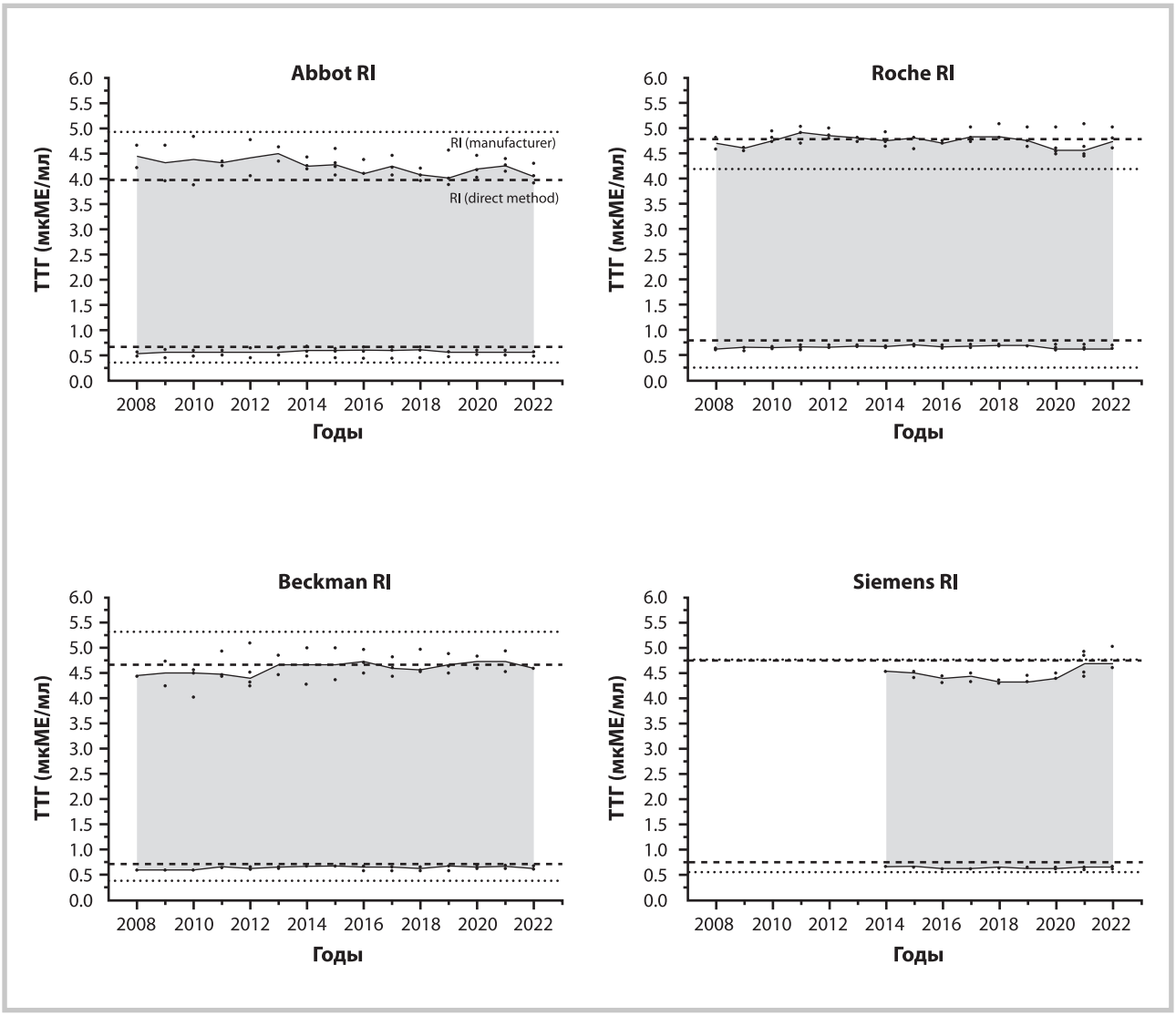


Рис. 1. **Референтные интервалы (РИ, RI) ТТГ, определенные непрямым способом для четырех производителей (Roche, Abbott, Beckman и Siemens) в период 2008—2022 гг.**

Отдельные лабораторные непрямые РИ обозначены одиночными точками. РИ, рассчитанные непрямым способом, определяются медианными годовыми НРП и ВРП, полученными отдельными лабораториями (серая зона). РИ, выбранные производителем, показаны точечными линиями. РИ, оцененные прямым методом, изображены пунктирными линиями.

Fig. 1. **Indirect TSH reference intervals (RIs) for the four manufacturers during the periods 2008—2022 (Roche, Abbott and Beckman) or 2014—2022 (Siemens)**

Individual laboratory indirect RLs are depicted as single dots. Indirect RIs are defined by the median yearly LRLs and URLs from the individual laboratories (grey area). Manufacturer opted RLs are depicted as dotted lines. RIs estimated by the direct method are depicted as dashed lines.

продемонстрировали долгосрочную устойчивость в течение изучаемого периода времени. Расчетные РИ FT4 от Abbott не продемонстрировали постоянной устойчивости. Вместо этого наблюдается снижение в период 2013—2015 гг., как и ожидалось на основе результатов предварительно проведенного анализа стабильности данных.

Сравнение непрямых и прямых референтных интервалов с референтными интервалами, рекомендованными производителем

В **табл. 2** (ТТГ) и **табл. 3** (FT4) показаны результаты сравнения РИ, определенных прямым и непрямым методом, и РИ, рекомендованных производителем. Для визуального сравнения рекомендованные производителем РИ и прямые РИ были включены в виде горизонтальных линий на **рис. 1, 2**. Оценка расчетных РИ ТТГ с помощью непрямого метода в возрастной группе 18—60 лет дала очень схожие результаты по сравнению с расчетом прямым методом с оптимальными или даже желательными значениями BR. Необходимость разделения РИ отражается в улучшенных BR, полученных непрямым методом, в возрастной группе 18—60 лет по сравнению с теми, которые были получены в возрастной группе 18—100 лет. РИ ТТГ, предложенный Siemens, показал хорошее согласие с определенными прямым способом РИ, при этом BR были ниже желательного уровня. Напротив,

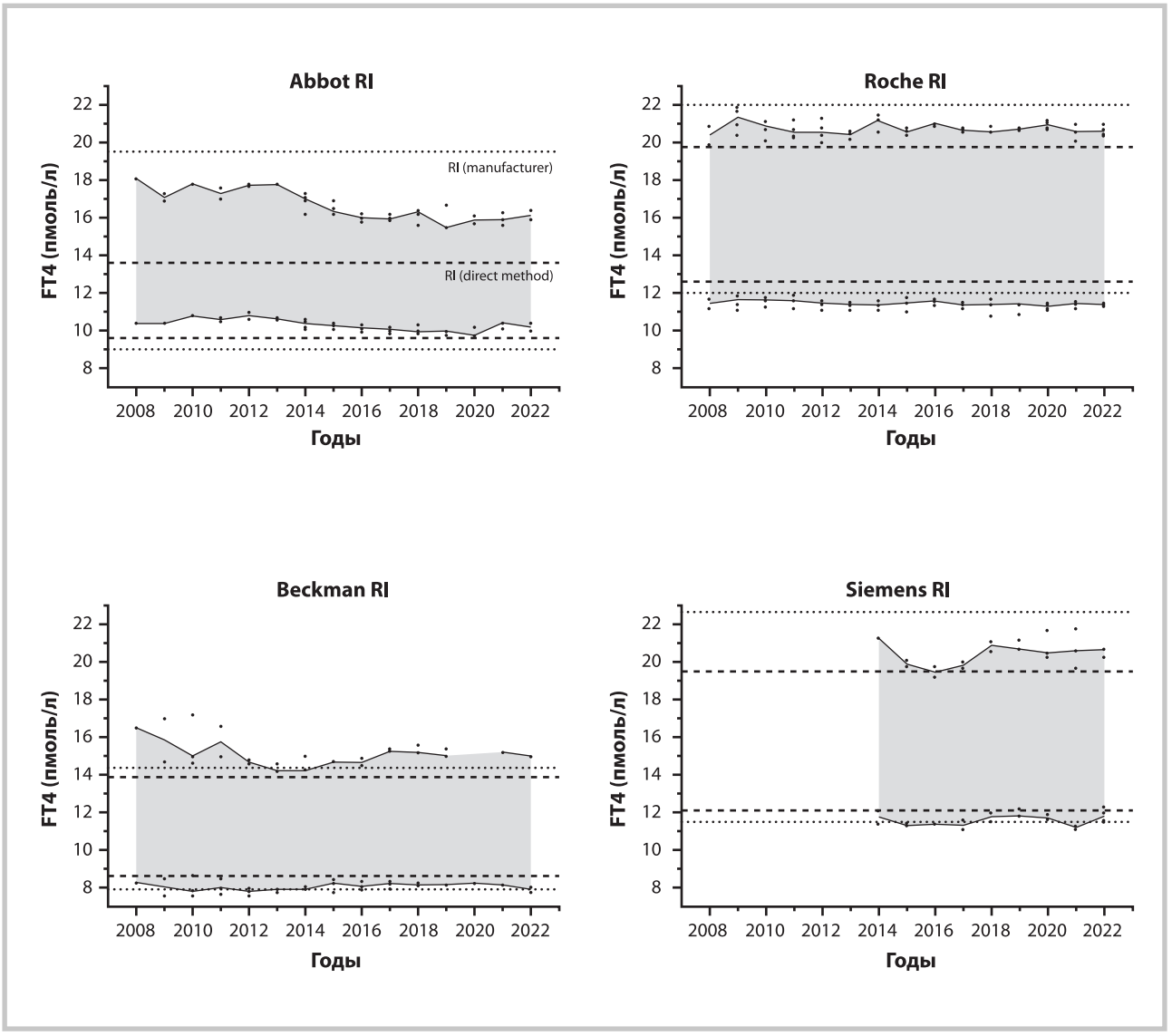


Рис. 2. **Референтные интервалы (РИ, RI) FT4, определенные непрямым способом для четырех производителей (Roche, Abbott, Beckman и Siemens) в период 2008—2022 гг.**

Отдельные лабораторные непрямые РИ обозначены одиночными точками. РИ, рассчитанные непрямым способом, определяются медианными годовыми НРП и ВРП, полученными отдельными лабораториями (серая зона). РИ, выбранные производителем, показаны точечными линиями. РИ, оцененные прямым методом, изображены пунктирными линиями.

Fig. 2. **Indirect FT4 reference intervals (RIs) for the four manufacturers (Roche, Abbott, Beckman and Siemens) during the time period 2008—2022.**

Individual laboratory indirect RLs are depicted as single dots. Indirect RIs are defined by the median yearly LRLs and URLs from the individual laboratories (grey area). Manufacturer opted RLs are depicted as dotted lines. RIs estimated by the direct method are depicted as dashed lines.

расчетные ТТГ, как предположили Roche, Abbott и Beckman, не превосходили минимальные BR по сравнению с определением прямым способом. Определение непрямым способом FT4 в возрастных группах 18—70 лет и 18—100 лет дало более широкие результаты по сравнению с прямым определением. Использование возрастной группы 18—70 лет лишь немного улучшило сравнение. Ни один из РП FT4, полученных непрямым методом, не превзошел минимальные BR, за исключением непрямого НРП Siemens FT4, который показал оптимальный BR. ВРП FT4, предложенные компаниями Roche, Abbott и Siemens, были слишком высокими по сравнению с таковыми, полученными прямым методом. НРП и ВРП для Beckman были ближе к полученным прямым методом, но не всегда превосходили минимальный BR (Abbott и Beckman).

Обсуждение

В этой статье описаны результаты надежности РИ ТТГ и FT4, неточности РИ ТТГ и FT4 из инструкций производителя и применимость современных непрямых методов для оценки РИ ТТГ и FT4. Продемонстрирована надежность расчета непрямым способом ТТГ производителей Roche, Abbott, Beckman и Siemens. Наши результаты также

Таблица 3. **Сравнение непрямых и прямых РИ FT4 с рекомендованными производителем РИ FT4**

Table 3. **Comparison of the indirect and direct FT4 RIs with the manufacturer-suggested FT4 RIs**

	Прямой метод (2003)		Непрямой метод (2022)												Рекомендованные производителем			
	19—68 лет		18—100 лет				18—70 лет				18–70 лет, ТТГ в пределах РИ							
	НРП (90% ДИ)	ВРП (90% ДИ)	НРП Me (IQR)	ВРП Me (IQR)	BR НРП	BR ВРП	НРП Me (IQR)	ВРП Me (IQR)	BR НРП	BR ВРП	НРП Me (IQR)	ВРП Me (IQR)	BR НРП	BR ВРП	НРП	ВРП	BR НРП	BR ВРП
Roche Cobas (*n*=4)	12,6 (12,1—13,1)	19,8 (19,3—20,2)	11,6 (11,3—12,0)	20,8 (20,5—21,2)	0,544	0,544	11,5 (11,4—11,9)	20,6 (20,4—21,0)	0,599	0,436	11,6 (11,4—12,0)	20,2 (19,7—20,9)	0,544	0,218	12	22	0,327	1,198
Abbott Alitity (*n*=3)	9,6 (9,4—9,9)	13,6 (13,3—13,8)	10,1 (10,1—10,5)	16,5 (15,9—16,7)	0,490	2,842	10,2[a] (9,9—10,4)	16,2[a] (15,9—16,6)	0,588	2,548	10,5[a] (10,3—10,6)	16,2[a] (15,9—16,5)	0,882	2,548	9,01[b]	19,5[b]	0,578	5,782
Siemens Atellica (*n*=4)	12,1 (11,4—12,9)	19,5 (17,9—21,1)	12,0[c] (11,7—12,3)	20,9[c] (20,2—22,2)	0,053	0,742	12,0[c] (11,6—12,3)	20,7[c] (20,3—20,7)	0,053	0,636	12,2 (11,8—12,5)	20,7 (20,0—21,4)	0,053	0,636	11,5	22,7	0,318	1,695
Beckman DxI (*n*=2)	8,6 (8,3—9,0)	13,9 (13,6—14,3)	7,9 (7,8—8,1)	15,3 (15,1—15,4)	0,518	1,035	7,9 (7,8—8,0	15,0 (14,0—15,0)	0,518	0,814	8,1 (7,9—8,3)	14,5 (14,3—14,7)	0,370	0,444	7,86	14,41	0,547	0,377

Примечание. РИ — референтный интервал. a — *n*=2, данные 2022 г. из лаборатории 9 были исключены после визуальной оценки модели RefineR; b — Abbott предоставляет РИ FT4 на основе центрального 99% интервала вместо обычного центрального 95% интервала. РИ Abbott FT4, основанный на центральном 95% интервале, будет составлять 10,3—17,8 пмоль/л. Коэффициенты смещения для этих РИ составляют 0,686 (НРП) и 4,077 (ВРП); c — *n*=3, данные 2022 г. из лаборатории 12 были исключены после визуальной проверки модели RefinR. $BR_{НРП}=\frac{НРП - НРП_0}{SD_{РИ}}$, $BR_{ВРП}=\frac{ВРП - ВРП_0}{SD_{РИ}}$, $SD_{РИ}=\frac{ВРП_0 - НРП_0}{3,92}$, минимальный BR<0,375, желательный BR<0,250, оптимальный BR<0,125. НРП — нижний референтный предел; ВРП — верхний референтный предел; BR — коэффициент смещения; ДИ — доверительный интервал.

показали надежность непрямого расчета FT4 производителей Roche, Beckman и Siemens, в то время как данные Abbott показали снижение ВРП FT4 за период 2013—2015 гг. Ранее A. Algeciras-Schimnich и соавт. предположили, что для оценки надежности результатов ТТГ и FT4 недостаточно проведения исследования межлотовой вариации [12, 13]. Это потенциально объясняет, почему снижение результатов Abbott могло быть упущено из виду.

Кроме того, в результате исследования выявлены неточности в значениях РИ ТТГ и FT4 в инструкциях производителя, не рекомендуется транслировать эти РИ в лаборатории из-за слишком узкого (Roche ТТГ), слишком широкого диапазона (Abbott ТТГ, Beckman ТТГ, Roche FT4 и Abbott FT4), слишком низких (Beckman FT4) или слишком высоких (Siemens) РИ. Как и ожидалось, НРП и ВРП Abbott, основанные на центральном 95% интервале (вместо центрального 99% интервала), больше напоминали прямые РИ, но были либо слишком низкими (ТТГ), либо слишком высокими (FT4).

Рассчитанные непрямым способом РИ для ТТГ, полученные в возрастной группе 18—60 лет, продемонстрировали превосходное согласие с результатами прямого метода. Расчетные РИ FT4, полученные для возрастной группы 18—70 лет, оказались шире, чем результаты прямого исследования. Причины этих неоптимальных РИ FT4 многосоставные и включают аналитические, межиндивидуальные и внутриндивидуальные вариации, сезонные колебания, вариабельность день ото дня и распределение результатов, все эти параметры различаются для ТТГ и FT4. Помимо этого, T. Ammer и соавт. показали, что непрямые методы могут превзойти прямой метод [7]. В целом рассчитанные непрямым способом РИ могут оказаться точнее по сравнению с рекомендованными производителем РИ FT4.

Примечателен факт, что ВРП FT4 производителя Abbott отличался почти на 6 пмоль/л от результатов, полученных прямым способом в 2023 г. С учетом 95% центрального интервала вместо 99% интервала результаты по-прежнему отличались более чем на 4 пмоль/л. Интересно, что непрямой метод также завысил ВРП FT4 на 3 пмоль/л. Поскольку другое недавнее исследование также показало разницу ВРП для Abbott FT4 примерно в 16 пмоль/л, возникло предположение, что эта разница вызвана вторым быстрым снижением результатов у верхней границы РИ — между результатами последних расчетов FT4 непрямыми методами (2022 г.), и сбором и измерением образцов с последующим расчетом прямым способом (2023 г.) [14].

После этого, в августе 2023 г., при исследовании в лаборатории 1 95 собранных и замороженых при 80°С в июне 2022 г., было обнаружено пропорциональное смещение в 12%. Для подтверждения этих результатов, образцы были дополнительно исследованы в эндокринологической лаборатории Амстердамского UMC с использованием последнего лота реагента Abbott FT4 в августе 2023 г. Была обнаружена значительная пропорциональная разница — в 16% между двумя лотами реагентов, что подтверждает обнаруженную отрицательную тенденцию и объясняет наблюдаемую разницу между ВРП, определенными прямым и непрямыми способами.

Объем включения данных для анализа из полного объема исторических данных ЛИС, необходимый для использования непрямого метода, на данный момент остается предметом споров. С. Ма и соавт. оценили два подхода к использованию непрямого метода для установления РИ для ТТГ и FT4 [14]. Они сравнили практически нефильтрованный набор данных пациентов, загруженный непосредственно из ЛИС, и набор данных пациентов, который соответствовал строгим критериям исключения. Оба метода дали схожие результаты, что говорит о возможности расчета РИ ТТГ и FT4 непрямым способом без использования дополнительных критериев исключения. В настоящей работе было решено исключить все результаты повторных измерений, поскольку F. Arzideh и соавт. показали, что использование всех данных без этапа фильтрации приводит к значительному смещению РП [15]. Полученные результаты подтверждают применимость непрямых методов с использованием минимальной фильтрации для установления или проверки РИ ТТГ и FT4. Подобный вывод уже был сделан в нескольких исследованиях, однако ни одно из этих исследований не соответствовало полноте настоящего исследования с данными всех четырех основных производителей, и до этого момента не была оценена надежность расчета РИ ТТГ и FT4 непрямым способом с использованием исторических данных.

R. Haeckel и соавт. использовали похожий подход для гармонизации РИ для 16 аналитов, включая ТТГ и FT4 [10]. Для ТТГ наличие независимого гармонизированного РИ было признано возможным в Канаде, тогда как результаты FT4 для конкретного метода существенно различались. Полученные результаты подтверждают различия FT4, но дополнительно обнаружены существенные различия между ВРП производителя Abbott (4,06 пмоль/л) и ВРП производителей Roche, Beckman и Siemens (4,73, 4,71 и 4,84 пмоль/л соответственно). Это означает, что мы еще далеки от гармонизированных, независимых от метода РИ для ТТГ и FT4, и ожидаем гармонизации от рабочей группы IFCC по стандартизации тестов функции щитовидной железы (WG-STFT) [16—19].

Уровни ТТГ и в меньшей степени FT4 демонстрируют возрастную зависимость, при этом диапазон РИ расширяется с возрастом. Является ли это повышение физиологическим проявлением дисфункции щитовидной железы или отражает наличие сопутствующих заболеваний, остается вопросом для дальнейших дискуссий [20]. Тем не менее это означает, что необходима правильная стратификация для адекватного сравнения РИ, определенных прямым и непрямым способом [4]. В другой статье было рассмотрено влияние возрастных РИ ТТГ и FT4 (H.I. Jansen и соавт., в стадии подготовки).

Причина, по которой соответствующие РИ важны для оценки функции щитовидной железы, была показана K.L. Coene и соавт. [21]. Они пришли к выводу, что РИ ТТГ и FT4 в лаборатории могут влиять на оценку состояния заболевания на национальном уровне, неверно указывая на субклиническое заболевание. Они показали, что это может быть вызвано большой вариацией концентраций ТТГ при определении в лаборатории и результатом применяемых РИ.

Были изучены диагнозы гипотиреоза и гипертиреоза (на основе первых измерений пациента) и подсчитано количество новых пациентов с заболеваниями щитовидной железы с помощью рекомендованных производителем РИ (0,35—4,94 мкМЕ/мл для ТТГ, 9,01—19,5 пмоль/л для FT4) по сравнению с расчетными РИ Abbott (0,57—4,06 мкМЕ/мл для ТТГ, 10,5—16,2 пмоль/л для FT4). Процент диагнозов манифестного гипотиреоза увеличился с 0,8% до 3,3%, а процент диагнозов манифестного гипертиреоза увеличился с 1,2% до 3,9%.

Исследование имеет несколько ограничений. Во-первых, невозможность дифференцировать небеременных и беременных женщин в наборах данных, несмо-

тря на известное влияние беременности на уровни ТТГ и FT4 [22]. Также невозможно оценить успешность исключения результатов повторных измерений. Кроме того, H.I. Jansen и соавт. продемонстрировали, что у пациентов, принимающих левотироксин, наблюдаются более высокие концентрации FT4 без полного подавления ТТГ, и неясно, классифицирует ли RefineR эти результаты как патологические или непатологические [23].

Выводы

Полученные результаты подчеркивают необходимость осторожности при внедрении рекомендованных производителем РИ для ТТГ и FT4 в лаборатории. Непрямое определение ТТГ и FT4 обеспечивает лучшую сопоставимость с определенными прямым способом РИ. Лабораториям следует установить свои собственные РИ, адаптированные к их конкретным методам и характеристикам популяции пациентов, или принять РИ из более надежного источника, чем инструкции производителя. Рекомендуется частая проверка РИ для выявления потенциальных отклонений в результатах и предотвращения гипердиагностики или гиподиагностики заболеваний щитовидной железы.

Авторы заявляют об отсутствии конфликта интересов.
The authors declare no conflicts of interest.

ЛИТЕРАТУРА/REFERENCES

1. CLSI. *Defining, establishing, and verifying reference intervals in the clinical laboratory; Approved guideline — Third Edition. CLSI document EP28-A3c*. Wayne PA: Clinical and Laboratory Standards Institute; 2008.
2. Ozcurumez MK, Haeckel R. Biological variables influencing the estimation of reference limits. *Scand J Clin Lab Invest.* 2018; 78(5):337-345. https://doi.org/10.1080/00365513.2018.1471617
3. Tate JR, Yen T, Jones GR. Transference and validation of reference intervals. *Clin Chem.* 2015;61:1012-1015. https://doi.org/10.1373/clinchem.2015.243055
4. Haeckel R, Wosniok W. The importance of correct stratifications when comparing directly and indirectly estimated reference intervals. *Clin Chem Lab Med.* 2021;59:1628-33. https://doi.org/10.1515/cclm-2021-0353
5. Clinical and Laboratory Standards Institute (CLSI). *Measurement procedure comparison and bias estimation using patient samples; approved guideline — third edition. CLSI document EP09-A3.* Wayne PA: CLSI; 2013.
6. Haeckel R, Wosniok W, Streichert T, Members of the Section Guide Limits of the DGKL. Review of potentials and limitations of indirect approaches for estimating reference limits/intervals of quantitative procedures in laboratory medicine. *J Lab Med.* 2021;45:35-53. https://doi.org/10.1515/labmed-2020-0131
7. Ammer T, Schützenmeister A, Prokosch HU, Zierk J, Rank CM, Rauh M. RIbench: A Proposed Benchmark for the Standardized Evaluation of Indirect Methods for Reference Interval Estimation. *Clin Chem*. 2022 Nov 03;68(11):1410-1424. https://doi.org/10.1093/clinchem/hvac142
8. Wosniok W, Haeckel R. A new indirect estimation of reference intervals: truncated minimum chi-square (TMC) approach. *Clin Chem Lab Med.* 2019;57:1933-1947. https://doi.org/10.1515/cclm-2018-1341
9. Ammer T, Schützenmeister A, Prokosch HU, Rauh M, Rank CM, Zierk J. refineR: A Novel Algorithm for Reference Interval Estimation from Real-World Data. *Sci Rep.* 2021 Aug 06;11(1):16023. https://doi.org/10.1038/s41598-021-95301-2
10. Haeckel R, Wosniok W, Arzideh F. Equivalence limits of reference intervals for partitioning of population data. Relevant differences of reference limits. *Laboratoriums Medizin*. 2016;40(3):199-205. https://doi.org/10.1515/labmed-2016-0002
11. Ozarda Y, Ichihara K, Jones G, Streichert T, Ahmadian R, IFCC, C-RIDL. Comparison of reference intervals derived by direct and indirect methods based on compatible datasets obtained in Turkey. *Clin Chim Acta.* 2021;520:186-195. https://doi.org/10.1016/j.cca.2021.05.030
12. Algeciras-Schimnich A, Bruns DE, Boyd JC, Bryant SC, La Fortune KA, Grebe SKG. Failure of current laboratory protocols to detect lot-to-lot reagent differences: findings and possible solutions. *Clin Chem*. 2013;59:1187-1194. https://doi.org/10.1373/clinchem.2013.205070
13. Katzman BM, Ness KM, Algeciras-Schimnich A. Evaluation of the CLSI EP26-A protocol for detection of reagent lot-to-lot differences. *Clin Biochem*. 2017;50:768-771. https://doi.org/10.1016/j.clinbiochem.2017.03.012
14. Ma C, Cheng X, Xue F, Li, X, Yin Y, Wu, J, et al. Validation of an approach using only patient big data from clinical laboratories to establish reference intervals for thyroid hormones based on data mining. *Clin Biochem*. 2020;80:25-30. https://doi.org/10.1016/j.clinbiochem.2020.03.012
15. Arzideh F, Özcürümez M, Albers E, Haeckel R, Strecihert T. Indirect estimation of reference intervals using first or last results and results from patients without repeated measurements. *J Lab Med*. 2021;45:103-109. https://doi.org/10.1515/labmed-2020-0149
16. Thienpont LM, Van Uytfanghe K, Van Houcke S, Das B, Faix JD, MacKenzie F, et al. A progress report of the IFCC committee for standardization of thyroid function tests. *Eur Thyroid J.* 2014;3:109-116. https://doi.org/10.1159/000358270
17. Thienpont LM, Van Uytfanghe Beastall G, Faix JD, Ieiri T, Miller WG, et al. Report of the IFCC working group for standardization of thyroid function tests; part 1: thyroid-stimulating hormone. *Clin Chem*. 2010;56:902-911. https://doi.org/10.1373/clinchem.2009.140178
18. Thienpont LM, Van Uytfanghe Beastall G, Faix JD, Ieiri T, Miller WG, et al. Report of the IFCC working group for standardization of thyroid function tests; part 2: free thyroxine and free triiodothyronine. *Clin Chem*. 2010;56:912-920. https://doi.org/10.1373/clinchem.2009.140194
19. Thienpont LM, Faix JD, Beastall G. Standardization of free thyroxine and harmonization of thyrotropin measurements: a request for input from endocrinologists and other physicians. *Thyroid*. 2015;25:1379-1380. https://doi.org/10.1089/thy.2015.0309
20. Raverot V, Bonjour M, du Payrat AJ, Perrin P, Roucher-Boulez F, Lasolle H, et al. Age- and sex-specific TSH upper-limit reference

intervals in the general French population: there is a need to adjust our actual practices. *J Clin Med.* 2020;9(3):792. https://doi.org/10.1530/endoabs.70.aep898

21. Coene KL, Demir AY, Broeren MAC, Verschuure P, Lentjes EGWM, Boer AK. Subclinical hypothyroidism: a "laboratory-induced" condition? *Eur J Endocrinol.* 2015;173:499-505. https://doi.org/10.1530/eje-15-0684
22. Osinga JAJ, Derakhshan A, Palomaki GE, Ashoor G, Männistö T, Maraka S, et al. TSH and FT4 reference intervals in pregnancy: a systematic review and individual participant data meta-analysis. *J Clin Endocrinol Metab* 2022;107:2925-2933. https://doi.org/10.1210/clinem/dgac425
23. Jansen HI, Bult MM, Bisschop PH, Boelen A, Heijboer AC, Hillebrand JJ. Increased fT4 concentrations in patients using levothyroxine without complete suppression of TSH. *Endocr Connect.* 2023;12:e220538. https://doi.org/10.1530/ec-22-0538

Поступила 05.09.2024
Received 05.09.2024
Принята к печати 11.10.2024
Accepted 11.10.2024