Marith van Schrojenstein Lantman*, Remco Grobben, Antonius E. van Herwaarden, Miranda van Berkel, Jeroen Schaap and Marc Thelen

To rule-in, or not to falsely rule-out, that is the question: evaluation of hs-cTnT EQA performance in light of the ESC-2020 guideline

https://doi.org/10.1515/cclm-2023-1226
Received October 30, 2023; accepted January 23, 2024;
published online February 7, 2024

Abstract

Objectives: To accurately evaluate non-ST-elevated acute cardiac syndrome (NSTE-ACS), the quality of high-sensitive cardiac troponin (hs-cTn) assays is of vital importance. The 2020 revision of the NSTE-ACS guideline includes clinical decision-limits (CDL's) to both rule-in and rule-out NSTE-ACS for most commercially available platforms, providing both 0/1 h and 0/2 h delta limits. Our study evaluated whether laboratories are able to meet the analytical performance specifications for imprecision (APS) for hs-cTnT.

Methods: Results from external quality assurance (EQA) in commutable samples were used to evaluate the current and historic performance of analyzers. The performance of analyzers that either passed or failed to comply with 0/1 h-APS were used on a real-world dataset of first hs-cTnT-values to simulate 10.000 samples of t=0, t=1 and t=2 h values with multiple delta's for all relevant CDL's. We compared the simulated values to the input values to obtain the percentage of aberrant results simulated.

Results: The majority of analyzers complies with APS for rule-in in 2022 (0/1 h: 90.4 % and 0/2 h: 100 %), compliance for the 0/1 h rule-out is still far from optimal (0/1 h: 30.7 %, 0/2 h: 75.4 %), with improving compliance over the past years (rule-in p=<0.0001, rule-out p=0.011, χ^2). Whilst 0/1 h-APS-passing analyzers have a minute risk to falsely rule-out patients whom should be ruled-in (0.0001 %), failing performance increases this risk to 2.1 % upon using 0/1 h CDL's. Here, adopting 0/2 h CDL's is favorable (0.01 %).

Conclusions: Laboratories that fail to meet hs-cTnT 0/1 h-APS should improve their performance to the required and achievable level. Until performance is reached clinics should adopt the 0/2 h CDL's.

Keywords: high-sensitivity cardiac troponin; diagnostic accuracy; analytical performance specifications; clinical impact; NST-ACS diagnosis; guideline

Introduction

For patients suspected of non ST-segment elevated acute cardiac syndrome (NSTE-ACS), the measurement and interpretation of high-sensitive cardiac troponin (hs-cTn) is the cornerstone of current diagnostic protocols [1]. The 2020 European Society for Cardiology (ESC) clinical guideline states that the adequate clinical intervention for these patients can be derived by interpreting the hs-cTn concentration at admission together with the delta of its serial assessment one or 2 h after presentation (Figure 1) [1]. The absolute and delta clinical decision limits (CDL's) mentioned in the 2020-ESC guideline specify at which hs-cTn concentrations an NSTE-ACS is to be ruled-in or ruled out, and depends on the platform used and the time difference between the serial measurements [1].

For adequate implementation of these CDLs in patient care, it is vital that the performance of the measurement procedure stays within the analytical performance specifications (APS) [2]. For hs-cTn, the APS for imprecision (following mentions of APS indicate APS for imprecision) are relatively strict, firstly because the CDL's mentioned are in the low measurement range, just above the limit of

***Corresponding author: Marith van Schrojenstein Lantman**, Department of Laboratory Medicine, Radboudumc, Nijmegen, The Netherlands; Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek (SKML), Nijmegen, The Netherlands; and Result Laboratorium, Amphia Hospital, Breda, The Netherlands, E-mail: Marith.vanSchrojensteinLantman@radboudumc.nl. https://orcid.org/0000-0002-5454-990X

Remco Grobben, Department of Cardiology, Amphia Hospital, Breda, The Netherlands

Antonius E. van Herwaarden and Miranda van Berkel, Department of Laboratory Medicine, Radboudumc, Nijmegen, The Netherlands

Jeroen Schaap, Department of Cardiology, Amphia Hospital, Breda, The Netherlands; and Dutch Network for Cardiovascular Research (WCN), Dutch Network for Cardiovascular Research (WCN), Utrecht, The Netherlands

Marc Thelen, Department of Laboratory Medicine, Radboudumc, Nijmegen, The Netherlands; Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek (SKML), Nijmegen, The Netherlands; and Result Laboratorium, Amphia Hospital, Breda, The Netherlands. https://orcid.org/0000-0003-1771-669X

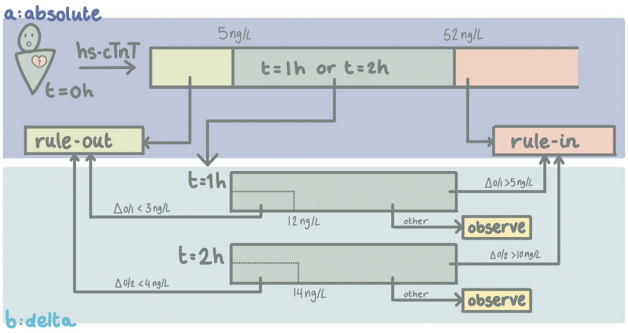

Figure 1: Clinical pathway of patient with a suspected NST-ACS. a. Upon the first measurement, a NST-ACS can be ruled out if the hs-cTnT value is lower than 5 ng/L, and ruled-in if hs-cTnT of t=0 exceeds 52 ng/L. b. Patients that cannot be ruled-in or out are sampled at timepoint t=1 h or t=2 h. The delta value between 0 and 1 h or 2 h is calculated. Patients with an initial hs-cTnT below 12 or 14 ng/L and a delta lower than 3 or 4 ng/L can be ruled out, and patients with a delta higher than 5 or 10 ng/L can be ruled in. Patients that do not fit into any of these categories are observed longer.

quantification. Thus, hs-cTn assays are only fit for intended use if low imprecision is present, which has only become feasible with the latest generation of hs-cTn assays. Secondly, the need for strict APS is vital as aberrant clinical decision-making has serious clinical consequences. On one hand, incorrect rule-in of NSTE-ACS results in unnecessary, and potentially harmful treatment, increases caregiver-workload and exerts stress on the patient and their family [3]. On the other hand, incorrect rule-out results in withholding or delay of adequate NSTE-ACS treatment, thereby putting patients at risk of increased morbidity and/or mortality [4]. In 2013, an international survey concluded a consensus among emergency department doctors about a maximally acceptable false-rule-out rate of NSTE-ACS patients of 1 % attributable to hs-cTn measurement imprecision [5].

In a previous study we have evaluated the ability of laboratories to meet the APS for imprecision for accurate NSTE-ACS diagnosis by using external quality assurance (EQA) data [6]. This study showed that, in 2015 and 2016, 18–60 % of laboratories complied with this APS for hs-cTn [6]. Since then, the ESC guideline was revised in 2020, with platform- and sampling-frequency specific cut-off values chosen in order to achieve 99 % clinical sensitivity for both 0/1 h and 0/2 h sampling time-points [1]. A critical difference between the 0/1 h and 0/2 h protocol are the consequences for required laboratory performance; the 0/2 h algorithm utilizes larger delta values compared to the 0/1 h algorithm (Figure 1), which allows for more lenient APS. Here, we used external quality assurance (EQA) data to investigate compliance with the ESC-2020-associated APS for imprecision with high-sensitivity measurement procedures and the resulting impact on NSTE-ACS associated clinical-decision making by modeling the imprecision on real-world patient data.

Materials and methods

Determining analyzer performance using external quality assurance data

As laboratories monitor their analytical performance by participating in EQA ring trials, the retrospective evaluation of these data is valuable and legit to establish the analytical performance of hs-cTn analyzers. We collected the EQA data from sample results of the years 2016–2022 from the Dutch EQA organizer (Foundation for Quality Assurance in Laboratory Medicine, Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek, short SKML). The EQA scheme of cardiac markers consists of 24 samples in the form of 12 blinded duplicates, which are prepared by pooling selected patient material [7]. In that program laboratories measure new EQA samples every two weeks. We included all data of hs-cTnT samples with a concentration at or below the direct rule-in threshold of 52 ng/L. We excluded administrative outliers, whereas analytical outliers without a known cause of bias were included in the data. Unfortunately, the limited number of hs-cTnI assay users (<10 users for the individual methods) was insufficient to allow statistically robust statements on their performance, and therefore the current evaluation is limited to hs-cTnT, only provided by Roche diagnostics.

The analytical performance was established per analyzer by adopting the multi sample evaluation (MUSE)-system of the SKML in simplified form [8]. Per year of participation in the EQA scheme, the imprecision of each analyzer is obtained by summing the residuals of a linear fit between the sample and the mean obtained value per method consensus group. For rule-in and rule-out, the residuals of samples within the relevant concentration ranges (i.e., <3–12 ng/L for rule-out and <3–52 ng/L for rule-in) were summed and divided by the degrees of freedom. The square root of these sums results in a standard deviation per analyzer with respect to the relevant concentration range for rule-in and rule-out.

The allowable performance specifications (APS) for imprecision of hs-cTn assays is derived by solving the delta change of rule-in and rule-out in the formula for reference change value (RCV, Eq. (1)). Herein, we utilized absolute standard deviation as the delta limits mentioned are absolute units. The within-person biological variation component (SDi) was derived from relative CVi data-studies in healthy subjects with sampling points around 1–2 h ([9–11], in meta-analysis of [12]), and was estimated to be 0.7 ng/L. A *Z*-value of 2.33 is appropriate to estimate a one-sided increase with 99 % certainty. Upon dividing the delta CDL's with the square root of two times 2.33, the APS for rule-in and rule-out of respectively 1.4 ng/L and 0.6 ng/L for 0/1 h sampling-points, and 3.0 ng/L and 1.0 ng/L for 0/2 h sampling-points were obtained. The imprecision estimates based on the residuals from regression were compared to the APS to determine whether analyzers complied with rule-in and rule-out APS for 0/1 h and/or 0/2 h.

$$\text{Reference Change Value (RCV)} = \sqrt{2} \times Z \times \sqrt{\text{SD}_a^2 + \text{SD}_i^2} \qquad (1)$$

Evaluating impact of imprecision on clinical decision-making with hs-cTnT by simulating imprecision on patient data

The impact of imprecision of hs-cTnT measurements on clinical decision-making was evaluated by modeling its impact on real patient data. To specify the input for that modeling, we first determined how the imprecision behaves as a function of the concentration (i.e., the precision profile). Two precision profiles were determined using duplicate variation data by generating two datasets: (1) PASS-APS was determined by using duplicate samples with a variance below the sample-median variance and (2) FAIL-APS was determined using the 25 % of samples with the highest duplicate standard deviation. For both datasets, a second-degree polynomic function was derived that predicted the variation per duplicate EQA sample as a function of the mean value of the duplicate samples (Supplementary 1). The PASS-APS and FAIL-APS precision profiles were subsequently used to model 10.000 t=0 measurements in-silico of a real-world patient dataset.

The patient data was derived from patients administered to the Amphia hospital, Breda, The Netherlands in the period from the 1st of January to the 26th of September 2022. All patient results were analyzed anonymously, for which we obtained a statement from the medical research ethnics committee from the board of directors of Amphia hospital, Breda, The Netherlands. The simulation was done by using exclusively the first hs-cTnT value obtained per patient. In the simulation, we assume that these values are a proxy for the 'true' value of each patient.

We evaluated the proportion of simulations that would result in aberrant clinical decisions in comparison to the original patient value by comparing the obtained results of the simulation with the absolute CDL's of the t=0 timepoint. Due to measurement uncertainty, four types of clinical errors can occur: (1) patients that are clinically observed which should be ruled-in (2) patients that are ruled-in which should be clinically observed (3) patients that are ruled-out which should be clinically observed and (4) patients that are held under observation who should be ruled-out. As no data was available for hs-cTnT values below 5 ng/L, no numeric hs-cTnT results were available to model the impact of analytical imprecision on error 4. We did evaluate clinical errors 1–3 by establishing the proportion of simulations that would be clinically misclassified in comparison to the 'true' value.

Simulations that resulted in 'observation' were subsequently modeled further to evaluate the effect of the PASS-APS and FAIL-APS precision on the misclassification rates for the multiple clinical decisions that are indicated by the 0/1 h and/or 0/2 h delta CDL's. We simulated 10.000 measurements of both a t=1 h and t=2 h sample with the imprecision of PASS-APS and FAIL-APS, using multiple delta values to evaluate the all clinical decisions possible. Empirical delta values of 0, 4 and 6 ng/L were used for the 1 h measurement, whereas delta values of 0, 9 and 11 ng/L for the 2 h measurement were used. The simulated 0/1 h and 0/2 h delta values were compared to the ESC-2020 CDL's (Figure 1) to determine the proportion of simulations that were ruled-in, observed, or ruled-out. A delta of 0 ng/L should lead to rule out in case of an t=0 hs-cTn value <12 or <14 ng/L (for resp. 0/1 h or 0/2 h). If initial hs-cTn values are above these limits, the intended clinical decision is to observe. The delta values of 4 and 9 ng/L (resp. for 0/1 h and 0/2 h) should result in the clinical decision to observe, whereas delta values of 6 and 11 ng/L (resp. for 0/1 h and 0/2 h) should result in the clinical decision to rule-in. The misclassification rate was evaluated by determining the proportion of simulated clinical decisions that were incongruent with the intended clinical decision of the input delta's.

Statistical analysis

The analyses were done in R version 4.1.3, using packages tidyverse 1.3.2 and data.table version 1.14.2 [13]. For the modeling of imprecision, we set the seed to 357. For the determination of statistical significance in analytical compliance rates over the years, a chi-squared test was performed.

Results

EQA performance of hs-cTnT analyzers

In 2022, 61 laboratories participated in the EQA scheme of cardiac markers with 114 analyzers with hs-cTnT. In 2022, six duplicate sample sets had a relevant concentration with respect to the ESC-2020-CDLs and were included in the evaluation. The performance of participating analyzers was assessed by evaluating the imprecision in the relevant concentration range for rule-in and rule-out (Figure 2). For rule-in, 90.4 and 100 % of the Roche analyzers complied with the APS of the 0/1 h algorithm and 0/2 h algorithm respectively. In contrast, for rule-out, 4.4 % of analyzers did not send in adequate information to determine the performance. Of the remaining 109 analyzers for which compliance could be assessed, only 30.7 % complied with 0/1 h, whereas the performance to the 0/2 h algorithm was more adequate (75.4 %).

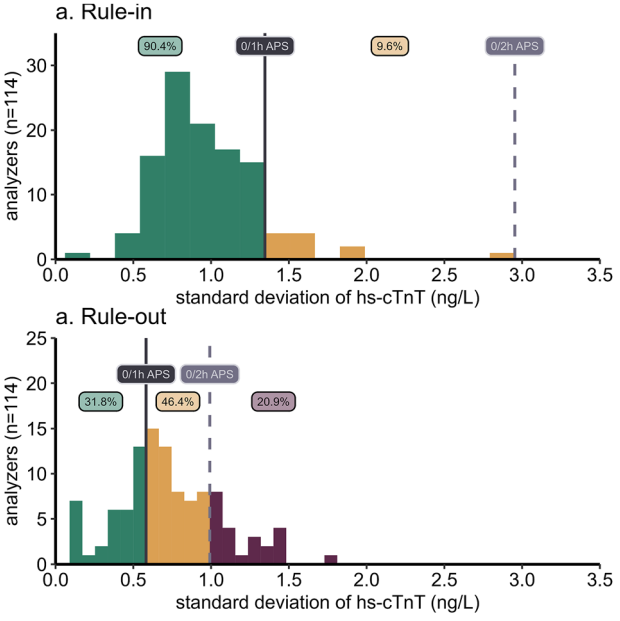

Figure 2: Analytical performance of hs-cTnT from analyzers in 2021. The number of analyzers that comply or fail with the 0/1 h and/or 0/2 h algorithm are depicted for a. Rule-in APS, and b. Rule-out APS. The allowable 0/1 h performance is shown as black line, whereas 0/2 h is shown as gray dotted line. Analyzers passing the rule-in and rule-out APS for 0/1 h were indicated in green, analyzers failing 0/1 h but passing 0/2 h in yellow, whereas analyzers failing 0/2 h are indicated in red.

We evaluated whether the hs-cTnT performance of Roche analyzers derived from the 2022 EQA data is representative for historical performance of these laboratories by establishing the share of analyzers that comply with 0/1 h and 0/2 h rule-in and rule-out APS over the period from 2016 to 2022 (Figure 3). In comparison to the data of van der Hagen et al., our data shows similar compliance for rule-in, and lower compliance for rule-out. For both rule-in as rule-out, an increase in the proportion of analyzers that perform within the APS is observed over the years (chi-squared p-values of resp. 0.011 and <0.0001 for rule-in and rule-out). In 2017, 79.2 % of analyzers met the 0/1 h APS for rule-in, vs. 90.4 % in 2022. For rule-out, we note that the proportion of analyzers with no results ranges over the years (4.4–29.0 %, 17.8 % in 2017). Of the analyzers that did send in results in the rule-out range, 12.0 and 60.2 % met 0/1 h and 0/2 h APS in 2017. Taken together, while we note a trend for improved performance of hs-cTnT to near-full compliance for rule-in, for rule-out still such a trend is missing with a substantial proportion of analyzers still fails to meet 0/1 h APS.

hs-cTnT assays performance in patients based on absolute cut-off

The consequences associated with meeting or failing the APS on the absolute cut-off levels were further investigated using a modeling approach with real-world patient data. In total, 3,289 unique hs-cTnT patient results were used for the simulation (Figure 4). Per hs-cTnT result, 10.000 estimates were modeled using the precision-profile of PASS-APS and FAIL-APS, and established the misclassification rates relative to the original patient value. Three types of misclassification were investigated (1) patients are clinically observed which should be ruled-in (2) patients that are ruled-in which should be clinically observed and (3) patients that are ruled-out which should be clinically observed. For all three types of misclassification, failing to meet 0/1 h APS resulted in higher misclassification rates (Table 1). In a 0/1 h APS-passing system, 1.1 % of patients are observed instead of ruled-in (vs. 2.2 % if the analyzer fails 0/1 h APS), 0.2 % are ruled-in instead of observed (vs. 0.4 % for failing 0/1 h APS), and 1.0 % are ruled-out instead of observed (vs. 5.1 % for failing 0/1 h APS). Thus, while in a 0/1 h APS-compliant system the main risk is a delay to rule-in, if 0/1 h APS fails the main risk is that patients are prematurely ruled out.

hs-cTnT assays performance in patients based on delta cut-off

Next, the effects of passing or failing the APS on misclassification rates associated with delta values are evaluated by modeling a second time-sample for both the 0/1 h and 0/2 h timepoints. Using both the PASS-APS and FAIL-APS precision profile, delta values of 0, 4 and 6 ng/L were modeled for the 0/1 h system, whereas for the 0/2 h system delta values of 0, 9 and 11 ng/L were modeled. Since 90 % of all patients in the

Figure 3: Proportion of analyzers that comply or fail with rule-in (A) and rule-out (B), as a function of time over the years.

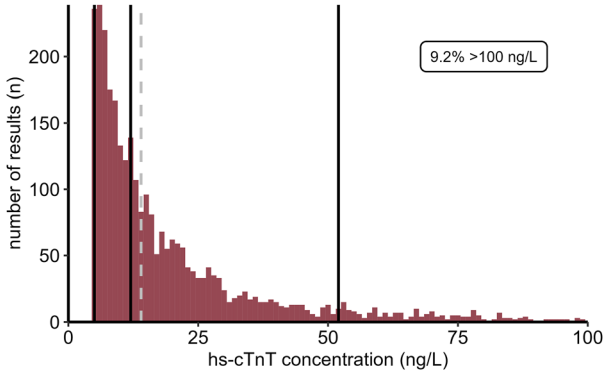

Figure 4: Distribution of hs-cTnT values used in the model, with absolute cut-off limits shown in black for 0/1 h, whereas the gray dotted line shows the 0/2 h limit for delta rule-out. 9.2 % of the total population had a cardiac troponin value >100 ng/L.

Table 1: Misclassification rates for clinical decisions associated with t=0 h NST-ACS, where the proportion of misclassified patients are displayed as percentages. The intended decision is displayed in green, whereas the decision made is displayed in red.

intended decision	decision made	PASS	FAIL
rule-in	→ observe	1.1%	2.2%
observe	→ rule-in	0.2%	0.4%
observe	→ rule-out	1.0%	5.1%

original data set had a delta below 12 ng/L (with 80.7 % having a delta <5 ng/L and 21.3 % a delta of 0 ng/L), the results of our simulation are relevant for a significant part of all patients. The misclassification rate was evaluated by determining the proportion of simulated clinical decisions that were incongruent with the intended clinical decision of the input delta upon comparing all delta values to the CDL's mentioned in the ESC-2020 guideline (Figure 1).

The clinical misclassification with most adverse effects is when patients that are supposed to be ruled-in are ruled-out. With adequate performance (PASS-APS), it is highly unlikely that patients are incorrectly ruled-out when 0/1 h CDL's are used (Table 2, Δ6 ng/L: 0.0001 %). However, upon modeling failing performance, we note a percentage as high as 2.1 % of patients requiring NSTE-ACS intervention would have been inadvertently ruled-out based on hs-cTnT. On the contrary, the 0/2 h algorithm was less affected by failing analytical performance (Table 2, Δ11 ng/L), as in this simulation 0.01 % of the to-be-ruled-in patients would have been ruled-out.

Next to falsely ruling-out an NSTE-ACS, the risk of other clinical errors are increased if 0/1 h CDL's are used in

Table 2: The proportion of misclassified patients with the 0/1 h and 0/2 h algorithm upon modeling three delta's (6, 4 and 0 ng/L for 0/1 h and 11, 9 and 0 ng/L for 0/2 h). Per row, the specific misclassification and proportion of patient simulations misclassified are shown as percentages. Percentages in the last column indicate with a green background for FAIL-APS if the adoption of the 0/2 h CDL reduces the misclassification risk to <1 % in comparison to the 0/1 h CDL with the same suboptimal performance. The intended decision is displayed in green, whereas the decision made is displayed in red.

		0/1h CDL			0/2h CDL		
intended decision	decision made	Δ	PASS	FAIL	Δ	PASS	FAIL
rule-in	→ rule-out	6	0.0001%	2.1%	11	0%	0.01%
rule-in	→ observe	6	6.7%	25.0%	11	7.8%	27.8%
observe	→ rule-out	4	1.1%	10.4%	9	0%	0.28%
observe	→ rule-in	4	28.0%	39.5%	9	29.7%	39.7%
observe	→ rule-out	0'	1.8%	3.3%	0''	1.4%	3.4%
observe	→ rule-in	0'	0.07%	6.5%	0''	0%	0.09%
rule-out	→ observe	0'	1.7%	24.2%	0''	1.4%	13.7%
rule-out	→ rule-in	0'	0%	3.4%	0''	0%	0.001%

' hs-cTnT(t0) <12 ng/L
'' hs-cTnT(t0) <14 ng/L

APS-failing systems (Table 2). Patients could be temporarily withheld from treatment as patients are observed instead of ruled-in (Table 2, Δ6 ng/L: 6.7 vs. 25.0 %). Additionally, patients that are ought to be observed have higher risk to be preliminarily ruled-out (Table 2, Δ0 ng/L: 1.8 vs. 3.3 %, Δ4 ng/L: 1.1 vs. 10.4 %), or pre-emptively ruled-in (Table 2, Δ0 ng/L: 0.07 vs. 6.5 %). Taken together, failing to meet 0/1 h APS negatively affects the ability to make consistently correct clinical decisions.

In comparison, the misclassification rates associated with the 0/2 h algorithm are less profoundly impacted if the analyzer fails to meet the 0/1 h APS as opposed to using 0/1 h CDL's (Table 2). Patients that are close to the rule-in CDL are less likely to be incorrectly ruled-out (Table 2, Δ9 ng/L: 0.28, vs. Δ4 ng/L: 10.4 %). Patients with a delta of 0 ng/L that ought to be observed are less likely to be ruled-in pre-emptively upon the use of the 0/2 h CDL's (Table 2, Δ0 ng/L: 0.09 vs. 6.5 %). Furthermore, patients with a delta of 0 ng/L that ought to be ruled-out are less likely to be held under prolonged observation (Table 2, Δ0 ng/L: 13.7 vs. 24.2 %), or falsely ruled-in (Table 2, Δ0 ng/L: 0.001 vs. 3.4 %). Thus, the 0/2 h algorithm is more robust against clinical errors due to insufficient analytical performance.

Discussion

Taken together, we show that laboratories using hs-cTnT are increasingly able to comply with the APS as set by the 0/1 h ESC-guideline for the diagnosis or rule-out of NST-ACS.

However, our modeling approach showed that the effect of inadequate performance is an unacceptable rise in false rule-out of patients that should have been ruled in (0.0001 vs. 2.1 %), a rate above the predetermined allowable <1 % missing rate as determined by expert emergency department doctor consensus [5]. For laboratories that fail to meet 0/1 h APS, adopting the newly introduced 0/2 h algorithm substantially reduces the risk of falsely ruling-out NSTE-ACS-patients based on hs-cTnT. Thus, to ensure adequate patient care and outcome, it is of vital importance to monitor both rule-in and rule-out performance of laboratories, with clear communication to clinicians who rely on the accuracy of the hs-cTn assays of their laboratory. Unfortunately, not all misclassifications can be reduced by adopting the 0/2 h-version, which necessitates laboratories to remain vigilant for aberrant performance and eager for improving performance until APS are met.

Our study makes use of external quality assurance (EQA) data, which is both a strength and a limitation of this study. The advantage of using EQA over data from commercial internal quality control (IQC) sample data is the possibility to assess performance of a large number of laboratories using results of materials that are in our case commutable and thereby representative for patient samples, which is not guaranteed for commercial IQC samples [14]. The downside of using EQA is that EQA is measured at lower frequency than IQC and thus the estimates of imprecision reflect the long-term imprecision. For patients, the hourly variation is the most pivotal source of variation that affects clinical decision making. However, we have compared the derived imprecisions from our study to the uncertainty as obtained by IQC (data not shown) and found similar imprecision rates. Thus, as sources of imprecision with a larger amplitude in timing contain all the variations that vary more frequently [15], the fact that daily and two-weekly evaluation of imprecision result in the similar estimates allows for the hypothesis that the hourly variation will take up a notable proportion of the imprecisions as evaluated by EQA.

Another limitation may be the lack of our knowledge on how individual laboratories apply our EQA material to generate the results which they send to us. Whereas most laboratories use the material to establish the performance of a single analyzer, other labs may participate as a so called virtual analyzer by presenting each EQA-sample to a different analyzer within a round (Survey 2021, data not shown) resulting laboratory performance rather than analyzer performance. Consequently, the between-analyzer variation is incorporated into the estimation of analytical imprecision. In clinical practice, the t=0 and t=1/2 h samples could be presented on different analyzers and thus the between-analyzer variation can augment the misclassification of patients. Laboratories should be aware of this risk and mitigate it with local performance specifications suitable for their local procedures. Taken together, the analytical performance as established by EQA is a good, but not perfect estimate of the hourly variability present in routine patient samples.

Furthermore, our study has certain assumptions in the modeling approach which should be mentioned. Our model assumes an $a + bx + cx^2$ precision profile and assumes that the clinical decisions associated with the delta samples are solely based on the delta thresholds mentioned in the guideline. In the interpretation of our data, we assume that a patient with a delta of 6 ng/L at 1 h has a delta of 11 ng/L at 2 h, which are both 1 unit above the respective CDL's. In practice, the increase in hs-cTnT may not be as straightforward, but the model does reflect the most stringent and critical situation. Since 28.9 % of all patients in our original data set had a delta in one of these worst-case scenarios, the misclassification rates obtained by the model cannot be trivialized. Additionally, we apply one-sided statistical analyses that judges the effect of either an increase or decrease, whereas the ESC-guideline could be interpreted to note the deltas as a 'change' (with undetermined direction) and consequently requiring a two-sided estimation of the reference change value. We feel that the clinical question in regards to an increase in hs-cTn is different from a decrease, as the former is suspected when patients present themselves rapidly to the ER upon noticing severe chest pain, whereas a decrease would be noted if a patient would have had an NSTE-ACS more than one day prior to presentation to the ER. The ESC guideline could benefit from incorporating this mindset into a future revision resulting in unequivocal wording about increase and/or decrease rather than change. In the meantime, the local implementation of the ESC-2020 guideline may vary between hospitals, and any model is unable to take this into account. Moreover, the specific cut-off values and deltas that are taken up in the ESC-2020 guideline have been debated as the thresholds mentioned in the current guideline was predominantly derived by a single group of investigators [16]. However, we have followed the CDL's of the ESC-2020 guideline to evaluate misclassifications associated with failing or meeting APS whilst utilizing the CDLs mentioned there. Any future CDL's that are (to be) established are ideally done so using a multi-centered approach, wherein the CDL's established should be compared to the analytical performance to evaluate the impact of analytical performance on clinical decisions made.

Lastly, we acknowledge that the choice of the number used for the hourly CVi used may in practice be different. Whilst we find the estimate of 0.7 adequate in terms of the

used studies, they are based on historical data and have not been re-established since. Considering the diurnal pattern of hs-cTnT [9] and the improved analytical performance over time, a re-establishment of short-term biological variation with sampling points <2 h could alter the numerical results of our RCV calculations, but not the conclusion of this paper: adequate performance allows for correct utility of the ESC-2020 0/1 h algorithm, whereas inadequate performance impedes the ability of ED-clinicians to adequately triage patients. Herein, application of the 0/2 h algorithm can improve decision-making while the laboratory works on improving its performance.

An unfortunate aspect of the current decision system is the rounding of hs-cTn values. Although it is good laboratory practice to round results adequately, rounding becomes problematic for low hs-cTn values that get incorporated into the calculation of the delta. We would suggest that the original unrounded raw troponin values from the laboratory instruments would be used in the calculation of the delta instead of the already rounded values. For example, the delta between two troponin values of 5.3 and 6.6 is 1.3 which would rule out the patient, but rounding these values to 5 and 7 results in a delta of two, which would not rule out the patient. Laboratories could help in this process by reporting the delta based on unrounded results together with rounded absolute values of the second troponin result (either 1 or 2 h after the first). Submitting unrounded results would at least be beneficial for EQA organizers as performance can more adequately be established.

The results of this study should motivate laboratories to monitor their performance of their hs-cTn assay for both rule-in and rule-out performance specifications and to engage in conversation with their requesters regarding the utilization of hs-cTnT measurements. Analytical performance is not just a theoretical laboratory-focused concept of essentialism, but has severe clinical decision impact, making it about consequentialism. It should encourage EQA providers to select sample concentrations per type of platform fit to assess this performance based on unrounded results and give insight into specific rule-in and rule-out performance and IVD-manufacturers to engage in post-marketing surveillance. Moreover, we advocate a dialog between cardiologists, emergency department doctors, and laboratories about using the 0/1 h or 0/2 h version of the ESC-2020 guideline. This implies a close collaboration between all to ensure adequate patient safety when measuring minute concentrations of molecules with major adversity resulting from getting results in the 'wrong' category.

It is important to realize that NSTE-ACS-associated clinical decision-making may also be carried out using serial measurements of high-sensitive cardiac troponin I (hs-cTnI) measurements [1]. Unfortunately, the distribution of analyzers in our EQA-scheme left us with insufficient users to evaluate the impact of failing analytical performance on clinical decision-making. Whilst the quantitative effect of failing performance on the magnitude of clinical error may be different for each hs-cTnI platform, similar conclusions could be anticipated for hs-cTnI. We endorse future research endeavors that utilize our methodology to elucidate the relationship between analytical performance and clinical decision-making using hs-cTnI data.

As the medical scientific community acquires more evidence regarding the prognostic relevance of hs-cTn measurements, we should note that any clinical decision-limit that may come out of future studies should be accompanied by clear APS which requires cardiologists to invite clinical chemists to participate in their guideline development to provide knowledge on the analytical performance and its sources of variation [3, 4, 17–19]. The same applies to evidence regarding the relationship between hs-cTn concentration and risk of (cardiac) death after (cardiac) surgery [20, 21], and the possible stratification of clinical decision-limits based on sex, kidney function, body mass index, lifestyle, and the use of decision-limits that are inferred by the exact time difference between the serial troponin measurements [22–26].

Our study shows that indirect modeling approaches can be used to elucidate the effects of real-world APS on clinical decision-making. Reversely, the simulation approach can be adopted to calculate the APS based on the patient data and clinical acceptable limit (in our case <1 % misclassification from rule-in to rule-out). We have carried out this exercise, which revealed that the previously mentioned imprecision of 0.9 ng/L for 0/1 h and 1.2 ng/L for 0/2 h were sufficient in meeting these criteria. We note that as many laboratories utilize coefficients of variation, the APS in terms of CV should be calculated for the concentration of the IQC material used. Previous studies have indicated that adopting a low level QC material near the limit of detection aids in maintaining adequate performance [27].

In conclusion, our study shows that while most laboratories comply (and are increasingly complying) with the strict APS associated with using 0/1 h CDL's in the evaluation of NST-ACS, the effects of not meeting these specifications are substantial. While compliant laboratories have a minute chance to produce results that cause a rule-out of patients that should be ruled-in, failing to meet the analytical requirements

increases this misclassification-rate above an acceptable level of 1 %. The implementation of the 0/2 h algorithm can drastically reduce the misclassification rates in analyzers that are unable to perform with 0/1 h performance specifications. However, for an adequate implementation of the 0/2 h algorithm, an active dialog between the laboratory and cardiology and emergency department specialists is vital.

Acknowledgments: Eline van der Hagen.
Research ethics: The local Institutional Review Board deemed the study exempt from review.
Informed consent: Not applicable.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: None declared.
Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Collet JP, Thiele H, Barbato E, Barthélémy O, Bauersachs J, Bhatt DL, et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: the Task Force for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). Eur Heart J 2020; 1–79. https://doi.org/10.1093/eurheartj/ehaa575.
2. Krintus M, Panteghini M. Judging the clinical suitability of analytical performance of cardiac troponin assays. Clin Chem Lab Med 2023;61: 801–10.
3. Mueller C, Boeddinghaus J, Nestelberger T. Downstream consequences of implementing high-sensitivity cardiac troponin: why indication and education matter. J Am Coll Cardiol 2021;77:3180–3.
4. Planer D, Mehran R, Ohman EM, White HD, Newman JD, Xu K, et al. Prognosis of patients with non–ST-segment–elevation myocardial infarction and nonobstructive coronary artery disease. Circ Cardiovasc Interv 2014;7:285–93.
5. Than M, Herbert M, Flaws D, Cullen L, Hess E, Hollander JE, et al. What is an acceptable risk of major adverse cardiac event in chest pain patients soon after discharge from the Emergency Department?: a clinical survey. Int J Cardiol 2013;166:752–4.
6. van der Hagen EAE, Cobbaert CM, Meijer R, Thelen MHM. Fast 0/1-h algorithm for detection of NSTEMI: are current high-sensitivity cardiac troponin assays fit for purpose? An EQA-based evaluation. Clin Chem Lab Med CCLM 2019;57:1999–2007.
7. SKML. SKML – scheme – cardiac markers. SKML; n.d. https://www.skml.nl/en/rondzendingen/overzicht/rondzending?id=82.
8. Thelen MHM, Jansen RTP, Weykamp CW, Steigstra H, Meijer R, Cobbaert CM. Expressing analytical performance from multi-sample evaluation in laboratory EQA. Clin Chem Lab Med CCLM 2017;55:1509–16.
9. Aakre KM, Røraas T, Petersen PH, Svarstad E, Sellevoll H, Skadberg Ø, et al. Weekly and 90-minute biological variations in cardiac troponin T and cardiac troponin I in hemodialysis patients and healthy controls. Clin Chem 2014;60:838–47.
10. Frankenstein L, Wu AH, Hallermayer K, Wians FH Jr., Giannitsis E, Katus HA. Biological variation and reference change value of high-sensitivity troponin T in healthy individuals during short and intermediate follow-up periods. Clin Chem 2011;57:1068–71.
11. Vasile VC, Saenger AK, Kroning JM, Jaffe AS. Biological and analytical variability of a novel high-sensitivity cardiac troponin T assay. Clin Chem 2010;56:1086–90.
12. Diaz-Garzon J, Fernandez-Calle P, Sandberg S, Özcürümez M, Bartlett WA, Coskun A, et al. Biological variation of cardiac troponins in health and disease: a systematic review and meta-analysis. Clin Chem 2021;67:256–64.
13. R Core Team. R: a language and environment for statistical computing. The R Foundation; 2022. https://www.R-project.org/.
14. Braga F, Pasqualetti S, Aloisio E, Panteghini M. The internal quality control in the traceability era. Clin Chem Lab Med 2020;59:291–300.
15. van Schrojenstein Lantman M, Çubukçu HC, Boursier G, Panteghini M, Bernabeu-Andreu FA, Milinkovic N, et al. An approach for determining allowable between reagent lot variation. Clin Chem Lab Med 2022;60: 681–8.
16. Apple FS, Collinson PO, Kavsak PA, Body R, Ordóñez-Llanos J, Saenger AK, et al. Getting cardiac troponin right: appraisal of the 2020 European Society of Cardiology guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation by the International Federation of Clinical Chemistry and Laboratory Medicine Committee on Clinical Applications of Cardiac Bio-Markers. Clin Chem 2021;67:730–5.
17. Kumar A, Patel DR, Menon V. Cardiac troponin testing: goodbye, “troponinemia”. Cleve Clin J Med 2020;87:483–4.
18. Chapman AR, Lee KK, McAllister DA, Cullen L, Greenslade JH, Parsonage W, et al. Association of high-sensitivity cardiac troponin I concentration with cardiac outcomes in patients with suspected acute coronary syndrome. JAMA 2017;318:1913–24.
19. Lazar DR, Lazar FL, Homorodean C, Cainap C, Focsan M, Cainap S, et al. High-sensitivity troponin: a review on characteristics, assessment, and clinical implications. Dis Markers 2022;2022:9713326.
20. Devereaux PJ, Lamy A, Chan MTV, Allard RV, Lomivorotov VV, Landoni G, et al. High-sensitivity troponin I after cardiac surgery and 30-day mortality. N Engl J Med 2022;386:827–36.
21. Chew MS, Saugel B, Lurati-Buse G. Perioperative troponin surveillance in major noncardiac surgery: a narrative review. Br J Anaesth 2023;130: 21–8.
22. Sandoval Y, Apple FS, Mahler SA, Body R, Collinson PO, Jaffe AS, et al. High-sensitivity cardiac troponin and the 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR guidelines for the evaluation and diagnosis of acute chest pain. Circulation 2022;146:569–81.
23. Gallacher PJ, Miller-Hodges E, Shah ASV, Farrah TE, Halbesma N, Blackmur JP, et al. High-sensitivity cardiac troponin and the diagnosis of myocardial infarction in patients with kidney impairment. Kidney Int 2022;102:149–59.
24. Horne KL, Taal MW, Selby NM. Room for improvement: diagnosing and managing acute coronary syndromes in persons with reduced eGFR. Kidney Int 2022;102:20–2.
25. Vavik V, Aakre KM, Ringdal Pedersen EK, Tveitevåg Svingen GF, Tell GS, Nygård O, et al. The association between serum high-sensitivity cardiac troponin T and acute myocardial infarction in patients with suspected

chronic coronary syndrome is modified by body mass index. Int J Cardiol Cardiovasc Risk Prev 2021;11:200109.
26. Aakre KM, Omland T. Physical activity, exercise and cardiac troponins: clinical implications. Prog Cardiovasc Dis 2019;62:108–15.
27. Aloisio E, Pasqualetti S, Dolci A, Panteghini M. Daily monitoring of a control material with a concentration near the limit of detection improves the measurement accuracy of highly sensitive troponin assays. Clin Chem Lab Med CCLM 2020;58:e29–31.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2023-1226).