

## Review

Mauro Panteghini\*, Alberto Dolci, Sarah Birindelli, Dominika Szoke, Elena Aloisio and Simone Caruso

# Pursuing appropriateness of laboratory tests: a 15-year experience in an academic medical institution

<https://doi.org/10.1515/cclm-2022-0683>

Received July 15, 2022; accepted August 8, 2022;

published online August 23, 2022

**Abstract:** Appropriateness in Laboratory Medicine has been the object of various types of interventions. From published experiences, it is now clear that to effectively manage the laboratory test demand it is recommended to activate evidence-based preventative strategies stopping inappropriate requests before they can reach the laboratory. To guarantee appropriate laboratory test utilization, healthcare institutions should implement and optimize a computerized provider order entry (CPOE), exploiting the potential of electronic requesting as “enabling factor” for reinforcing appropriateness and sustaining its effects over time. In our academic institution, over the last 15 years, our medical laboratory has enforced various interventions to improve test appropriateness, all directly or indirectly based on CPOE use. The following types of intervention were implemented: (1) applying specific recommendations supported by monitoring by CPOE as well as a continuous consultation with clinicians (tumour markers); (2) removing outdated tests and avoiding redundant duplications (cardiac markers, pancreatic enzymes); (3) order restraints to selected wards and gating policy (procalcitonin, B-type natriuretic peptide, homocysteine); (4) reflex testing (bilirubin fractions, free prostate-specific antigen, aminotransferases, magnesium in hypocalcemia); and (5) minimum retesting interval (D-Dimer, vitamin

B<sub>12</sub>, C-reactive protein,  $\gamma$ -glutamyltranspeptidase). In this paper, we reviewed these interventions and summarized their outcomes primarily related to the changes in total test volumes and cost savings, without neglecting patient safety. Our experience confirmed that laboratory professionals have an irreplaceable role as “stewards” in designing, implementing, evaluating, and maintaining interventions focused to improving test appropriateness.

**Keywords:** computerised physician order entry; laboratory medicine; laboratory test utilization; test ordering.

## Introduction

Advising on the optimal use of laboratory tests to improve the clinical effectiveness and patient outcome is one of the main tasks of laboratory professionals [1]. They should take advantage of seeing the entire process of test utilization, which is crucial to identifying and prioritizing efforts to improve ordering strategies [2]. However, practising test appropriateness is not easy as this necessitates updated knowledge supporting changes in policies and diagnostic procedures. As described by Salinas and co-workers [3], the initial step is to identify which laboratory test is used inappropriately and in which patient population. Afterwards, effective interventions should be selected, and their impact monitored and evaluated through suitable quality indicators. Initiatives to manage upstream demand and downstream interpretation of laboratory tests are crucial to defining the mission of medical laboratories [4].

## Effective interventions to improve appropriateness in laboratory medicine

A wide body of literature has described actions and strategies that can be employed to improve test appropriateness. Tools

---

\*Corresponding author: **Mauro Panteghini**, UOC Patologia Clinica, ASST Fatebenefratelli-Sacco, Via G.B. Grassi 74, Milan, Italy; and Dipartimento di Scienze Biomediche e Cliniche, Università degli Studi, Milan, Italy, E-mail: mauro.panteghini@unimi.it

**Alberto Dolci**, UOC Patologia Clinica, ASST Fatebenefratelli-Sacco, Milan, Italy; and Dipartimento di Scienze Biomediche e Cliniche, Università degli Studi, Milan, Italy. <https://orcid.org/0000-0002-9833-7538>

**Sarah Birindelli, Dominika Szoke, Elena Aloisio and Simone Caruso**, UOC Patologia Clinica, ASST Fatebenefratelli-Sacco, Milan, Italy. <https://orcid.org/0000-0003-0769-3156> (D. Szoke). <https://orcid.org/0000-0001-6311-5700> (E. Aloisio)

**Table 1:** Types of intervention strategies proposed to reduce the inappropriate use of laboratory tests.

Intervention category		Activity type	Strength
Educational	Pre-test	Local recommendation	Low
		Advice for clinicians	Medium
		Training and update meetings	Medium
	Post-test	Audits	Low
		Feedback reports on data, activities, and costs	Low
		Formal admonitions	Medium
Information technology-based	Test removal from order entry	High	
	Reflex testing	High	
	Constrains for minimum retesting interval	High	
	Restrains to selected wards	High	
	Gating policy	High	

used for this purpose can be divided into two main categories: educational strategies and information technology (IT)-based interventions (Table 1) [5]. In 2018, a systematic review summarized the available evidence regarding the effectiveness of these interventions [6]. Three activities for improving the practice supporting appropriate laboratory test utilization were recommended: (1) the use of a well-designed computerized provider order entry (CPOE), (2) the use of reflex testing, and (3) the implementation of combined practices (because of a type of ‘summation effect’). On the contrary, there was insufficient evidence for recommending all the remaining types of investigated interventions, such as education, use of expert systems, and feedbacks. Fryer and Smellie summarized in advance this evidence when they stated that to effectively manage demands for laboratory tests and reduce inappropriate requesting it is mandatory the activation of preventative strategies stopping inappropriate requests before they could reach the laboratory [7]. In line with this belief, we previously recommended that healthcare institutions should fully

exploit the potential of electronic requesting acting as “enabling factor” for reinforcing educational messages and sustaining their effects over time [8]. While modifying CPOE can be labor-intensive, and time and resources required to develop interventions (e.g., for establishing how to modify order sets or how a new order form should be designed) can be significant, once such modifications are agreed and implemented, their efficacy is basically high as all requestors receive the intervention upon ordering. Poor functionality within IT systems, however, may limit the effective implementation of such strategies [9].

### Developing and applying recommendations

Educational measures alone are the weakest and the least lasting approach over time, and recommendations alone have little impact and are insufficient agents of change [10]. To become effective and finally achieve efficacy of

care, recommendations should be blown up at the local level into more practical rules, extensively stating how, when, where and why specific tests should or should not be ordered. Furthermore, the release of local recommendations to guide laboratory test ordering is not enough to curb the excess of requests and maintain the appropriateness if it is not supported by strict monitoring by CPOE as well as a continuous consultation with requesting clinicians [11].

## Removing outdated tests and avoiding redundant duplications

Laboratories should revise their menus for obsolete tests on a regular basis. Removing tests that offer little incremental information would save money, avoid additional investigations arising from incidental and clinically irrelevant abnormalities, and improve the risk to benefit ratio. On the other hand, there are pairs of tests that are jointly ordered in many laboratories, although in most cases there are no pathophysiological reasons for associating these requests, making this routine procedure an important source of redundant duplications [8].

## Reflex testing

A reflex test is a laboratory test performed exclusively following certain results of the first related test. The approach is based on pre-determined criteria that, if are met by first-level test results, trigger the reflex test itself. The reflex test is therefore performed automatically, without the need of a second order. Benefits are the reduction in unnecessary second level test measurements and the no need of additional specimen collection, promoting timeliness in driving medical reasoning and decision. Usually, these types of interventions have a high efficacy because they constitute an all-in-one evaluation, governed by IT, that differs from the same tests when requested individually and opposes to the practice of ‘shotgun’ testing, in which more tests are performed simultaneously [8]. In some cases, reflex testing may also prevent test underutilization [12].

## Constrains for minimum retesting interval (MRI)

Several tests are believed to be too frequently ordered and the introduction of MRI has been advocated [13–15]. MRI interventions concern the periodicity of retesting (i.e., the

minimum time before a test should be repeated, based on the properties of the test and the clinical situation in which it is used) [16]. Criteria for determining MRI are based on different measurand characteristics, which include biological properties, analytical aspects, treatment needs, monitoring requirements, and established clinical guidance. Each test implanted with this type of appropriateness intervention cannot be requested more than once in an established period. This approach has a high effectiveness because the block is plugged directly into the CPOE and cannot be circumvented by ordering physicians unless laboratory specialists decide to force the IT rule and restore the rejected request in case of proven clinical need.

## Restraints to selected wards and gating policy

Applying restricted policies of test utilization, though vetting of their orders (the so-called gating policy) is an additional important option that laboratories can promote to preserve the cost-benefit, particularly for complex and costly tests with proven utility only in specific medical conditions. This intervention is extremely effective because the only way for clinicians to obtain the test is to contact laboratory specialists for justifying the request and override automated rules [8].

## The experience in our academic institution

### Setting

The Clinical Pathology Unit (CPU) of the ASST Fatebenefratelli-Sacco in Milan (a multidisciplinary teaching institution affiliated with the University of Milan) is a regional accredited medical laboratory serving a network of four public hospitals located within the city and through northern hinterland (two general hospitals with all major adult specialties and two infant-maternity hospitals), according to the hub-and-spoke model. With a staff of 15 people specialized in laboratory medicine and a junior staff of eight trainees, it performs approximately 3 million tests per year, including tests for clinical wards, intensive care units (ICU), emergency departments (ED), outpatients, and two affiliated retirement homes. During the last decade, one of the most important changes related to CPU has been the creation of a core laboratory (core-lab) structure in each hospital setting using total laboratory automation (TLA) for performing first-line tests. Restructuring CPU by creating local core-lab facilities has

permitted to consolidate a Laboratory Medicine Department by increasing workload efficiency, but, more importantly, has provided the occasion to create a decision making-based laboratory department, characterised by a very short turnaround time (TAT) for all tests performed in the core-labs and by satellite laboratory sections devoted to performing tests requiring specialised knowledge (e.g., for protein diagnostics, oncology, haematology, etc.), with more fruitful cooperation with care teams for specific medical conditions, better allowing performed tests to effectively work in the right clinical setting [17].

## Information system background and summary of interventions

For any institution trying to achieve optimum patient care (and control costs), the availability of an appropriate information system is critical [18]. Indeed, the majority of the interventions reviewed in this paper followed the introduction of a new CPOE (Galileo, Dedalus), which took place in 2010 (Table 2). The introduction of the new CPOE allowed CPU to effectively link the educational strategies for the correct use of laboratory tests, consisting of local recommendations and interdepartmental audits, with direct interventions adapting

IT to the improvement of demands. In parallel, to support changes of laboratory test management, a new laboratory information system (LIS) (DnLab, Dedalus) was introduced, which contributed to enrich the list of appropriateness interventions, by using blocks and constraints, together with a substantial improvement in data storage and statistical analyses, vital for performing a careful evaluation and monitoring of appropriateness intervention performance over time. Last but not least, both analyzer software and the laboratory middleware (AlinIQ AMS, Abbott Diagnostics) through a query program language (QPL) had also a central role in supporting interventions such as a reflex testing, permitting their execution in a fully automated way. As the labour time of laboratory professionals is costly, softwares have a primary role in supporting their work for improving test utilization management.

## Detection of inappropriate requests and evaluation metrics

Usually, it is challenging to assess the true number of tests inappropriately requested, because this would imply the evaluation of each single clinical case with subsequent

**Table 2:** Summary of the intervention strategies implemented in our academic healthcare system to reduce the inappropriate use of laboratory tests.

Type of intervention	Laboratory test	Year of introduction (setting)
Applying local recommendations together with continuous consultation with requestors	Tumour markers	2007 (inpatients)
Removing obsolete tests	CK-MB activity, myoglobin	2005
	Total $\alpha$ -amylase	2005
	CK-MB mass	2016
Reflex testing	Bilirubin reflex	2009 (inpatients) 2015 (outpatients)
	PSA reflex	2009 (inpatients) 2014 (outpatients)
	ALT reflex	2011 (inpatients)
	Icteric index reflex	2016 (inpatients)
	Calcium reflex	2019 (inpatients)
	Minimum retesting interval constraints	Vitamin B <sub>12</sub>
Restrains to selected wards and gating policy	C-reactive protein	2018 (inpatients)
	D-Dimer	2018 (inpatients)
	Procalcitonin	2018 (intensive care units)
	$\gamma$ -glutamyl transferase	2018 (inpatients)
	NT-proBNP	2005 (emergency departments)
	Procalcitonin	2008 (intensive care units, pediatrics)
	Homocysteine	2013 (stroke unit, pediatrics)
	Pancreatic amylase	2014 (outpatients)

verification of the real impact that a given test has brought to patient care, in relation to disease diagnosis, prognosis and/or a more general outcome. To try to quantify the phenomenon of inappropriate requests, CPU has, over the years, constantly performed evaluations and institutional audits to review and assess the status of use of performed tests, enforcing and putting in the field, on a case-by-case basis, various countermeasures with the aim of stemming prescriptive inappropriateness. It is noteworthy to remember that the review of laboratory performance required co-operation with functional areas outside the laboratory, providing the basis of a cooperative venture among medical specialty fields.

Table 2 summarizes the main interventions together with the year of their introduction. Once implemented, the measures were continuous.

Strategies were monitored through two main process indicators:

- (1) Changes in total test volumes.
- (2) Economic impact (cost savings).

In addition, to assess the potential detrimental impact of the interventions on patient care, we monitored the reporting of possible situations or harmful consequences for patients through our institutional quality system.

Costs were computed by multiplying the reagent price for a given test (other fixed laboratory fees were not included) by the number of tests in the corresponding category. Savings were defined as the difference, if any, between the before and after intervention expenses.

## Ethics

This article retrospectively presents processes and outcomes of demand management strategies, which were carried out in routine care. No interventions were performed solely for study purposes, and all data were fully anonymized. Therefore, there was no requirement for the Ethics Committee approval.

## Result description

### Tumour markers

In 2006, in partnership with oncologists, CPU developed local recommendations on the correct use of tumour markers (TM) in hospitalized patients [19]. The main scope was to improve TM prescriptive appropriateness, harmonizing their use by specifying the clinical characteristics of each TM in relation to different malignancies. As

a rule, we established a maximum of two TM requests in the same order except for well documented clinical situations. Through the LIS, TM requests not compliant with agreed-on recommendations were automatically identified and, before making any final decision related to their (in) appropriateness, orders were discussed with the requesting clinician. After one year, we experienced a large decrease in the number of ordered tests (in average, -55%), without any negative clinical impact. In terms of reagent costs, CPU saved € 38,229 per year [20]. Six years later, CPU performed a second 21-month (2012–2014) audit by specifically checking all requests containing more than two TM, blocked and discussed with ordering clinicians. 3.6% of total requests still exceeded two TMs, containing a median of 3 (up to 7) TM. Consultations led to withdraw 43.3% of TM requests, which were removed by laboratory specialists because they did not meet appropriateness criteria [11]. This showed that to curb the excess of requests and maintain the TM appropriateness for a long time (Table 3), local guidelines should be supported by strict daily monitoring by laboratory professionals as well as a continuous consultation with requesting clinicians. In this study, we also tried to estimate the impact and the dimension of TM request inappropriateness. By considering the epidemiological estimates of tumour incidence and prevalence in Milan (2010 data), the expected total number of TM tests per year amounted to ~126,000, if TM ordering would fulfil international recommendations. By comparing this estimate with the number of TM tests yearly performed in Milan medical laboratories (~350,000), we estimated that performed TM tests exceeded the justified ones by approximately three-fold. This regional-based estimate was later confirmed at the national level [21].

**Table 3:** Percent variations of tumour marker requests, obtained one year after the introduction of local recommendations (2007) and six years later by continuous consultation with requesting clinicians (2012–2014).

Marker	Request change after recommendation introduction	Request change after consultation
α-Fetoprotein	-68%	-65%
Carcinoembryonic antigen	-43%	-25%
CA 125	-65%	-53%
CA 15.3	-60%	-48%
CA 19.9	-61%	-35%
Neuron-specific enolase	-66%	-43%
Human chorionic gonadotropin	-12%	-68%

## Cardiac markers

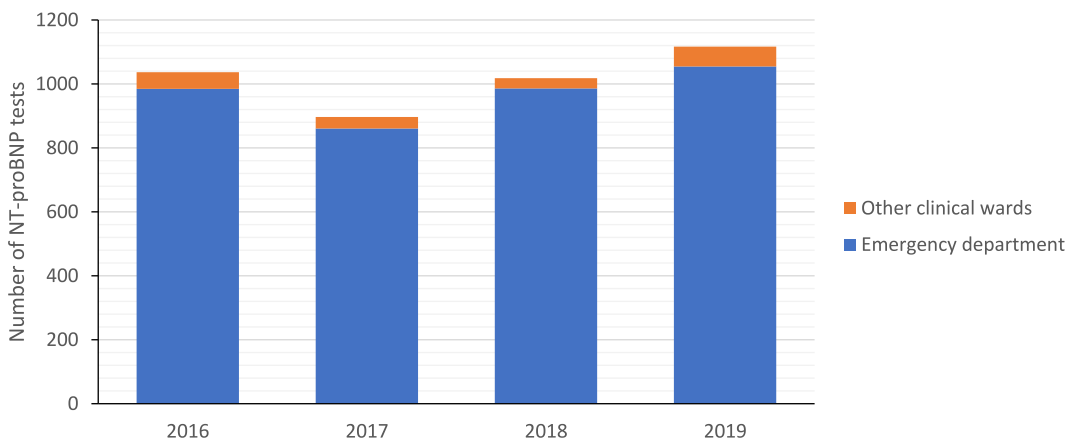
Today, the measurement of cardiac troponins represents the biochemical “gold standard” that is central to the new millennium’s diagnostic criteria for acute myocardial infarction (AMI) [22]. The availability of highly sensitive troponin assays has markedly shortened the time to rule out or rule in AMI and has improved the prognostic assessment of critical patients in clinical contexts different from acute coronary syndrome [23]. Due to their superior diagnostic accuracy, cardiac troponins should definitively replace other traditional ‘cardiac’ markers, such as creatine kinase (CK) and its MB isoenzyme or myoglobin, for diagnosing suspected AMI [24]. Accordingly, laboratory professionals have a central role in removing from the menu these tests that have become obsolete and useless, contributing to reduce possible confusion in data interpretation and patient management.

By experiencing an indiscriminate and unjustified use of all tests included in the so-called ‘cardiac profile’ and the lack of adherence to whatever protocol concerning the evidence-based use of cardiac markers, in 2005, our healthcare institution introduced new strategies for optimal use of cardiac markers by practice guidelines devised by a multidisciplinary team [25]. The new strategies brought relevant changes in the protocols for cardiac marker requests, with the elimination of obsolete tests (i.e., myoglobin and measurement of CK-MB activity) and the introduction of new procedures for requesting and interpreting cardiac troponin for AMI rule-in/out and risk stratification, and CK-MB mass for infarct size estimation and detection of post-percutaneous coronary intervention injury. One year later, a comprehensive audit evaluating the guideline effectiveness on test utilization and costs showed a decrease of –33% in the total number of tests. Myoglobin and CK-MB (as activity) were completely

abolished, whereas CK-MB mass showed a –87.7% reduction. Testing costs were reduced by € 104,871 per year [26]. The audit was repeated after two years from the recommendation release, showing a further decrease (–6.5%) in the number of tests still available (troponin and CK-MB mass) [27]. This proved that the objectives of containing the test number and costs were fully maintained.

As reported above, in 2005 we introduced two main changes related to the CK-MB test. Firstly, we removed the CK-MB activity assay from the menu, letting only CK-MB mass available. Secondly, we limited the use of this test to only two specific clinical cardiology situations (see above). In the following decade (2005–2015), a scientific debate started however, supporting the full replacement of this test with troponin, as CK-MB was considered to provide no incremental information [28–30]. This was translated in a ‘Choosing Wisely’ troponin-only recommendation released in 2015. Accordingly, in 2016 we eliminated CK-MB from the laboratory portfolio, indicating to users to employ troponin for any scope in cardiology. After CK-MB removal, no complaints were received from clinicians and, more importantly, no discernible negative effects on clinical care were detected. The cost analysis showed a further € 3,000 per year savings.

The main indication of B-type natriuretic peptide testing is to find out the aetiology of dyspnoea of unknown causes. In the acute setting, using BNP or NT-proBNP may therefore facilitate shorter ED visit durations [31]. For this application, B-type natriuretic peptide should only be measured once per acute episode. We included this recommendation in our local guideline, also limiting the NT-proBNP free availability to the ED only [25]. For other clinical wards, NT-proBNP requests should be preventively approved by laboratory specialists, who should be contacted by clinical requestors if symptoms are suggestive of heart failure. This kept basically unchanged the number of performed tests (Figure 1) in a setting



**Figure 1:** Number of NT-proBNP tests performed annually in our institution for inpatients (including those admitted to the Emergency Department) over the 2016–2019 period.



(hospitalized patients) where the risk of overuse has been frequently reported [32].

## Pancreatic enzymes

Today, lipase measurement in serum is the recommended laboratory test to diagnose acute pancreatitis. Particularly, increase of serum lipase activity to greater than three times the upper reference limit, in the absence of renal failure, is a more specific diagnostic finding than increases in serum  $\alpha$ -amylase activity [33]. Therefore, it has been recommended that lipase should definitively replace  $\alpha$ -amylase as the initial diagnostic test for acute pancreatitis in the ED [34]. The specificity of  $\alpha$ -amylase for the diagnosis of acute pancreatitis is low because increased values are also found in several acute intra-abdominal disorders and in a number of extra-pancreatic conditions, including macroamylasemia and some cancers. Lack of specificity of total  $\alpha$ -amylase measurement has promoted the direct measurement of pancreatic amylase isoenzyme (P-AMY) instead of total enzyme activity for the differential diagnosis of patients with acute abdominal pain [35]. Accordingly, in 2005 CPU did the first improvement step in the field of pancreatic enzymes by replacing the measurement of total  $\alpha$ -amylase with P-AMY. In a simulation, we estimated the economic impact of this replacement in our ED setting, accounting to approximately € 130,000 per year savings (Table 4) [36]. In the meantime, studies confirmed the superiority of lipase in comparison to P-AMY in terms of diagnostic performance [37]. Therefore, in a following step, CPU eliminated the P-AMY measurement from the list of ED tests keeping only lipase available. Surprisingly, one year after the change implementation, total test requests (P-AMY + lipase) were halved. The expected dramatic decrease of P-AMY (−98.5%) was only partly compensated by the increase in lipase requests (+65%). As no significant change occurred in numbers of patients admitted to ED

with suspected pancreatitis in the two evaluated periods, the decrease in test number only reflected the change in the test availability in ED setting. Unlike P-AMY, possibly used as a screening test in unselected patients, the use of lipase appeared to be more correctly restricted to a symptomatic population with suspected pancreatitis. Owing to the reduction of total test number, reagent costs were also reduced by € 18,500 per year [38]. The last step concerning pancreatic enzymes involved all hospitalized patients for whom the possibility to request both lipase and P-AMY was abolished in 2014 when core-lab installations were started, as the pathophysiologic evidence showed incontrovertibly that obtaining of both enzymes for diagnosis of pancreatic disease represented an unwarranted duplication without any diagnostic advantage [39, 40]. The cost analysis showed a further € 3,300 per year savings.

## D-Dimer

D-Dimer value lies in its high sensitivity for hypercoagulability conditions, so that concentrations below appropriate cut-offs may exclude venous thromboembolism pathologies, with a markedly elevated negative predictive value [41]. To improve the appropriateness of D-Dimer requests, in 2014 our healthcare institution issued a local guidance that had an immediate effect by reducing the number of ordered tests in the following year (−45%). However, this educational intervention mitigated test inappropriateness only for a relatively short period, as in 2016 D-Dimer requests started to raise again (+22% when compared with 2015). Therefore, in 2018 we decided to introduce a CPOE-based intervention by implementing a 24 h block on the periodicity of retesting. If considered clinically appropriate, the laboratory could, however, be contacted by clinical requestors to support the earlier retesting request. After this further corrective action, D-Dimer requests showed a −24% decrease after one year, which was further consolidated in 2019 (−15% vs. 2018). A saving of ~30,000 €/year due to the improved prescription appropriateness was obtained, about one third of which just attributable to the control of retesting periodicity [42].

**Table 4:** Estimated economic outcome of the replacement of total  $\alpha$ -amylase (AMY) with pancreatic AMY isoenzyme measurement in the Emergency Department (ED) of our healthcare system.

No. of patients performing pancreatic AMY in ED per year	22,500	Estimated number of total AMY false positive results per year <sup>a</sup>	2,700
Annual increase of reagent costs for pancreatic AMY vs. total AMY	€ 6,450	Annual savings in second-level diagnostic procedures (e.g., abdominal ultrasound) <sup>b</sup>	€ 135,000

<sup>a</sup>12% of total AMY measurements as estimated by Panteghini M et al. [35]. <sup>b</sup>Assuming a cost of € 50.00 per procedure.

## Procalcitonin

Procalcitonin is an expensive test, which can burden medical laboratories' budgets if requests spin out of control. Not rarely, clinicians believe in the absolute diagnostic ability of procalcitonin to detect bacterial sepsis, but the literature evidence shows that its diagnostic power is

however limited. On the other hand, using procalcitonin to optimize antibiotic therapies in critically ill patients can be cost-effective, but only if there is high adherence to proposed algorithms for antibiotic stewardship [43]. Evidence also exists that procalcitonin may be useful in paediatrics, especially in children with suspected meningitis, even if some confounding factors, such as the physiologically higher concentrations in newborns with less than 72 h of life, should be correctly recognized [44]. In line with this evidence, CPU offered an unrestricted procalcitonin testing only to ICUs (as an aid in decision for continuing or stopping antibiotics) and paediatric wards. For all other clinical wards, procalcitonin requests should be preventively approved by laboratory specialists, contacted by clinical requestors to discuss about the clinical suspicion supporting the procalcitonin request in addition to other already available tests (e.g., C-reactive protein [CRP]). In this strictly controlled situation, during 2017 we recorded however a +85% increase in the number of procalcitonin determinations, causing the laboratory's test budget to exceed maximum expenditure limits. The contributors to procalcitonin testing increase were ICUs, where intensivists were often unwilling to interrupt antimicrobial therapies based on laboratory results leading to a situation where a series of procalcitonin measurements were done uselessly, therefore literally throwing "money down the drain" [45]. Interestingly, procalcitonin measurements performed for non-ICU adult wards requiring approval from laboratory professionals, which accounted for about one test/day, remained unchanged. To bring the situation back under control, at the beginning of 2018 the following measures were undertaken: (a) a standard comment to the procalcitonin report was introduced to alert intensivists when the 80% decrease from peak value was reached, which should be the cue to stop both antibiotic administration and procalcitonin testing; (b) a 24-h MRI was implemented, based on the analyte's half-life; and (c) an update of the internal guidelines for antibiotic stewardship was released asking for a higher adherence to the algorithm, and educational seminars organized. These integrated interventions obtained a -31% decrease of procalcitonin tests one year later, which was also maintained unchanged in 2019. The corresponding savings were 9,000 € per year [46].

## Homocysteine

Homocysteine testing is appropriate in case of suspected homocystinuria (an inherited disorder of the metabolism of the amino acid methionine), in patients with previous venous

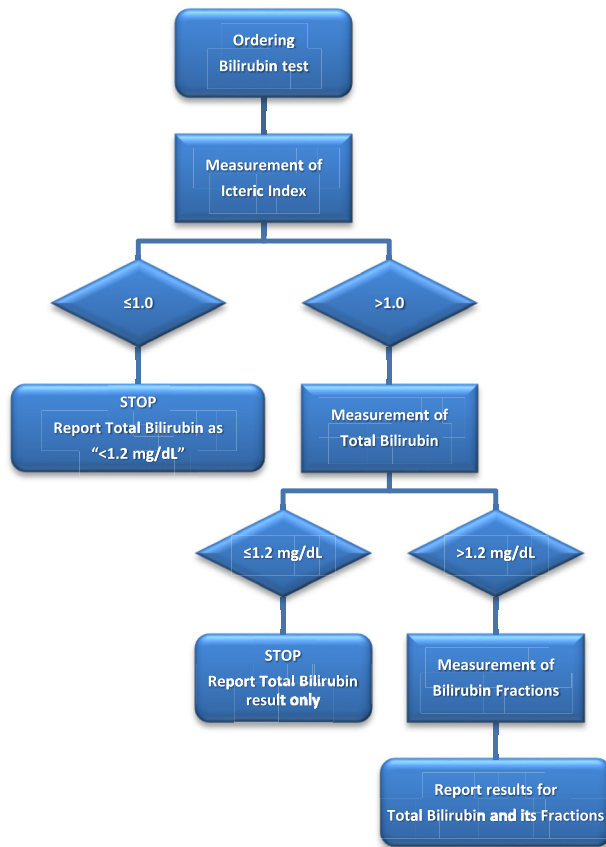
or arterial thromboembolism and in patients with severe hyperhomocysteinemia treated with B-complex vitamins. Conversely, the measurement of plasma homocysteine is not recommended for cardiovascular disease screening in the general population [47]. In 2012, our institution spent more than 50,000 € to purchase reagents for homocysteine determination, putting this test in the second place of the most expensive ones among those performed by our laboratory. This was mainly due to the inappropriate use of homocysteine testing for detecting hypercoagulation in poorly selected subjects. Therefore, in 2013 CPU introduced a CPOE-based restriction for the test request, allowing to order it freely only by some specialized wards (e.g., the Stroke Unit). Most hospital wards may obtain the test only after the requesting clinician obtained authorization by laboratory specialists. Accordingly, there was a reduction in test requests from 1,430 in 2012 to 261 three years after (-81.7%), with an estimated savings of 36,900 € per year.

## Reflex tests

Measurement of bilirubin fractions in serum is part of a consolidated reflex test, in which this happens only when the total bilirubin concentration is higher than the upper reference limit [48]. To further expand the appropriateness of total bilirubin ordering, as initially proposed by Salinas et al. [49], in 2016 CPU introduced for hospitalized patients an additional reflex test using the assessment of icteric index as front-line test for the identification of blood samples with abnormal total bilirubin concentrations, which are the only ones that need the measurement of total bilirubin (Figure 2) [50]. The application of an optimal cut-off for icteric index that reliably identifies abnormal total bilirubin concentrations (which must be validated according to the employed analytical platform) allowed the accurate "zero-cost" detection of samples with normal total bilirubin concentrations, with a sensitivity  $\geq 99\%$  for discriminating between specimens with high or normal TB, with a false negative rate of 0.1% [51]. This avoided direct measurements in ~40% of bilirubin orders in our clinical setting resulting in economic savings of about 5,000 € per year. We should not forget that for obtaining optimal results with this approach, the photometric determination of icteric index should be subjected to a structured quality assessment as all other laboratory tests [52].

Although the role of free prostate-specific antigen (PSA) measurement in serum as second level investigation to minimize total PSA false-positive results is still in dispute, available guidelines recommend its determination only when the total PSA concentrations range between 3 and 10  $\mu\text{g/L}$  [53]. By auditing free PSA requests in our institution, in 2006





**Figure 2:** The two sequential reflex test strategy adopted in our institution when the ‘Bilirubin’ test is requested. Note that the icteric index cut-off is measuring system dependent, i.e. 1.0 is the validated optimal cut-off for the Abbott Alinity c platform currently employed in our laboratories.

we reported that only 15% of those requests complied with this recommendation, with an estimated economic waste for our health care system of ~50,000 € per year [54]. These data supported the activation of a reflex test allowing free PSA determination only when total PSA fell within the recommended concentration range and labelling as “inappropriate” the free PSA requests in samples with total PSA out of the recommended limits. This reflex testing was first introduced in 2009 for inpatients and five years later for outpatients. In 2019, we audited data showing that the free PSA reflex testing worked quite well in decreasing the free PSA inappropriateness, 96% being the reached rate of appropriately measured free PSA [55].

The two aminotransferases are a pair of tests jointly ordered in many laboratories when hepatocellular damage is suspected, although there are no pathophysiological reasons for associating these requests, making this approach an important source of redundant duplications [8]. To limit the inappropriate aspartate aminotransferase

(AST) testing, it has been recommended that laboratories should offer it as a reflex test only in samples with abnormal alanine aminotransferase (ALT) results [33]. In 2011, after consultation with clinicians, CPU deleted from the order entry panel the AST request and introduced an automatic ALT reflex test. If considered clinically appropriate, the laboratory could, however, be contacted by clinical requestors to support the direct AST request in addition to ALT. However, during the following 10-year experience, no extra requests for AST determination to supplement diagnosis were registered, and no detrimental situations for patients were reported. Overall, a 90% reduction of AST requests was obtained without any negative impact on patient safety. The strategy resulted in an average savings of reagent costs of 5,000 € per year [56].

The association between hypomagnesemia and hypocalcemia is well documented, often making the latter refractory if the former is not recognized [57]. Salinas and co-workers proposed the adoption of a reflex test automatically adding serum magnesium determination to samples with severe hypocalcemia to be applied in an ED setting [58]. Similarly, in 2019 we introduced for all clinical wards the automatic reflex addition of magnesium to serum samples with severe hypocalcemia, defined as a serum calcium concentration <7.0 mg/dL. During the following 16 months (July 2019–October 2020), the introduction of this hypocalcemia reflex test was able to detect 10% of all cases of mild hypomagnesemia (i.e., serum magnesium <1.4 mg/dL) and 22% of all cases of severe hypomagnesemia (i.e., serum magnesium <1.0 mg/dL) in our hospitalized patients. More importantly, the reflex approach induced clinicians to administer magnesium sulphate as replacement therapy in 40% of these patients, two-third of which showing severe hypomagnesemia [59]. The workload aggravation in terms of number of tests and costs was very limited, corresponding to ~1% of the total magnesium determinations.

## MRI constraints

Requests for vitamin B<sub>12</sub> (B12) measurement should aim to define a possible vitamin deficiency in specific patient categories [60]. After suspecting an overutilization of this test, in 2014 CPU introduced an IT-based constraint for B12 monitoring, based on a recommended MRI of 6 months, to reduce unnecessary B12 testing [61, 62]. The introduction of this IT rule on CPOE for inpatients immediately had a marked effect with a 49% reduction of B12 requests during 2015, resulting in 5,800 € savings of reagent costs. In the following years, B12 requests assumed a plateau behaviour with minor variations.

CRP is one of the most ordered tests, covering 5–6% of all requests annually made in our institution and some Authors have considered this test as a typical example of the impact of unmanaged demand [63]. An automated rejection rule based on MRI for serum CRP was reported as a sustainable method for reducing unnecessary test repeat [64]. Therefore, in 2018 an IT-based rule was developed by CPU to block CRP requests made within a 24 h time window of an initial request. After implementing this policy, requests for inpatients decreased from 53,536 in 2017 (before the block implementation) to 50,803 in 2019 after the MRI policy introduction (–5.1%). By the way, the percentage of total CRP workload rejected in our institution was quite similar to that found in studies applying the same approach (–5.9%) [16].

Over the years, our institution recorded a significant increase in requests for  $\gamma$ -glutamyl transferase (GGT). It was suggested that poor understanding by clinicians of the timeframe of GGT changes in pathophysiology may partially contribute to this increased demand [65]. Up to 20% of GGT repeats have been related to a too short retesting time [16]. Considering an average half-life of GGT in serum of ~100 h, an MRI of 4 days could be advocated [33]. Using a more conservative MRI of 36 h, in 2018 we introduced a CPOE block to prevent early GGT repetitions. This led to a reduction of –8.5% GGT requests during 2019.

## Discussion

Failure in laboratory test ordering and result utilization are major contributors to diagnostic errors [66]. Therefore, the adoption of targeted interventions for improving test appropriateness is of critical importance. The long journey of our CPU, operating in an academic hospital network in Milan, with the aim to improve the test appropriateness has included more than 15 years of relentless research of scientific evidence, performance of audits, and implementation of interventions. With the aim of achieving a better management of many laboratory tests offered to clinician users, this work has required the active participation of the whole laboratory team by using their skills and improve their knowledge for identifying test utilization issues and applying appropriate countermeasures without putting patient safety at risk. As a further option of an optimally structured LIS, it was also possible to analyse the economic trend over time. By considering the annual cumulative saving projections following the introduction of the various appropriateness interventions described in this paper, this evaluation revealed that the cumulative monetary savings

exceeded 3,000,000 € in the 15 years considered (from 2005 to 2019).

Not all performed tests can be involved in appropriateness interventions. Given the enormous amount of work required for the assessment and implementation of an appropriateness intervention, combined with low laboratory costs for some daily measured biomarkers, some tests are usually not included in any appropriateness pathway. However, medical appropriateness implementation is a field that is always rich in possibilities, with the aim of improving laboratory cost-effectiveness profile and to reach the clinical awareness on the correct use of laboratory tests. Furthermore, the ever-changing scientific and technological progress makes this study branch an environment in continuous evolution, which should be constantly monitored and renewed to keep its effectiveness intact.

## Study limitations

Due to different evaluation criteria and settings, the presented results may not be generalizable or comparable. Our institutional initiatives focussed on the improvement of appropriateness of laboratory test use and on reduction of unnecessary testing and increase of cost-effectiveness regarding a list of tests where the audits indicated serious problems of overuse. The impact on patient clinical outcomes was just indirectly evaluated by verifying if detrimental situations or harmful consequences for patients were reported through the institutional quality system. Although no negative reports were received, it remains unclear to what extent patient outcomes are linked to obtained reductions in laboratory testing. In addition, as previously pointed out [67], our double role of intervention developers and of evaluators of the process efficacy may be a potential source of bias.

## Conclusions

Cadamuro et al. encouraged medical laboratories to publish their strategies in use to manage tests inappropriateness, with the aim that all involved figures may profit from these experiences [68]. Accordingly, in this paper, we presented our extensive experience that has two major aspects of novelty. Firstly, based on previously published studies, only interventions with demonstrated high strength of efficacy were implemented. Secondly, those interventions were applied in daily practice (and not just for the limited time of a specific study) and monitored throughout years of

application. The quality of reported evidence was therefore enhanced by showing the impact of interventions over longer periods of time, while previously the long-term sustainability of results was often questioned [69]. Finally, our experience confirmed that laboratorians have an irreplaceable role in designing, implementing, evaluating, and maintaining interventions focused to improving test appropriateness. Pursuing and maintaining appropriate test requests is a daily achievement with important bioethical aspects that should not be forgotten [70].

**Research funding:** None declared.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Competing interests:** Authors state no conflict of interest.

**Informed consent:** Not applicable.

**Ethical approval:** Not applicable.

## References

- Ferraro S, Braga F, Panteghini M. Laboratory medicine in the new healthcare environment. *Clin Chem Lab Med* 2016;54:523–33.
- Baird GS, Jackson BR, Dickerson J, Sikaris K, Croal B, Laposata M. What's new in laboratory test utilization management? *Clin Chem* 2018;64:994–1000.
- Salinas M, Flores E, López-Garrigós M, Leiva-Salinas C. Laboratory test inappropriateness: lessons revisited and clarified in seven questions. *J Lab Precis Med* 2018;3:34.
- Plebani M, Panteghini M. Promoting clinical and laboratory interaction by harmonization. *Clin Chim Acta* 2014;432:15–21.
- Eaton KP, Levy K, Soong C, Pahwa AK, Petrilli C, Ziemba JB, et al. Evidence-based guidelines to eliminate repetitive laboratory testing. *JAMA Intern Med* 2017;177:1833–9.
- Rubinstein M, Hirsch R, Bandyopadhyay K, Madison B, Taylor T, Ranne A, et al. Effectiveness of practices to support appropriate laboratory test utilization: a laboratory medicine best practices systematic review and meta-analysis. *Am J Clin Pathol* 2018;149:197–221.
- Fryer AA, Smellie WS. Managing demand for laboratory tests: a laboratory toolkit. *J Clin Pathol* 2013;66:62–72.
- Ferraro S, Panteghini M. The role of laboratory in ensuring appropriate test requests. *Clin Biochem* 2017;50:555–61.
- McCall SJ, Souers RJ, Blond B, Massie L. Physician satisfaction with clinical laboratory services: a College of American Pathologists Q-Probes study of 81 institutions. *Arch Pathol Lab Med* 2016;140:1098–103.
- Rosenberg A, Agiro A, Gottlieb M, Barron J, Brady P, Liu Y, et al. Early trends among seven recommendations from the choosing wisely campaign. *JAMA Intern Med* 2015;175:1913–20.
- Ferraro S, Mozzi R, Panteghini M. Tumor marker ordering: do not lose control: a prospective clinical trial. *Am J Clin Pathol* 2015;144:649–58.
- Salinas M, Puche CM, Lopez-Garrigos M, Flores E, Leiva-Salinas C. Serum magnesium, the forgotten test: intervention from the laboratory to identify emergency department patients with hypomagnesemia. *Clin Chim Acta* 2019;493:S699.
- Morgen EK, Naugler C. Inappropriate repeats of six common tests in a Canadian city: a population cohort study within a laboratory informatics framework. *Am J Clin Pathol* 2015;144:704–12.
- Chami N, Simons JE, Sweetman A, Don-Wauchope AC. Rates of inappropriate laboratory test utilization in Ontario. *Clin Biochem* 2017;50:822–7.
- Hueth KD, Jackson BR, Schmidt RL. An audit of repeat testing at an academic medical center: consistency of order patterns with recommendations and potential cost savings. *Am J Clin Pathol* 2018;150:27–33.
- Lang T. Minimum retesting intervals in practice: 10 years experience. *Clin Chem Lab Med* 2020;59:39–50.
- Dolci A, Giavarina D, Pasqualetti S, Szoke D, Panteghini M. Total laboratory automation: do stat tests still matter? *Clin Biochem* 2017;50:605–11.
- Hughes AEO, Jackups R. Clinical decision support for laboratory testing. *Clin Chem* 2022;68:402–12.
- Panteghini M, Piazza E, Dolci A, Scapellato L, Luraschi P. Strategie per l'impiego ottimale dei biomarcatori in oncologia: raccomandazioni e protocolli operativi. *Biochim Clin* 2006;30:210–23.
- Dolci A, Scapellato L, Panteghini M. One-year audit of the implementation of recommendations for optimal use of tumor biomarkers in a university hospital. *Biochim Clin* 2008;32:181–5.
- Gion M, Peloso L, Trevisiol C, Squarcina E, Zappa M, Fabricio AS. An epidemiology-based model as a tool to monitor the outbreak of inappropriateness in tumor marker requests: a national scale study. *Clin Chem Lab Med* 2016;54:473–82.
- Ferraro S, Panteghini M. Laboratory medicine as the science that underpins medicine: the “high-sensitivity” troponin paradigm. *Clin Chem Lab Med* 2015;53:653–64.
- Ferraro S, Dolci A, Panteghini M. Fast track protocols using highly sensitive troponin assays for ruling out and ruling in non-ST elevation acute coronary syndrome. *Clin Chem Lab Med* 2017;55:1683–9.
- Dolci A, Panteghini M. The exciting story of cardiac biomarkers: from retrospective detection to gold diagnostic standard for acute myocardial infarction and more. *Clin Chim Acta* 2006;369:179–87.
- Panteghini M, Belloni A, Botta M, Dominici R, Luraschi P, Magenta M, et al. Strategie per l'impiego ottimale dei biomarcatori cardiaci: raccomandazioni e protocolli operativi. *Biochim Clin* 2005;29:419–28.
- Dominici R, Dolci A, Valente C, Infusino I, Panteghini M. One-year audit of the implementation of protocols for optimal use of cardiac biomarkers in an Italian university hospital. *Clin Chem Lab Med* 2007;45:S229.
- Dominici R, Dolci A, Valente C, Panteghini M. Monitoraggio dell'efficacia di strategie per l'impiego ottimale dei biomarcatori cardiaci a due anni dall'introduzione. *Biochim Clin* 2008;32:462.
- Saenger AK, Jaffe AS. Requiem for a heavyweight: the demise of creatine kinase-MB. *Circulation* 2008;118:2200–6.
- Singh G, Baweja PS. Creatine kinase-MB: the journey to obsolescence. *Am J Clin Pathol* 2014;141:415–9.
- Alvin MD, Jaffe AS, Ziegelstein RC, Trost JC. Eliminating creatine kinase-myocardial band testing in suspected acute coronary syndrome: a value-based quality improvement. *JAMA Intern Med* 2017;177:1508–12.

31. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, ESC Scientific Document Group, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J* 2021;42:3599–726.
32. Lum G. Evaluation of a protocol to control utilization of B-type natriuretic peptide testing. *Am J Clin Pathol* 2006;126:190–4.
33. Panteghini M, Bais R. Serum enzymes. In: Rifai N, Horvath AR, Wittwer CT, editors. *Tietz fundamentals of clinical chemistry and molecular diagnostics*, 8th ed. St. Louis: Elsevier Saunders; 2012: 313–30 pp.
34. Lippi G, Panteghini M, Bernardini S, Bonfanti L, Carraro P, Casagrande I, et al. Laboratory testing in the emergency department: an Italian society of clinical biochemistry and clinical molecular biology (SIBioC) and academy of emergency medicine and care (AcEMC) consensus report. *Clin Chem Lab Med* 2018;56:1655–9.
35. Panteghini M, Ceriotti F, Pagani F, Secchiero S, Zaninotto M, Franzini C for the Italian Society of Clinical Biochemistry and Clinical Molecular Biology (SIBioC) Working Group on Enzymes. Recommendations for the routine use of pancreatic amylase measurement instead of total amylase for the diagnosis and monitoring of pancreatic pathology. *Clin Chem Lab Med* 2002;40:97–100.
36. Valente C, Dominici R, Panteghini M. Amilasi pancreatica vs. amilasi totale: relazione fra impiego del test e outcome. *Biochim Clin* 2006;30:378.
37. Vissers RJ, Abu-Laban RB, McHugh DF. Amylase and lipase in the emergency department evaluation of acute pancreatitis. *J Emerg Med* 1999;17:1027–37.
38. Dominici R, Valentinfusino CI, Panteghini M. Replacing pancreatic amylase (P-AMY) with lipase in emergency department (ED) evaluation of acute pancreatitis: a six-month experience. *Clin Chem Lab Med* 2007;45:S287.
39. Barbieri JS, Riggio JM, Jaffe R. Amylase testing for abdominal pain and suspected acute pancreatitis. *J Hosp Med* 2016;11:366–8.
40. Ismail OZ, Bhayana V. Lipase or amylase for the diagnosis of acute pancreatitis? *Clin Biochem* 2017;50:1275–80.
41. Crawford F, Andras A, Welch K, Sheares K, Keeling D, Chappell FM. D-dimer test for excluding the diagnosis of pulmonary embolism. *Cochrane Database Syst Rev* 2016;8:CD010864.
42. Caruso S, Szoke D, Birindelli S, Falvella FS, Dolci A, Panteghini M. Improving D-dimer testing appropriateness by controlling periodicity of retesting: prevention is better than cure. *Clin Chem Lab Med* 2022;60:e175–6.
43. Caruso S, Szoke D, Panteghini M. ‘Penelope test:’ a practical instrument for checking appropriateness of laboratory tests. *Clin Chem Lab Med* 2022;60:1342–9.
44. Aloisio E, Dolci A, Panteghini M. Procalcitonin: between evidence and critical issues. *Clin Chim Acta* 2019;496:7–12.
45. Fung AWS, Beriault D, Diamandis EP, Burnham CD, Dorman T, Downing M, et al. The role of procalcitonin in diagnosis of sepsis and antibiotic stewardship: opportunities and challenges. *Clin Chem* 2017;63:1436–41.
46. Aloisio E, Dolci A, Panteghini M. Managing post-analytical phase of procalcitonin testing in intensive care unit improves the request appropriateness. *Clin Chim Acta* 2019;493:M442.
47. Szoke D, Dolci A, Russo U, Panteghini M. Determination of plasma homocysteine: recommendations for test requesting. *Biochim Clin* 2014;38:234–7.
48. VanWagner LB, Green RM. Evaluating elevated bilirubin levels in asymptomatic adults. *JAMA* 2015;313:516–7.
49. Salinas M, López-Garrigós M, Lugo J, Gutiérrez M, Flors L, Leiva-Salinas C. Diagnostic accuracy of icteric index to detect abnormal total bilirubin values. *J Clin Pathol* 2012;65:928–33.
50. Szoke D, Braga F, Valente C, Panteghini M. Measurement of icteric index as approach to detect abnormal total bilirubin values. *J Clin Pathol* 2013;66:1095–7.
51. Pasqualetti S, Szoke D, Valente C, Panteghini M. Suitability of icteric index as front-line test for the identification of blood samples with abnormal total bilirubin concentrations. *Clin Chem Lab Med* 2015;53:S1139.
52. Aloisio E, Carnevale A, Pasqualetti S, Birindelli S, Dolci A, Panteghini M. Implementation of an internal quality control programme for the photometric determination of icteric index. *J Clin Pathol* 2018;71:851–2.
53. Sturgeon CM, Duffy MJ, Stenman UH, Lilja H, Brünner N, Chan DW, et al. National Academy of Clinical Biochemistry. National academy of clinical biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers. *Clin Chem* 2008;54:e11–79.
54. Dolci A, Scapellato L, Panteghini M. Appropriately delle richieste di PSA libero in un servizio di Medicina di Laboratorio ospedaliero. *Biochim Clin* 2006;30:362.
55. Ferraro S, Caruso S, Panteghini M. Reflex testing of free prostate-specific antigen as effective health care policy. *Arch Pathol Lab Med* 2019;143:1045.
56. Birindelli S, Pasqualetti S, Panteghini M. Offering aspartate aminotransferase as a reflex test: an easy but effective way to improve appropriateness of laboratory requests. *Am J Clin Pathol* 2018;149:456–7.
57. Ahmed F, Mohammed A. Magnesium: the forgotten electrolyte—a review on hypomagnesemia. *Med Sci (Basel)* 2019;7:56.
58. Salinas M, Flores E, López-Garrigós M, Puche C, Leiva-Salinas C. Automatic laboratory interventions to unmask and treat hypomagnesemia in the emergency department. *Clin Biochem* 2020;75:48–52.
59. Pasqualetti S, Chibireva M, Birindelli S, Panteghini M. Automatic reflex tool to detect and treat hypomagnesemia. *Clin Biochem* 2021;92:89.
60. Ferraro S, Mozzi R, Panteghini M. Tracing a roadmap for vitamin B<sub>12</sub> testing using the health technology assessment approach. *Clin Chem Lab Med* 2014;52:767–77.
61. McHugh J, Afghan R, O’Brien E, Kennedy P, Leahy M, O’Keeffe D. Impact of the introduction of guidelines on vitamin B<sub>12</sub> testing. *Clin Chem* 2012;58:471–2.
62. Pelloso M, Basso D, Padoan A, Fogar P, Plebani M. Computer-based-limited and personalised education management maximise appropriateness of vitamin D, vitamin B12 and folate retesting. *J Clin Pathol* 2016;69:777–83.
63. Hutton HD, Drummond HS, Fryer AA. The rise and fall of C-reactive protein: managing demand within clinical biochemistry. *Ann Clin Biochem* 2009;46:155–8.
64. Waldron JL, Ford C, Dobie D, Danks G, Humphrey R, Rolli A, et al. An automated minimum retest interval rejection rule reduces repeat CRP workload and expenditure, and influences clinician requesting behaviour. *J Clin Pathol* 2014;67:731–3.
65. Hawkins R. Unnecessary repeat gamma glutamyl transferase requesting. *Clin Chem* 2017;63:S284.

66. Epner PL, Gans JE, Graber ML. When diagnostic testing leads to harm: a new outcomes-based approach for laboratory medicine. *BMJ Qual Saf* 2013;22(2 Suppl):ii6–10.
67. Roshanov PS, You JJ, Dhaliwal J, Koff D, Mackay JA, Weise-Kelly L, et al. CCDS Systematic Review Team. Can computerized clinical decision support systems improve practitioners' diagnostic test ordering behavior? a decision-maker-researcher partnership systematic review. *Implement Sci* 2011;6:88.
68. Cadamuro J, Ibarz M, Cornes M, Nybo M, Haschke-Becher E, von Meyer A, et al. Managing inappropriate utilization of laboratory resources. *Diagnosis (Berl)* 2019;6:5–13.
69. Bindraban RS, Ten Berg MJ, Naaktgeboren CA, Kramer MHH, Van Solinge WW, Nanayakkara PWB. Reducing test utilization in hospital settings: a narrative review. *Ann Lab Med* 2018;38:402–12.
70. Watine J. Overusing laboratory tests: why is it unethical? *J Lab Precis Med* 2017;2:47.