

BMJ Open Quality Improving laboratory test utilisation at the multihospital Yale New Haven Health System

Roa Harb,¹ David Hajdasz,² Marie L Landry,³ L Scott Sussman⁴

To cite: Harb R, Hajdasz D, Landry ML, *et al*. Improving laboratory test utilisation at the multihospital Yale New Haven Health System. *BMJ Open Quality* 2019;**8**:e000689. doi:10.1136/bmjopen-2019-000689

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-000689>).

Received 11 March 2019

Revised 26 July 2019

Accepted 20 August 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Laboratory Medicine, Yale School of Medicine, New Haven, Connecticut, USA

²Clinical Redesign, Office of Strategy Management, Yale New Haven Health System, New Haven, Connecticut, USA

³Departments of Laboratory Medicine and Medicine, Yale School of Medicine, New Haven, Connecticut, USA

⁴Clinical Redesign, Department of Medicine, Yale New Haven Health System, New Haven, Connecticut, USA

Correspondence to

Dr L Scott Sussman;
Scott.Sussman@ynhh.org

ABSTRACT

Background Waste persists in healthcare and negatively impacts patients. Clinicians have direct control over test ordering and ongoing international efforts to improve test utilisation have identified multifaceted approaches as critical to the success of interventions. Prior to 2015, Yale New Haven Health lacked a coherent strategy for laboratory test utilisation management.

Methods In 2015, a system-wide laboratory formulary committee was formed at Yale New Haven Health to manage multiple interventions designed to improve test utilisation. We report here on specific interventions conducted between 2015 and 2017 including reduction of (1) obsolete or misused testing, (2) duplicate orders, and (3) daily routine lab testing. These interventions were driven by a combination of modifications to computerised physician order entry, test utilisation dashboards and physician education. Measurements included test order volume, blood savings and cost savings.

Results Testing for a number of obsolete/misused analytes was eliminated or significantly decreased depending on alert rule at order entry. Hard stops significantly decreased duplicate testing and educational sessions significantly decreased daily orders of routine labs and increased blood savings but the impact waned over time for select groups. In total, we realised approximately \$100 000 of cost savings during the study period.

Conclusion Through a multifaceted approach to utilisation management, we show significant reductions in low-value clinical testing that have led to modest but significant savings in both costs and patients' blood.

INTRODUCTION

Practitioners contribute to rising American healthcare costs through overutilisation of laboratory testing^{1 2} which grew by 91% between 2000 and 2016, more than any other clinical service.³ Importantly, overutilisation may lead to patient harm by instigating investigative cascades that include additional laboratory testing, unnecessary imaging and invasive procedures.⁴ As a result, several international organisations have issued test utilisation recommendations in the Choosing Wisely campaign⁴ and multiple academic institutions have introduced utilisation management programmes.⁵

Utilisation management typically comprises modifications to electronic ordering systems, targeted education and feedback on ordering behaviour.^{6–12} Often these tools¹³ are governed by a laboratory formulary committee (LFC) that curates the testing formulary.^{13–18} Despite reported success in reducing unnecessary testing in specific areas, there is no universally effective approach for utilisation management.⁵ In that regard, the test utilisation literature is not different from other quality improvement (QI) interventions which show heterogeneous outcomes,^{19–23} most significantly attributed to the effects of context.^{24–27}

This paper describes a set of interventions customised to the local context and informed by input from end-users and other institutions that were developed and refined to improve test utilisation across Yale New Haven Health (YNHH) between 2015 and 2017. This study meets the Yale University Institutional Review Board (100 CH.9 Clinical Quality Improvement) requirements and is reported according to the Standards for Quality Improvement Reporting Excellence guidelines for healthcare improvement.²⁸

METHODS

Framework

Until 2015, YNHH lacked a coherent strategy for laboratory test utilisation (online supplementary figure 1). We identified optimal test utilisation for high value care as our overarching goal and devised a multifaceted approach informed by programme theory and based on relevant literature (online supplementary figure 2). To ensure buy-in, promote culture change and sustain improved utilisation we engaged a diverse, transdisciplinary group, created an LFC, modified electronic order entry, designed utilisation dashboards and provided end-user targeted education. Feedback was continuously sought throughout the process and, with the data collected on each intervention's

performance, informed our strategy as part of the Plan-Do-Study-Act cycle described by the Institute for Healthcare Improvement.

Patient and public involvement

Patients were involved throughout this work as LFC members and were the impetus for each intervention to reduce low value care. Specifically, they provided input on the burdens of repeat inpatient phlebotomy.

Context

This QI programme occurred at YNH, a health system composed of four hospitals and a physician group (see online supplementary table 1). The system uses Epic (Epic Systems, Verona, WI) for its electronic health record and laboratory information system. Over 85% of laboratory tests are ordered electronically and the majority of these orders are placed by residents and fellows in the graduate medical education programmes at three of the four delivery networks.

Overview of intervention

In November 2015, we launched a system-wide LFC cosponsored by the Department of Laboratory Medicine Chair and the YNH Chief Medical Officer and cochaired by the Laboratory Medicine Vice-Chair and the Senior Medical Director of Clinical Operations. Unlike many LFCs, which have fewer than 10 members,¹⁷ our committee has 40 members including clinicians, laboratorians, finance professionals, information technology professionals and patient advisors. The LFC is charged to improve test utilisation while mitigating the negative effects of alert fatigue.^{29–32} An overall timeline of interventions is presented in online supplementary table 2 and select LFC proposals (online supplementary figure 3) are detailed in the following two sections.

Reduction of low-value testing utilisation

We sought to reduce low-value clinical testing at YNH beginning with creatine kinase-MB (CK-MB) and the free thyroxine panel (FTP), which included total thyroxine and an index-based method for determining free thyroxine (FT4). These tests were targeted first because they were commonly ordered at our health system despite evidence of low clinical utility.^{33–37} For FTP, we implemented an electronic hard stop that guided practitioners to FT4 as the appropriate alternative test. Following consultation with local subject matter experts, we determined that CK-MB testing was needed for a subset of ongoing coronary device trials, so we first implemented an alternate test alert (ATA) (online supplementary figure 4) that recommended troponins for clinical testing but allowed practitioners to order CK-MB and, 2 months later, converted this order to a research-only test requiring a study protocol number attestation.

In contrast to FTP and CK-MB, free triiodothyronine (FT3) continues to play a role in clinical testing, however it is less reliable than total T3.³⁸ Therefore, we created ATAs for FT3 and two other low-value thyroid function

tests, T3 uptake (T3U) and reverse T3 (RT3), that offered practitioners the option of test cancellation but also allowed original order placement.

Reduction of clinically unnecessary and wasteful repeat testing

Duplicate testing within a specified time period is both common and associated with adverse clinical effects.^{39–40} We implemented an inpatient duplicate check system that included lookback periods ranging from 30 min to once in a lifetime (see online supplementary tables 3 and 4) as determined by national guidelines and local subject matter experts. Hard stop alerts displayed the most recent test result and included a customer service phone number for potential situations where duplicate testing was clinically indicated. By calling customer service, practitioners obtained a code that allowed authorised duplicate test ordering through manual entry of the test name, tube type and the reason for the duplicate order (see online supplementary figure 5).

Similar to duplicate testing, routine labs including complete blood count (CBC) and basic or comprehensive metabolic panels (BMP/CMP) are commonly overused, especially in the inpatient setting.^{2–41–43} Checking routine labs for 10 days has been reported to phlebotomise approximately 1 unit of blood, potentially necessitating red blood cell transfusions.⁴⁴ We focused on reducing routine lab testing on two inpatient services at Yale New Haven Hospital. The average daily census of the hospitalist service is 360 patients and there are 26 rounding teams composed of pairs of attending physicians and advanced practice providers. A separate teaching service consists of attending physicians supervising two sets of intern/resident pairs. We delivered month-long educational sessions on mindful ordering to both hospitalist and resident teams. These sessions were designed to be one-time events and were reinforced by the creation of a routine labs dashboard (see online supplementary figure 6) to collect data on CBC, BMP and CMP ordering patterns (table 1). Weekly emails celebrating practitioners who ordered most wisely, team dashboard review at monthly in-person meetings and optional dashboard subscriptions supplemented the educational component and aimed to promote the laboratory ordering cultural change.

Measures and statistical analysis

We used Helix, Yale's customised data warehouse system, to create dashboards that extracted laboratory order data from Epic (see online supplementary figure 7) and allowed filtering by test name, delivery network, department, provider and specialty. For each intervention, we excluded the implementation month to separate the preintervention and the postintervention periods. For CK-MB and thyroid function tests we obtained total monthly test orders 6 months before intervention and 6 months after intervention. For duplicate checks, we collected total weekly duplicate orders normalised to inpatient census at Yale New Haven Hospital between 5

Table 1 Definition of routine labs

Routine* labs rate	Days when either CBC+BMP or CBC+CMP is ordered/total patient-days
Targeted labs rate	Days when 0 or 1 routine lab is ordered/total patient-days
Lab-free days rate	Days when no lab is ordered/total patient-days
Blood saved [†]	Lab-free days rate × 6.5 mL × 1000 patient-days/total patient-days

*Routine labs include complete blood count (CBC), basic metabolic panel (BMP) and complete metabolic panel (CMP).

†At Yale New Haven Health (YNHH), laboratory draws for CBC and BMP/CMP average 3.5 and 3 mL, respectively.

January 2014 and 30 December 2017. For routine labs, we collected percentages of patient-days when both CBC and BMP/CMP were ordered for each week between 28 December 2014 and 1 July 2017. We also monitored two types of savings: blood savings calculated from lab-free days and normalised to 1000 patient-days and cost savings calculated from material and labour savings on specific interventions.

CK-MB and thyroid function test data were plotted as individual control charts (JMP, SAS Institute) as the optimal way to depict changes in order volume after intervention. Individual control charts in JMP software use the average (\bar{x}) and the process SD (σ) estimated by the moving range to determine lower ($\bar{x}-3\sigma$) and upper ($\bar{x}+3\sigma$) control limits. Mann-Whitney test was performed to detect significant median differences (CK-MB and thyroid function tests) using GraphPad Prism V.7 (GraphPad Software). In contrast, interventions for duplicate labs occurred at multiple points in time and routine labs, which involved educational sessions and continuous feedback, were not discrete. Therefore, we performed analysis of covariance to determine the effects of both time and intervention on orders before and after intervention for duplicate labs, routine labs and blood savings using JMP software. A probability value ($p \leq 0.05$) was considered statistically significant.

RESULTS

Effects of test eliminations and alerts on low-value testing

Comparison between the 6-month period immediately before and after the elimination of FTP and the

conversion of CK-MB into a research-only test revealed a 97% and 98% decrease in orders, respectively (table 2). Comparison between the 6-month period immediately before and after the ATA implementation for FT3, T3U and RT3 revealed a 41%, 52% and 14% decrease in test orders, respectively (table 2).

Orders for FTP decreased abruptly at the time of intervention and ceased shortly thereafter since the test was eliminated (figure 1A), while CK-MB orders asymptotically approached zero as testing for research studies was maintained (figure 1B). Interestingly, CK-MB orders dropped by approximately 42% (figure 1B, arrow) after the ATA was implemented but before the test was converted to research only. Factors that contributed to persistent CK-MB orders for clinical purposes after the conversion include its presence in a CK total/CK-MB panel (subsequently removed), time needed to communicate with the geographically diverse outreach community and paper requisitions which circumvented the electronic decision support.

FT3, T3U and RT3 decreased less dramatically (figure 1C–E) since practitioners could continue with their original order. Nonetheless, there were significant reductions in FT3 and T3U orders demonstrating the ATA impact. The reduction in RT3 orders was not significant, likely due to the low volume at baseline. Notably, there was no significant difference in FT4 orders during the same period (online supplementary figure 8 and table 2), suggesting that fluctuating patient census did not explain the FT3 and T3U results.

Table 2 Changes in test order volume after intervention

Test	Time period	Median test order volume difference* (95% CI)	% Change	P value†
CK-MB	June 2016±6 months	1764 (916 to 1866)	-98	0.002
FTP	January 2016±6 months	882 (718 to 910)	-97	0.002
FT3	April 2016±6 months	209 (81 to 281)	-41	0.002
T3U	April 2016±6 months	37 (30 to 50)	-52	0.002
RT3	April 2016±6 months	3 (3 to 11)	-14	0.4
FT4	April 2016±6 months	54 (424 to 962)	+11	0.6

*Hodges-Lehmann absolute differences between preintervention and postintervention medians were calculated and reported with their 95% CIs.

†Mann-Whitney non-parametric analysis was performed to detect significant differences.

CK-MB, creatine kinase-MB; FT3, free triiodothyronine; FT4, free thyroxine; FTP, free thyroxine panel; RT3, reverse triiodothyronine; T3U, triiodothyronine uptake.

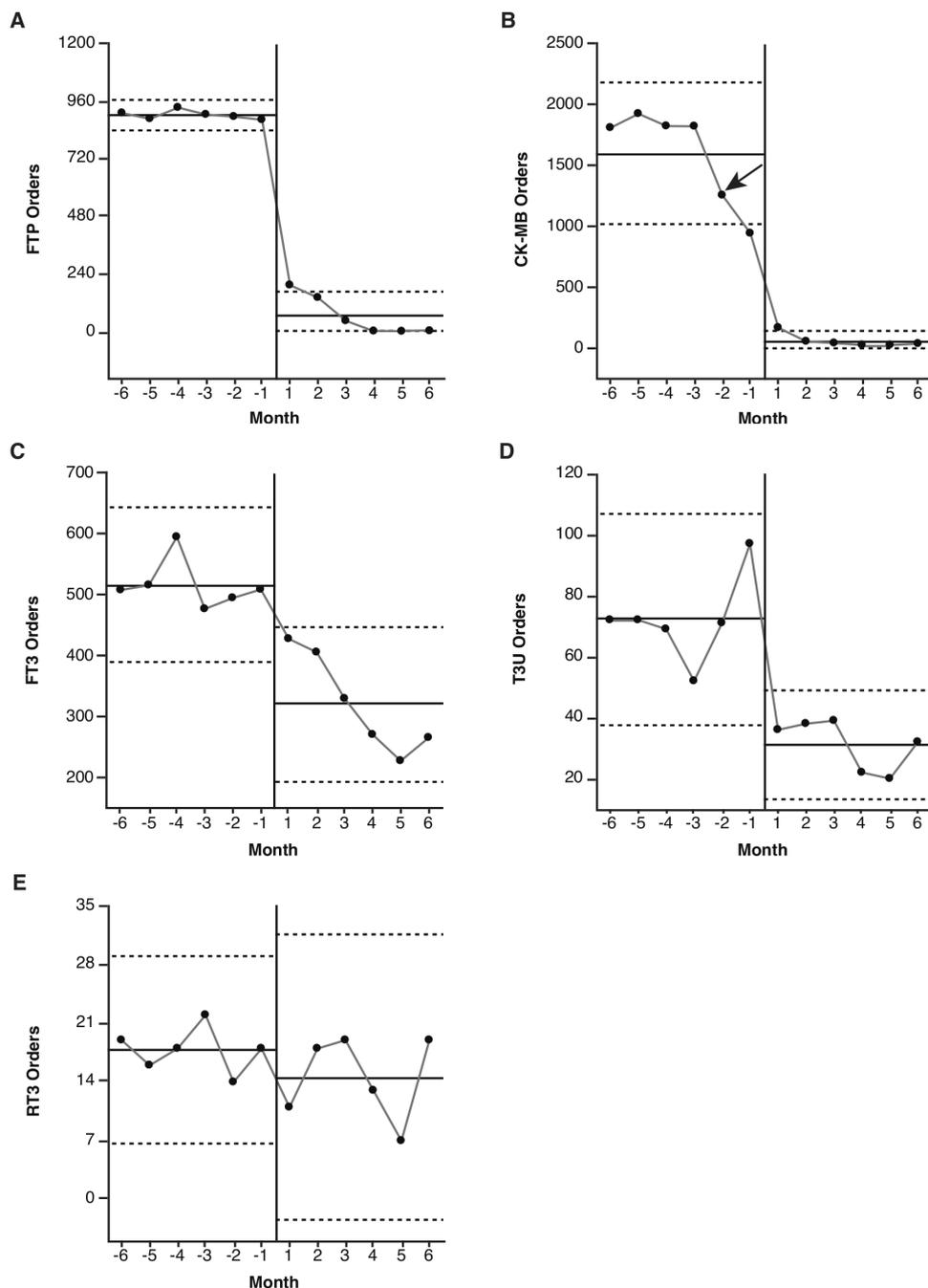


Figure 1 Individual control charts depicting test orders before and after intervention at Yale New Haven Health System. (A) FTP orders before (1 July 2015 to 31 December 2015) and after (1 February 2016 to 31 July 2016) panel were eliminated from the lab menu on 22 January 2016. (B) Total CK-MB orders before (1 December 2015 to 30 May 2016) and after (1 July 2016 to 31 December 2016) CK-MB were converted into a research-only test on 22 June 2016. Arrow corresponds to April 2016 when alternate test alert for CK-MB was implemented. (C–E) FT3, T3U and RT3 orders before (1 October 2015 to 31 March 2016) and after (1 May 2016 to 30 October 2016) alternate test alert were created on 1 April 2016. (A–E) Data are expressed as order volume for each month. Solid horizontal lines represent data average and dotted horizontal lines represent the lower and upper control limits. Calendar months of interventions are removed from analysis. CK-MB, creatine kinase-MB; FTP, free thyroxine panel; FT3, free triiodothyronine; RT3, reverse triiodothyronine; T3U, triiodothyronine uptake.

Effects of hard stops and physician education on frequency of testing

Electronic duplicate checks were initiated in October 2015. Analysis of weekly orders (figure 2A) prior to and after this date revealed that time (F ratio=85.3, $p<0.0001$) and intervention (F ratio=10.2, $p=0.002$) significantly

decreased the level of duplicate orders but there was no interaction (F ratio=2.2, $p=0.1$). This suggests that although the intervention was effective, it did not accelerate a pre-existing trend for reduced duplicate orders. Several unexpected workarounds allowed staff to circumvent active controls. For example, some practitioners

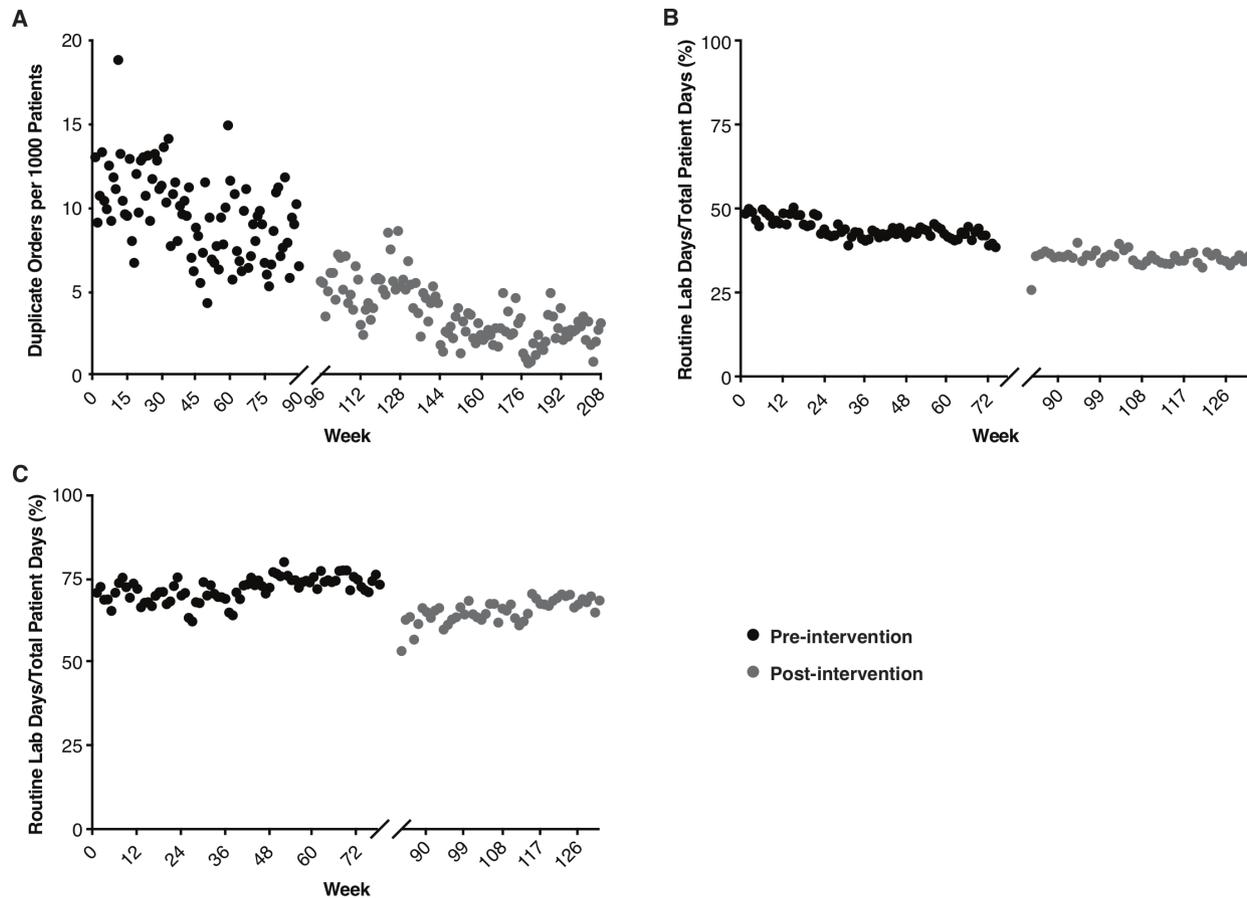


Figure 2 Duplicate and routine orders before and after intervention on the inpatient service at Yale New Haven Hospital. (A) Weekly duplicate orders over the period 5 January 2014 to 24 December 2017. The period of intervention spanning 27 September 2015 to 31 October 2015 is removed from the analysis. Data are expressed as orders per 1000 patients. (B) Hospitalists' weekly CBC and BMP/CMP orders over the period 28 December 2014 to 25 June 2017. The period of educational intervention spanning 29 May 2016 to 2 July 2016 is removed from the analysis. Data are expressed as per cent-day routine labs were ordered out of total patient-days for each week. (C) Residents' weekly CBC and BMP/CMP orders over the period 28 December 2014 to 25 June 2017. The period of educational intervention spanning 26 June 2016 to 30 July 2016 is removed from the analysis. Data are expressed as per cent-day routine labs were ordered out of total patient-days for each week.

recognised that the clinical laboratory would process orders submitted on paper requisitions. Others resorted to duplicate orders as the quickest way to view previously ordered test results. Another process flaw was that add-on testing did not trigger a duplicate check. Since these behaviours did not occur at discrete points in time and were reflected in the data as the dashboard tracked orders as opposed to resulted tests, it was not possible to isolate improvements in the performance of the duplicate checks when these behaviours were corrected.

Educational sessions for hospitalists and residents took place in June and July 2016 (figure 2B,C). Analysis of routine labs prior to and after these sessions revealed that time (F ratio=39.7, $p<0.0001$) and intervention (F ratio=16.8, $p<0.0001$) significantly decreased the level of hospitalist daily lab orders but there was no interaction (F ratio=2.6, $p=0.1$). On the other hand, time (F ratio=55.8, $p<0.0001$) and intervention (F ratio=127.6, $p<0.0001$) significantly decreased the level of resident daily labs and there was a significant interaction (F ratio=6.1, $p=0.01$). This suggests that the educational sessions were effective but did not accelerate a pre-existing trend

for reduced daily lab orders by hospitalists. Educational sessions were also effective for residents but only in the short term as the direction of the interaction demonstrated that residents tended to return towards baseline levels over time.

An unintended consequence occurred as practitioners ordered routine labs more mindfully. Requests by nurses to place routine orders overnight increased as the night team was unaware that lab holidays were intentional. Therefore, we created a 'No Labs Needed' order that displayed for a specific date to improve communication between the different shifts.

Impact of interventions on blood and cost savings

Analysis of blood saved from lab-free days revealed that time (F ratio=22.4, $p<0.0001$) but not intervention (F ratio=1.8, $p=0.19$) significantly increased the level of hospitalist blood savings but there was also a significant interaction (F ratio=5.3, $p=0.02$) manifesting as decreased rate of blood savings after intervention (figure 3A). On the other hand, both time (F ratio=11.3, $p=0.001$) and intervention (F ratio=30.3, $p<0.0001$) significantly increased the level

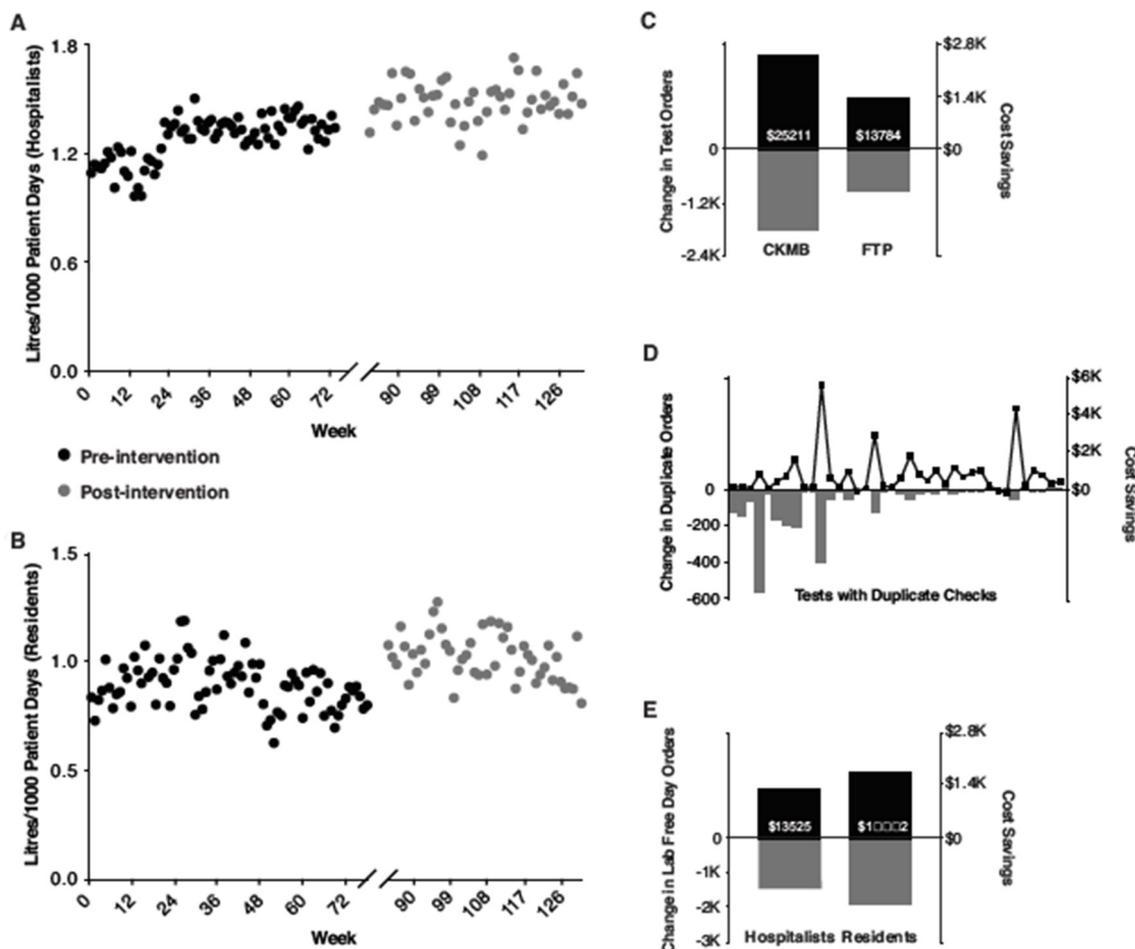


Figure 3 Blood and cost savings due to interventions at Yale New Haven Hospital. (A) Hospitalists' blood savings over the period 28 December 2014 to 25 June 2017. The period of educational intervention spanning 29 May 2016 to 2 July 2016 is removed from the analysis. Data are expressed as litres saved per 1000 patient-days. (B) Residents' blood savings over the period 28 December 2014 to 25 June 2017. The period of educational intervention spanning 26 June 2016 to 30 July 2016 is removed from the analysis. Data are expressed as litres saved per 1000 patient-days. (C) Cost savings for creatine kinase-MB (CK-MB) and free thyroxine panel (FTP). Grey columns represent the change in test orders and black columns represent the resultant cost savings between the immediate postintervention and the preintervention 12-month periods. (D) Cost savings for duplicate checks. Grey columns represent the change in duplicate orders and black circles represent the resultant cost savings between the late postintervention period (1 January 2017 to 30 December 2017) and the preintervention period (5 January 2014 to 3 January 2015). (E) Cost savings for routine labs. Grey columns represent the change in lab-free orders and black columns represent the resultant cost savings between the late postintervention periods (3 July 2016 to 2 July 2017 for hospitalists and 31 July 2016 to 30 July 2017 for residents) and the preintervention periods (31 May 2015 to 29 May 2016 for hospitalists and 28 June 2015 to 26 June 2016 for residents).

of blood savings by residents (figure 3B) but there was no interaction (F ratio=1.4, $p=0.24$).

In addition to blood savings, cost savings accrued from improved test utilisation. For CK-MB and FTP, respective cost savings of \$25 211 and \$13 784 occurred in the 1-year period after intervention (figure 3C). At Yale New Haven Hospital, duplicate checks led to cost savings of \$29 519 when comparing the late postintervention period to the preintervention period (figure 3D). Importantly, only 4 of the 38 tests surveyed were ordered more than five times a week prior to intervention which could explain the modest cost savings (see online supplementary table 5). Cost savings of \$13 525–\$16 558 for hospitalists and \$17 841–\$21 843 for residents were attained from lab-free days by comparing the immediate postintervention periods to the immediate preintervention periods (figure 3E). Prior to intervention,

hospitalists' higher rate of lab-free patient-days accounted for \$139 125–\$170 329 in cost savings compared with residents. The range of these savings reflects the cost of ordering a BMP versus CMP as part of daily labs.

DISCUSSION

In this work, we present a QI initiative to optimise laboratory test utilisation in a multihospital health system. We show significant reductions in obsolete and misused test orders, unnecessary duplicate orders and daily routine labs. These interventions have led to modest but significant savings in both costs and patients' blood.

Test-specific intervention outcomes

The degree of reductions in test orders depended on the specific utilisation management tools employed.

ATAs for CK-MB, FT3 and T3U allowed practitioners to continue with the original order and caused significant but modest reductions in order volume after intervention. In contrast, changing CK-MB to a research-only test and creating a hard stop for FTP virtually eliminated orders. These findings corroborate previous studies that have shown significant test volume reduction with hard stop interventions^{45–47} and less impact with soft stops.^{43 46 48–50} This could be attributable to alert fatigue⁵¹ since the per cent reductions following soft stops tend to be similar regardless of test type which suggests an underlying systematic process. Our research attestation requirement for CK-MB was not a hard stop yet, interestingly, it performed as one in our system. This could be explained by the fact that CK-MB's utility was already declining in clinical practice; therefore, caution should be exercised in applying similar interventions to more popular tests.⁵² The lack of significant reductions in RT3 orders suggests that careful thought must be given for the decision to allocate time and resources to tests that are low volume at baseline. In conclusion, test-specific interventions require consideration for both perceived clinical utility and test volume before intervention to identify the optimal decision support.

Duplicate check intervention outcomes

Hard stops generated significant reductions in duplicate orders affirming their efficacy as a test utilisation tool.^{11 46 53} The authorised duplicate order, which involved both a phone call to the laboratory and a new electronic order, was enough of a hurdle to cause significant reductions in non-clinically indicated orders. Similar to others,⁵⁴ we found that there was a preintervention trend for reduced duplicate test orders given the significant effect of time. Multiple factors may contribute for this baseline trend including internal motivations that reflect an increased interest in improving test utilisation across academic centres and healthcare institutions, though this may be optimistic.⁵⁵ More likely explanations include the menu for duplicate checks which focused on non-controversial targets and the conservative lookback intervals aimed at ensuring consensus across the health system and eliminating inadvertent patient harm or inconvenience. Therefore, testing that is clinically redundant likely remains uncaptured by our checks and reflects that need to balance the goals and incentives of stakeholders with the desire to decrease duplicate testing.

Daily lab intervention outcomes

Although the educational and feedback intervention was generally effective for both residents and hospitalists, there were important differences. In both groups, there was an immediate reduction in daily lab orders after intervention that either continued a pre-existing trend (hospitalists) or reverted back to baseline (residents). In both groups, blood savings were higher after intervention but they stabilised over time for hospitalists, possibly reflecting maximal attainable savings at baseline. Others

have shown similar reductions in daily lab orders with education and feedback-based interventions^{8 10 56} and more enduring reductions with significant cost savings with a multilevel approach that included education and feedback.¹⁰ Our results are consistent with a recent meta-analysis that did not find sufficient evidence to recommend education alone to improve test utilisation despite a generally favourable effect on the number of tests ordered.⁵³

Residents often attribute their daily lab orders to the perceived expectations of their attending physicians which could account for the accelerated return to baseline following the immediate improvement after intervention. Future educational approaches need to target these perceptions and clarify the attending physicians' expectations for daily lab orders. Educational sessions focusing on each group separately are unlikely to sustain behavioural modifications, particularly for residents as demonstrated by our data. Furthermore, dedicating the time and human resources required for educational interventions should take into consideration the significant effects of time in both groups which argue for a pre-existing trend in judicious routine lab orders.

Financial impact

Using a conservative direct cost approach generated a modest cost savings of just under \$100 000 for our interventions. Cost savings reported in previous studies have varied greatly depending on baseline performance and interventions used.^{43 46 47 54} Our cost savings were influenced by similar factors. These include choice of target tests and lookback periods for duplicate checks, restricting these checks to inpatients and calculating the resultant cost savings for a single hospital (Yale New Haven Hospital) given the lack of standardisation of material and labour costs across the system. Less discernible but arguably more significant reasons include the fact that our interventions occurred in a large health system with multiple stakeholders. The focus on consensus and durability necessitated compromises (inpatient vs outpatient or ATAs vs hard stops) that moderated our interventions' impact. Furthermore, laboratory direct costs are a fraction of the costs of diagnostic imaging or procedures⁵⁷ and true savings from test stewardship are likely to be substantial due to the cascading effects of test results on medical decision-making.

One intriguing finding is that hospitalists saved over \$130 000 in routine lab orders as compared with residents before intervention which far supersedes the savings for either group following educational sessions. This suggests that interventions that build on inherent practices or systems at an institution may have significantly more impact on behaviour modifications or cost savings than those previously reported.

Limitation

This study took place at a single health system in the Northeast United States. The most significant limitation

is the degree of uncertainty in ascribing the improvements seen to our interventions. Since interventions were introduced at various points in time and the data collected over a long period in a dynamic clinical environment, it is conceivable that some of the outcomes were influenced by other factors. Our statistical analyses aligned the improvement metrics with the timing of the interventions, however there is also a significant effect of time. It is unlikely that a QI study similar to ours could be conducted under more controlled conditions, but it is possible to implement and compare discrete projects at different hospitals within the system to elucidate significant context-specific factors. This could also counteract another downside of the lack of site-specific interventions, namely the potential reduction in the overall magnitude of impact.

CONCLUSIONS

Roth and Lee argue against recreating ways to organically improve utilisation.⁵⁸ This work at YNH builds on successes from other institutions and applies lessons learnt to our local context. Institutions may opt for varied interventions to improve laboratory test stewardship, but we found that an integrated approach across people, systems and technology coupled with ongoing stakeholder engagement key to sustained success. An LFC with clinical and laboratorian representation is the ideal foundation and can enhance coordination across multiple hospitals.

Much work lies ahead for aligning test utilisation recommendations with actual practice, especially in the area of reference testing. We plan to apply our framework to other areas including diagnostic imaging, non-operating room procedures and non-clinical areas such as the supply chain.

Acknowledgements We thank Dr. Frank Davidoff (Geisel School of Medicine, Dartmouth, New Hampshire) for critical review of this manuscript. We thank Dr. Stephanie Miller (Quinnipiac University, Connecticut) for help with the initial draft of this manuscript.

Contributors RH compiled and analysed the data. DH and LSS analysed the data. RH, DH, MLL and LSS wrote and edited the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study meets the Yale University Institutional Review Board (100 CH.9 Clinical Quality Improvement) requirements.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Berwick DM, Hackbarth AD. Eliminating waste in US health care. *JAMA* 2012;307:1513–6.
2. Zhi M, Ding EL, Theisen-Toupal J, *et al*. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS One* 2013;8:e78962.
3. MedPAC. *Medicare payment Advisory Commission data book*, 2016.
4. Cassel CK, Guest JA. Choosing wisely: helping physicians and patients make smart decisions about their care. *JAMA* 2012;307:1801–2.
5. Bindraban RS, Ten Berg MJ, Naaktgeboren CA, *et al*. Reducing test utilization in hospital settings: a narrative review. *Ann Lab Med* 2018;38:402–12.
6. Feldman LS, Shihab HM, Thiemann D, *et al*. Impact of providing fee data on laboratory test ordering: a controlled clinical trial. *JAMA Intern Med* 2013;173:903.
7. Eaton KP, Levy K, Soong C, *et al*. Evidence-based guidelines to eliminate repetitive laboratory testing. *JAMA Intern Med* 2017;177:1833.
8. Iams W, Heck J, Kapp M, *et al*. A multidisciplinary Housestaff-Led initiative to safely reduce daily laboratory testing. *Acad Med* 2016;91:813–20.
9. Sedrak MS, Myers JS, Small DS, *et al*. Effect of a price transparency intervention in the electronic health record on clinician ordering of inpatient laboratory tests: the price randomized clinical trial. *JAMA Intern Med* 2017;177:939–45.
10. Vidyarthi AR, Hamill T, Green AL, *et al*. Changing resident test ordering behavior: a multilevel intervention to decrease laboratory utilization at an academic medical center. *Am J Med Qual* 2015;30:81–7.
11. Procop GW, Yerian LM, Wyllie R, *et al*. Duplicate laboratory test reduction using a clinical decision support tool. *Am J Clin Pathol* 2014;141:718–23.
12. Kobewka DM, Ronksley PE, McKay JA, *et al*. Influence of educational, audit and feedback, system based, and incentive and penalty interventions to reduce laboratory test utilization: a systematic review. *Clin Chem Lab Med* 2015;53:157–83.
13. Baird G. The laboratory test utilization management toolbox. *Biochem Med* 2014;24:223–34.
14. Jackson BR. Laboratory formularies. *Clinica Chimica Acta* 2014;427:151–3.
15. Kim JY, Dzik WH, Dighe AS, *et al*. Utilization management in a large urban academic medical center: a 10-year experience. *Am J Clin Pathol* 2011;135:108–18.
16. Wilson ML. Decreasing inappropriate laboratory test utilization: controlling costs and improving quality of care. *Am J Clin Pathol* 2015;143:614–6.
17. Warren JS. Laboratory test utilization program: structure and impact in a large academic medical center. *Am J Clin Pathol* 2013;139:289–97.
18. Snozek C, Kaleta E, Hernandez JS. Management structure: establishing a laboratory utilization program and tools for utilization management. *Clinica Chimica Acta* 2014;427:118–22.
19. Grimshaw JM, Shirran L, Thomas R, *et al*. Changing provider behavior: an overview of systematic reviews of interventions. *Med Care* 2001;39(8 Suppl 2):li2–45.
20. Grimshaw J, Thomas R, MacLennan G, *et al*. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004;8:1–72.
21. Blumenthal D, Kilo CM. A report card on continuous quality improvement. *Milbank Q* 1998;76:511:625–48.
22. Schouten LMT, Hulscher MEJL, van Everdingen JJE, *et al*. Evidence for the impact of quality improvement Collaboratives: systematic review. *BMJ* 2008;336:1491–4.
23. Øvretveit J, Bate P, Cleary P, *et al*. Quality Collaboratives: lessons from research. *Qual Saf Health Care* 2002;11:345–51.
24. Lomas J. Using research to inform healthcare managers' and policy makers' questions: from summative to interpretive synthesis. *Hcpol* 2005;1:55–71.
25. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. *Med J Aust* 2004;180(6 Suppl):S57–60.
26. Kaplan HC, Provost LP, Froehle CM, *et al*. The model for understanding success in quality (MUSIQ): building a theory of context in healthcare quality improvement. *BMJ Qual Saf* 2012;21:13–20.
27. Roshanov PS, You JJ, Dhaliwal J, *et al*. Can computerized clinical decision support systems improve practitioners' diagnostic test ordering behavior? A decision-maker-researcher partnership systematic review. *Implementation Science* 2011;6.

28. Ogrinc G, Davies L, Goodman D, *et al.* SQUIRE 2.0 (standards for quality improvement reporting excellence): revised publication guidelines from a detailed consensus process. *BMJ Qual Saf* 2016;25:986–92.
29. van der Sijs H, Mulder A, van Gelder T, *et al.* Drug safety alert generation and overriding in a large Dutch University medical centre. *Pharmacoepidemiol Drug Saf* 2009;18:941–7.
30. Carspecken CW, Sharek PJ, Longhurst C, *et al.* A clinical case of electronic health record drug alert fatigue: consequences for patient outcome. *Pediatrics* 2013;131:e1970–3.
31. Phansalkar S, van der Sijs H, Tucker AD, *et al.* Drug–drug interactions that should be non-interruptive in order to reduce alert fatigue in electronic health records. *J Am Med Inform Assoc* 2013;20:489–93.
32. Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. *J Am Med Inform Assoc* 2012;19:e145–8.
33. Thygesen K, Alpert JS, White HD, *et al.* Universal definition of myocardial infarction. *Circulation* 2007;116:2634–53.
34. Kavsak PA, MacRae AR, Newman AM, *et al.* Effects of contemporary troponin assay sensitivity on the utility of the early markers myoglobin and CKMB isoforms in evaluating patients with possible acute myocardial infarction. *Clinica Chimica Acta* 2007;380:213–6.
35. Saenger AK, Jaffe AS. Requiem for a heavyweight: the demise of creatine kinase-MB. *Circulation* 2008;118:2200–6.
36. Reichlin T, Hochholzer W, Bassetti S, *et al.* Early diagnosis of myocardial infarction with sensitive cardiac troponin assays. *N Engl J Med* 2009;361:858–67.
37. Rifai N, Horvath AR, Wittwer C. *Tietz textbook of clinical chemistry and molecular diagnostics*. Sixth edition. St. Louis, Missouri: Elsevier, 2018.
38. Ross DS, Burch HB, Cooper DS, *et al.* 2016 American thyroid association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. *Thyroid* 2016;26:1343–421.
39. Tosiri P, Kanitsap N, Kanitsap A. Approximate iatrogenic blood loss in medical intensive care patients and the causes of anemia. *J Med Assoc Thai* 2010;93 Suppl 7:S271–6.
40. Lyon AW, Chin AC, Slotsve GA, *et al.* Simulation of repetitive diagnostic blood loss and onset of iatrogenic anemia in critical care patients with a mathematical model. *Comput Biol Med* 2013;43:84–90.
41. Ezzie ME, Aberegg SK, O'Brien JM. Laboratory testing in the intensive care unit. *Crit Care Clin* 2007;23:435–65.
42. Ng VL. Utilization management in the core laboratory. *Clin Chim Acta* 2014;427:154–7.
43. Krasowski MD, Chudzik D, Dolezal A, *et al.* Promoting improved utilization of laboratory testing through changes in an electronic medical record: experience at an academic medical center. *BMC Med Inform Decis Mak* 2015;15:11.
44. Thavendiranathan P, Bagai A, Ebidia A, *et al.* Do blood tests cause anemia in hospitalized patients? the effect of diagnostic phlebotomy on hemoglobin and hematocrit levels. *J Gen Intern Med* 2005;20:520–4.
45. Seppänen K, Kauppila T, Pitkälä K, *et al.* Altering a computerized laboratory test order form rationalizes ordering of laboratory tests in primary care physicians. *Int J Med Inform* 2016;86:49–53.
46. Procop GW, Keating C, Stagno P, *et al.* Reducing duplicate testing: a comparison of two clinical decision support tools. *Am J Clin Pathol* 2015;143:623–6.
47. Lippi G, Brambilla M, Bonelli P, *et al.* Effectiveness of a computerized alert system based on re-testing intervals for limiting the inappropriateness of laboratory test requests. *Clin Biochem* 2015;48:1174–6.
48. Levick DL, Stern G, Meyerhoefer CD, *et al.* "Reducing unnecessary testing in a CPOE system through implementation of a targeted CDS intervention". *BMC Med Inform Decis Mak* 2013;13:43.
49. Moyer AM, Saenger AK, Willrich M, *et al.* Implementation of clinical decision support rules to reduce repeat measurement of serum ionized calcium, serum magnesium, and N-terminal pro-B-type natriuretic peptide in intensive care unit inpatients. *Clin Chem* 2016;62:824–30.
50. Love SA, McKinney ZJ, Sandoval Y, *et al.* Electronic medical record-based performance improvement project to document and reduce excessive cardiac troponin testing. *Clin Chem* 2015;61:498–504.
51. Ash JS, Sittig DF, Campbell EM, *et al.* Some unintended consequences of clinical decision support systems. *AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium*, 2007:26–30.
52. Jackups R. The Promise-and Pitfalls-of computerized provider alerts for laboratory test ordering. *Clin Chem* 2016;62:791–2.
53. Rubinstein M, Hirsch R, Bandyopadhyay K, *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.
54. Pageler NM, Franzon D, Longhurst CA, *et al.* Embedding time-limited laboratory orders within computerized provider order entry reduces laboratory utilization. *Pediatr Crit Care Med* 2013;14:413–9.
55. Kerr EA, Kullgren JT, Saini SD. Choosing wisely: how to fulfill the promise in the next 5 years. *Health Aff* 2017;36:2012–8.
56. Neilson EG, Johnson KB, Rosenbloom ST, *et al.* The impact of peer management on test-ordering behavior. *Ann Intern Med* 2004;141:196–204.
57. Baird GS. The choosing wisely initiative and laboratory test stewardship. *Diagnosis* 2019;6:15–23.
58. Roth AM, Lee TH. Health care needs less innovation and more imitation. *Harvard Business Review* 2014.