Original Article

A bayesian approach to laboratory utilization management

Ronald G. Hauser, Brian R. Jackson[1,3], Brian H. Shirts[2]

Department of Laboratory Medicine, Yale University School of Medicine, New Haven, Connecticut, [1]Department of Pathology, University of Utah, [3]ARUP Laboratories, Salt Lake City, Utah, [2]Department of Laboratory Medicine, University of Washington, Seattle, Washington, USA

E-mail: *Dr. Ronald G. Hauser - ronald.hauser@yale.edu
*Corresponding author

Abstract

Background: Laboratory utilization management describes a process designed to increase healthcare value by altering requests for laboratory services. A typical approach to monitor and prioritize interventions involves audits of laboratory orders against specific criteria, defined as rule-based laboratory utilization management. This approach has inherent limitations. First, rules are inflexible. They adapt poorly to the ambiguity of medical decision-making. Second, rules judge the context of a decision instead of the patient outcome allowing an order to simultaneously save a life and break a rule. Third, rules can threaten physician autonomy when used in a performance evaluation. **Methods:** We developed an alternative to rule-based laboratory utilization. The core idea comes from a formula used in epidemiology to estimate disease prevalence. The equation relates four terms: the prevalence of disease, the proportion of positive tests, test sensitivity and test specificity. When applied to a laboratory utilization audit, the formula estimates the prevalence of disease (pretest probability [PTP]) in the patients tested. The comparison of PTPs among different providers, provider groups, or patient cohorts produces an objective evaluation of laboratory requests. We demonstrate the model in a review of tests for enterovirus (EV) meningitis. **Results:** The model identified subpopulations within the cohort with a low prevalence of disease. These low prevalence groups shared demographic and seasonal factors known to protect against EV meningitis. This suggests too many orders occurred from patients at low risk for EV. **Conclusion:** We introduce a new method for laboratory utilization management programs to audit laboratory services.

Key words: Delivery of health care, efficiency, guideline adherence, health care, organization, physicians' practice patterns, process assessment (health care), quality assurance, utilization review

INTRODUCTION

The question can be raised, who benefits from which test, when, where, and at what cost?[1] From journals devoted to such diverse areas of medical practice including emergency medicine to medical education, coagulation, quality improvement, and HIV, a continued interest in the appropriateness of laboratory testing has existed from the 1940s to the present day.[2-7] Pathologists and others who evaluate laboratory test utilization management, generally echo a common theme of improving patient care and decreasing medical costs.[8-12]

The dominant method used to evaluate test utilization management involves a retrospective comparison of clinical practice guidelines to actual decisions.[7] We refer to this method as a rule-based approach because of its reliance on rules such as "patients taking warfarin should have at least one prothrombin time/international normalized ratio test within 60 days of beginning the drug."[13] Because these rules have minimal ambiguity, they translate easily into database queries, which eliminate the need for chart review, an opinionated, time-consuming and costly endeavor.[14] In addition healthcare administrators, generally find such rules easy to implement because of their black-and-white interpretation, origin in empirical trials, and buy-in from physicians, whose professional societies and expert panels participate in their development.

However, rule-based utilization management has limits, primarily due to its inability to monitor utilization in the ambiguous gray-zones of medical decision-making. Rule-based utilization management does not always perfectly reflect recognized guidelines and physicians, who at times see the simplicity of rules as overgeneralized and unrealistic, may justifiably be reluctant to follow strict rules.[15] A limit on the number of well-researched, clear-cut medical decisions may explain the difficulty of rule-based utilization management programs like the Centers for Medicare and Medicaid Services' Clinical Quality Measures, or HEDIS to expand past a few hundred rules, even as the lower estimate on the number of medical scenarios exceeds this number by multiple orders of magnitude.[16]

Rules may become more numerous and complex, but the rules-based approach shares many of the same limitations as another long-standing issue in the evaluation of medical decisions, the determination of pretest probability (PTP).[17-22] Both attempt to probe the black-box that is the patient-physician relationship, whose nuances may not easily translate to structured fields in a database. A sufficiently advanced rule system may obviate the need for an autonomous physician, many of whom may have already left medical practice because autonomy is key to their job satisfaction.[23] Perhaps more importantly, utilization management rules, like pre-test probability, focus on the pretext of a decision rather than its outcome. An alternative approach would evaluate the outcome of patterns of behavior over time, not a single decision formally separated from its complete context.

We sought to create a method for test utilization management that shares the strengths of the rule-based approach, namely its empiricism, simplicity and avoidance of manual chart review, while ameliorating at least some of its limitations, such as its difficulty with ambiguous decisions, infringement on provider autonomy, and emphasis on the ordering decision rather than its outcome.

METHODS

Method Explanation

Our approach to test utilization management closely follows medical decision-making theory.[24] In medical decision making the optimal strategy when faced with a diagnostic dilemma is to choose the option with the highest utility: nonintervention, test, or treatment. The PTP of disease in the patient under evaluation informs the choice. Figure 1 shows a prototypical example of the utility curve for each choice across a range of PTP. At low PTP, the patient has a low probability of disease, and nonintervention maximizes utility. The patient likely has the disease at high PTP and will benefit most from treatment. The patient's disease state has the most uncertainty and the test has the greatest expected utility between the extreme values of PTP. Along the continuum of PTP the three decisions (nonintervention, test, and treat) form two points of equivalent utility (nonintervention-test and test-treat). We define these two cutoff points as PTP_{Low} and PTP_{High} [Figure 1].

The physician's estimate of the patient probability of disease (PTP_{Est}) compared to PTP_{Low} and PTP_{High} determines their choice. The decision to not intervene ($PTP_{Est} < PTP_{Low}$), test ($PTP_{Low} \leq PTP_{Est} \leq PTP_{High}$), and treat ($PTP_{High} < PTP_{Est}$) becomes as simple as evaluating an inequality. Our model takes an identical approach by defining the appropriate use of a test as an estimate of PTP between PTP_{Low} and PTP_{High}. If we obtain the constant values of PTP_{Low} and PTP_{High} from cost-efficacy analysis, an estimate of the physician's PTP PTP_{Est} would allows us to determine their decision-making strategy.

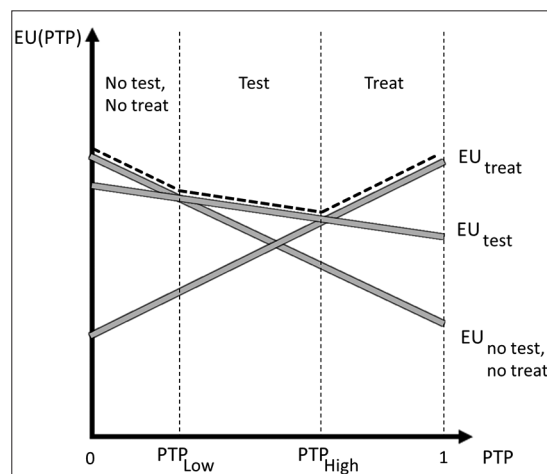To determine the value of PTP_{Est} in the patient

Figure 1: Deciding to test, a graph of pretest probability (PTP) versus expected utility. Utility curves for treatment, test, and nonintervention are shown as gray solid lines. The dotted line represents the maximum utility of the available choices. The cutoff PTPs were labeled PTP_{Low} and PTP_{High}. PTP: Pretest probability, EU: Expected utility

population tested, whose value falls between PTP_{Low} and PTP_{High} for appropriate testing, we defer to the Rogan and Gladen equation.[25] The Rogan and Gladen equation provides an unbiased estimate of the prevalence of disease as a function of three parameters: The proportion of tests indicating disease (t), the sensitivity of the test (α) and its specificity (β).[25] The variance of PTP_{Est} relates to the sample size (N).[25] Thus, as the sample size approaches a large number, PTP_{Est} approaches the expected value of the true PTP.

$$PTP_{Est} = \frac{t + \beta - 1}{\alpha + \beta - 1}, Var(PTP_{Est}) = \frac{t\ (1 - t)}{N\ (\alpha + \beta - 1)^2}$$

Although the original paper does not mention a connection between the Rogan and Gladen equation and Bayes theorem, they can be demonstrated to be mathematically equivalent [Appendix 1].

Method Application

Our model relies on cost-efficacy analysis to determine the PTP values (PTP_{Low}, PTP_{High}) for decision making. It is applicable whenever a reasonable cost efficacy analysis estimate is available. In its most general application one would evaluate the testing practices for an entire institution by determining if test outcomes are consistent with PTPs in the cost effective range. Actual PTPs and expected cost-effective ideal ranges would be reported as feedback to ordering providers. This feedback does not specify which specific ordering actions were noncompliant, but instead provides general feedback relating test orders to the results of those tests, allowing clinicians to re-calibrate their internal PTPs.

There are several logical alternative applications of this method. Instead of stratifying the test results by ordering provider, they could be stratified by other variables, including disease risk factors or patient care setting with different PTP cutoffs for each group. For example, a provider who works in a clinic and an intensive care unit could receive separate feedback for each setting. For comparisons with smaller test volumes, the estimate of PTP is not biased because it normalizes test volume by using the positive rate (number of positives/total). However, the variance of PTP_{Est} decreases with increasing sample size, which makes the conclusion of appropriate or inappropriate utilization management more robust with large samples.

As physicians or health systems incorporate feedback about PTP in the context of their risk aversion strategies and situation-specific exposure to clinical scenarios, this feedback will inform their future decision making strategy without assigning blame to a specific clinical decision. Our model does not attempt to quantify the specific factors involved in the decision-making, but rather infers the provider's decision-making strategy from the estimate of PTP. In summary, we present a Bayesian approach

to test utilization management that relates closely to decision-making theory.

Example Method Application

We sought to observe the appropriate or inappropriate utilization of a diagnostic test when applied to a cohort stratified by disease risk. This allowed us to analyze how the decision to test varied by patient risk factors. We selected a disease with known risk factors and its corresponding diagnostic test. Enterovirus (EV) causes seasonal meningitis that primarily affects infants during summer and autumn (June to November).[26] A common test for its diagnosis is real-time reverse transcriptase polymerase chain reaction (PCR) for EV RNA from cerebrospinal fluid (CSF) (EV-PCR). We obtained the test results from a cohort of patients, stratified them by age and month of testing, and determined if appropriate or inappropriate testing occurred with our model for each stratum. We then looked for patterns of utilization across the risk-stratified cohort to qualitatively determine the effect of population risk on the physician's decision to test.

Variables: Positive rate and disease risk factors

Two tertiary-care hospitals (Yale-New Haven Hospital, New Haven CT and the University of Washington Medical Center, St. Louis, MO) and one national reference laboratory (ARUP laboratories, Salt Lake City, UT) retrospectively contributed three data elements for each EV-PCR test performed between 2010 and 2012: The test results, patient age at the time of testing, and the month of order. Each institution employed a similar test methodology. The study used anonymized data and was determined to be human subjects exempt at participating institutions.

Constants: Low pretest probability, high pretest probability, sensitivity, and specificity

We obtained sensitivity, specificity, PTP_{Low}, and PTP_{High} through a systematic literature review. The review searched Medline with "EV [Mesh] NOT polio NOT poliovirus + English-only + Humans-only + Journal categories: Core clinical journals." An author (RH) screened the resulting publications in two steps, first by title and abstract then by full text, to identify cost-efficacy studies with PTP_{Low}, PTP_{High}, sensitivity and specificity.

Analysis

We analyzed the data according to our model in steps. First, we stratified the test results by patient age and month of the order. The categories for age were < 1, 1, 2, 3–10, 11–20, 21–30, 31–40, 41–50, and > 50 years. We arrived at these age categories by modification of an age interval published by the Center for Disease Control on EV.[26] Second, we calculated the PTP (PTP_{Est}) for each stratum. Third, we compared PTP_{Est} to cost-effectiveness recommendations for testing (PTP_{Low}, PTP_{High}). We

labeled as appropriate utilization each stratum with PTP_{Est} between PTP_{Low} and PTP_{High}. Otherwise, the stratum received the label of inappropriate utilization. We performed the analysis separately for each individual site and again with data combined from all three sites.

Sensitivity analysis

To explore the robustness of the outcome, we repeated the analysis after modifying the test performance and cohort stratification. We altered the performance of the test by an increase in both sensitivity and specificity, and, in a separate trial, decease in both sensitivity and specificity. We also changed the data stratification by aggregating by season instead of month, and again by aggregating age at different intervals: <1, 1–4, 5–9, 10–19, 20–44, ≥45 years.[26]

RESULTS

Constants: Low Pretest Probability, High Pretest Probability, Sensitivity and Specificity

To identify the constants required by our model, we screened the 642 publications returned by our Medline search. We filtered the publications to 14 with the title and abstract review. We found one cost efficacy study after full-text review.[27] The study assessed the use of EV-PCR in infants with fever and CSF pleocytosis admitted from an emergency department. It estimated sensitivity at 95%, specificity at 99%, and the PTP_{Low} to exist between 5.9% (PTP_{Low1}) and 12.8% (PTP_{Low2}). It did not provide a value for PTP_{High}.

Comparison to Recommendations

We collected 16,648 samples from Yale-New Haven ($n = 1197$), University of Washington Medical Center ($n = 1000$) and ARUP laboratories ($n = 14451$). ARUP samples originated from 608 different hospitals and clinics across the United States. After stratifying the samples by age and season, we calculated the utilization management metrics of volume, positive rate, and PTP_{Est} as shown in Table 1. We label each strata as overuse ($PTP_{Est} < PTP_{Low1} = 5.9\%$), equivocal ($5.9\% = PTP_{Low1} \leq PTP_{Est} \leq PTP_{Low2} = 12.8\%$), or not overuse ($12.8\% = PTP_{Low2} < PTP_{Est}$) in Figure 2. Overall we observed testing with overuse or equivocal benefit in 81 of the 108 stratum (70.8% of total tests).

Testing for EV occurred in low risk groups including people age 21 or older and the months of December to May. People 21 or older represented 58.2% of total tests, had a positive rate of 3.9%, a PTP_{Est} of 3.1%, and 44 of their 48 stratum (54.2% of total tests) received a label of overuse or equivocal. Tests ordered from December to May represented 41.9% of total tests, had a positive rate of 4.5%, a PTP_{Est} of 3.7%, and 50 of their 54 (40.5% of total tests) stratum received a label of overuse or equivocal. In contrast, the highest value testing occurred in people

under 21 years of age and between June and November. They produced 28.0% of total tests, had a positive rate of 6.6%, a PTP_{Est} of 5.9%, and 11 of their 30 stratum (1.7% of total tests) had a label of overuse or equivocal.

We noticed an exception to the general age trend at 1-year of age. 1-year-old represented 2.1% ($n = 357$) of total tests, had a positive rate of 0.6%, a PTP_{Est} of 0%, and all 12 monthly stratum received a label of overuse or equivocal. We observed the trends for season and age, including the trend for 1-year of age, across the three sites.

Sensitivity Analysis

For sensitivity analysis, we altered several constant parameters [Table 2]. To increase the performance of the test, we changed the sensitivity from 95% to 100% and specificity from 97% to 100%. We also decreased test performance by decreasing sensitivity from 95% to 92% and specificity from 99% to 97%. We changed the stratification of order date from month to season, and the stratification of age (original: <1, 1, 2, 3–10, 11–20, 21–30, 31–40, 41–50, >50; altered: <1, 1–4, 5–9, 10–19, 20–44, ≥45). We observed the largest change in tests labeled overuse or equivocal, a 2.7% decrease (32% to 29.3%) when altering the age stratification. The alternative age stratification grouped 1-year-old, a stratum shown to have low-value testing, with age one to four.

DISCUSSION

We have created a unique approach to providing feedback on physician's ordering habits. Our method derives from decision-making theory, and we base the conclusion of appropriate or inappropriate testing on cost-efficacy analysis. Unlike a rule-based approach, which dictates a particular physician decision in a given medical context, this method evaluates patterns of utilization allowing the physician to tailor their decision-making

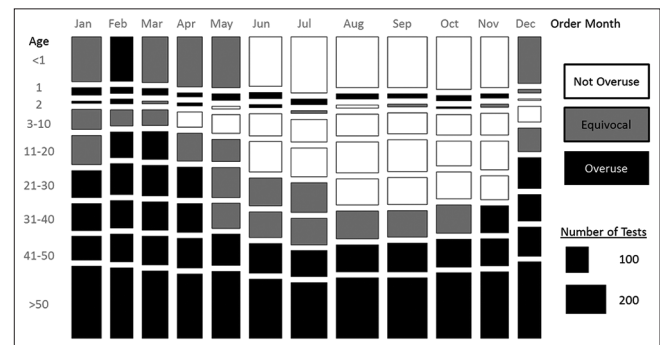**Figure 2: Comparison of clinician behavior to cost-efficacy analysis recommendations.** We compared the pretest probability estimates (PTP_{Est}) from Table 1 to cost efficacy recommendations (PTP_{Low1}, PTP_{Low2}) to determine overuse ($PTP_{Est} < PTP_{Low1} = 5.9\%$), equivocal benefit ($5.9\% = PTPLow1 \leq PTPEst \leq PTPLow2 = 12.8\%$), or not overuse ($12.8\% = PTP_{Low2} < PTP_{Est}$)

Table 1: Test results, positive rate and PTPEst by age and month. For each cell the top line represents positive results/total, and the second line the positive rate. We calculated the PTPEst in the third line with a sensitivity of 95% and a specificity of 99% . PTP: Pretest probability

	Jan (%)	Feb (%)	Mar (%)	Apr (%)	May (%)	Jun (%)	Jul (%)	Aug (%)	Sep (%)	Oct (%)	Nov (%)	Dec (%)
<1	21/235	10/176	16/210	21/227	30/258	45/289	89/363	102/384	95/360	77/331	42/255	18/192
	(8.94)	(5.68)	(7.62)	(9.25)	(11.63)	(15.57)	(24.52)	(26.56)	(26.39)	(23.26)	(16.47)	(9.38)
	8.44	4.98	7.04	8.78	11.31	15.50	25.02	27.19	27.01	23.68	16.46	8.91
1	0/40	0/24	0/31	0/18	0/34	0/35	1/36	0/39	0/30	0/33	0/22	1/15
	(0)	(0)	(0)	(0)	(0)	(0)	(2.78)	(0)	(0)	(0)	(0)	(6.67)
	0	0	0	0	0	0	1.89	0	0	0	0	6.03
2	0/11	0/18	1/13	0/14	2/12	1/19	2/16	5/24	2/19	0/12	1/15	1/5
	(0)	(0)	(7.69)	(0)	(16.67)	(5.26)	(12.50)	(20.83)	(10.53)	(0)	(6.67)	(20.0)
	0	0	7.12	0	16.67	4.54	12.23	21.10	10.13	0	6.03	20.21
3-10	12/107	7/65	8/71	11/69	26/95	26/128	54/146	67/164	63/148	47/127	41/110	10/65
	(11.21)	(10.77)	(11.27)	(15.94)	(27.37)	(20.31)	(36.99)	(40.85)	(42.57)	(37.01)	(37.27)	(15.38)
	10.87	10.39	10.92	15.90	28.05	20.55	38.28	42.40	44.22	38.31	38.59	15.30
11-20	11/153	3/102	4/130	9/126	13/112	34/175	33/187	73/233	54/226	38/157	33/144	10/97
	(7.19)	(2.94)	(3.08)	(7.14)	(11.61)	(19.43)	(17.65)	(31.33)	(23.89)	(24.20)	(22.92)	(10.31)
	6.58	2.07	2.21	6.53	11.28	19.60	17.71	32.27	24.36	24.68	23.32	9.90
21-30	5/143	1/123	1/139	7/138	12/151	19/164	21/190	36/198	35/189	27/169	17/120	5/127
	(3.50)	(0.81)	(0.72)	(5.07)	(7.95)	(11.59)	(11.05)	(18.18)	(18.52)	(15.98)	(14.17)	(3.94)
	2.66	0	0	4.33	7.39	11.26	10.69	18.28	18.64	15.93	14.01	3.12
31-40	3/140	1/108	2/132	5/130	13/126	11/149	19/170	27/210	19/189	20/177	8/135	4/110
	(2.14)	(0.93)	(1.52)	(3.85)	(10.32)	(7.38)	(11.18)	(12.86)	(10.05)	(11.30)	(5.93)	(3.64)
	1.22	0	0.55	3.03	9.91	6.79	10.83	12.61	9.63	10.96	5.24	2.80
41-50	1/124	0/111	0/132	3/133	1/159	2/173	7/166	11/203	8/202	7/174	2/135	0/119
	(0.81)	(0)	(0)	(2.26)	(0.63)	(1.16)	(4.22)	(5.42)	(3.96)	(4.02)	(1.48)	(0)
	0	0	0	1.34	0	0.17	3.42	4.70	3.15	3.22	0.51	0
>50	1/380	1/280	0/310	0/291	1/338	2/343	2/359	3/454	2/426	2/410	5/333	0/314
	(0.26)	(0.36)	(0)	(0)	(0.30)	(0.58)	(0.56)	(0.66)	(0.47)	(0.49)	(1.50)	(0)
	0	0	0	0	0	0	0	0	0	0	0.53	0

Table 2: Sensitivity analysis of percent overuse

Sensitivity (%)	Specificity (%)	Time of year groups	Age groups	Overuse (%)	Equivocal (%)	Not overuse (%)
95	99	Month	Control	50.0	20.8	29.3
100	100	Month	Control	49.2	20.3	30.5
92	97	Month	Control	55.0	16.5	28.5
95	99	Month	CDC	45.7	22.3	32.0
95	99	Season	Control	48.5	20.0	31.5

Control age groups: <1, 1, 2, 3-10, 11-20, 21-30, 31-40, 41-50, >50. Center for Disease Control (CDC) age groups: <1, 1-4, 5-9, 10-19, 20-44, ≥45[26]

strategy to the clinical context. It also differs from a rule-based utilization management in that it evaluates the consequence of the decision to test, specifically the positive or negative result, rather than the context in which the physician made the choice. Because it uses retrospective test results, the information needed by the model exists in nearly all clinical laboratories without the need for manual chart review.

The method for utilization management presented here demonstrates a basic concept upon which future authors could expand. For example, we focused on a prototypical test, a test with a binary result and a

diagnostic interpretation. A variation of our model may apply to tests informing prognosis or monitoring for side-effects of therapy. Another model variation could measure the significance of sample size differences among the strata as well as uncertainty in the sensitivity and specificity.[28] For a situation where a cost-efficacy analysis does not exist, our approach could be adapted to identify variations in utilization among groups.

Like other statistical models, random variation may influence the model's conclusions. For example, a physician could test a population with a high PTP of

disease, but a possibility exists where no patient has a positive result. Thus, the PTP_{Est} would appear low because it depends on the proportion of positive results. The physician would also appear to over-utilize the test. Fortunately, the probability of this scenario decreases as the sample size increases. The increase in sample size decreases the variance of PTP_{Est}.

Each patient in the cohort of tested patients has an individual PTP, and the distribution of the PTP among the cohort represents a second important factor in the interpretation of the model. In a large sample, the PTP_{Est} reflects the average PTP of the patients tested, but it does not provide information on the distribution of the patients' PTP. If therefore $PTP_{Est} < PTP_{Low}$ in a large population, it would be incorrect to conclude the majority of tests occurred in patients with too low of PTP. This statement assumes PTP_{Est} represents the population median, which it does not. In contrast, the conclusion could state the average PTP of the cohort had a value less than the lower limit of the suggested PTP range.

The systematic literature review to find a cost-efficacy analysis for EV-PCR found one paper that may overestimate the benefit of EV-PCR by supposing certain assumptions. First, the paper assumes a positive EV-PCR can rule-out bacterial meningitis faster than the gold standard diagnosis for bacterial meningitis, CSF culture. Second, it proposes EV-PCR had a turnaround time of 1-day. Third, patients with confirmed EV meningitis were assumed discharged in 1-day. When evaluated at one of the study's sites, three cases of positive CSF culture had a result within 1-day, while only half (15/30) of the positive cases of EV meningitis met the 1-day turnaround time and 1-day discharge time. The submission of a sample to a reference laboratory, as occurred in the majority of samples in this study, would presumably require longer than 1-day.

Because our method views decisions over time, not in the moment, it can separate the utilization metric (PTP_{Est}) from the judgment of the metric (i.e. appropriate, inappropriate). The modularity of our approach allows a future cost-efficacy study to re-evaluate our conclusions with different PTP cutoff values (PTP_{Low}, PTP_{High}). Similarly, a health care center could retrospectively trend its utilization knowing only its test results.

We demonstrate our approach to utilization on a large retrospective cohort of patients tested by real-time reverse transcriptase EV-PCR. The sample, gathered over multiple years from sites across the United States, represents one of the largest published EV cohorts. By stratifying the cohort across age and season, we determined specific patient subgroups receiving low-value testing. As the next step, we could then provide quantitative feedback to inform physician decision making on individual patients with population-based concerns of resource allocation.

REFERENCES

1. Lundberg GD. The modern clinical laboratory; Justification, scope, and directions. JAMA 1975;232:528-9.
2. Kochert E, Goldhahn L, Hughes I, Gee K, Stahlman B. Cost-effectiveness of routine coagulation testing in the evaluation of chest pain in the ED. Am J Emerg Med 2012;30:2034-8.
3. Eisenberg JM. An educational program to modify laboratory use by house staff. J Med Educ 1977;52:578-81.
4. Arnason T, Wells PS, Forster AJ. Appropriateness of diagnostic strategies for evaluating suspected venous thromboembolism. Thromb Haemost 2007;97:195-201.
5. Sikka R, Sweis R, Kaucky C, Kulstad E. Automated dispensing cabinet alert improves compliance with obtaining blood cultures before antibiotic administration for patients admitted with pneumonia. Jt Comm J Qual Patient Saf 2012;38:224-8.
6. Eyawo O, Fernandes KA, Brandson EK, Palmer A, Chan K, Lima VD, et al. Suboptimal use of HIV drug resistance testing in a universal health-care setting. AIDS Care 2011;23:42-51.
7. Hauser RG, Shirts BH. Do we now know what inappropriate laboratory utilization is? An expanded systematic review of laboratory clinical audits. Am J Clin Pathol 2014;141:774-83.
8. Juday T, Tang H, Harris M, Powers AZ, Kim E, Hanna GJ. Adherence to chronic hepatitis B treatment guideline recommendations for laboratory monitoring of patients who are not receiving antiviral treatment. J Gen Intern Med 2011;26:239-44.
9. Favaloro EJ, Mohammed S, Pati N, Ho MY, McDonald D. A clinical audit of congenital thrombophilia investigation in tertiary practice. Pathology 2011;43:266-72.
10. Cassel CK, Guest JA. Choosing wisely: Helping physicians and patients make smart decisions about their care. JAMA 2012;307:1801-2.
11. Kim JY, Dzik WH, Dighe AS, Lewandrowski KB. Utilization management in a large urban academic medical center: A 10-year experience. Am J Clin Pathol 2011;135:108-18.
12. Rang M. The Ulysses syndrome. Can Med Assoc J 1972;106:122-3.
13. Hirsh J, Guyatt G, Albers GW, Harrington R, Schünemann HJ, American College of Chest Physician. Antithrombotic and thrombolytic therapy: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest 2008;133:110S-2S.
14. Pawlson LG, Scholle SH, Powers A. Comparison of administrative-only versus administrative plus chart review data for reporting HEDIS hybrid measures. Am J Manag Care 2007;13:553-8.
15. van Wijk MA, van der Lei J, Mosseveld M, Bohnen AM, van Bemmel JH. Compliance of general practitioners with a guideline-based decision support system for ordering blood tests. Clin Chem 2002;48:55-60.
16. Sorace J, Wong HH, Worrall C, Kelman J, Saneinejad S, MaCurdy T. The complexity of disease combinations in the Medicare population. Popul Health Manag 2011;14:161-6.
17. Dolan JG, Bordley DR, Mushlin AI. An evaluation of clinicians' subjective prior probability estimates. Med Decis Making 1986;6:216-23.
18. Phelps MA, Levitt MA. Pretest probability estimates: A pitfall to the clinical utility of evidence-based medicine? Acad Emerg Med 2004;11:692-4.
19. Cahan A, Gilon D, Manor O, Paltiel O. Probabilistic reasoning and clinical decision-making: Do doctors overestimate diagnostic probabilities? QJM 2003;96:763-9.
20. Elstein AS. Heuristics and biases: Selected errors in clinical reasoning. Acad Med 1999;74:791-4.
21. Lyman GH, Balducci L. The effect of changing disease risk on clinical reasoning. J Gen Intern Med 1994;9:488-95.
22. Lyman GH, Balducci L. Overestimation of test effects in clinical judgment. J Cancer Educ 1993;8:297-307.
23. Williams ES, Konrad TR, Scheckler WE, Pathman DE, Linzer M, McMurray JE, et al. Understanding physicians' intentions to withdraw from practice: The role of job satisfaction, job stress, mental and physical health 2001. Health Care Manage Rev 2010;35:105-15.
24. Felder S, Mayrhofer T. Medical Decision Making – A Health Economics Primer. Berlin: Springer-Verlag; 2011.

25. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. Am J Epidemiol 1978;107:71-6.
26. Khetsuriani N, Lamonte-Fowlkes A, Oberst S, Pallansch MA, Centers for Disease Control and Prevention. Enterovirus surveillance – United States, 1970-2005. MMWR Surveill Summ 2006;55:1-20.
27. Nigrovic LE, Chiang VW. Cost analysis of enteroviral polymerase chain reaction in infants with fever and cerebrospinal fluid pleocytosis. Arch Pediatr Adolesc Med 2000;154:817-21.
28. Diggle PJ. Estimating prevalence using an imperfect test. Epidemiol Res Int 2011;2011:5.

APPENDIX, DERIVATION OF THE ROGAN-GLADEN EQUATION FROM BAYES THEOREM

Bayes Theorem

$P \ (Disease) \ P \ (Positive \ test \mid Disease) = P \ (Positive \ test) P \ (Disease \mid Positive \ Test)$

- Substitution: $P(Positive \ Test \mid Disease) \equiv Sensitivity$
- Substitution: $PPV \equiv P(Disease \mid Positive \ Test)$

$$P(Disease) \ Sensitivity = P(Positive \ Test) \ PPV$$

Substitution: $PPV \equiv \dfrac{True\,Positive}{Total\,Positive} = \left[\dfrac{P(Disease)\,Sensitivity}{P(Disease)\,Sensitivity + (1 - Specificity)(1 - P[Disease])} \right]$

$P(Disease) \ Sensitivity = P(Positive\,Test) \left[\dfrac{P(Disease)\,Sensitivity}{P(Disease)\,Sensitivity + (1 - Specificity)(1 - P[Disease])} \right]$

- Solve for P $(Disease)$ $\quad P(Disease) = \dfrac{Specificity + P(Positive\,Result) - 1}{Specificity + Sensitivity - 1}$