BMJ open

Molecular diagnostics require a Bayesian approach

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001804
Article Type:	Research
Date Submitted by the Author:	06-Sep-2012
Complete List of Authors:	Zilberberg, Marya; EviMed Research Group, LLC, ; University of Massachusetts, Shorr, Andrew; Washington Hospital Center, Medicine
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology, Health services research, Evidence based practice, Diagnostics
Keywords:	EPIDEMIOLOGY, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Diagnostic microbiology < INFECTIOUS DISEASES, Epidemiology < INFECTIOUS DISEASES, Adult intensive & critical care < INTENSIVE & CRITICAL CARE

Molecular diagnostics require a Bayesian approach

Marya D. Zilberberg, MD, MPH*[1,2],
[1]EviMed Research Group, LLC, Goshen, MA
[2]University of Massachusetts, Amherst, MA
*Corresponding author
President & CEO, Evi*Med* Research Group, LLC
Adjunct Associate Professor of Epidemiology, University of Massachusetts School of Public Health and Health Sciences
PO Box 303
Goshen, MA 01032
(413)-268-6381
(413)-268-3416 fax
evimedgroup@gmail.com

Andrew F. Shorr, MD, MPH[3]
[3]Washington Hospital Center, Washington, DC
Associate Section Director, Pulmonary Critical Care, Washington Hospital Center
Associate Professor of Medicine, Georgetown University
110 Irving Street, NW
Washington, DC 20010
(202)-877-2998
(202)-782-9032 fax
Andrew.Shorr@gmail.com

Competing interests: The authors have nothing to disclose. The authors received no funding for this study. All authors have completed the Unified Competing Interest form.
Guarantor statement: Dr. Zilberberg has had access to all the data and guarantees its integrity.
IRB statement: The study is a model and did not require an Ethics Committee review
Author contributions:
Dr. Zilberberg conceived, designed and executed the study, data collection and analyses, drafted the manuscript and revised it for important intellectual content.
Dr. Shorr refined the study idea, design and interpretation, and revised the manuscript for important intellectual content.

Word count: 2,562
References: 19
Tables and Figures: 4

Abstract

Objectives. Traditional microbiology identification takes 48-72 hours to complete. This lag forces clinicians to rely on broad-spectrum empiric coverage. To address this gap, manufacturers are developing rapid molecular diagnostics (RMD). We hypothesized that RMD's accuracy is more dependent upon population risk of harboring the culprit pathogen than to their sensitivity and specificity.
Design. A mathematical model
Setting and Participants. We used the range of risks (5%-50%) for methicillin-resistant *Staphylococcus aureus* (MRSA) among patients hospitalized with complicated skin and skin structure infections (cSSSI), pneumonia, or sepsis.
Interventions. None
Main outcome measures. We modeled the impact of changing a test's characteristics on its positive (PPV) and negative (NPV) predictive values, and hence the risk of over- or under-treatment, within strata of an organism's population prevalence. MRSA diagnostics provided assumptions for the test sensitivity and specificity (95%-99%). Scenarios with low sensitivity and specificity (90%), and best- and worst-case scenarios normalized to the annual universe of populations of interest, were examined.
Results. With a low prevalence (5%) and high test specificity, the PPV was 84%. Conversely, with 50% prevalence and 95% test specificity the PPV rose to ≥95%. Even when the test's specificity and sensitivity were both 90%, in a high-risk population both PPV and NPV were ~90%. In the worst-case scenario, 150,000 patients with cSSSI, pneumonia and sepsis annually were at risk for inappropriate treatment, 91% of these at risk for over-treatment. In the best-case scenario, 81% of 18,000 patients at risk for inappropriate coverage were subject to over-treatment.
Conclusions. Although promising for limiting exposure to excessive antimicrobial coverage, RMDs alone will not solve the issue of inappropriate, and particularly over-, treatment. Increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a test with near-perfect characteristics. The healthcare community must build robust evidence and information technology infrastructure to guide appropriate use of such testing.

Article summary

Article focus

- Traditional microbiology identification takes 48-72 hours to complete, which forces clinicians to rely on broad-spectrum empiric coverage.
- To address this gap, manufacturers are developing rapid molecular diagnostics (RMD).
- It is unclear what impact RMDs may have in different population on overdiagnosis and overtreatment

Key messages

- Although promising for limiting exposure to excessive antimicrobial coverage, RMDs alone will not solve the issue of inappropriate, and particularly over-, treatment.
- Increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a test with near-perfect sensitivity and specificity.

Strengths and limitations

- As a mathematical model, our study relies on the accuracy of estimates in the literature, which predisposes our computations to greater uncertainty.
- The model is transparent
- Our findings span a wide range of plausible epidemiology.
- The data underscore the need to understand local pathogen patterns, the recognition of which should drive decisions about the utility of these powerful molecular diagnostics.

Introduction

Despite the fact that antibiotics represent a relatively recent advance in medicine, antibiotic resistant bacteria are now common in both the hospital and the community. Antibiotic misuse and abuse represent a key driver of the increasing prevalence in antibiotic resistance (1, 2). The spread of antimicrobial resistance has similarly created a vicious cycle where clinicians repeatedly reach for extended spectrum agents in order to address the current patterns of resistance while potentially worsening them for the future. Underlying this practice approach has been the general unavailability of reliable, rapid diagnostics to help establish the etiology of an infection. Indeed, traditional phenotypic microbiology methods take 48 to 72 hours to identify an organism when present and to determine the antibiotic susceptibility profile. Without a prompt means for either including or excluding potentially resistant pathogens, clinicians frequently have no alternative but to rely on broad-spectrum options for empiric therapy. Such approach is currently warranted, given the extensive data documenting that delayed and inappropriate antibiotic treatment increases the risk for mortality and prolongs the duration of hospitalization (3-9). However, rapid and accurate diagnosis should diminish the uncertainty and help target the culprit organisms without straying into the extremes of overly narrow or overly broad coverage.

To fill this diagnostic gap, several manufacturers are engaged in developing rapid diagnostic modalities that incorporate recent advances in molecular techniques relying on genotyping the organisms. Indeed, some of these technologies are able to arrive at the microbiologic diagnosis in as little as 2 hours, a critical period for tailoring treatment

(10). Such improvement in shortening the diagnostic time is invaluable, particularly given these tests' ostensible accuracy.

At the same time, one must exercise caution because these tests are not 100% accurate. And while manufacturers strive for ever-increasing sensitivity and specificity for their tests, a more fruitful area of investigation may be learning to identify characteristics of specific populations in whom these tests may prove to be most helpful for targeting and tailoring treatment. In other words, the central clinical question may revolve not around issues of sensitivity and specificity intrinsic to the test, but rather around the positive (PPV) and negative (NPV) predictive values associated with these newer tools in populations with various levels of risk for the organisms in question. Irrespective of the sensitivity and specificity, if the PPV and NPV are not sufficiently high, then these new tests may not help clinicians either to withhold unnecessarily broad coverage or to tailor it shortly after the results return.

We hypothesized that even under conditions where such rapid diagnostic tests had near-perfect sensitivity and specificity, the population-specific risk for having a particular organism would represent a crucial consideration in driving diagnostic accuracy. That is failure to consider the pre-test probability of these organisms in the population screened would undermine the potential value of rapid diagnostic tests. To address this question we developed a model simulation evaluating the application of these assays, and relied upon publicly available data to populate our analysis.

Methods

We developed a mathematical model simulating the impact of changing a test's characteristics on its accuracy within several strata of population risk for a particular

organism. All the inputs were extracted from publicly available data. The primary outcome of interest was the potential magnitude for over-diagnosis of a particular pathogen, or the proportion of false positive tests under the varying assumptions. We were specifically interested in the false positive rates, since these cases are the ones most likely to receive overly broad treatment when it is not indicated. Such overly broad treatment represents a key clinical endpoint since it exposes the patient and the healthcare system to adverse consequences individually and as a group. As a secondary endpoint we examined the overall inaccuracy of the test in various scenarios, defined as the sum of the false positive and false negative results as a proportion of the total population.

The model was based on the approximate risks of methicillin-resistant *Staphylococcus aureus* (MRSA) among three distinct hospitalized populations: 1) complicated skin and skin structure infections (cSSSI) (11-13), 2) pneumonia (11, 14), and 3) sepsis (11). We sought the most generalizable estimates for at least two factors out of the following three, using the available data to calculate the third when necessary: 1) total volume of hospitalizations for each of the diseases of interest, 2) proportion of the total volume represented by MRSA, and 3) total number of MRSA infections in each disease category (11, 14, 15).

For consistency, the assumptions for the corresponding test characteristics mimicked those from MRSA diagnostics (16). To derive estimates for positive (PPV) and negative predictive value (NPV) for a plausible range of test characteristics, we developed four hypothetical testing situations: 1) Test A, with the sensitivity and specificity of 95%, 2) Test B, with the sensitivity 99% and specificity 95%, 3) Test C, with the sensitivity 95% and specificity 99%, and 4) Test D, with the sensitivity and

specificity 99%. To explore how deviations from the average sensitivity and specificity metrics may impact the accuracy of identification, we conducted sensitivity analyses assuming 90% sensitivity and specificity. Based on the range of MRSA risk estimates in the populations of interest (i.e., cSSSI, pneumonia, and sepsis), we varied the prevalence estimates from 5% to 50%, and calculated the PPV and NPV for each of the intermediate values.

We additionally performed best- and worst-case scenario simulations for each of the populations in question. Thus, for the worst-case scenario where all variables were biased against the novel rapid diagnostic assay, we utilized as inputs the highest disease volume and lowest disease prevalence, along with the lowest test sensitivity and specificity values. Skewing the inputs in this fashion provides a potential estimate of the extent and impact of misclassification when all assumptions are shifted so as to constrain the potential value of the rapid diagnostic test in question. Conversely, for the best-case scenario, we input the lowest disease volume and the highest disease prevalence, along with the highest test sensitivity and test specificity. For both of these analyses, the total annual universe of specific disease hospitalizations in the US was used. We utilized these values to estimate the total numbers of potential cases within each population that would be over-treated (i.e., treated for MRSA when no MRSA is present), under-treated (i.e., not treated for MRSA when MRSA is present) and treated inappropriately (i.e., either over- or undertreated).

Both the values for sensitivity and specificity and disease risk were rounded in order to ease computational presentation. Volumes and prevalence of MRSA in the disease states of interest were extracted from several large surveys available in the public

domain (Table 1). Thus, for the MRSA cSSSI volumes we relied on a study by Klein, which quantified these hospitalizations in 2005 (11). The proportions of cSSSI in which MRSA is the offending pathogen derived from two recent epidemiologic studies of cSSSI hospitalizations in the US (12, 13). The volume of pneumonia hospitalizations was extracted from the American Lung Association's 2010 data, and the proportion represented by MRSA from a large and representative database analysis by Kollef and colleagues (14, 15). Finally, we relied on the Agency's for Healthcare Research and Quality recent statistical brief quantifying the burden of hospitalizations with sepsis, while the Klein study provided the proportion likely caused by MRSA (11, 16).

Results

The input assumptions and their sources are presented in Table 1. The estimated prevalence of MRSA ranges from approximately 5% in sepsis to nearly 50% in cSSSI, while the prevalence of MRSA in pneumonia falls between those extremes. Under the conditions of lowest prevalence (5%) along with the average test specificity of 95%, the PPV reaches only 50% (Figure 1). Improving the specificity by nearly 5% to 99% without altering the disease prevalence results in a moderate improvement in the PPV to approximately 84%. Alternatively, a change of a similar magnitude in the PPV occurs, when the prevalence of disease increases from 5% to the 10%-20% range, even as the specificity remains anchored at 95% (Figure 1). The PPV further improves as the prevalence of disease approaches 50%. Notably, at the extremes of disease prevalence and test specificity, the relative improvement in test accuracy is numerically greater when the prevalence is increased while holding the specificity constant (PPV 95.0% and NPV 95.2% for Tests A and B, prevalence 50%) as compared to a scenario where one

modulates the test specificity and maintains the prevalence constant (PPV 83.3% and NPV 83.9% for Tests C and D, prevalence 5%). Put another way, the net change in PPV is maximized based on moderate changes in disease prevalence as opposed to alterations in test sensitivity. As for the NPV, a rise in sensitivity from 95% to 99% does not yield substantial alterations in the value. Essentially, the NPV is already quite high, no matter what the prevalence of resistance in the population. Conversely, the NPV suffers only modestly in the populations where disease prevalence is highest compared to those with the lowest disease prevalence (Figure 1).

The sensitivity analysis in which we assume that both the sensitivity and specificity of the test equal 90% is illustrated in Figure 2. At the lowest prevalence of disease, this specificity affords an unacceptably low PPV (32.1%), while the NPV remains high, exceeding 99%. As the prevalence of the disease rises in the target population, while the test's specificity and sensitivity remain fixed at 90%, the PPV and NPV converge at 90%, indicating a major improvement in the PPV without dramatically compromising the NPV.

Best- and worst-case scenario estimates of the total annual pool of patients at risk for MRSA infection in cSSSI, pneumonia and sepsis demonstrate that the potential for over-treatment far exceeds that for under-treatment (Table 2). Focusing on sepsis as an example, for the worst-case calculation we assumed 1,141,000 sepsis hospitalizations annually, a 5% MRSA prevalence, along with test characteristics of 95% sensitivity and 95% specificity. These parameters resulted in 57,050 potential cases of inappropriate treatment reflecting the sum of subjects classified as falsely positive or negative. Of these misclassified subjects, 54,198 (95%) represent those at risk for over-treatment.

Conversely, under the best-case assumptions of a high MRSA prevalence (10%) in sepsis (n=727,000), and a test with near-perfect sensitivity and specificity (both 99%), only 7,270 individuals are at risk for inappropriate treatment with 6,543 (90%) being over-treated (Table 2).

Overall, under the worst-case assumptions for all three of the conditions of interest, over 150,000 patients annually with these three conditions may be treated inappropriately, with overtreatment accounting for 136,000 (91%) of this cohort. Under the best circumstances, among the more than 18,000 patients treated potentially inappropriately, nearly 15,000 (81%) may be subjected to over-treatment (Table 2).

Discussion

We have demonstrated explicitly that organism prevalence is an important driver of the accuracy of rapid molecular diagnostic tests even when their sensitivity and specificity are near perfect. Additionally, we have shown that although improving the theoretical test's specificity results in greater accuracy, one enhances accuracy even more by restricting test utilization to a population at an increased risk for infection with the pathogen in question. In other words, increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a marginally superior rapid diagnostic test with near-perfect specificity. Finally, although promising for limiting exposure to excessive antimicrobial coverage, molecular diagnostics are still likely to result in a substantial amount of inappropriate treatment. The vast majority (over 90%) of such inappropriate coverage is due to over-treatment in scenarios where the test is applied irrespective of considerations of the prevalence of a resistant pathogen.

Our data have several important implications. First, as manufacturers, regulators and clinicians consider what tests may have superior characteristics, it is important for all stake holders to engage in defining the appropriate populations for testing with these novel technologies. Our data clearly demonstrate that rather than expending resources for every laboratory to elevate their sensitivity and specificity to close to 100%, the more fruitful effort may be to develop algorithms to identify those patient populations at high risk for the disease being tested. This is particularly true given that marginal enhancements in sensitivity and specificity often come at the cost of substantial financial investments. Second, raising the sensitivity of these technologies even beyond the current levels may be pursuing diminishing returns, given the already high NPV. That is, even when the sensitivity is no higher than 90%, the negative predictive value reaches very high levels (over 95%) in the setting of moderate pre-test probability for disease.

Third, and possibly most important, by using genotyping as opposed to phenotyping employed in the traditional microbiology laboratory methods, molecular diagnostics promise to result in sensitivity values that far exceed those of the traditional techniques. The flip side of this optimization in sensitivity is a blunting in specificity, whereby it may become unclear whether the identified organism is indeed the cause of the clinical condition. Our data indicate that the true need in diagnostic testing lies not in further optimization of sensitivity, but in improving the specificity of the results.

Because improvements in one by necessity lead to detriments in the other, future directions in molecular diagnostics require thoughtful planning. We have clearly shown that, in order to live up to the promise of improved targeting of antibiotic treatment, such planning must include careful consideration of the populations in whom molecular

diagnostic techniques are appropriate. In fact the most important lesson from our simulation is that we need to develop algorithms to help identify patients belonging to populations with a high risk for the particular pathogen. If a predictive algorithm is able to enrich the population to be tested to the disease prevalence between 30% and 40%, both PPV and NPV will be moved into a useful range even when the test's sensitivity and specificity are both well below 100%. With the advent of health informatics and the massive growth in computing ability, turning reams of patient data into predictive equations is a clearly needed functionality. Already several computing systems are addressing this need, and the trend should continue with the input from all stakeholders (18, 19).

Our study has a number of limitations. The most important limitation is that it is merely a mathematical model, and, as such, by necessity relies on the accuracy of estimates in the literature. The fact that some of the papers we used for deriving our assumptions themselves were modeling exercises (11), predisposes our computations to greater uncertainty. This, however, does not negate our findings that span a wide range of plausible epidemiology. Furthermore, our model underscores the need to understand local pathogen patterns, the recognition of which should drive decisions about the utility of these powerful molecular diagnostics.

In summary, molecular diagnostics promise to streamline identification and treatment of many infectious diseases. While the emergence of these powerful technologies is a positive development, we need to attend to developing algorithms to aid in selecting appropriate patients for their use. Indiscriminate application of molecular diagnostics to all-comers presenting with signs of an infection without consideration for

pre-test probability of disease is likely to result in a great deal of antimicrobial overtreatment. This will then only accelerate the current trajectory of escalating resistance. In conjunction with developing these important technologies, it is incumbent upon the healthcare community to build robust evidence and information technology infrastructure to guide appropriate use of such testing.

Acknowledgements. No one other than the listed authors participated in the study. Dr. Zilberberg had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Dr. Zilberberg participated in the conception and design of the study, data anaysisn and interpretation and drafting and revising of the manuscript for important intellectual content. Dr. Shorr participated in the conception of the study, data interpretation and revising of the manuscript for important intellectual content.

Data sharing:

The model and the inputs are quite transparent and mostly can be found in the paper. If there is a need for further information, we are happy to provide it.

References

1. Dellit TH, Owens RC, McGowan JE, Jr., Gerding DN, Weinstein RA, Burke JP et al. Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. Clin Infect Dis 2007; 44(2):159-177

2. Spellberg B, Guidos R, Gilbert D, Bradley J, Boucher HW, Scheld WM et al. The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. Clin Infect Dis 2008; 46(2):155-164

3. Micek ST, Kollef KE, Reichley RM, Roubinian N, Kollef MH. Health care-associated pneumonia and community-acquired pneumonia: a single-center experience. Antimicrob Agents Chemother 2007; 51(10):3568-3573

4. Schramm GE, Johnson JA, Doherty JA, Micek ST, Kollef MH. Methicillin-resistant Staphylococcus aureus sterile-site infection: The importance of appropriate initial antimicrobial treatment. Crit Care Med 2006; 34(8):2069-2074.

5. Ibrahim EH, Sherman G, Ward S, Fraser VJ, Kollef MH. The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting. Chest 2000; 118(1):146-155.

6. Alvarez-Lerma F. Modification of empiric antibiotic treatment in patients with pneumonia acquired in the intensive care unit. ICU-Acquired Pneumonia Study Group. Intensive Care Med 1996; 22(5):387-394.

7. Iregui M, Ward S, Sherman G, Fraser VJ, Kollef MH. Clinical importance of delays in the initiation of appropriate antibiotic treatment for ventilator-associated pneumonia. Chest 2002; 122(1):262-268.

8. Zilberberg MD, Shorr AF, Micek ST, Mody SH, Kollef MH. Antimicrobial therapy escalation and hospital mortality among patients with health-care-associated pneumonia: a single-center experience. Chest 2008; 134(5):963-968.

9. Zilberberg MD, Shorr AF, Micek ST, Chen J, Ramsey AM, Hoban AP et al. Hospitalizations with healthcare-associated complicated skin and skin structure infections: impact of inappropriate empiric therapy on outcomes. J Hosp Med 2010; 5(9):535-540.

10. Peterson LR, Liesenfeld O, Woods CW et al. Multicenter evaluation of the LightCycler methicillin-resistant *Staphylococcus aureus* (MRSA) advanced test as a rapid method for detection of MRSA in nasal surveillance swabs. J Clin Microbiol 2010;48:1661-6

11. Klein E, Smith DL, Laxminarayan R. Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus*, United States, 1999–2005. Emerg Infect Dis 2007;13:1840-6

12. Lipsky BA, Weigelt JA, Gupta V, Killian A, Peng MM. Skin, soft tissue, bone, and joint infections in hospitalized patients: epidemiology and microbiological, clinical, and economic outcomes. Infect Control Hosp Epidemiol 2007;28:1290-8

13. Jenkins TC, Sabel AL, Sarcone EE, et al. Skin and soft tissue infections requiring hospitalization at an academic medical center: Opportunities for antimicrobial stewardship. Clin Infect Dis 2010;51:895-903

14. Kollef MH, Shorr A, Tabak YP, et al. Epidemiology and outcomes of health-care-associated pneumonia: Results from a large US database of culture-positive pneumonia. Chest 2005;128:3854-3862

15. American Lung Association. Trends in Pneumonia and Influenza Morbidity and Mortality, 2010. Available at http://www.lungusa.org/finding-cures/our-research/trend-reports/pi-trend-report.pdf, accessed January 17, 2012

16. Hall MJ, Williams SN, DeFrances CJ, Golosinskiy A. Inpatient care for septicemia or sepsis: A challenge for patients and hospitals. NCHS data brief, no 62. Hyattsville, MD: National Center for Health Statistics. 2011. Available at http://www.cdc.gov/nchs/data/databriefs/db62.pdf, accessed February 1, 2012

17. BD molecular diagnostic solutions. BD GeneOhm Staph SR Assay package insert. Available at http://www.bd.com/geneohm/english/products/pdfs/staphsr_pkginsert.pdf, accessed February 15, 2012

18. Pepitone J. IBM's 'Jeopardy' computer lands health care job. CNNMoney 2011 Sep 12, available at http://money.cnn.com/2011/09/12/technology/ibm_watson_health_care/index.htm , accessed February 15, 2012

19. Terry K. Futuristic Clinical Decision Support Tool Catches On. InformationWeek Healthcare 2012 Jan 27, available at http://informationweek.com/news/healthcare/clinical-systems/232500603, accessed February 15, 2012

Table 1. Annual hospitalization volumes

Infection type	Annual volume	MRSA prevalence (%)	MRSA volume
cSSSI	434,227-1,211,863	15.3%[12]-42.7[13]	185,415[11]
Pneumonia	651,000[15]	5.6%-14.3%[14]	36,540[11]-93,093
Sepsis	727,000-1,141,000[16]	4.9%-7.7%	56,246[11]

MRSA=methicillin-resistant *Staphylococcus aureus*; cSSSI=complicated skin and skin structure infection

Table 2. Best- and worst-case scenario simulations for each disease group

	Over-treated	Under-treated	Treated inappropriately
Best-case scenario			
cSSSI	2,600	1,733	4,333
Pneumonia	5,534	977	6,510
Sepsis	6,543	727	7,270
Total	**14,676**	**3,437**	**18,113**
Worst-case scenario			
cSSSI	51,389	9,069	60,458
Pneumonia	30,923	1,628	32,550
Sepsis	54,198	2,853	57,050
Total	**136,509**	**13,549**	**150,058**

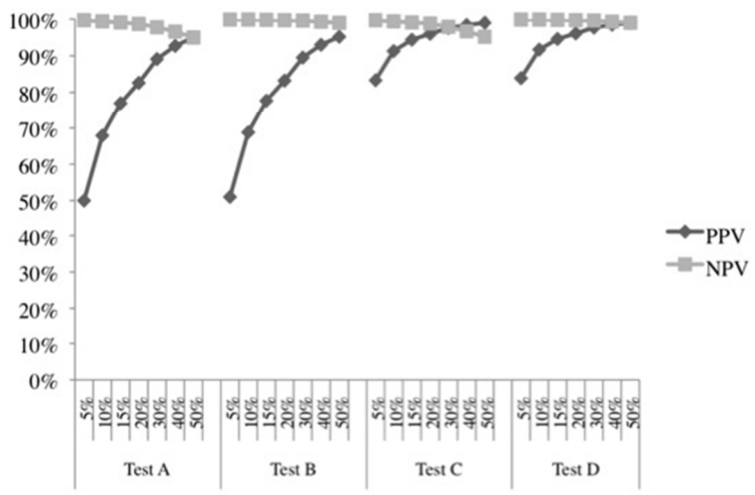
cSSSI=complicated skin and skin structure infection

Figure 1. Positive and negative predictive values of a test with the given sensitivity and specificity, stratified by population disease prevalence*

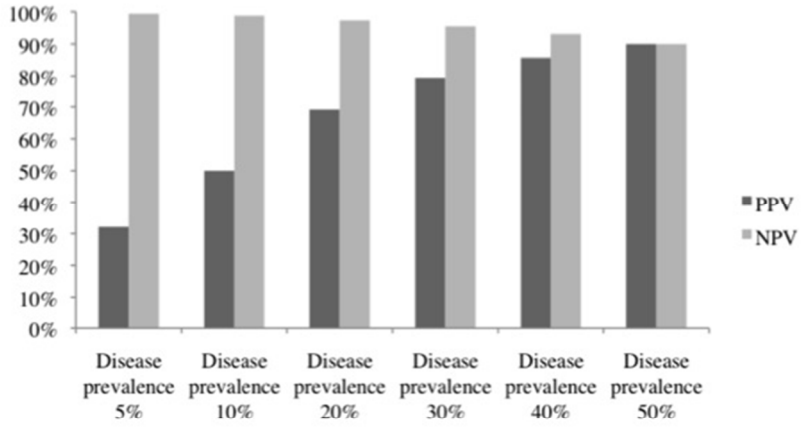
*Percentages along the X-axis represent disease prevalence strata
PPV = positive predictive value; NPV = negative predictive value
Test A: sensitivity = 95%, specificity = 95%; Test B: sensitivity = 99%, specificity = 95%;
Test C: sensitivity = 95%, specificity = 99%; Test D: sensitivity = 99%, specificity = 99%

Figure 2. Sensitivity analysis under the conditions of test sensitivity and specificity equaling 90%

PPV = positive predictive value; NPV = negative predictive value



254x190mm (72 x 72 DPI)



254x190mm (72 x 72 DPI)

BMJ open

Impact of prior probabilities of MRSA as an infectious agent on the accuracy of the emerging molecular diagnostic tests: A model simulation

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-001804.R1
Article Type:	Research
Date Submitted by the Author:	16-Nov-2012
Complete List of Authors:	Zilberberg, Marya; EviMed Research Group, LLC, ; University of Massachusetts, Shorr, Andrew; Washington Hospital Center, Medicine
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology, Health services research, Evidence based practice, Diagnostics
Keywords:	EPIDEMIOLOGY, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Diagnostic microbiology < INFECTIOUS DISEASES, Epidemiology < INFECTIOUS DISEASES, Adult intensive & critical care < INTENSIVE & CRITICAL CARE

Impact of prior probabilities of MRSA as an infectious agent on the accuracy of the emerging molecular diagnostic tests: A model simulation

Marya D. Zilberberg, MD, MPH*[1,2],
[1]EviMed Research Group, LLC, Goshen, MA
[2]University of Massachusetts, Amherst, MA
*Corresponding author
President & CEO, Evi*Med* Research Group, LLC
Adjunct Associate Professor of Epidemiology, University of Massachusetts School of Public Health and Health Sciences
PO Box 303
Goshen, MA 01032
(413)-268-6381
(413)-268-3416 fax
evimedgroup@gmail.com

Andrew F. Shorr, MD, MPH[3]
[3]Washington Hospital Center, Washington, DC
Associate Section Director, Pulmonary Critical Care, Washington Hospital Center
Associate Professor of Medicine, Georgetown University
110 Irving Street, NW
Washington, DC 20010
(202)-877-2998
(202)-782-9032 fax
Andrew.Shorr@gmail.com

Competing interests: The authors have nothing to disclose. The authors received no funding for this study. All authors have completed the Unified Competing Interest form.
Guarantor statement: Dr. Zilberberg has had access to all the data and guarantees its integrity.
IRB statement: The study is a model and did not require an Ethics Committee review
Author contributions:
Dr. Zilberberg conceived, designed and executed the study, data collection and analyses, drafted the manuscript and revised it for important intellectual content.
Dr. Shorr refined the study idea, design and interpretation, and revised the manuscript for important intellectual content.

Word count: 2,690
References: 19

Tables and Figures: 4

Abstract

Objectives. Traditional microbiology identification takes 48-72 hours to complete. This lag forces clinicians to rely on broad-spectrum empiric coverage. To address this gap, manufacturers are developing rapid molecular diagnostics (RMD). We hypothesized that RMD's accuracy is more dependent upon population risk of harboring the culprit pathogen than to their sensitivity and specificity.
Design. A mathematical model
Setting and Participants. We used the range of risks (5%-50%) for methicillin-resistant *Staphylococcus aureus* (MRSA) among patients hospitalized with complicated skin and skin structure infections (cSSSI), pneumonia, or sepsis.
Main outcome measures. We modeled the impact of changing a test's characteristics on its positive (PPV) and negative (NPV) predictive values, and hence the risk of over- or under-treatment, within strata of an organism's population prevalence. MRSA diagnostics provided assumptions for the test sensitivity and specificity (95%-99%). Scenarios with low sensitivity and specificity (90%), and best- and worst-case scenarios normalized to the annual universe of populations of interest, were examined.
Results. With a low prevalence (5%) and high test specificity, the PPV was 84%. Conversely, with 50% prevalence and 95% test specificity the PPV rose to ≥95%. Even when the test's specificity and sensitivity were both 90%, in a high-risk population both PPV and NPV were ~90%. In the worst-case scenario, 150,000 patients with cSSSI, pneumonia and sepsis annually were at risk for inappropriate treatment, 91% of these at risk for over-treatment. In the best-case scenario, 81% of 18,000 patients at risk for inappropriate coverage were subject to over-treatment.
Conclusions. Although promising for limiting exposure to excessive antimicrobial coverage, RMDs alone will not solve the issue of inappropriate, and particularly over-, treatment. Increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a test with near-perfect characteristics. The healthcare community must build robust evidence and information technology infrastructure to guide appropriate use of such testing.

Article summary

Article focus

- Traditional microbiology identification takes 48-72 hours to complete, which forces clinicians to rely on broad-spectrum empiric coverage.
- To address this gap, manufacturers are developing rapid molecular diagnostics (RMD).
- It is unclear what impact RMDs may have in different population on overdiagnosis and overtreatment

Key messages

- Although promising for limiting exposure to excessive antimicrobial coverage, RMDs alone will not solve the issue of inappropriate, and particularly over-, treatment.
- Increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a test with near-perfect sensitivity and specificity.

Strengths and limitations

- As a mathematical model, our study relies on the accuracy of estimates in the literature, which predisposes our computations to greater uncertainty.
- The model is transparent.
- Our findings span a wide range of plausible epidemiology.
- The data underscore the need to understand local pathogen patterns, the recognition of which should drive decisions about the utility of these powerful molecular diagnostics.

Introduction

Despite the fact that antibiotics represent a relatively recent advance in medicine, antibiotic resistant bacteria are now common in both the hospital and the community. Antibiotic misuse and abuse represent a key driver of the increasing prevalence in antibiotic resistance (1, 2). The spread of antimicrobial resistance has similarly created a vicious cycle where clinicians repeatedly reach for extended spectrum agents in order to address the current patterns of resistance while potentially worsening them for the future. Underlying this practice approach has been the general unavailability of reliable, rapid diagnostics to help establish the etiology of an infection. Indeed, traditional phenotypic microbiology methods take 48 to 72 hours to identify an organism when present and to determine the antibiotic susceptibility profile. Without a prompt means for either including or excluding potentially resistant pathogens, clinicians frequently have no alternative but to rely on broad-spectrum options for empiric therapy. Such approach is currently warranted, given the extensive data documenting that delayed and inappropriate antibiotic treatment increases the risk for mortality and prolongs the duration of hospitalization (3-9). However, rapid and accurate diagnosis should diminish the uncertainty and help target the culprit organisms without straying into the extremes of overly narrow or overly broad coverage.

To fill this diagnostic gap, several manufacturers are engaged in developing rapid diagnostic modalities that incorporate recent advances in molecular techniques relying on genotyping the organisms. Indeed, some of these technologies are able to arrive at the microbiologic diagnosis in as little as 2 hours, a critical period for tailoring treatment

(10). Such improvement in shortening the diagnostic time is invaluable, particularly given these tests' ostensible accuracy.

At the same time, one must exercise caution because these tests are not 100% accurate. And while manufacturers strive for ever-increasing sensitivity and specificity for their tests, a more fruitful area of investigation may be learning to identify characteristics of specific populations in whom these tests may prove to be most helpful for targeting and tailoring treatment. In other words, the central clinical question may revolve not around issues of sensitivity and specificity intrinsic to the test, but rather around the positive (PPV) and negative (NPV) predictive values associated with these newer tools in populations with various levels of risk for the organisms in question. This approach fits in with the Bayesian decision making, whereby the prior probability of an event informs the interpretation of the diagnostic data. Irrespective of the sensitivity and specificity, if the PPV and NPV are not sufficiently high, then these new tests may not help clinicians either to withhold unnecessarily broad coverage or to tailor it shortly after the results return.

We hypothesized that even under conditions where such rapid diagnostic tests had near-perfect sensitivity and specificity, the population-specific risk for having a particular organism would represent a crucial consideration in driving diagnostic accuracy. That is failure to consider the pre-test probability of these organisms in the population screened would undermine the potential value of rapid diagnostic tests. To address this question we developed a model simulation evaluating the application of these assays, and relied upon publicly available data to populate our analysis.

Methods

We developed a mathematical model simulating the impact of changing a test's characteristics on its accuracy within several strata of population risk for a particular organism. All the inputs were extracted from publicly available data. The primary outcome of interest was the potential magnitude for over-diagnosis of a particular pathogen, or the proportion of false positive tests under the varying assumptions. We were specifically interested in the false positive rates, since these cases are the ones most likely to receive overly broad treatment when it is not indicated. Such overly broad treatment represents a key clinical endpoint since it exposes the patient and the healthcare system to adverse consequences individually and as a group. As a secondary endpoint we examined the overall inaccuracy of the test in various scenarios, defined as the sum of the false positive and false negative results as a proportion of the total population.

The model was based on the approximate risks of methicillin-resistant *Staphylococcus aureus* (MRSA) among three distinct hospitalized populations: 1) complicated skin and skin structure infections (cSSSI) (11-13), 2) pneumonia (11, 14), and 3) sepsis (11). We sought the most generalizable estimates for at least two factors out of the following three, using the available data to calculate the third when necessary: 1) total volume of hospitalizations for each disease of interest, 2) proportion of the total volume represented by MRSA, and 3) total number of MRSA infections in each disease category (11, 14, 15).

For consistency, the assumptions for the corresponding test characteristics mimicked those from MRSA diagnostics (16). To derive estimates for positive (PPV) and negative predictive values (NPV) for a plausible range of test characteristics, we developed four hypothetical testing situations: 1) Test A, with the sensitivity and

specificity of 95%, 2) Test B, with the sensitivity 99% and specificity 95%, 3) Test C, with the sensitivity 95% and specificity 99%, and 4) Test D, with the sensitivity and specificity 99%. To explore how deviations from the average sensitivity and specificity metrics may impact the accuracy of identification, we conducted sensitivity analyses assuming 90% sensitivity and specificity. Based on the range of MRSA risk estimates in the populations of interest (i.e., cSSSI, pneumonia, and sepsis), we varied the prevalence estimates from 5% to 50%, and calculated the PPV and NPV for each of the intermediate values.

We additionally performed best- and worst-case scenario simulations for each population in question. Thus, for the worst-case scenario where all variables were biased against the novel rapid diagnostic assay, we utilized as inputs the highest disease volume and lowest disease prevalence, along with the lowest test sensitivity and specificity values. Skewing the inputs in this fashion provides a potential estimate of the extent and impact of misclassification when all assumptions are shifted so as to constrain the potential value of the rapid diagnostic test in question. Conversely, for the best-case scenario, we input the lowest disease volume and the highest disease prevalence, along with the highest test sensitivity and test specificity. For both of these analyses, the total annual universe of specific disease hospitalizations in the US was used. We utilized these values to estimate the total numbers of potential cases within each population that would be over-treated (i.e., treated for MRSA when no MRSA is present), under-treated (i.e., not treated for MRSA when MRSA is present) and treated inappropriately (i.e., either over- or undertreated).

Both the values for sensitivity and specificity and disease risk were rounded in order to ease computational presentation. Volumes and prevalence of MRSA in the disease states of interest were extracted from several large surveys available in the public domain (Table 1). Thus, for the MRSA cSSSI volumes we relied on a study by Klein, which quantified these hospitalizations in 2005 (11). The proportions of cSSSI in which MRSA is the offending pathogen derived from two recent epidemiologic studies of cSSSI hospitalizations in the US (12, 13). The volume of pneumonia hospitalizations was extracted from the American Lung Association's 2010 data, and the proportion represented by MRSA from a large and representative database analysis by Kollef and colleagues (14, 15). Finally, we relied on the Agency's for Healthcare Research and Quality recent statistical brief quantifying the burden of hospitalizations with sepsis, while the Klein study provided the proportion likely caused by MRSA (11, 16).

Results

The input assumptions and their sources are presented in Table 1. The estimated prevalence of MRSA ranges from approximately 5% in sepsis to nearly 50% in cSSSI, while the prevalence of MRSA in pneumonia falls between those extremes. Under the conditions of lowest prevalence (5%) along with the average test specificity of 95%, the PPV reaches only 50% (Figure 1). Improving the specificity by nearly 5% to 99% without altering the disease prevalence results in a moderate improvement in the PPV to approximately 84%. Alternatively, a change of a similar magnitude in the PPV occurs, when the prevalence of disease increases from 5% to the 10%-20% range, even as the specificity remains anchored at 95% (Figure 1). The PPV further improves as the prevalence of disease approaches 50%. Notably, at the extremes of disease prevalence

and test specificity, the relative improvement in test accuracy is numerically greater when the prevalence is increased while holding the specificity constant (PPV 95.0% and NPV 95.2% for Tests A and B, prevalence 50%) as compared to a scenario where one modulates the test specificity and maintains the prevalence constant (PPV 83.3% and NPV 83.9% for Tests C and D, prevalence 5%). Put another way, the net change in PPV is maximized based on moderate changes in disease prevalence as opposed to alterations in test sensitivity. As for the NPV, a rise in sensitivity from 95% to 99% does not yield substantial alterations in the value. Essentially, the NPV is already quite high, no matter what the prevalence of resistance in the population. Conversely, the NPV suffers only modestly in the populations where disease prevalence is highest compared to those with the lowest disease prevalence (Figure 1).

The sensitivity analysis in which we assume that both the sensitivity and specificity of the test equal 90% is illustrated in Figure 2. At the lowest prevalence of disease, this specificity affords an unacceptably low PPV (32.1%), while the NPV remains high, exceeding 99%. As the prevalence of the disease rises in the target population, while the test's specificity and sensitivity remain fixed at 90%, the PPV and NPV converge at 90%, indicating a major improvement in the PPV without dramatically compromising the NPV.

Best- and worst-case scenario estimates of the total annual pool of patients at risk for MRSA infection in cSSSI, pneumonia and sepsis demonstrate that the potential for over-treatment far exceeds that for under-treatment (Table 2). Focusing on sepsis as an example, for the worst-case calculation we assumed 1,141,000 sepsis hospitalizations annually, a 5% MRSA prevalence, along with test characteristics of 95% sensitivity and

95% specificity. These parameters resulted in 57,050 potential cases of inappropriate treatment reflecting the sum of subjects classified as falsely positive or negative. Of these misclassified subjects, 54,198 (95%) represent those at risk for over-treatment. Conversely, under the best-case assumptions of a high MRSA prevalence (10%) in sepsis (n=727,000), and a test with near-perfect sensitivity and specificity (both 99%), only 7,270 individuals are at risk for inappropriate treatment with 6,543 (90%) being over-treated (Table 2).

Overall, under the worst-case assumptions for all three of the conditions of interest, over 150,000 patients annually with these three conditions may be treated inappropriately, with overtreatment accounting for 136,000 (91%) of this cohort. Under the best circumstances, among the more than 18,000 patients treated potentially inappropriately, nearly 15,000 (81%) may be subjected to over-treatment (Table 2).

Discussion

We have demonstrated explicitly that organism prevalence is an important driver of the accuracy of rapid molecular diagnostic tests even when their sensitivity and specificity are near perfect. Additionally, we have shown that although improving the theoretical test's specificity results in greater accuracy, one enhances accuracy even more by restricting test utilization to a population at an increased risk for infection with the pathogen in question. In other words, increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a marginally superior rapid diagnostic test with near-perfect specificity. In fact, given the already high NPV, the new molecular diagnostics have the potential to limit the use of empiric broad-spectrum coverage substantially. However, although

promising for limiting exposure to excessive antimicrobial coverage, molecular diagnostics are still likely to result in a substantial amount of inappropriate treatment. The vast majority (over 90%) of such inappropriate coverage is due to over-treatment in scenarios where the test is applied irrespective of considerations of the prevalence of a resistant pathogen.

Our data have several important implications. First, as manufacturers, regulators and clinicians consider what tests may have superior characteristics, it is important for all stakeholders to engage in defining the appropriate populations for testing with these novel technologies. Our data clearly demonstrate that rather than expending resources for every laboratory to elevate their sensitivity and specificity to close to 100%, the more fruitful effort may be to develop algorithms to identify those patient populations at high risk for the disease being tested. This is particularly true given that marginal enhancements in sensitivity and specificity often come at the cost of substantial financial investments. Second, raising the sensitivity of these technologies even beyond the current levels may be pursuing diminishing returns, given the already high NPV. That is, even when the sensitivity is no higher than 90%, the negative predictive value reaches very high levels (over 95%) in the setting of moderate pre-test probability for disease.

Third, and possibly most important, by using genotyping as opposed to phenotyping employed in the traditional microbiology laboratory methods, molecular diagnostics promise to result in sensitivity values that far exceed those of the traditional techniques. The flip side of this optimization in sensitivity is a blunting in specificity, whereby it may become unclear whether the identified organism is indeed the cause of

the clinical condition. Our data indicate that the true need in diagnostic testing lies not in further optimization of sensitivity, but in improving the specificity of the results.

Because improvements in one by necessity lead to detriments in the other, future directions in molecular diagnostics require thoughtful planning. We have clearly shown that, in order to live up to the promise of improved targeting of antibiotic treatment, such planning must include careful consideration of the populations in whom molecular diagnostic techniques are appropriate. In fact the most important lesson from our simulation is that we need to develop algorithms to help identify patients belonging to populations with a high risk for the particular pathogen. If a predictive algorithm is able to enrich the population to be tested to the disease prevalence between 30% and 40%, both PPV and NPV will be moved into a useful range even when the test's sensitivity and specificity are both well below 100%. With the advent of health informatics and the massive growth in computing ability, turning reams of patient data into predictive equations is a clearly needed functionality. Already several computing systems are addressing this need, and the trend should continue with the input from all stakeholders (18, 19).

Our study has a number of limitations. The most important limitation is that it is merely a mathematical model, and, as such, by necessity relies on the accuracy of estimates in the literature. The fact that some of the papers we used for deriving our assumptions themselves were modeling exercises (11), predisposes our computations to greater uncertainty. This, however, does not negate our findings that span a wide range of plausible epidemiology. Furthermore, our model underscores the need to understand local

pathogen patterns, the recognition of which should drive decisions about the utility of these powerful molecular diagnostics.

In summary, molecular diagnostics promise to streamline identification and treatment of many infectious diseases. While the emergence of these powerful technologies is a positive development, we need to attend to developing algorithms to aid in selecting appropriate patients for their use. Indiscriminate application of molecular diagnostics to all-comers presenting with signs of an infection without consideration for pre-test probability of disease is likely to result in a great deal of antimicrobial overtreatment. This will then only accelerate the current trajectory of escalating resistance. In conjunction with developing these important technologies, it is incumbent upon the healthcare community to build robust evidence and information technology infrastructure to guide appropriate use of such testing.

Acknowledgements. No one other than the listed authors participated in the study. Dr. Zilberberg had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Dr. Zilberberg participated in the conception and design of the study, data anaysisn and interpretation and drafting and revising of the manuscript for important intellectual content. Dr. Shorr participated in the conception of the study, data interpretation and revising of the manuscript for important intellectual content.

References

1. Dellit TH, Owens RC, McGowan JE, Jr., Gerding DN, Weinstein RA, Burke JP et al. Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. Clin Infect Dis 2007; 44(2):159-177

2. Spellberg B, Guidos R, Gilbert D, Bradley J, Boucher HW, Scheld WM et al. The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. Clin Infect Dis 2008; 46(2):155-164

3. Micek ST, Kollef KE, Reichley RM, Roubinian N, Kollef MH. Health care-associated pneumonia and community-acquired pneumonia: a single-center experience. Antimicrob Agents Chemother 2007; 51(10):3568-3573

4. Schramm GE, Johnson JA, Doherty JA, Micek ST, Kollef MH. Methicillin-resistant Staphylococcus aureus sterile-site infection: The importance of appropriate initial antimicrobial treatment. Crit Care Med 2006; 34(8):2069-2074.

5. Ibrahim EH, Sherman G, Ward S, Fraser VJ, Kollef MH. The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting. Chest 2000; 118(1):146-155.

6. Alvarez-Lerma F. Modification of empiric antibiotic treatment in patients with pneumonia acquired in the intensive care unit. ICU-Acquired Pneumonia Study Group. Intensive Care Med 1996; 22(5):387-394.

7. Iregui M, Ward S, Sherman G, Fraser VJ, Kollef MH. Clinical importance of delays in the initiation of appropriate antibiotic treatment for ventilator-associated pneumonia. Chest 2002; 122(1):262-268.

8. Zilberberg MD, Shorr AF, Micek ST, Mody SH, Kollef MH. Antimicrobial therapy escalation and hospital mortality among patients with health-care-associated pneumonia: a single-center experience. Chest 2008; 134(5):963-968.

9. Zilberberg MD, Shorr AF, Micek ST, Chen J, Ramsey AM, Hoban AP et al. Hospitalizations with healthcare-associated complicated skin and skin structure infections: impact of inappropriate empiric therapy on outcomes. J Hosp Med 2010; 5(9):535-540.

10. Peterson LR, Liesenfeld O, Woods CW et al. Multicenter evaluation of the LightCycler methicillin-resistant *Staphylococcus aureus* (MRSA) advanced test as a rapid method for detection of MRSA in nasal surveillance swabs. J Clin Microbiol 2010;48:1661-6

11. Klein E, Smith DL, Laxminarayan R. Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus*, United States, 1999–2005. Emerg Infect Dis 2007;13:1840-6

12. Lipsky BA, Weigelt JA, Gupta V, Killian A, Peng MM. Skin, soft tissue, bone, and joint infections in hospitalized patients: epidemiology and microbiological, clinical, and economic outcomes. Infect Control Hosp Epidemiol 2007;28:1290-8

13. Jenkins TC, Sabel AL, Sarcone EE, et al. Skin and soft tissue infections requiring hospitalization at an academic medical center: Opportunities for antimicrobial stewardship. Clin Infect Dis 2010;51:895-903

14. Kollef MH, Shorr A, Tabak YP, et al. Epidemiology and outcomes of health-care-associated pneumonia: Results from a large US database of culture-positive pneumonia. Chest 2005;128:3854-3862

15. American Lung Association. Trends in Pneumonia and Influenza Morbidity and Mortality, 2010. Available at http://www.lungusa.org/finding-cures/our-research/trend-reports/pi-trend-report.pdf, accessed January 17, 2012

16. Hall MJ, Williams SN, DeFrances CJ, Golosinskiy A. Inpatient care for septicemia or sepsis: A challenge for patients and hospitals. NCHS data brief, no 62. Hyattsville, MD: National Center for Health Statistics. 2011. Available at http://www.cdc.gov/nchs/data/databriefs/db62.pdf, accessed February 1, 2012

17. BD molecular diagnostic solutions. BD GeneOhm Staph SR Assay package insert. Available at http://www.bd.com/geneohm/english/products/pdfs/staphsr_pkginsert.pdf, accessed February 15, 2012

18. Pepitone J. IBM's 'Jeopardy' computer lands health care job. CNNMoney 2011 Sep 12, available at http://money.cnn.com/2011/09/12/technology/ibm_watson_health_care/index.htm , accessed February 15, 2012

19. Terry K. Futuristic Clinical Decision Support Tool Catches On. InformationWeek Healthcare 2012 Jan 27, available at http://informationweek.com/news/healthcare/clinical-systems/232500603, accessed February 15, 2012

Table 1. Annual hospitalization volumes

Infection type	Annual volume	MRSA prevalence (%)	MRSA volume
cSSSI	434,227-1,211,863	15.3%[12]-42.7[13]	185,415[11]
Pneumonia	651,000[15]	5.6%-14.3%[14]	36,540[11]-93,093
Sepsis	727,000-1,141,000[16]	4.9%-7.7%	56,246[11]

MRSA=methicillin-resistant *Staphylococcus aureus*; cSSSI=complicated skin and skin structure infection

Table 2. Best- and worst-case scenario simulations for each disease group

	Over-treated	Under-treated	Treated inappropriately
Best-case scenario			
cSSSI	2,600	1,733	4,333
Pneumonia	5,534	977	6,510
Sepsis	6,543	727	7,270
Total	**14,676**	**3,437**	**18,113**
Worst-case scenario			
cSSSI	51,389	9,069	60,458
Pneumonia	30,923	1,628	32,550
Sepsis	54,198	2,853	57,050
Total	**136,509**	**13,549**	**150,058**

cSSSI=complicated skin and skin structure infection

Figure 1. Positive and negative predictive values of a test with the given sensitivity and specificity, stratified by population disease prevalence*

*Percentages along the X-axis represent disease prevalence strata
PPV = positive predictive value; NPV = negative predictive value
Test A: sensitivity = 95%, specificity = 95%; Test B: sensitivity = 99%, specificity = 95%;
Test C: sensitivity = 95%, specificity = 99%; Test D: sensitivity = 99%, specificity = 99%

Figure 2. Sensitivity analysis under the conditions of test sensitivity and specificity equaling 90%

PPV = positive predictive value; NPV = negative predictive value

~~Molecular diagnostics require a Bayesian approach~~Impact of prior probabilities of MRSA as an infectious agent on the accuracy of the emerging molecular diagnostic tests: A model simulation

Marya D. Zilberberg, MD, MPH*[1,2],
[1]EviMed Research Group, LLC, Goshen, MA
[2]University of Massachusetts, Amherst, MA
*Corresponding author
President & CEO, Evi*Med* Research Group, LLC
Adjunct Associate Professor of Epidemiology, University of Massachusetts School of Public Health and Health Sciences
PO Box 303
Goshen, MA 01032
(413)-268-6381
(413)-268-3416 fax
evimedgroup@gmail.com

Andrew F. Shorr, MD, MPH[3]
[3]Washington Hospital Center, Washington, DC
Associate Section Director, Pulmonary Critical Care, Washington Hospital Center
Associate Professor of Medicine, Georgetown University
110 Irving Street, NW
Washington, DC 20010
(202)-877-2998
(202)-782-9032 fax
Andrew.Shorr@gmail.com

Competing interests: The authors have nothing to disclose. The authors received no funding for this study. All authors have completed the Unified Competing Interest form.
Guarantor statement: Dr. Zilberberg has had access to all the data and guarantees its integrity.
IRB statement: The study is a model and did not require an Ethics Committee review
Author contributions:
Dr. Zilberberg conceived, designed and executed the study, data collection and analyses, drafted the manuscript and revised it for important intellectual content.
Dr. Shorr refined the study idea, design and interpretation, and revised the manuscript for important intellectual content.

Word count: 2,690~~562~~

References: 19
Tables and Figures: 4

Abstract

Objectives. Traditional microbiology identification takes 48-72 hours to complete. This lag forces clinicians to rely on broad-spectrum empiric coverage. To address this gap, manufacturers are developing rapid molecular diagnostics (RMD). We hypothesized that RMD's accuracy is more dependent upon population risk of harboring the culprit pathogen than to their sensitivity and specificity.

Design. A mathematical model

Setting and Participants. We used the range of risks (5%-50%) for methicillin-resistant *Staphylococcus aureus* (MRSA) among patients hospitalized with complicated skin and skin structure infections (cSSSI), pneumonia, or sepsis.

~~**Interventions.** None~~

Main outcome measures. We modeled the impact of changing a test's characteristics on its positive (PPV) and negative (NPV) predictive values, and hence the risk of over- or under-treatment, within strata of an organism's population prevalence. MRSA diagnostics provided assumptions for the test sensitivity and specificity (95%-99%). Scenarios with low sensitivity and specificity (90%), and best- and worst-case scenarios normalized to the annual universe of populations of interest, were examined.

Results. With a low prevalence (5%) and high test specificity, the PPV was 84%. Conversely, with 50% prevalence and 95% test specificity the PPV rose to ≥95%. Even when the test's specificity and sensitivity were both 90%, in a high-risk population both PPV and NPV were ~90%. In the worst-case scenario, 150,000 patients with cSSSI, pneumonia and sepsis annually were at risk for inappropriate treatment, 91% of these at risk for over-treatment. In the best-case scenario, 81% of 18,000 patients at risk for inappropriate coverage were subject to over-treatment.

Conclusions. Although promising for limiting exposure to excessive antimicrobial coverage, RMDs alone will not solve the issue of inappropriate, and particularly over-, treatment. Increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a test with near-perfect characteristics. The healthcare community must build robust evidence and information technology infrastructure to guide appropriate use of such testing.

Article summary

Article focus

- Traditional microbiology identification takes 48-72 hours to complete, which forces clinicians to rely on broad-spectrum empiric coverage.
- To address this gap, manufacturers are developing rapid molecular diagnostics (RMD).
- It is unclear what impact RMDs may have in different population on overdiagnosis and overtreatment

Key messages

- Although promising for limiting exposure to excessive antimicrobial coverage, RMDs alone will not solve the issue of inappropriate, and particularly over-, treatment.
- Increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a test with near-perfect sensitivity and specificity.

Strengths and limitations

- As a mathematical model, our study relies on the accuracy of estimates in the literature, which predisposes our computations to greater uncertainty.
- The model is transparent.
- Our findings span a wide range of plausible epidemiology.
- The data underscore the need to understand local pathogen patterns, the recognition of which should drive decisions about the utility of these powerful molecular diagnostics.

Introduction

Despite the fact that antibiotics represent a relatively recent advance in medicine, antibiotic resistant bacteria are now common in both the hospital and the community. Antibiotic misuse and abuse represent a key driver of the increasing prevalence in antibiotic resistance (1, 2). The spread of antimicrobial resistance has similarly created a vicious cycle where clinicians repeatedly reach for extended spectrum agents in order to address the current patterns of resistance while potentially worsening them for the future. Underlying this practice approach has been the general unavailability of reliable, rapid diagnostics to help establish the etiology of an infection. Indeed, traditional phenotypic microbiology methods take 48 to 72 hours to identify an organism when present and to determine the antibiotic susceptibility profile. Without a prompt means for either including or excluding potentially resistant pathogens, clinicians frequently have no alternative but to rely on broad-spectrum options for empiric therapy. Such approach is currently warranted, given the extensive data documenting that delayed and inappropriate antibiotic treatment increases the risk for mortality and prolongs the duration of hospitalization (3-9). However, rapid and accurate diagnosis should diminish the uncertainty and help target the culprit organisms without straying into the extremes of overly narrow or overly broad coverage.

To fill this diagnostic gap, several manufacturers are engaged in developing rapid diagnostic modalities that incorporate recent advances in molecular techniques relying on genotyping the organisms. Indeed, some of these technologies are able to arrive at the microbiologic diagnosis in as little as 2 hours, a critical period for tailoring treatment

(10). Such improvement in shortening the diagnostic time is invaluable, particularly given these tests' ostensible accuracy.

At the same time, one must exercise caution because these tests are not 100% accurate. And while manufacturers strive for ever-increasing sensitivity and specificity for their tests, a more fruitful area of investigation may be learning to identify characteristics of specific populations in whom these tests may prove to be most helpful for targeting and tailoring treatment. In other words, the central clinical question may revolve not around issues of sensitivity and specificity intrinsic to the test, but rather around the positive (PPV) and negative (NPV) predictive values associated with these newer tools in populations with various levels of risk for the organisms in question. This approach fits in with the Bayesian decision making, whereby the prior probability of an event informs the interpretation of the diagnostic data. Irrespective of the sensitivity and specificity, if the PPV and NPV are not sufficiently high, then these new tests may not help clinicians either to withhold unnecessarily broad coverage or to tailor it shortly after the results return.

We hypothesized that even under conditions where such rapid diagnostic tests had near-perfect sensitivity and specificity, the population-specific risk for having a particular organism would represent a crucial consideration in driving diagnostic accuracy. That is failure to consider the pre-test probability of these organisms in the population screened would undermine the potential value of rapid diagnostic tests. To address this question we developed a model simulation evaluating the application of these assays, and relied upon publicly available data to populate our analysis.

Methods

We developed a mathematical model simulating the impact of changing a test's characteristics on its accuracy within several strata of population risk for a particular organism. All the inputs were extracted from publicly available data. The primary outcome of interest was the potential magnitude for over-diagnosis of a particular pathogen, or the proportion of false positive tests under the varying assumptions. We were specifically interested in the false positive rates, since these cases are the ones most likely to receive overly broad treatment when it is not indicated. Such overly broad treatment represents a key clinical endpoint since it exposes the patient and the healthcare system to adverse consequences individually and as a group. As a secondary endpoint we examined the overall inaccuracy of the test in various scenarios, defined as the sum of the false positive and false negative results as a proportion of the total population.

The model was based on the approximate risks of methicillin-resistant *Staphylococcus aureus* (MRSA) among three distinct hospitalized populations: 1) complicated skin and skin structure infections (cSSSI) (11-13), 2) pneumonia (11, 14), and 3) sepsis (11). We sought the most generalizable estimates for at least two factors out of the following three, using the available data to calculate the third when necessary: 1) total volume of hospitalizations for each ~~of the~~ disease~~s~~ of interest, 2) proportion of the total volume represented by MRSA, and 3) total number of MRSA infections in each disease category (11, 14, 15).

For consistency, the assumptions for the corresponding test characteristics mimicked those from MRSA diagnostics (16). To derive estimates for positive (PPV) and negative predictive values (NPV) for a plausible range of test characteristics, we developed four hypothetical testing situations: 1) Test A, with the sensitivity and

specificity of 95%, 2) Test B, with the sensitivity 99% and specificity 95%, 3) Test C, with the sensitivity 95% and specificity 99%, and 4) Test D, with the sensitivity and specificity 99%. To explore how deviations from the average sensitivity and specificity metrics may impact the accuracy of identification, we conducted sensitivity analyses assuming 90% sensitivity and specificity. Based on the range of MRSA risk estimates in the populations of interest (i.e., cSSSI, pneumonia, and sepsis), we varied the prevalence estimates from 5% to 50%, and calculated the PPV and NPV for each of the intermediate values.

We additionally performed best- and worst-case scenario simulations for each ~~of the~~ population~~s~~ in question. Thus, for the worst-case scenario where all variables were biased against the novel rapid diagnostic assay, we utilized as inputs the highest disease volume and lowest disease prevalence, along with the lowest test sensitivity and specificity values. Skewing the inputs in this fashion provides a potential estimate of the extent and impact of misclassification when all assumptions are shifted so as to constrain the potential value of the rapid diagnostic test in question. Conversely, for the best-case scenario, we input the lowest disease volume and the highest disease prevalence, along with the highest test sensitivity and test specificity. For both of these analyses, the total annual universe of specific disease hospitalizations in the US was used. We utilized these values to estimate the total numbers of potential cases within each population that would be over-treated (i.e., treated for MRSA when no MRSA is present), under-treated (i.e., not treated for MRSA when MRSA is present) and treated inappropriately (i.e., either over- or undertreated).

Both the values for sensitivity and specificity and disease risk were rounded in order to ease computational presentation. Volumes and prevalence of MRSA in the disease states of interest were extracted from several large surveys available in the public domain (Table 1). Thus, for the MRSA cSSSI volumes we relied on a study by Klein, which quantified these hospitalizations in 2005 (11). The proportions of cSSSI in which MRSA is the offending pathogen derived from two recent epidemiologic studies of cSSSI hospitalizations in the US (12, 13). The volume of pneumonia hospitalizations was extracted from the American Lung Association's 2010 data, and the proportion represented by MRSA from a large and representative database analysis by Kollef and colleagues (14, 15). Finally, we relied on the Agency's for Healthcare Research and Quality recent statistical brief quantifying the burden of hospitalizations with sepsis, while the Klein study provided the proportion likely caused by MRSA (11, 16).

Results

The input assumptions and their sources are presented in Table 1. The estimated prevalence of MRSA ranges from approximately 5% in sepsis to nearly 50% in cSSSI, while the prevalence of MRSA in pneumonia falls between those extremes. Under the conditions of lowest prevalence (5%) along with the average test specificity of 95%, the PPV reaches only 50% (Figure 1). Improving the specificity by nearly 5% to 99% without altering the disease prevalence results in a moderate improvement in the PPV to approximately 84%. Alternatively, a change of a similar magnitude in the PPV occurs, when the prevalence of disease increases from 5% to the 10%-20% range, even as the specificity remains anchored at 95% (Figure 1). The PPV further improves as the prevalence of disease approaches 50%. Notably, at the extremes of disease prevalence

and test specificity, the relative improvement in test accuracy is numerically greater when the prevalence is increased while holding the specificity constant (PPV 95.0% and NPV 95.2% for Tests A and B, prevalence 50%) as compared to a scenario where one modulates the test specificity and maintains the prevalence constant (PPV 83.3% and NPV 83.9% for Tests C and D, prevalence 5%). Put another way, the net change in PPV is maximized based on moderate changes in disease prevalence as opposed to alterations in test sensitivity. As for the NPV, a rise in sensitivity from 95% to 99% does not yield substantial alterations in the value. Essentially, the NPV is already quite high, no matter what the prevalence of resistance in the population. Conversely, the NPV suffers only modestly in the populations where disease prevalence is highest compared to those with the lowest disease prevalence (Figure 1).

The sensitivity analysis in which we assume that both the sensitivity and specificity of the test equal 90% is illustrated in Figure 2. At the lowest prevalence of disease, this specificity affords an unacceptably low PPV (32.1%), while the NPV remains high, exceeding 99%. As the prevalence of the disease rises in the target population, while the test's specificity and sensitivity remain fixed at 90%, the PPV and NPV converge at 90%, indicating a major improvement in the PPV without dramatically compromising the NPV.

Best- and worst-case scenario estimates of the total annual pool of patients at risk for MRSA infection in cSSSI, pneumonia and sepsis demonstrate that the potential for over-treatment far exceeds that for under-treatment (Table 2). Focusing on sepsis as an example, for the worst-case calculation we assumed 1,141,000 sepsis hospitalizations annually, a 5% MRSA prevalence, along with test characteristics of 95% sensitivity and

95% specificity. These parameters resulted in 57,050 potential cases of inappropriate treatment reflecting the sum of subjects classified as falsely positive or negative. Of these misclassified subjects, 54,198 (95%) represent those at risk for over-treatment. Conversely, under the best-case assumptions of a high MRSA prevalence (10%) in sepsis (n=727,000), and a test with near-perfect sensitivity and specificity (both 99%), only 7,270 individuals are at risk for inappropriate treatment with 6,543 (90%) being over-treated (Table 2).

Overall, under the worst-case assumptions for all three of the conditions of interest, over 150,000 patients annually with these three conditions may be treated inappropriately, with overtreatment accounting for 136,000 (91%) of this cohort. Under the best circumstances, among the more than 18,000 patients treated potentially inappropriately, nearly 15,000 (81%) may be subjected to over-treatment (Table 2).

Discussion

We have demonstrated explicitly that organism prevalence is an important driver of the accuracy of rapid molecular diagnostic tests even when their sensitivity and specificity are near perfect. Additionally, we have shown that although improving the theoretical test's specificity results in greater accuracy, one enhances accuracy even more by restricting test utilization to a population at an increased risk for infection with the pathogen in question. In other words, increasing pre-test probability as a strategy to minimize antibiotic abuse results in more accurate patient classification than does developing a marginally superior rapid diagnostic test with near-perfect specificity. In fact, given the already high NPV, the new molecular diagnostics have the potential to limit the use of empiric broad-spectrum coverage substantially. ~~Finally~~However,

although promising for limiting exposure to excessive antimicrobial coverage, molecular diagnostics are still likely to result in a substantial amount of inappropriate treatment. The vast majority (over 90%) of such inappropriate coverage is due to over-treatment in scenarios where the test is applied irrespective of considerations of the prevalence of a resistant pathogen.

Our data have several important implications. First, as manufacturers, regulators and clinicians consider what tests may have superior characteristics, it is important for all stake-holders to engage in defining the appropriate populations for testing with these novel technologies. Our data clearly demonstrate that rather than expending resources for every laboratory to elevate their sensitivity and specificity to close to 100%, the more fruitful effort may be to develop algorithms to identify those patient populations at high risk for the disease being tested. This is particularly true given that marginal enhancements in sensitivity and specificity often come at the cost of substantial financial investments. Second, raising the sensitivity of these technologies even beyond the current levels may be pursuing diminishing returns, given the already high NPV. That is, even when the sensitivity is no higher than 90%, the negative predictive value reaches very high levels (over 95%) in the setting of moderate pre-test probability for disease.

Third, and possibly most important, by using genotyping as opposed to phenotyping employed in the traditional microbiology laboratory methods, molecular diagnostics promise to result in sensitivity values that far exceed those of the traditional techniques. The flip side of this optimization in sensitivity is a blunting in specificity, whereby it may become unclear whether the identified organism is indeed the cause of

the clinical condition. Our data indicate that the true need in diagnostic testing lies not in further optimization of sensitivity, but in improving the specificity of the results.

Because improvements in one by necessity lead to detriments in the other, future directions in molecular diagnostics require thoughtful planning. We have clearly shown that, in order to live up to the promise of improved targeting of antibiotic treatment, such planning must include careful consideration of the populations in whom molecular diagnostic techniques are appropriate. In fact the most important lesson from our simulation is that we need to develop algorithms to help identify patients belonging to populations with a high risk for the particular pathogen. If a predictive algorithm is able to enrich the population to be tested to the disease prevalence between 30% and 40%, both PPV and NPV will be moved into a useful range even when the test's sensitivity and specificity are both well below 100%. With the advent of health informatics and the massive growth in computing ability, turning reams of patient data into predictive equations is a clearly needed functionality. Already several computing systems are addressing this need, and the trend should continue with the input from all stakeholders (18, 19).

Our study has a number of limitations. The most important limitation is that it is merely a mathematical model, and, as such, by necessity relies on the accuracy of estimates in the literature. The fact that some of the papers we used for deriving our assumptions themselves were modeling exercises (11), predisposes our computations to greater uncertainty. This, however, does not negate our findings that span a wide range of plausible epidemiology. Furthermore, our model underscores the need to understand local

pathogen patterns, the recognition of which should drive decisions about the utility of these powerful molecular diagnostics.

In summary, molecular diagnostics promise to streamline identification and treatment of many infectious diseases. While the emergence of these powerful technologies is a positive development, we need to attend to developing algorithms to aid in selecting appropriate patients for their use. Indiscriminate application of molecular diagnostics to all-comers presenting with signs of an infection without consideration for pre-test probability of disease is likely to result in a great deal of antimicrobial overtreatment. This will then only accelerate the current trajectory of escalating resistance. In conjunction with developing these important technologies, it is incumbent upon the healthcare community to build robust evidence and information technology infrastructure to guide appropriate use of such testing.

Acknowledgements. No one other than the listed authors participated in the study. Dr. Zilberberg had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Dr. Zilberberg participated in the conception and design of the study, data anaysisn and interpretation and drafting and revising of the manuscript for important intellectual content. Dr. Shorr participated in the conception of the study, data interpretation and revising of the manuscript for important intellectual content.

References

1. Dellit TH, Owens RC, McGowan JE, Jr., Gerding DN, Weinstein RA, Burke JP et al. Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. Clin Infect Dis 2007; 44(2):159-177

2. Spellberg B, Guidos R, Gilbert D, Bradley J, Boucher HW, Scheld WM et al. The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. Clin Infect Dis 2008; 46(2):155-164

3. Micek ST, Kollef KE, Reichley RM, Roubinian N, Kollef MH. Health care-associated pneumonia and community-acquired pneumonia: a single-center experience. Antimicrob Agents Chemother 2007; 51(10):3568-3573

4. Schramm GE, Johnson JA, Doherty JA, Micek ST, Kollef MH. Methicillin-resistant Staphylococcus aureus sterile-site infection: The importance of appropriate initial antimicrobial treatment. Crit Care Med 2006; 34(8):2069-2074.

5. Ibrahim EH, Sherman G, Ward S, Fraser VJ, Kollef MH. The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting. Chest 2000; 118(1):146-155.

6. Alvarez-Lerma F. Modification of empiric antibiotic treatment in patients with pneumonia acquired in the intensive care unit. ICU-Acquired Pneumonia Study Group. Intensive Care Med 1996; 22(5):387-394.

7. Iregui M, Ward S, Sherman G, Fraser VJ, Kollef MH. Clinical importance of delays in the initiation of appropriate antibiotic treatment for ventilator-associated pneumonia. Chest 2002; 122(1):262-268.

8. Zilberberg MD, Shorr AF, Micek ST, Mody SH, Kollef MH. Antimicrobial therapy escalation and hospital mortality among patients with health-care-associated pneumonia: a single-center experience. Chest 2008; 134(5):963-968.

9. Zilberberg MD, Shorr AF, Micek ST, Chen J, Ramsey AM, Hoban AP et al. Hospitalizations with healthcare-associated complicated skin and skin structure infections: impact of inappropriate empiric therapy on outcomes. J Hosp Med 2010; 5(9):535-540.

10. Peterson LR, Liesenfeld O, Woods CW et al. Multicenter evaluation of the LightCycler methicillin-resistant *Staphylococcus aureus* (MRSA) advanced test as a rapid method for detection of MRSA in nasal surveillance swabs. J Clin Microbiol 2010;48:1661-6

11. Klein E, Smith DL, Laxminarayan R. Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus*, United States, 1999–2005. Emerg Infect Dis 2007;13:1840-6

12. Lipsky BA, Weigelt JA, Gupta V, Killian A, Peng MM. Skin, soft tissue, bone, and joint infections in hospitalized patients: epidemiology and microbiological, clinical, and economic outcomes. Infect Control Hosp Epidemiol 2007;28:1290-8

13. Jenkins TC, Sabel AL, Sarcone EE, et al. Skin and soft tissue infections requiring hospitalization at an academic medical center: Opportunities for antimicrobial stewardship. Clin Infect Dis 2010;51:895-903

14. Kollef MH, Shorr A, Tabak YP, et al. Epidemiology and outcomes of health-care-associated pneumonia: Results from a large US database of culture-positive pneumonia. Chest 2005;128:3854-3862

15. American Lung Association. Trends in Pneumonia and Influenza Morbidity and Mortality, 2010. Available at http://www.lungusa.org/finding-cures/our-research/trend-reports/pi-trend-report.pdf, accessed January 17, 2012

16. Hall MJ, Williams SN, DeFrances CJ, Golosinskiy A. Inpatient care for septicemia or sepsis: A challenge for patients and hospitals. NCHS data brief, no 62. Hyattsville, MD: National Center for Health Statistics. 2011. Available at http://www.cdc.gov/nchs/data/databriefs/db62.pdf, accessed February 1, 2012

17. BD molecular diagnostic solutions. BD GeneOhm Staph SR Assay package insert. Available at http://www.bd.com/geneohm/english/products/pdfs/staphsr_pkginsert.pdf, accessed February 15, 2012

18. Pepitone J. IBM's 'Jeopardy' computer lands health care job. CNNMoney 2011 Sep 12, available at http://money.cnn.com/2011/09/12/technology/ibm_watson_health_care/index.htm , accessed February 15, 2012

19. Terry K. Futuristic Clinical Decision Support Tool Catches On. InformationWeek Healthcare 2012 Jan 27, available at http://informationweek.com/news/healthcare/clinical-systems/232500603, accessed February 15, 2012

Table 1. Annual hospitalization volumes

Infection type	Annual volume	MRSA prevalence (%)	MRSA volume
cSSSI	434,227-1,211,863	15.3%[12]-42.7[13]	185,415[11]
Pneumonia	651,000[15]	5.6%-14.3%[14]	36,540[11]-93,093
Sepsis	727,000-1,141,000[16]	4.9%-7.7%	56,246[11]

MRSA=methicillin-resistant *Staphylococcus aureus*; cSSSI=complicated skin and skin structure infection

Table 2. Best- and worst-case scenario simulations for each disease group

	Over-treated	Under-treated	Treated inappropriately
Best-case scenario			
cSSSI	2,600	1,733	4,333
Pneumonia	5,534	977	6,510
Sepsis	6,543	727	7,270
Total	**14,676**	**3,437**	**18,113**
Worst-case scenario			
cSSSI	51,389	9,069	60,458
Pneumonia	30,923	1,628	32,550
Sepsis	54,198	2,853	57,050
Total	**136,509**	**13,549**	**150,058**

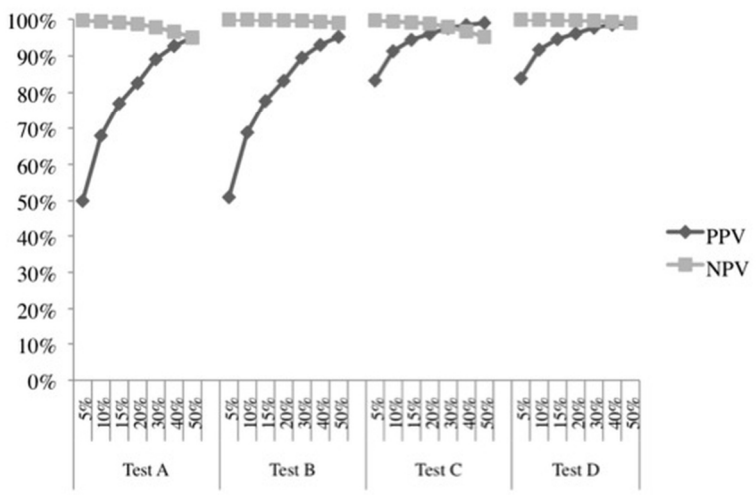
cSSSI=complicated skin and skin structure infection

Figure 1. Positive and negative predictive values of a test with the given sensitivity and specificity, stratified by population disease prevalence*

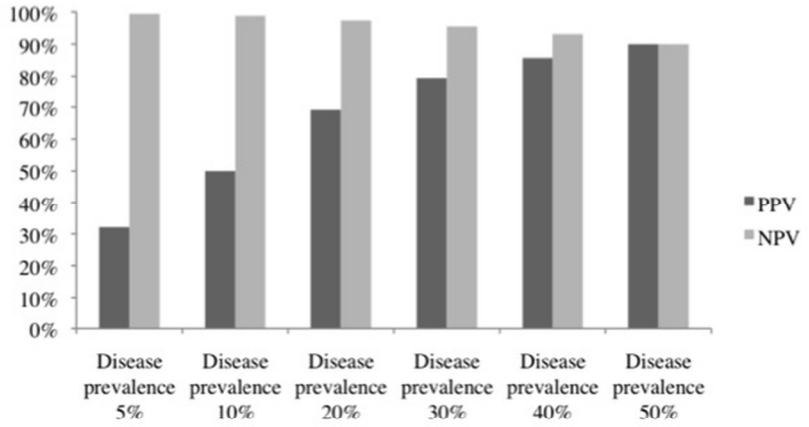
*Percentages along the X-axis represent disease prevalence strata
PPV = positive predictive value; NPV = negative predictive value
Test A: sensitivity = 95%, specificity = 95%; Test B: sensitivity = 99%, specificity = 95%;
Test C: sensitivity = 95%, specificity = 99%; Test D: sensitivity = 99%, specificity = 99%

Figure 2. Sensitivity analysis under the conditions of test sensitivity and specificity equaling 90%

PPV = positive predictive value; NPV = negative predictive value



119x90mm (300 x 300 DPI)



119x90mm (300 x 300 DPI)