



Molecular diagnostics require a Bayesian approach

Journal:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts

Molecular diagnostics require a Bayesian approach

Marya D. Zilberberg, MD, MPH^{*1,2},

¹EviMed Research Group, LLC, Goshen, MA

²University of Massachusetts, Amherst, MA

*Corresponding author

President & CEO, EviMed 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, MPH³

³Washington 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

Copyright © 2012 Marya Zilberberg

Dr. Zilberberg, the Corresponding Author, has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJ PGL products and sublicenses such use and exploit all subsidiary rights, as set out in our license.

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 $\geq 95\%$. Even when the test's specificity and sensitivity were both 90%, in a high-risk population both PPV and NPV were $\sim 90\%$. 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

1
2
3 (10). Such improvement in shortening the diagnostic time is invaluable, particularly
4
5 given these tests' ostensible accuracy.
6
7

8 At the same time, one must exercise caution because these tests are not 100%
9 accurate. And while manufacturers strive for ever-increasing sensitivity and specificity
10 for their tests, a more fruitful area of investigation may be learning to identify
11 characteristics of specific populations in whom these tests may prove to be most helpful
12 for targeting and tailoring treatment. In other words, the central clinical question may
13 revolve not around issues of sensitivity and specificity intrinsic to the test, but rather
14 around the positive (PPV) and negative (NPV) predictive values associated with these
15 newer tools in populations with various levels of risk for the organisms in question.
16
17 Irrespective of the sensitivity and specificity, if the PPV and NPV are not sufficiently
18 high, then these new tests may not help clinicians either to withhold unnecessarily broad
19 coverage or to tailor it shortly after the results return.
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34 We hypothesized that even under conditions where such rapid diagnostic tests had
35 near-perfect sensitivity and specificity, the population-specific risk for having a particular
36 organism would represent a crucial consideration in driving diagnostic accuracy. That is
37 failure to consider the pre-test probability of these organisms in the population screened
38 would undermine the potential value of rapid diagnostic tests. To address this question
39 we developed a model simulation evaluating the application of these assays, and relied
40 upon publicly available data to populate our analysis.
41
42
43
44
45
46
47
48
49

50 **Methods**

51
52
53 We developed a mathematical model simulating the impact of changing a test's
54 characteristics on its accuracy within several strata of population risk for a particular
55
56
57
58
59
60

1
2
3 organism. All the inputs were extracted from publicly available data. The primary
4
5 outcome of interest was the potential magnitude for over-diagnosis of a particular
6
7 pathogen, or the proportion of false positive tests under the varying assumptions. We
8
9 were specifically interested in the false positive rates, since these cases are the ones most
10
11 likely to receive overly broad treatment when it is not indicated. Such overly broad
12
13 treatment represents a key clinical endpoint since it exposes the patient and the healthcare
14
15 system to adverse consequences individually and as a group. As a secondary endpoint we
16
17 examined the overall inaccuracy of the test in various scenarios, defined as the sum of the
18
19 false positive and false negative results as a proportion of the total population.
20
21
22
23

24
25 The model was based on the approximate risks of methicillin-resistant
26
27 *Staphylococcus aureus* (MRSA) among three distinct hospitalized populations: 1)
28
29 complicated skin and skin structure infections (cSSSI) (11-13), 2) pneumonia (11, 14),
30
31 and 3) sepsis (11). We sought the most generalizable estimates for at least two factors out
32
33 of the following three, using the available data to calculate the third when necessary: 1)
34
35 total volume of hospitalizations for each of the diseases of interest, 2) proportion of the
36
37 total volume represented by MRSA, and 3) total number of MRSA infections in each
38
39 disease category (11, 14, 15).
40
41
42

43
44 For consistency, the assumptions for the corresponding test characteristics
45
46 mimicked those from MRSA diagnostics (16). To derive estimates for positive (PPV) and
47
48 negative predictive value (NPV) for a plausible range of test characteristics, we
49
50 developed four hypothetical testing situations: 1) Test A, with the sensitivity and
51
52 specificity of 95%, 2) Test B, with the sensitivity 99% and specificity 95%, 3) Test C,
53
54 with the sensitivity 95% and specificity 99%, and 4) Test D, with the sensitivity and
55
56
57
58
59
60

1
2
3 specificity 99%. To explore how deviations from the average sensitivity and specificity
4 metrics may impact the accuracy of identification, we conducted sensitivity analyses
5
6 assuming 90% sensitivity and specificity. Based on the range of MRSA risk estimates in
7
8 the populations of interest (i.e., cSSSI, pneumonia, and sepsis), we varied the prevalence
9
10 estimates from 5% to 50%, and calculated the PPV and NPV for each of the intermediate
11
12 values.
13
14
15

16
17 We additionally performed best- and worst-case scenario simulations for each of
18
19 the populations in question. Thus, for the worst-case scenario where all variables were
20
21 biased against the novel rapid diagnostic assay, we utilized as inputs the highest disease
22
23 volume and lowest disease prevalence, along with the lowest test sensitivity and
24
25 specificity values. Skewing the inputs in this fashion provides a potential estimate of the
26
27 extent and impact of misclassification when all assumptions are shifted so as to constrain
28
29 the potential value of the rapid diagnostic test in question. Conversely, for the best-case
30
31 scenario, we input the lowest disease volume and the highest disease prevalence, along
32
33 with the highest test sensitivity and test specificity. For both of these analyses, the total
34
35 annual universe of specific disease hospitalizations in the US was used. We utilized these
36
37 values to estimate the total numbers of potential cases within each population that would
38
39 be over-treated (i.e., treated for MRSA when no MRSA is present), under-treated (i.e.,
40
41 not treated for MRSA when MRSA is present) and treated inappropriately (i.e., either
42
43 over- or undertreated).
44
45
46
47
48
49

50 Both the values for sensitivity and specificity and disease risk were rounded in
51
52 order to ease computational presentation. Volumes and prevalence of MRSA in the
53
54 disease states of interest were extracted from several large surveys available in the public
55
56
57
58
59
60

1
2
3 domain (Table 1). Thus, for the MRSA cSSSI volumes we relied on a study by Klein,
4 which quantified these hospitalizations in 2005 (11). The proportions of cSSSI in which
5 MRSA is the offending pathogen derived from two recent epidemiologic studies of cSSSI
6 hospitalizations in the US (12, 13). The volume of pneumonia hospitalizations was
7 extracted from the American Lung Association's 2010 data, and the proportion
8 represented by MRSA from a large and representative database analysis by Kollef and
9 colleagues (14, 15). Finally, we relied on the Agency's for Healthcare Research and
10 Quality recent statistical brief quantifying the burden of hospitalizations with sepsis,
11 while the Klein study provided the proportion likely caused by MRSA (11, 16).
12
13
14
15
16
17
18
19
20
21
22
23
24

25 Results

26
27 The input assumptions and their sources are presented in Table 1. The estimated
28 prevalence of MRSA ranges from approximately 5% in sepsis to nearly 50% in cSSSI,
29 while the prevalence of MRSA in pneumonia falls between those extremes. Under the
30 conditions of lowest prevalence (5%) along with the average test specificity of 95%, the
31 PPV reaches only 50% (Figure 1). Improving the specificity by nearly 5% to 99%
32 without altering the disease prevalence results in a moderate improvement in the PPV to
33 approximately 84%. Alternatively, a change of a similar magnitude in the PPV occurs,
34 when the prevalence of disease increases from 5% to the 10%-20% range, even as the
35 specificity remains anchored at 95% (Figure 1). The PPV further improves as the
36 prevalence of disease approaches 50%. Notably, at the extremes of disease prevalence
37 and test specificity, the relative improvement in test accuracy is numerically greater when
38 the prevalence is increased while holding the specificity constant (PPV 95.0% and NPV
39 95.2% for Tests A and B, prevalence 50%) as compared to a scenario where one
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 modulates the test specificity and maintains the prevalence constant (PPV 83.3% and
4 NPV 83.9% for Tests C and D, prevalence 5%). Put another way, the net change in PPV
5 is maximized based on moderate changes in disease prevalence as opposed to alterations
6 in test sensitivity. As for the NPV, a rise in sensitivity from 95% to 99% does not yield
7 substantial alterations in the value. Essentially, the NPV is already quite high, no matter
8 what the prevalence of resistance in the population. Conversely, the NPV suffers only
9 modestly in the populations where disease prevalence is highest compared to those with
10 the lowest disease prevalence (Figure 1).
11
12
13
14
15
16
17
18
19
20
21

22 The sensitivity analysis in which we assume that both the sensitivity and
23 specificity of the test equal 90% is illustrated in Figure 2. At the lowest prevalence of
24 disease, this specificity affords an unacceptably low PPV (32.1%), while the NPV
25 remains high, exceeding 99%. As the prevalence of the disease rises in the target
26 population, while the test's specificity and sensitivity remain fixed at 90%, the PPV and
27 NPV converge at 90%, indicating a major improvement in the PPV without dramatically
28 compromising the NPV.
29
30
31
32
33
34
35
36
37
38

39 Best- and worst-case scenario estimates of the total annual pool of patients at risk
40 for MRSA infection in cSSSI, pneumonia and sepsis demonstrate that the potential for
41 over-treatment far exceeds that for under-treatment (Table 2). Focusing on sepsis as an
42 example, for the worst-case calculation we assumed 1,141,000 sepsis hospitalizations
43 annually, a 5% MRSA prevalence, along with test characteristics of 95% sensitivity and
44 95% specificity. These parameters resulted in 57,050 potential cases of inappropriate
45 treatment reflecting the sum of subjects classified as falsely positive or negative. Of
46 these misclassified subjects, 54,198 (95%) represent those at risk for over-treatment.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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.

1
2
3 Our data have several important implications. First, as manufacturers, regulators
4 and clinicians consider what tests may have superior characteristics, it is important for all
5 stake holders to engage in defining the appropriate populations for testing with these
6 novel technologies. Our data clearly demonstrate that rather than expending resources for
7 every laboratory to elevate their sensitivity and specificity to close to 100%, the more
8 fruitful effort may be to develop algorithms to identify those patient populations at high
9 risk for the disease being tested. This is particularly true given that marginal
10 enhancements in sensitivity and specificity often come at the cost of substantial financial
11 investments. Second, raising the sensitivity of these technologies even beyond the current
12 levels may be pursuing diminishing returns, given the already high NPV. That is, even
13 when the sensitivity is no higher than 90%, the negative predictive value reaches very
14 high levels (over 95%) in the setting of moderate pre-test probability for disease.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 Third, and possibly most important, by using genotyping as opposed to
32 phenotyping employed in the traditional microbiology laboratory methods, molecular
33 diagnostics promise to result in sensitivity values that far exceed those of the traditional
34 techniques. The flip side of this optimization in sensitivity is a blunting in specificity,
35 whereby it may become unclear whether the identified organism is indeed the cause of
36 the clinical condition. Our data indicate that the true need in diagnostic testing lies not in
37 further optimization of sensitivity, but in improving the specificity of the results.
38
39
40
41
42
43
44
45
46
47

48 Because improvements in one by necessity lead to detriments in the other, future
49 directions in molecular diagnostics require thoughtful planning. We have clearly shown
50 that, in order to live up to the promise of improved targeting of antibiotic treatment, such
51 planning must include careful consideration of the populations in whom molecular
52
53
54
55
56
57
58
59
60

1
2
3 diagnostic techniques are appropriate. In fact the most important lesson from our
4
5 simulation is that we need to develop algorithms to help identify patients belonging to
6
7 populations with a high risk for the particular pathogen. If a predictive algorithm is able
8
9 to enrich the population to be tested to the disease prevalence between 30% and 40%,
10
11 both PPV and NPV will be moved into a useful range even when the test's sensitivity and
12
13 specificity are both well below 100%. With the advent of health informatics and the
14
15 massive growth in computing ability, turning reams of patient data into predictive
16
17 equations is a clearly needed functionality. Already several computing systems are
18
19 addressing this need, and the trend should continue with the input from all stakeholders
20
21 (18, 19).
22
23
24
25
26

27 Our study has a number of limitations. The most important limitation is that it is
28
29 merely a mathematical model, and, as such, by necessity relies on the accuracy of
30
31 estimates in the literature. The fact that some of the papers we used for deriving our
32
33 assumptions themselves were modeling exercises (11), predisposes our computations to
34
35 greater uncertainty. This, however, does not negate our findings that span a wide range of
36
37 plausible epidemiology. Furthermore, our model underscores the need to understand local
38
39 pathogen patterns, the recognition of which should drive decisions about the utility of
40
41 these powerful molecular diagnostics.
42
43
44
45

46 In summary, molecular diagnostics promise to streamline identification and
47
48 treatment of many infectious diseases. While the emergence of these powerful
49
50 technologies is a positive development, we need to attend to developing algorithms to aid
51
52 in selecting appropriate patients for their use. Indiscriminate application of molecular
53
54 diagnostics to all-comers presenting with signs of an infection without consideration for
55
56
57
58
59
60

1
2
3 pre-test probability of disease is likely to result in a great deal of antimicrobial
4
5 overtreatment. This will then only accelerate the current trajectory of escalating
6
7 resistance. In conjunction with developing these important technologies, it is incumbent
8
9 upon the healthcare community to build robust evidence and information technology
10
11 infrastructure to guide appropriate use of such testing.
12
13

14
15
16
17 **Acknowledgements.** No one other than the listed authors participated in the study. Dr.
18
19 Zilberberg had full access to all of the data in the study and takes responsibility for the
20
21 integrity of the data and the accuracy of the data analysis.
22
23

24
25 Dr. Zilberberg participated in the conception and design of the study, data analysis and
26
27 interpretation and drafting and revising of the manuscript for important intellectual
28
29 content. Dr. Shorr participated in the conception of the study, data interpretation and
30
31 revising of the manuscript for important intellectual content.
32
33

34
35
36 **Data sharing:**
37

38
39 The model and the inputs are quite transparent and mostly can be found in the paper. If
40
41 there is a need for further information, we are happy to provide it.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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
<i>Best-case scenario</i>			
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
<i>Worst-case scenario</i>			
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%

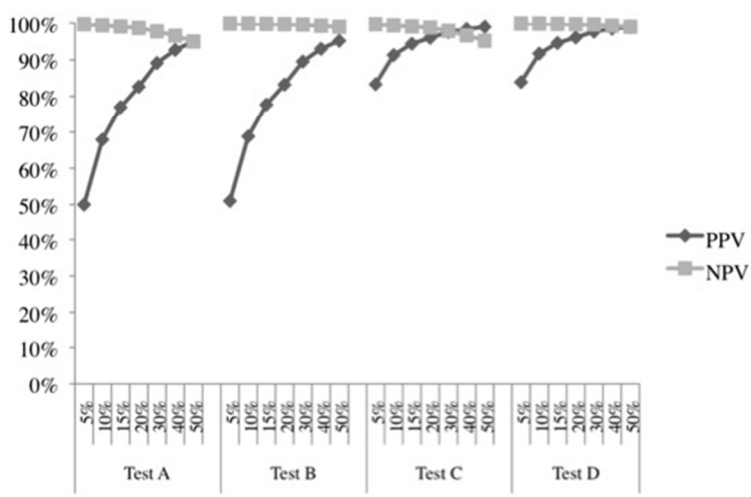
For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Figure 2. Sensitivity analysis under the conditions of test sensitivity and specificity**
4 **equaling 90%**
5
6
7

8 PPV = positive predictive value; NPV = negative predictive value
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

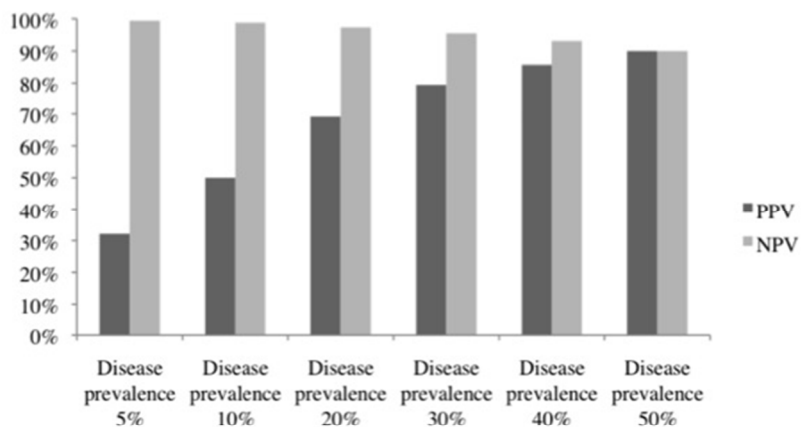
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



254x190mm (72 x 72 DPI)

Review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



254x190mm (72 x 72 DPI)

Review only



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

Journal:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts

1
2
3 **Impact of prior probabilities of MRSA as an infectious agent on the accuracy of the**
4 **emerging molecular diagnostic tests: A model simulation**
5
6

7 **Marya D. Zilberberg, MD, MPH^{*1,2},**

8 ¹EviMed Research Group, LLC, Goshen, MA

9 ²University of Massachusetts, Amherst, MA

10 *Corresponding author

11 President & CEO, EviMed Research Group, LLC

12 Adjunct Associate Professor of Epidemiology, University of Massachusetts School of
13 Public Health and Health Sciences

14 PO Box 303

15 Goshen, MA 01032

16 (413)-268-6381

17 (413)-268-3416 fax

18 evimedgroup@gmail.com
19
20
21

22 **Andrew F. Shorr, MD, MPH³**

23 ³Washington Hospital Center, Washington, DC

24 Associate Section Director, Pulmonary Critical Care, Washington Hospital Center

25 Associate Professor of Medicine, Georgetown University

26 110 Irving Street, NW

27 Washington, DC 20010

28 (202)-877-2998

29 (202)-782-9032 fax

30 Andrew.Shorr@gmail.com
31
32
33

34 **Copyright** © 2012 Marya Zilberberg

35 Dr. Zilberberg, the Corresponding Author, has the right to grant on behalf of all authors
36 and does grant on behalf of all authors, an exclusive license on a worldwide basis to the
37 BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ
38 editions and any other BMJ PGL products and sublicenses such use and exploit all
39 subsidiary rights, as set out in our license.
40

41 **Competing interests:** The authors have nothing to disclose. The authors received no
42 funding for this study. All authors have completed the Unified Competing Interest form.

43 **Guarantor statement:** Dr. Zilberberg has had access to all the data and guarantees its
44 integrity.
45

46 **IRB statement:** The study is a model and did not require an Ethics Committee review

47 **Author contributions:**

48 Dr. Zilberberg conceived, designed and executed the study, data collection and analyses,
49 drafted the manuscript and revised it for important intellectual content.

50 Dr. Shorr refined the study idea, design and interpretation, and revised the manuscript for
51 important intellectual content.
52
53
54

55 **Word count:** 2,690

56 **References:** 19
57
58
59
60

1
2
3 **Tables and Figures: 4**
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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 $\geq 95\%$. Even when the test's specificity and sensitivity were both 90%, in a high-risk population both PPV and NPV were $\sim 90\%$. 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

1
2
3 (10). Such improvement in shortening the diagnostic time is invaluable, particularly
4
5 given these tests' ostensible accuracy.
6
7

8 At the same time, one must exercise caution because these tests are not 100%
9 accurate. And while manufacturers strive for ever-increasing sensitivity and specificity
10 for their tests, a more fruitful area of investigation may be learning to identify
11 characteristics of specific populations in whom these tests may prove to be most helpful
12 for targeting and tailoring treatment. In other words, the central clinical question may
13 revolve not around issues of sensitivity and specificity intrinsic to the test, but rather
14 around the positive (PPV) and negative (NPV) predictive values associated with these
15 newer tools in populations with various levels of risk for the organisms in question. This
16 approach fits in with the Bayesian decision making, whereby the prior probability of an
17 event informs the interpretation of the diagnostic data. Irrespective of the sensitivity and
18 specificity, if the PPV and NPV are not sufficiently high, then these new tests may not
19 help clinicians either to withhold unnecessarily broad coverage or to tailor it shortly after
20 the results return.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 We hypothesized that even under conditions where such rapid diagnostic tests had
39 near-perfect sensitivity and specificity, the population-specific risk for having a particular
40 organism would represent a crucial consideration in driving diagnostic accuracy. That is
41 failure to consider the pre-test probability of these organisms in the population screened
42 would undermine the potential value of rapid diagnostic tests. To address this question
43 we developed a model simulation evaluating the application of these assays, and relied
44 upon publicly available data to populate our analysis.
45
46
47
48
49
50
51
52
53
54

55 **Methods**

56
57
58
59
60

1
2
3 We developed a mathematical model simulating the impact of changing a test's
4 characteristics on its accuracy within several strata of population risk for a particular
5 organism. All the inputs were extracted from publicly available data. The primary
6 outcome of interest was the potential magnitude for over-diagnosis of a particular
7 pathogen, or the proportion of false positive tests under the varying assumptions. We
8 were specifically interested in the false positive rates, since these cases are the ones most
9 likely to receive overly broad treatment when it is not indicated. Such overly broad
10 treatment represents a key clinical endpoint since it exposes the patient and the healthcare
11 system to adverse consequences individually and as a group. As a secondary endpoint we
12 examined the overall inaccuracy of the test in various scenarios, defined as the sum of the
13 false positive and false negative results as a proportion of the total population.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

29 The model was based on the approximate risks of methicillin-resistant
30 *Staphylococcus aureus* (MRSA) among three distinct hospitalized populations: 1)
31 complicated skin and skin structure infections (cSSSI) (11-13), 2) pneumonia (11, 14),
32 and 3) sepsis (11). We sought the most generalizable estimates for at least two factors out
33 of the following three, using the available data to calculate the third when necessary: 1)
34 total volume of hospitalizations for each disease of interest, 2) proportion of the total
35 volume represented by MRSA, and 3) total number of MRSA infections in each disease
36 category (11, 14, 15).
37
38
39
40
41
42
43
44
45
46
47

48 For consistency, the assumptions for the corresponding test characteristics
49 mimicked those from MRSA diagnostics (16). To derive estimates for positive (PPV) and
50 negative predictive values (NPV) for a plausible range of test characteristics, we
51 developed four hypothetical testing situations: 1) Test A, with the sensitivity and
52
53
54
55
56
57
58
59
60

1
2
3 specificity of 95%, 2) Test B, with the sensitivity 99% and specificity 95%, 3) Test C,
4
5 with the sensitivity 95% and specificity 99%, and 4) Test D, with the sensitivity and
6
7 specificity 99%. To explore how deviations from the average sensitivity and specificity
8
9 metrics may impact the accuracy of identification, we conducted sensitivity analyses
10
11 assuming 90% sensitivity and specificity. Based on the range of MRSA risk estimates in
12
13 the populations of interest (i.e., cSSSI, pneumonia, and sepsis), we varied the prevalence
14
15 estimates from 5% to 50%, and calculated the PPV and NPV for each of the intermediate
16
17 values.
18
19
20
21

22 We additionally performed best- and worst-case scenario simulations for each
23
24 population in question. Thus, for the worst-case scenario where all variables were biased
25
26 against the novel rapid diagnostic assay, we utilized as inputs the highest disease volume
27
28 and lowest disease prevalence, along with the lowest test sensitivity and specificity
29
30 values. Skewing the inputs in this fashion provides a potential estimate of the extent and
31
32 impact of misclassification when all assumptions are shifted so as to constrain the
33
34 potential value of the rapid diagnostic test in question. Conversely, for the best-case
35
36 scenario, we input the lowest disease volume and the highest disease prevalence, along
37
38 with the highest test sensitivity and test specificity. For both of these analyses, the total
39
40 annual universe of specific disease hospitalizations in the US was used. We utilized these
41
42 values to estimate the total numbers of potential cases within each population that would
43
44 be over-treated (i.e., treated for MRSA when no MRSA is present), under-treated (i.e.,
45
46 not treated for MRSA when MRSA is present) and treated inappropriately (i.e., either
47
48 over- or undertreated).
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Both the values for sensitivity and specificity and disease risk were rounded in
4
5 order to ease computational presentation. Volumes and prevalence of MRSA in the
6
7 disease states of interest were extracted from several large surveys available in the public
8
9 domain (Table 1). Thus, for the MRSA cSSSI volumes we relied on a study by Klein,
10
11 which quantified these hospitalizations in 2005 (11). The proportions of cSSSI in which
12
13 MRSA is the offending pathogen derived from two recent epidemiologic studies of cSSSI
14
15 hospitalizations in the US (12, 13). The volume of pneumonia hospitalizations was
16
17 extracted from the American Lung Association's 2010 data, and the proportion
18
19 represented by MRSA from a large and representative database analysis by Kollef and
20
21 colleagues (14, 15). Finally, we relied on the Agency's for Healthcare Research and
22
23 Quality recent statistical brief quantifying the burden of hospitalizations with sepsis,
24
25 while the Klein study provided the proportion likely caused by MRSA (11, 16).
26
27
28
29
30
31

32 **Results**

33
34 The input assumptions and their sources are presented in Table 1. The estimated
35
36 prevalence of MRSA ranges from approximately 5% in sepsis to nearly 50% in cSSSI,
37
38 while the prevalence of MRSA in pneumonia falls between those extremes. Under the
39
40 conditions of lowest prevalence (5%) along with the average test specificity of 95%, the
41
42 PPV reaches only 50% (Figure 1). Improving the specificity by nearly 5% to 99%
43
44 without altering the disease prevalence results in a moderate improvement in the PPV to
45
46 approximately 84%. Alternatively, a change of a similar magnitude in the PPV occurs,
47
48 when the prevalence of disease increases from 5% to the 10%-20% range, even as the
49
50 specificity remains anchored at 95% (Figure 1). The PPV further improves as the
51
52 prevalence of disease approaches 50%. Notably, at the extremes of disease prevalence
53
54
55
56
57
58
59
60

1
2
3 and test specificity, the relative improvement in test accuracy is numerically greater when
4
5 the prevalence is increased while holding the specificity constant (PPV 95.0% and NPV
6
7 95.2% for Tests A and B, prevalence 50%) as compared to a scenario where one
8
9 modulates the test specificity and maintains the prevalence constant (PPV 83.3% and
10
11 NPV 83.9% for Tests C and D, prevalence 5%). Put another way, the net change in PPV
12
13 is maximized based on moderate changes in disease prevalence as opposed to alterations
14
15 in test sensitivity. As for the NPV, a rise in sensitivity from 95% to 99% does not yield
16
17 substantial alterations in the value. Essentially, the NPV is already quite high, no matter
18
19 what the prevalence of resistance in the population. Conversely, the NPV suffers only
20
21 modestly in the populations where disease prevalence is highest compared to those with
22
23 the lowest disease prevalence (Figure 1).
24
25
26
27
28

29
30 The sensitivity analysis in which we assume that both the sensitivity and
31
32 specificity of the test equal 90% is illustrated in Figure 2. At the lowest prevalence of
33
34 disease, this specificity affords an unacceptably low PPV (32.1%), while the NPV
35
36 remains high, exceeding 99%. As the prevalence of the disease rises in the target
37
38 population, while the test's specificity and sensitivity remain fixed at 90%, the PPV and
39
40 NPV converge at 90%, indicating a major improvement in the PPV without dramatically
41
42 compromising the NPV.
43
44
45

46
47 Best- and worst-case scenario estimates of the total annual pool of patients at risk
48
49 for MRSA infection in cSSSI, pneumonia and sepsis demonstrate that the potential for
50
51 over-treatment far exceeds that for under-treatment (Table 2). Focusing on sepsis as an
52
53 example, for the worst-case calculation we assumed 1,141,000 sepsis hospitalizations
54
55 annually, a 5% MRSA prevalence, along with test characteristics of 95% sensitivity and
56
57
58
59
60

1
2
3 95% specificity. These parameters resulted in 57,050 potential cases of inappropriate
4 treatment reflecting the sum of subjects classified as falsely positive or negative. Of
5
6 these misclassified subjects, 54,198 (95%) represent those at risk for over-treatment.
7
8
9
10 Conversely, under the best-case assumptions of a high MRSA prevalence (10%) in sepsis
11
12 (n=727,000), and a test with near-perfect sensitivity and specificity (both 99%), only
13
14 7,270 individuals are at risk for inappropriate treatment with 6,543 (90%) being over-
15
16 treated (Table 2).
17
18

19
20 Overall, under the worst-case assumptions for all three of the conditions of
21
22 interest, over 150,000 patients annually with these three conditions may be treated
23
24 inappropriately, with overtreatment accounting for 136,000 (91%) of this cohort. Under
25
26 the best circumstances, among the more than 18,000 patients treated potentially
27
28 inappropriately, nearly 15,000 (81%) may be subjected to over-treatment (Table 2).
29
30

31 **Discussion**

32
33
34 We have demonstrated explicitly that organism prevalence is an important driver
35
36 of the accuracy of rapid molecular diagnostic tests even when their sensitivity and
37
38 specificity are near perfect. Additionally, we have shown that although improving the
39
40 theoretical test's specificity results in greater accuracy, one enhances accuracy even more
41
42 by restricting test utilization to a population at an increased risk for infection with the
43
44 pathogen in question. In other words, increasing pre-test probability as a strategy to
45
46 minimize antibiotic abuse results in more accurate patient classification than does
47
48 developing a marginally superior rapid diagnostic test with near-perfect specificity. In
49
50 fact, given the already high NPV, the new molecular diagnostics have the potential to
51
52 limit the use of empiric broad-spectrum coverage substantially. However, although
53
54
55
56
57
58
59
60

1
2
3 promising for limiting exposure to excessive antimicrobial coverage, molecular
4
5
6 diagnostics are still likely to result in a substantial amount of inappropriate treatment. The
7
8 vast majority (over 90%) of such inappropriate coverage is due to over-treatment in
9
10 scenarios where the test is applied irrespective of considerations of the prevalence of a
11
12 resistant pathogen.
13

14
15 Our data have several important implications. First, as manufacturers, regulators
16
17 and clinicians consider what tests may have superior characteristics, it is important for all
18
19 stakeholders to engage in defining the appropriate populations for testing with these
20
21 novel technologies. Our data clearly demonstrate that rather than expending resources for
22
23 every laboratory to elevate their sensitivity and specificity to close to 100%, the more
24
25 fruitful effort may be to develop algorithms to identify those patient populations at high
26
27 risk for the disease being tested. This is particularly true given that marginal
28
29 enhancements in sensitivity and specificity often come at the cost of substantial financial
30
31 investments. Second, raising the sensitivity of these technologies even beyond the current
32
33 levels may be pursuing diminishing returns, given the already high NPV. That is, even
34
35 when the sensitivity is no higher than 90%, the negative predictive value reaches very
36
37 high levels (over 95%) in the setting of moderate pre-test probability for disease.
38
39
40
41
42

43
44 Third, and possibly most important, by using genotyping as opposed to
45
46 phenotyping employed in the traditional microbiology laboratory methods, molecular
47
48 diagnostics promise to result in sensitivity values that far exceed those of the traditional
49
50 techniques. The flip side of this optimization in sensitivity is a blunting in specificity,
51
52 whereby it may become unclear whether the identified organism is indeed the cause of
53
54
55
56
57
58
59
60

1
2
3 the clinical condition. Our data indicate that the true need in diagnostic testing lies not in
4
5 further optimization of sensitivity, but in improving the specificity of the results.
6
7

8 Because improvements in one by necessity lead to detriments in the other, future
9
10 directions in molecular diagnostics require thoughtful planning. We have clearly shown
11
12 that, in order to live up to the promise of improved targeting of antibiotic treatment, such
13
14 planning must include careful consideration of the populations in whom molecular
15
16 diagnostic techniques are appropriate. In fact the most important lesson from our
17
18 simulation is that we need to develop algorithms to help identify patients belonging to
19
20 populations with a high risk for the particular pathogen. If a predictive algorithm is able
21
22 to enrich the population to be tested to the disease prevalence between 30% and 40%,
23
24 both PPV and NPV will be moved into a useful range even when the test's sensitivity and
25
26 specificity are both well below 100%. With the advent of health informatics and the
27
28 massive growth in computing ability, turning reams of patient data into predictive
29
30 equations is a clearly needed functionality. Already several computing systems are
31
32 addressing this need, and the trend should continue with the input from all stakeholders
33
34
35
36
37
38
39 (18, 19).

40
41 Our study has a number of limitations. The most important limitation is that it is
42
43 merely a mathematical model, and, as such, by necessity relies on the accuracy of
44
45 estimates in the literature. The fact that some of the papers we used for deriving our
46
47 assumptions themselves were modeling exercises (11), predisposes our computations to
48
49 greater uncertainty. This, however, does not negate our findings that span a wide range of
50
51 plausible epidemiology. Furthermore, our model underscores the need to understand local
52
53
54
55
56
57
58
59
60

1
2
3 pathogen patterns, the recognition of which should drive decisions about the utility of
4
5 these powerful molecular diagnostics.
6
7

8 In summary, molecular diagnostics promise to streamline identification and
9
10 treatment of many infectious diseases. While the emergence of these powerful
11
12 technologies is a positive development, we need to attend to developing algorithms to aid
13
14 in selecting appropriate patients for their use. Indiscriminate application of molecular
15
16 diagnostics to all-comers presenting with signs of an infection without consideration for
17
18 pre-test probability of disease is likely to result in a great deal of antimicrobial
19
20 overtreatment. This will then only accelerate the current trajectory of escalating
21
22 resistance. In conjunction with developing these important technologies, it is incumbent
23
24 upon the healthcare community to build robust evidence and information technology
25
26 infrastructure to guide appropriate use of such testing.
27
28
29
30
31
32
33

34 **Acknowledgements.** No one other than the listed authors participated in the study. Dr.
35
36 Zilberberg had full access to all of the data in the study and takes responsibility for the
37
38 integrity of the data and the accuracy of the data analysis.
39

40 Dr. Zilberberg participated in the conception and design of the study, data analysis and
41
42 interpretation and drafting and revising of the manuscript for important intellectual
43
44 content. Dr. Shorr participated in the conception of the study, data interpretation and
45
46 revising of the manuscript for important intellectual content.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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
<i>Best-case scenario</i>			
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
<i>Worst-case scenario</i>			
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

1
2
3 **Figure 1. Positive and negative predictive values of a test with the given sensitivity**
4 **and specificity, stratified by population disease prevalence***
5
6
7
8

9 *Percentages along the X-axis represent disease prevalence strata

10 PPV = positive predictive value; NPV = negative predictive value

11 Test A: sensitivity = 95%, specificity = 95%; Test B: sensitivity = 99%, specificity = 95%;

12 Test C: sensitivity = 95%, specificity = 99%; Test D: sensitivity = 99%, specificity = 99%

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

PPV = positive predictive value; NPV = negative predictive value

For peer review only

1
2
3 **Molecular diagnostics require a Bayesian approach****Impact of prior probabilities of**
4 **MRSA as an infectious agent on the accuracy of the emerging molecular diagnostic**
5 **tests: A model simulation**
6
7

8 **Marya D. Zilberberg, MD, MPH^{*1,2},**

9 ¹EviMed Research Group, LLC, Goshen, MA

10 ²University of Massachusetts, Amherst, MA

11 *Corresponding author

12 President & CEO, EviMed Research Group, LLC

13 Adjunct Associate Professor of Epidemiology, University of Massachusetts School of
14 Public Health and Health Sciences

15 PO Box 303

16 Goshen, MA 01032

17 (413)-268-6381

18 (413)-268-3416 fax

19 evimedgroup@gmail.com
20
21
22

23 **Andrew F. Shorr, MD, MPH³**

24 ³Washington Hospital Center, Washington, DC

25 Associate Section Director, Pulmonary Critical Care, Washington Hospital Center

26 Associate Professor of Medicine, Georgetown University

27 110 Irving Street, NW

28 Washington, DC 20010

29 (202)-877-2998

30 (202)-782-9032 fax

31 Andrew.Shorr@gmail.com
32
33
34

35 **Copyright** © 2012 Marya Zilberberg

36 Dr. Zilberberg, the Corresponding Author, has the right to grant on behalf of all authors
37 and does grant on behalf of all authors, an exclusive license on a worldwide basis to the
38 BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ
39 editions and any other BMJ PGL products and sublicenses such use and exploit all
40 subsidiary rights, as set out in our license.
41

42 **Competing interests:** The authors have nothing to disclose. The authors received no
43 funding for this study. All authors have completed the Unified Competing Interest form.

44 **Guarantor statement:** Dr. Zilberberg has had access to all the data and guarantees its
45 integrity.
46

47 **IRB statement:** The study is a model and did not require an Ethics Committee review

48 **Author contributions:**

49 Dr. Zilberberg conceived, designed and executed the study, data collection and analyses,
50 drafted the manuscript and revised it for important intellectual content.

51 Dr. Shorr refined the study idea, design and interpretation, and revised the manuscript for
52 important intellectual content.
53
54
55

56 **Word count:** 2,690562
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References: 19
Tables and Figures: 4

For peer review only

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 $\geq 95\%$. Even when the test's specificity and sensitivity were both 90%, in a high-risk population both PPV and NPV were $\sim 90\%$. 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

1
2
3 (10). Such improvement in shortening the diagnostic time is invaluable, particularly
4
5 given these tests' ostensible accuracy.
6
7

8 At the same time, one must exercise caution because these tests are not 100%
9
10 accurate. And while manufacturers strive for ever-increasing sensitivity and specificity
11
12 for their tests, a more fruitful area of investigation may be learning to identify
13
14 characteristics of specific populations in whom these tests may prove to be most helpful
15
16 for targeting and tailoring treatment. In other words, the central clinical question may
17
18 revolve not around issues of sensitivity and specificity intrinsic to the test, but rather
19
20 around the positive (PPV) and negative (NPV) predictive values associated with these
21
22 newer tools in populations with various levels of risk for the organisms in question. This
23
24 approach fits in with the Bayesian decision making, whereby the prior probability of an
25
26 event informs the interpretation of the diagnostic data. Irrespective of the sensitivity and
27
28 specificity, if the PPV and NPV are not sufficiently high, then these new tests may not
29
30 help clinicians either to withhold unnecessarily broad coverage or to tailor it shortly after
31
32 the results return.
33
34
35
36
37

38 We hypothesized that even under conditions where such rapid diagnostic tests had
39
40 near-perfect sensitivity and specificity, the population-specific risk for having a particular
41
42 organism would represent a crucial consideration in driving diagnostic accuracy. That is
43
44 failure to consider the pre-test probability of these organisms in the population screened
45
46 would undermine the potential value of rapid diagnostic tests. To address this question
47
48 we developed a model simulation evaluating the application of these assays, and relied
49
50 upon publicly available data to populate our analysis.
51
52
53
54

55 **Methods**

56
57
58
59
60

1
2
3
4 We developed a mathematical model simulating the impact of changing a test's
5 characteristics on its accuracy within several strata of population risk for a particular
6 organism. All the inputs were extracted from publicly available data. The primary
7 outcome of interest was the potential magnitude for over-diagnosis of a particular
8 pathogen, or the proportion of false positive tests under the varying assumptions. We
9 were specifically interested in the false positive rates, since these cases are the ones most
10 likely to receive overly broad treatment when it is not indicated. Such overly broad
11 treatment represents a key clinical endpoint since it exposes the patient and the healthcare
12 system to adverse consequences individually and as a group. As a secondary endpoint we
13 examined the overall inaccuracy of the test in various scenarios, defined as the sum of the
14 false positive and false negative results as a proportion of the total population.
15
16
17
18
19
20
21
22
23
24
25
26
27
28

29 The model was based on the approximate risks of methicillin-resistant
30 *Staphylococcus aureus* (MRSA) among three distinct hospitalized populations: 1)
31 complicated skin and skin structure infections (cSSSI) (11-13), 2) pneumonia (11, 14),
32 and 3) sepsis (11). We sought the most generalizable estimates for at least two factors out
33 of the following three, using the available data to calculate the third when necessary: 1)
34 total volume of hospitalizations for each ~~of the~~ diseases of interest, 2) proportion of the
35 total volume represented by MRSA, and 3) total number of MRSA infections in each
36 disease category (11, 14, 15).
37
38
39
40
41
42
43
44
45
46
47

48 For consistency, the assumptions for the corresponding test characteristics
49 mimicked those from MRSA diagnostics (16). To derive estimates for positive (PPV) and
50 negative predictive values (NPV) for a plausible range of test characteristics, we
51 developed four hypothetical testing situations: 1) Test A, with the sensitivity and
52
53
54
55
56
57
58
59
60

1
2
3 specificity of 95%, 2) Test B, with the sensitivity 99% and specificity 95%, 3) Test C,
4
5 with the sensitivity 95% and specificity 99%, and 4) Test D, with the sensitivity and
6
7 specificity 99%. To explore how deviations from the average sensitivity and specificity
8
9 metrics may impact the accuracy of identification, we conducted sensitivity analyses
10
11 assuming 90% sensitivity and specificity. Based on the range of MRSA risk estimates in
12
13 the populations of interest (i.e., cSSSI, pneumonia, and sepsis), we varied the prevalence
14
15 estimates from 5% to 50%, and calculated the PPV and NPV for each of the intermediate
16
17 values.
18
19
20
21

22 We additionally performed best- and worst-case scenario simulations for each of
23
24 the populations in question. Thus, for the worst-case scenario where all variables were
25
26 biased against the novel rapid diagnostic assay, we utilized as inputs the highest disease
27
28 volume and lowest disease prevalence, along with the lowest test sensitivity and
29
30 specificity values. Skewing the inputs in this fashion provides a potential estimate of the
31
32 extent and impact of misclassification when all assumptions are shifted so as to constrain
33
34 the potential value of the rapid diagnostic test in question. Conversely, for the best-case
35
36 scenario, we input the lowest disease volume and the highest disease prevalence, along
37
38 with the highest test sensitivity and test specificity. For both of these analyses, the total
39
40 annual universe of specific disease hospitalizations in the US was used. We utilized these
41
42 values to estimate the total numbers of potential cases within each population that would
43
44 be over-treated (i.e., treated for MRSA when no MRSA is present), under-treated (i.e.,
45
46 not treated for MRSA when MRSA is present) and treated inappropriately (i.e., either
47
48 over- or undertreated).
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Both the values for sensitivity and specificity and disease risk were rounded in
4
5 order to ease computational presentation. Volumes and prevalence of MRSA in the
6
7 disease states of interest were extracted from several large surveys available in the public
8
9 domain (Table 1). Thus, for the MRSA cSSSI volumes we relied on a study by Klein,
10
11 which quantified these hospitalizations in 2005 (11). The proportions of cSSSI in which
12
13 MRSA is the offending pathogen derived from two recent epidemiologic studies of cSSSI
14
15 hospitalizations in the US (12, 13). The volume of pneumonia hospitalizations was
16
17 extracted from the American Lung Association's 2010 data, and the proportion
18
19 represented by MRSA from a large and representative database analysis by Kollef and
20
21 colleagues (14, 15). Finally, we relied on the Agency's for Healthcare Research and
22
23 Quality recent statistical brief quantifying the burden of hospitalizations with sepsis,
24
25 while the Klein study provided the proportion likely caused by MRSA (11, 16).
26
27
28
29
30
31

32 **Results**

33
34 The input assumptions and their sources are presented in Table 1. The estimated
35
36 prevalence of MRSA ranges from approximately 5% in sepsis to nearly 50% in cSSSI,
37
38 while the prevalence of MRSA in pneumonia falls between those extremes. Under the
39
40 conditions of lowest prevalence (5%) along with the average test specificity of 95%, the
41
42 PPV reaches only 50% (Figure 1). Improving the specificity by nearly 5% to 99%
43
44 without altering the disease prevalence results in a moderate improvement in the PPV to
45
46 approximately 84%. Alternatively, a change of a similar magnitude in the PPV occurs,
47
48 when the prevalence of disease increases from 5% to the 10%-20% range, even as the
49
50 specificity remains anchored at 95% (Figure 1). The PPV further improves as the
51
52 prevalence of disease approaches 50%. Notably, at the extremes of disease prevalence
53
54
55
56
57
58
59
60

1
2
3 and test specificity, the relative improvement in test accuracy is numerically greater when
4
5 the prevalence is increased while holding the specificity constant (PPV 95.0% and NPV
6
7 95.2% for Tests A and B, prevalence 50%) as compared to a scenario where one
8
9 modulates the test specificity and maintains the prevalence constant (PPV 83.3% and
10
11 NPV 83.9% for Tests C and D, prevalence 5%). Put another way, the net change in PPV
12
13 is maximized based on moderate changes in disease prevalence as opposed to alterations
14
15 in test sensitivity. As for the NPV, a rise in sensitivity from 95% to 99% does not yield
16
17 substantial alterations in the value. Essentially, the NPV is already quite high, no matter
18
19 what the prevalence of resistance in the population. Conversely, the NPV suffers only
20
21 modestly in the populations where disease prevalence is highest compared to those with
22
23 the lowest disease prevalence (Figure 1).
24
25
26
27
28

29 The sensitivity analysis in which we assume that both the sensitivity and
30
31 specificity of the test equal 90% is illustrated in Figure 2. At the lowest prevalence of
32
33 disease, this specificity affords an unacceptably low PPV (32.1%), while the NPV
34
35 remains high, exceeding 99%. As the prevalence of the disease rises in the target
36
37 population, while the test's specificity and sensitivity remain fixed at 90%, the PPV and
38
39 NPV converge at 90%, indicating a major improvement in the PPV without dramatically
40
41 compromising the NPV.
42
43
44
45

46 Best- and worst-case scenario estimates of the total annual pool of patients at risk
47
48 for MRSA infection in cSSSI, pneumonia and sepsis demonstrate that the potential for
49
50 over-treatment far exceeds that for under-treatment (Table 2). Focusing on sepsis as an
51
52 example, for the worst-case calculation we assumed 1,141,000 sepsis hospitalizations
53
54 annually, a 5% MRSA prevalence, along with test characteristics of 95% sensitivity and
55
56
57
58
59
60

1
2
3 95% specificity. These parameters resulted in 57,050 potential cases of inappropriate
4 treatment reflecting the sum of subjects classified as falsely positive or negative. Of
5
6 these misclassified subjects, 54,198 (95%) represent those at risk for over-treatment.
7
8
9
10 Conversely, under the best-case assumptions of a high MRSA prevalence (10%) in sepsis
11
12 (n=727,000), and a test with near-perfect sensitivity and specificity (both 99%), only
13
14 7,270 individuals are at risk for inappropriate treatment with 6,543 (90%) being over-
15
16 treated (Table 2).
17
18

19
20 Overall, under the worst-case assumptions for all three of the conditions of
21
22 interest, over 150,000 patients annually with these three conditions may be treated
23
24 inappropriately, with overtreatment accounting for 136,000 (91%) of this cohort. Under
25
26 the best circumstances, among the more than 18,000 patients treated potentially
27
28 inappropriately, nearly 15,000 (81%) may be subjected to over-treatment (Table 2).
29
30

31 Discussion

32
33
34 We have demonstrated explicitly that organism prevalence is an important driver
35
36 of the accuracy of rapid molecular diagnostic tests even when their sensitivity and
37
38 specificity are near perfect. Additionally, we have shown that although improving the
39
40 theoretical test's specificity results in greater accuracy, one enhances accuracy even more
41
42 by restricting test utilization to a population at an increased risk for infection with the
43
44 pathogen in question. In other words, increasing pre-test probability as a strategy to
45
46 minimize antibiotic abuse results in more accurate patient classification than does
47
48 developing a marginally superior rapid diagnostic test with near-perfect specificity. In
49
50 fact, given the already high NPV, the new molecular diagnostics have the potential to
51
52 limit the use of empiric broad-spectrum coverage substantially. Finally However,
53
54
55
56
57
58
59
60

1
2
3 although promising for limiting exposure to excessive antimicrobial coverage, molecular
4
5
6 diagnostics are still likely to result in a substantial amount of inappropriate treatment. The
7
8
9 vast majority (over 90%) of such inappropriate coverage is due to over-treatment in
10
11 scenarios where the test is applied irrespective of considerations of the prevalence of a
12
13 resistant pathogen.

14
15 Our data have several important implications. First, as manufacturers, regulators
16
17 and clinicians consider what tests may have superior characteristics, it is important for all
18
19
20 stake-holders to engage in defining the appropriate populations for testing with these
21
22 novel technologies. Our data clearly demonstrate that rather than expending resources for
23
24 every laboratory to elevate their sensitivity and specificity to close to 100%, the more
25
26 fruitful effort may be to develop algorithms to identify those patient populations at high
27
28 risk for the disease being tested. This is particularly true given that marginal
29
30
31 enhancements in sensitivity and specificity often come at the cost of substantial financial
32
33 investments. Second, raising the sensitivity of these technologies even beyond the current
34
35 levels may be pursuing diminishing returns, given the already high NPV. That is, even
36
37 when the sensitivity is no higher than 90%, the negative predictive value reaches very
38
39 high levels (over 95%) in the setting of moderate pre-test probability for disease.
40
41
42

43
44 Third, and possibly most important, by using genotyping as opposed to
45
46 phenotyping employed in the traditional microbiology laboratory methods, molecular
47
48 diagnostics promise to result in sensitivity values that far exceed those of the traditional
49
50 techniques. The flip side of this optimization in sensitivity is a blunting in specificity,
51
52 whereby it may become unclear whether the identified organism is indeed the cause of
53
54
55
56
57
58
59
60

1
2
3 the clinical condition. Our data indicate that the true need in diagnostic testing lies not in
4 further optimization of sensitivity, but in improving the specificity of the results.
5
6

7
8 Because improvements in one by necessity lead to detriments in the other, future
9 directions in molecular diagnostics require thoughtful planning. We have clearly shown
10 that, in order to live up to the promise of improved targeting of antibiotic treatment, such
11 planning must include careful consideration of the populations in whom molecular
12 diagnostic techniques are appropriate. In fact the most important lesson from our
13 simulation is that we need to develop algorithms to help identify patients belonging to
14 populations with a high risk for the particular pathogen. If a predictive algorithm is able
15 to enrich the population to be tested to the disease prevalence between 30% and 40%,
16 both PPV and NPV will be moved into a useful range even when the test's sensitivity and
17 specificity are both well below 100%. With the advent of health informatics and the
18 massive growth in computing ability, turning reams of patient data into predictive
19 equations is a clearly needed functionality. Already several computing systems are
20 addressing this need, and the trend should continue with the input from all stakeholders
21 (18, 19).
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 Our study has a number of limitations. The most important limitation is that it is
41 merely a mathematical model, and, as such, by necessity relies on the accuracy of
42 estimates in the literature. The fact that some of the papers we used for deriving our
43 assumptions themselves were modeling exercises (11), predisposes our computations to
44 greater uncertainty. This, however, does not negate our findings that span a wide range of
45 plausible epidemiology. Furthermore, our model underscores the need to understand local
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 pathogen patterns, the recognition of which should drive decisions about the utility of
4
5 these powerful molecular diagnostics.
6
7

8 In summary, molecular diagnostics promise to streamline identification and
9
10 treatment of many infectious diseases. While the emergence of these powerful
11
12 technologies is a positive development, we need to attend to developing algorithms to aid
13
14 in selecting appropriate patients for their use. Indiscriminate application of molecular
15
16 diagnostics to all-comers presenting with signs of an infection without consideration for
17
18 pre-test probability of disease is likely to result in a great deal of antimicrobial
19
20 overtreatment. This will then only accelerate the current trajectory of escalating
21
22 resistance. In conjunction with developing these important technologies, it is incumbent
23
24 upon the healthcare community to build robust evidence and information technology
25
26 infrastructure to guide appropriate use of such testing.
27
28
29
30
31
32
33

34 **Acknowledgements.** No one other than the listed authors participated in the study. Dr.
35
36 Zilberberg had full access to all of the data in the study and takes responsibility for the
37
38 integrity of the data and the accuracy of the data analysis.
39

40 Dr. Zilberberg participated in the conception and design of the study, data analysis and
41
42 interpretation and drafting and revising of the manuscript for important intellectual
43
44 content. Dr. Shorr participated in the conception of the study, data interpretation and
45
46 revising of the manuscript for important intellectual content.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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
<i>Best-case scenario</i>			
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
<i>Worst-case scenario</i>			
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%

For peer review only

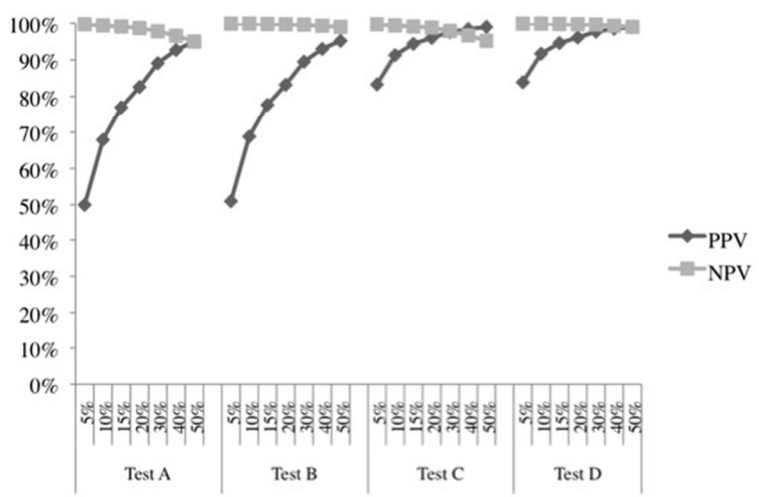
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Figure 2. Sensitivity analysis under the conditions of test sensitivity and specificity**
4 **equaling 90%**
5
6
7

8 PPV = positive predictive value; NPV = negative predictive value
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

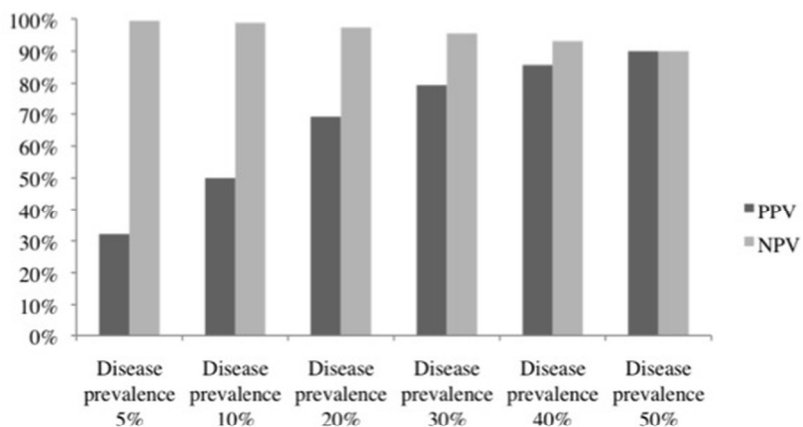
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



119x90mm (300 x 300 DPI)

Review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



119x90mm (300 x 300 DPI)

Review only