Automated DNA Sequence-Based Early Warning System for the Detection of Methicillin-Resistant *Staphylococcus aureus* Outbreaks

Alexander Mellmann¹, Alexander W. Friedrich¹, Nicole Rosenkötter², Jörg Rothgäenger³, Helge Karch¹, Ralf Reintjes², Dag Harmsen⁴*

¹ Institute for Hygiene, University Hospital Münster, Münster, Germany, ² Hamburg University of Applied Sciences, Hamburg, Germany, ³ Ridom GmbH, Würzburg, Germany, ⁴ Department of Periodontology, University Hospital Münster, Münster, Germany

Competing Interests: JR, HK, and DH have declared a potential conflict of interest. JR and DH are the developers of the Ridom StaphType software mentioned in the manuscript. The software is distributed and sold by the company Ridom GmbH that is partially owned by them. HK is a scientific advisor of this company. All other authors have declared that no competing interests exist.

Author Contributions: AM, AWF, HK, and DH designed the study. AM, AWF, NR, JR, RR, and DH analyzed the data. AM, AWF, NR, HK, RR, and DH contributed to writing the paper. HK supervised the clinical sample and data collection. AWF headed the ICP team. AM did the laboratory identification and strain analysis. JR and DH developed the Ridom StaphType software. AM and DH wrote the main part of the paper. All authors gave useful comment on the analysis of data and text of the manuscript.

Academic Editor: Mervyn Singer, University College London, United Kingdom


Received: May 12, 2005
Accepted: October 28, 2005
Published: January 10, 2006

DOI: 10.1371/journal.pmed.0030033

Copyright: © 2006 Mellmann et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: CI, confidence interval; ICP, infection control professional; MLST, multi-locus sequence typing; MRSA, methicillin-resistant *Staphylococcus aureus*; NPV, negative predictive value; PFGE, pulsed-field gel electrophoresis; PPV, positive predictive value; spa, *S. aureus* protein A gene; UHM, University Hospital Münster

* To whom correspondence should be addressed. E-mail: dharmsen@uni-muenster.de

**Abstract**

Background

The detection of methicillin-resistant *Staphylococcus aureus* (MRSA) usually requires the implementation of often rigorous infection-control measures. Prompt identification of an MRSA epidemic is crucial for the control of an outbreak. In this study we evaluated various early warning algorithms for the detection of an MRSA cluster.

Methods and Findings

Between 1998 and 2003, 557 non-replicate MRSA strains were collected from staff and patients admitted to a German tertiary-care university hospital. The repeat region of the *S. aureus* protein A (*spa*) gene in each of these strains was sequenced. Using epidemiological and typing information for the period 1998–2002 as reference data, clusters in 2003 were determined by temporal-scan test statistics. Various early warning algorithms (frequency, clonal, and infection control professionals [ICP] alerts) were tested in a prospective analysis for the year 2003. In addition, a newly implemented automated clonal alert system of the Ridom StaphType software was evaluated.

A total of 549 of 557 MRSA were typeable using *spa* sequencing. When analyzed using scan test statistics, 42 out of 175 MRSA in 2003 formed 13 significant clusters (*p* < 0.05). These clusters were used as the “gold standard” to evaluate the various algorithms. Clonal alerts (*spa* typing and epidemiological data) were 100% sensitive and 95.2% specific. Frequency (epidemiological data only) and ICP alerts were 100% and 62.1% sensitive and 47.2% and 97.3% specific, respectively. The difference in specificity between clonal and ICP alerts was not significant. Both methods exhibited a positive predictive value above 80%.

Conclusions

Rapid MRSA outbreak detection, based on epidemiological and *spa* typing data, is a suitable alternative for classical approaches and can assist in the identification of potential sources of infection.
Introduction

In the United States alone, infections acquired in hospitals affect 2 million patients, account for half of all major hospital complications, and result in annual costs of more than $4.5 billion [1]. Staphylococcus aureus is the leading cause of these nosocomial infections that include a wide range of diseases such as endocarditis, septicaemia, skin infections, soft tissue infections, and bone infections [2]. Strains resistant to methicillin, in particular, have become a major concern in the hospital environment because of the high mortality rate and the stringent hygienic requirements needed for patients who are harboring a methicillin-resistant S. aureus (MRSA) [3,4]. Moreover, since the emergence of strains that are insensitive or have reduced sensitivity to glycopeptides, there is a real danger of infections spreading that have even greater drug resistance [5].

Analysis of laboratory test results and patients’ charts are the methods usually used to identify outbreaks. However, the manual review of laboratory test results is time-consuming and resource-intensive. Electronic analysis of data can help identify suspicious patterns of disease and antimicrobial resistance [6], but such sentinel methods are rarely used in clinical practice. The typing of MRSA isolates, not only from clinical specimens, but also from surveillance cultures, is necessary for the elucidation of possible transmission routes. Because the procedures are slow and laborious, molecular typing (e.g., pulsed-field gel electrophoresis [PFGE]) is usually used a posteriori to track the course of nosocomial infections in an already established outbreak. Furthermore, PFGE requires great efforts to harmonize protocols and is therefore only partially successful in generating reproducible results [7]. In order to improve the speed of typing, DNA sequence-based approaches, such as the multi-locus sequence typing (MLST), are becoming more frequently used [8]. However, MLST is not suitable for routine surveillance of MRSA because of the high costs involved and the low discriminatory power compared to PFGE. Frenay et al., who were the first to use a single-locus sequence typing method for S. aureus, employed the sequence of the polymorphic region X of the S. aureus protein A gene (spa) for typing [9]. Since then, numerous studies evaluated this variable number of tandem repeat targets as quite suitable for short-term epidemiological applications, e.g., [10–13]. Because of the paucity in software for repeat identification and lack of a consensus in assigning spa type names, the wide-spread use of the method was hampered for years until the recent introduction of the Ridom StaphType software [14]. With this software, the spa sequences are analyzed automatically and linked to a database integrated with epidemiological information. A universal nomenclature is achieved by synchronization with a central server that assigns new spa types for all users (http://www.spaserver.ridom.de).

The aim of the study reported here was therefore to analyze the utility of a spa sequence-based, automatic early warning algorithm to detect MRSA clusters in hospitals and to compare this approach with classical surveillance techniques. We hypothesized that the automated system, once established, can complement and even replace the labor-intensive traditional methods used for cluster identification.

Methods

Setting

Between 1998 and 2003, a total of 557 non-replicate MRSA isolates were collected at the University Hospital Münster (UHM), Germany, a 1,480-bed tertiary-care teaching facility. In 2003, there were approximately 43,000 annual admissions to the hospital where the mean length of stay was 9.8 d. The prevalence of patients with MRSA colonizations and infections was taken as the annual number of persons harboring MRSA (×100) divided by the total number of admissions at UHM [15]. The baseline for calculation of the relative risk was the year 1998.

Surveillance and Infection Control Measures

All new MRSA cases were monitored prospectively by infection control professionals (ICP) from the day when MRSA was first identified until hospital discharge. Information on each patient was obtained by reviewing medical records and laboratory data and holding telephone interviews with the attending physician. Subsequently, the ICP decided if a transmission event was likely and if further investigation was necessary. In more detail, the following infection control measures were implemented: (i) As recommended in the guidelines of the Robert Koch Institute (Berlin, Germany), all patients infected or colonized with MRSA were placed in contact isolation until the time of discharge or until eradication could be documented in three consecutive sets of negative surveillance cultures (separated by at least 24 h). MRSA surveillance cultures included swabs of several body sites (nose, groin, skin lesions, inguinal, perineal, and axillary swabs). In the case of infected patients, samples were taken from the site of infection. (ii) All patients known to have been previously colonized or infected with MRSA were isolated on re-admission to UHM and surveillance swabs were obtained. Negative surveillance cultures were mandatory in order to terminate contact isolation. (iii) Clinical microbiology laboratory results were monitored daily for the occurrence of specimens containing MRSA. (iv) spa typing of MRSA isolates, as performed since 2002, were carried out directly after detection of a new MRSA isolate. (v) Colonized patients were treated with nasal mupirocin ointment for 5 d and daily chlorhexidine body washes were applied. In the case of patients remaining in the hospital after eradication, weekly surveillance cultures were recommended over a 4-wk period, and then at monthly intervals to detect possible re-colonization. (vi) To detect MRSA colonization and cross-transmission, surveillance cultures were obtained from roommates as soon as a new MRSA patient was identified. (vii) Staff were screened when nosocomial transmissions were suspected and at intervals as a surveillance method on high-risk wards. (viii) Hospital staff found MRSA-positive patients were suspended from work on the wards until the successful eradication of MRSA could be documented. (ix) Systematic surveillance cultures at the time of admission and on a weekly basis thereafter were begun in 2002 in wards caring for high-risk patients, e.g., intensive care units [15,16]. Colonization and infection were defined in accordance with the Centers for Disease Control and Prevention criteria [17].

Microbiology and Molecular Typing

The strain collection consisted of MRSA from various clinical sources (e.g., blood cultures and wound infections)
and included surveillance cultures from patients and staff. Of all clinical S. aureus isolates, 6.4% exhibited methicillin resistance in 2003. For species identification, every strain was tested with API ID 32 Staph (bioMérieux, Marci l’Etoile, France) and for the presence of free coagulase. The presence of the meca gene responsible for methicillin resistance was confirmed using PCR [18]. The sequence of the short sequence repeat region of the spa gene encoding the S. aureus protein A was determined in 557 strains [14]. The primers spa-1113f (5’- TAA AGA CGA TCC TTC GGT GAC C –3’) and spa-1514r (5’- CAG CAG TAG TGC CGT TTG CT –3’) were used for spa amplification and Taq Cycle sequencing. DNA sequences were obtained with an ABI Prism 3100 Avant Genetic Analyzer (Applied Biosystems, Foster City, California, United States) and analyzed with the Ridom StaphType software version 1.5 beta (Ridom GmbH, Würzburg, Germany) incorporating the newly added automated early warning system (“clonal alerts”) for MRSA cluster detection [14]. Typability, discriminatory index, and the 95% confidence interval (CI) of the discriminatory index were calculated using the procedures published previously [19,20].

Retrospective Temporal-Scan Test Statistics

To evaluate the various early warning algorithms, we performed scan test statistics using the epidemiological and typing information from 1998 to 2002 as historical data to determine MRSA clusters in 2003 [21,22]. Temporal-scan statistics evaluates whether an apparent cluster of disease is unlikely to occur by chance alone. Thereby, the test determines a likelihood \( p \)-value for an observed number of cases appearing in a window of fixed width as the window is moved along the time axis studied (2003). Observed and expected cases, the latter calculated using the historical data (1998–2002), were compared with a null hypothesis that states cases occur at random, evaluated against the alternative hypothesis that states cases cluster in certain time periods. In this evaluation, a Poisson distribution was assumed because a positive MRSA finding is a rare and irregular event. Clusters of two or more infected/colonized patients or colonized staff on the same ward or wards in close contact (e.g., interdisciplinary intensive care units) occurring within a 2-wk window and harboring the same MRSA isolate according to the spa typing, were identified as significant at the 5% level. These statistically confirmed clusters were then used as the “gold standard” for comparing the various alert mechanisms. Non-significant clusters were considered to be sporadic occurrences.

Early Warning Algorithms

Every MRSA isolate obtained in 2003 was examined in a prospective analysis by applying descriptive epidemiologic parameters such as time, place, and person. When two or more MRSA isolates were detected within a 2-wk window on the same ward or on wards having close contact, the resulting alert was regarded as a “frequency alert” and added to a “frequency cluster.” If MRSA isolates also shared an identical spa type, the allocation to “clonal alerts” and associated “clonal clusters” was triggered. An ICP, which is a panel consisting of two physicians and four infection control nurses who meet weekly and hold additional meetings when an outbreak occurs, rate the findings as “ICP alerts” and “ICP clusters,” respectively. When feasible, the area of surveillance is widened and an investigation initiated. The ICP uses microbial data and data from patients’ charts to reach their decisions but are blind to the occurrence of an outbreak on the basis of spa typing results.

Statistical Analysis

Sensitivity, specificity, positive and negative predictive values (PPV, NPV), and pre-test probability were determined as described by Sackett et al. [23]. The pre-test probability is defined as the proportion with the target disorder (MRSA cluster) in the population at risk (MRSA positive) at a specific time interval. Two-tailed, 95% CIs were calculated to assess sensitivity, specificity, PPV, and NPV using a normal approximation for the pertinent (binomial) distribution. The chi-square distribution, with one degree of freedom, was used to determine the significance of the differences in these parameters.

Results

Table 1 summarizes the important epidemiological indicators for MRSA at the UHM. The overall prevalence of MRSA cases was 0.17 per 100 admissions and the relative risk of acquiring MRSA increased 4-fold during the study period. The annual number of patients with MRSA bacteremia reached a peak in 2003 with six patients. The average turnaround time for spa typing under routine laboratory conditions was 2.4 d. Of the 557 MRSA isolates tested, 549 (98.6%) could be typed using spa sequencing, and the eight strains, which could not be typed, were excluded from the analysis. A total of 79 different spa types were identified in samples collected for the period 1998–2003.

### Table 1. Important Epidemiological Indicators for MRSA at the UHM (1998–2003)

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of MRSA</th>
<th>Prevalence</th>
<th>Relative Risk of MRSA Acquisition (95% CI)</th>
<th>Number of Patients with MRSA Bacteremia</th>
<th>Number of spa Types</th>
<th>Typability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>51</td>
<td>0.11</td>
<td>reference</td>
<td>1</td>
<td>16</td>
<td>0.879 (0.829–0.929)/100</td>
</tr>
<tr>
<td>1999</td>
<td>62</td>
<td>0.17</td>
<td>1.6 (1.1–2.3)</td>
<td>0</td>
<td>19</td>
<td>0.895 (0.861–0.928)/96.3</td>
</tr>
<tr>
<td>2000</td>
<td>56</td>
<td>0.12</td>
<td>1.1 (0.8–1.7)</td>
<td>2</td>
<td>22</td>
<td>0.873 (0.804–0.942)/100</td>
</tr>
<tr>
<td>2001</td>
<td>55</td>
<td>0.13</td>
<td>1.1 (0.8–1.7)</td>
<td>1</td>
<td>28</td>
<td>0.950 (0.926–0.973)/98.2</td>
</tr>
<tr>
<td>2002</td>
<td>134</td>
<td>0.31</td>
<td>2.8 (2.0–3.9)</td>
<td>1</td>
<td>31</td>
<td>0.897 (0.870–0.925)/100</td>
</tr>
<tr>
<td>2003</td>
<td>179</td>
<td>0.42</td>
<td>3.8 (2.8–5.3)</td>
<td>6</td>
<td>35</td>
<td>0.859 (0.824–0.894)/97.8</td>
</tr>
<tr>
<td>1998–2003</td>
<td>557</td>
<td>0.17</td>
<td>-</td>
<td>11</td>
<td>79</td>
<td>0.918 (0.908–0.929)/98.6</td>
</tr>
</tbody>
</table>

Di, discriminatory index.
Table 2. Most Frequent *spa* Types and Epidemiological Background (PFGE, MLST) at the UHM and in Germany

<table>
<thead>
<tr>
<th><em>spa</em> Type</th>
<th>Cases at the UHM, 2003 (%)</th>
<th>Cases at the UHM, 1998–2003 (%)</th>
<th>% Clones in Germany, 2003/1998–2003</th>
<th>Concordant MLST ST&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Comment&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>t003</td>
<td>53 (29.6)</td>
<td>87 (15.6)</td>
<td>24.5/10.1 (together with t002)</td>
<td>ST-5, ST-225</td>
<td>Rhine Hesse MRSA (subclone), EMRSA-3, New York clone, CC5</td>
</tr>
<tr>
<td>t032</td>
<td>30 (16.7)</td>
<td>58 (10.4)</td>
<td>29.0/20.1 (together with t026)</td>
<td>ST-22</td>
<td>Barnim MRSA (prototype and subclone), EMRSA-15, prototype of ST-22, CC22</td>
</tr>
<tr>
<td>t001</td>
<td>19 (10.6)</td>
<td>68 (12.2)</td>
<td>13.8/26.6 (together with t023 and t041)</td>
<td>ST-5, ST-222, ST-228</td>
<td>Southern German MRSA (prototype and subclone), Rhine Hesse MRSA (subclone), EMRSA-3, New York clone, CC5</td>
</tr>
<tr>
<td>t004</td>
<td>14 (7.8)</td>
<td>82 (14.7)</td>
<td>13.8/22.5 (together with t026 and t038)</td>
<td>ST-45</td>
<td>Berlin MRSA (prototype), USA600 ORSA II, USA600 ORSA IV, CC45</td>
</tr>
<tr>
<td>t008</td>
<td>10 (5.6)</td>
<td>29 (5.2)</td>
<td>0.1/2.9 (together with t051)</td>
<td>ST-8, ST-247, ST-250, ST-254</td>
<td>Northern German MRSA (subclone), USA300 ORSA IV (cMRSA in the US), Archaic/Iberian, ST250 ORSA I, CC8</td>
</tr>
<tr>
<td>t038</td>
<td>5 (2.8)</td>
<td>10 (1.8)</td>
<td>0.0/1.8 (see t003)</td>
<td>ST-45</td>
<td>Berlin MRSA (subclone NRW), CC45</td>
</tr>
<tr>
<td>t002</td>
<td>4 (2.2)</td>
<td>13 (2.3)</td>
<td>-</td>
<td>ST-5, ST-231</td>
<td>Rhine Hesse MRSA (prototype), EMRSA-3, New York clone, Japan clone, Pediatric, USA100 ORSA II, USA800 ORSA IV, ST 5 ORSA I, CC5</td>
</tr>
<tr>
<td>NT</td>
<td>4 (2.2)</td>
<td>8 (1.4)</td>
<td>-</td>
<td>NT</td>
<td></td>
</tr>
<tr>
<td>t044</td>
<td>2 (1.1)</td>
<td>8 (1.4)</td>
<td>Data not available</td>
<td>ST-80</td>
<td>cMRSA ( luk&lt;sup&gt;–&lt;/sup&gt;/ luk&lt;sup&gt;+&lt;/sup&gt; ) widely disseminated in Europe</td>
</tr>
<tr>
<td>t009</td>
<td>1 (0.6)</td>
<td>22 (3.9)</td>
<td>0.5/4.6</td>
<td>ST-254</td>
<td>Hannover MRSA, EMRSA-10, CC8</td>
</tr>
<tr>
<td>t018</td>
<td>1 (0.6)</td>
<td>8 (1.4)</td>
<td>Data not available</td>
<td>ST-30, ST-36, ST-38</td>
<td>Prototype of ST-36, EMRSA-16, USA200 ORSA II, CC30</td>
</tr>
<tr>
<td>t037</td>
<td>0 (0.0)</td>
<td>19 (3.4)</td>
<td>0.0/0.3</td>
<td>ST-239, ST-240, ST-241</td>
<td>Vienna MRSA, Brazilian/Hungarian, ST239 ORSA III, ST240 ORSA III, EMRSA-1, −4, −7, −9, −11, CC8/239</td>
</tr>
<tr>
<td>t051</td>
<td>0 (0.0)</td>
<td>1 (0.2)</td>
<td>(see t008)</td>
<td>ST-247</td>
<td>Northern German MRSA (prototype and subclone), Archaic/Iberian, ST247 ORSA I, EMRSA-5, CC8</td>
</tr>
<tr>
<td>t041</td>
<td>0 (0.0)</td>
<td>3 (0.5)</td>
<td>(see t001)</td>
<td>ST-111, ST-228</td>
<td>Southern German MRSA, CC5</td>
</tr>
<tr>
<td>t026</td>
<td>0 (0.0)</td>
<td>1 (0.2)</td>
<td>(see t004)</td>
<td>ST-45, ST-47</td>
<td>Berlin MRSA (subclone), CC45</td>
</tr>
<tr>
<td>t023</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>(see t001)</td>
<td>ST-228</td>
<td>Southern German MRSA, CC5</td>
</tr>
<tr>
<td>t190</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>Data not available</td>
<td>ST-8</td>
<td>EMRSA-2, −6, CC8</td>
</tr>
<tr>
<td>Other</td>
<td>36 (20.2)</td>
<td>140 (25.1)</td>
<td>18.3/12.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Predominant ST in bold.

<sup>b</sup>German clone as defined by PFGE and MLST CC.

Other <s>36 (20.2)</s> (25.1) 18.3/12.9

Other clusters were located on a single ward, whereas seven other clusters were located at sites distributed throughout a clinic.

In the prospective analysis of all MRSA in 2003 using the various alert procedures, there were 106 frequency alerts attributable to 13 significant clusters (Table 4). A total of 36 clonal alerts, comprising 20 clonal clusters, were triggered by the early warning system. The ICP called 22 ICP alerts corresponding to nine ICP clusters, but in only five clusters (the two largest clusters, numbers five and seven were included) was the recognition of an existing outbreak and the need for further investigation correct. The four other clusters arose from false alerts by the ICP. In Table 4, the alerts triggered by the various methods are categorized as true or false alerts using the alerts for the 13 significant “true” clusters. The sensitivity, specificity, PPV, and NPV, and where appropriate, the 95% CIs for the various alert methods, are displayed in Table 5. Because of the high number of false-positive frequency alerts (n = 77), the specificity of the frequency and clonal methods (47.2% and 95.2%, respectively) differed considerably. The ICP alerts had the highest specificity, but the number of false-negative alerts (eight of 13 confirmed clusters were missed) led to the lowest sensitivity (62.1%). Given a pre-test probability of 24%, the PPV (same as the post-test probability) of the ICP and clonal alerts was above 80%, whereas the PPV of the frequency alert was only 27.4%. There were no significant differences in specificity and PPV between clonal and ICP alerts. Frequency alerts were
We showed that the feasibility and speed with which it was possible to carry out spa typing was highly satisfactory. The discriminatory power, however, was lower than previously reported, probably because only a local strain collection was analyzed [12,13]. Although not examined by us, the high intra- and inter-laboratory reproducibility of 100% and the robustness of the method have recently been documented (Aires-de-Sousa et al., unpublished data). Moreover, there is a high concordance of results between spa and PFGE, microarray and MLST [11,12]. The practicability of using spa in short-term epidemiological studies has been questioned because differences in PCR amplicon sizes in related strains was thought to imply instability in the target gene [28]. In the meantime, however, there has been a plethora of publications demonstrating the value of spa in the investigation of MRSA outbreaks, e.g., [10,14]. Moreover, it has recently been shown that spa data not only contain information on short-term, but also long-term evolutionary events, as observed in whole repeat duplications and deletions [12,29]. Because of the steady fall in the cost of DNA sequencing and an average hands-on time of only 20 min per sample (determination of both strands of DNA and processing ten samples in parallel), this technique is within the capability of even small laboratories [30].

The present study has compared three early warning algorithms for the detection of nosocomial MRSA outbreaks before limited clusters of preventable MRSA transmissions develop into larger outbreaks. The evaluation of an early warning system, however, is difficult because there is no accepted “gold standard” and it is likely that no system will be completely reliable [31]. Therefore, we chose to combine epidemiological and molecular typing data with statistical analysis to provide an objective measure of performance between the varied approaches. In this approach, by definition, the sensitivity and NPV of the frequency and clonal alert methods will always be 100%. This means that the ICP method will give results that are less sensitive, or at best, of only equal sensitivity to those of the automated methods. Infection and colonization with MRSA were given equal status since both can lead to further transmissions. The practicability of using spa typing data accumulated since 1998 enables the significance of only equal sensitivity to those of the automated methods.

In this paper we have presented a new method for prospective MRSA outbreak surveillance in a hospital that uses case and molecular typing data. Historically, MRSA outbreak detection in hospitals has relied on the watchful eyes of physicians and other health-care workers. However, the increasing availability of timely electronic surveillance and molecular typing data raises the possibility of earlier outbreak detection and intervention if suitable analytic methods are found.

Germany belongs to a group of Western European countries with an intermediate level of MRSA (approximately 20% of all S. aureus diagnosed in laboratories are MRSA positive). However, the isolation rate has increased significantly in recent years [27]. Although the MRSA laboratory isolation rate in UHM of 6.4% in 2003 is still rather low in comparison with other German hospitals, the relative risk of acquiring MRSA within this hospital facility rose significantly during the study period (Table 1). Furthermore, the absolute risk will also probably rise because of epidemiological pressure and the rising prevalence of MRSA in Germany as a whole (Table 2 and Figure 1). It is clear that control of MRSA is a pressing concern where new concepts are needed, and therefore we studied spa typing in combination with an automatic early warning algorithm to detect MRSA clusters at UHM.

Discussion

In this paper we have presented a new method for prospective MRSA outbreak surveillance in a hospital that uses case and molecular typing data. Historically, MRSA outbreak detection in hospitals has relied on the watchful eyes of physicians and other health-care workers. However, the increasing availability of timely electronic surveillance and molecular typing data raises the possibility of earlier outbreak detection and intervention if suitable analytic methods are found.

Germany belongs to a group of Western European countries with an intermediate level of MRSA (approximately 20% of all S. aureus diagnosed in laboratories are MRSA positive). However, the isolation rate has increased significantly in recent years [27]. Although the MRSA laboratory isolation rate in UHM of 6.4% in 2003 is still rather low in comparison with other German hospitals, the relative risk of acquiring MRSA within this hospital facility rose significantly during the study period (Table 1). Furthermore, the absolute risk will also probably rise because of epidemiological pressure and the rising prevalence of MRSA in Germany as a whole (Table 2 and Figure 1). It is clear that control of MRSA is a pressing concern where new concepts are needed, and therefore we studied spa typing in combination with an automatic early warning algorithm to detect MRSA clusters at UHM.
mented this approach successfully [33,34]. A similar approach has also been used in hospitals, e.g., using 2-fold standard deviation and monthly increase algorithms for detecting clusters of nosocomial infections [31]. In using the 2-fold standard deviation algorithm, the threshold for a suspected outbreak is defined as the mean of all previous cases per unit time plus two standard deviations. The monthly increase algorithm triggers an alert if there is either a 100% increase in the number of observed cases in the current month compared to the monthly totals for the two previous months, or a 50% increase over a three-month period. In the case of MRSA, however, infections/colonizations occur infrequently and irregularly. Applying the 2-fold standard deviation and monthly increase algorithms to our data (including typing data) resulted in delayed alerts for cluster detection indicating that they were insufficiently sensitive (unpublished data).

In order to improve the detection of MRSA clusters and to avoid delay, we applied the first of the two definitions of an outbreak mentioned above, i.e., two or more cases of infection/colonization that are linked epidemiologically, because discriminatory typing details of cases were rapidly available.

The ability of the current procedures in hospitals to prevent nosocomial infections and to recognize nosocomial outbreaks often depends on the manual review of laboratory results and surveillance by the ICP. However, this review process is resource-intensive because duplicate isolates must be eliminated, results must be correlated with patients’ charts, patient locations within the hospital must be tracked, and related events must be correlated and monitored [6]. Not surprising, many minor transmissions of MRSA infections amenable to intervention go undiscovered. In this study, only five of the 13 “true” clusters were detected as clusters by visual screening of laboratory reports by the ICP (Table 3 and 4). However, the ICP alerts had the highest PPV (81.8%) because of their high specificity. On the other hand, frequency alerts with a sensitivity of 100% detected every cluster. However, the high number of false-positive alerts giving the lowest PPV (27.4%) clearly demonstrates that this method is unsuitable (also shown by the significant differences in specificity and PPV compared with clonal and ICP alerts). Clonal alerts combine the best of both methods, i.e., high specificity and high sensitivity with a PPV comparable to that of ICP (no significant difference). In comparison to ICP, only a few more false-positive alerts were triggered, and more clusters, especially the smaller ones, were detected (Table 5). The data also indicated that surveillance conducted in the laboratory has the advantage in that clusters occurring throughout the hospital can be identified at a single, central data collection point. Further advantages are its speed (<3 d after detection of MRSA) and the portable nature of data generated by spa typing permitting the differentiation between outbreaks and pseudo-outbreaks and the central coordination of a suitable response in real-time [14].

### Table 3. Significant MRSA Clusters in 2003 (Chronologically Ordered) Identified by Scan Test Statistics (p < 0.05)

<table>
<thead>
<tr>
<th>Cluster Number</th>
<th>Clinic/Ward</th>
<th>Cluster Length (d)</th>
<th>spa Type</th>
<th>Number of Cluster Isolates</th>
<th>Total Number of Strains Isolated (Same Clinic/Ward) with the Same spa Type (1998 to 2003)</th>
<th>p-Value of Temporal Clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B1</td>
<td>5</td>
<td>t008</td>
<td>2</td>
<td>2</td>
<td>0.009</td>
</tr>
<tr>
<td>2a</td>
<td>C1</td>
<td>1</td>
<td>t003</td>
<td>2</td>
<td>8</td>
<td>0.03</td>
</tr>
<tr>
<td>3</td>
<td>C1</td>
<td>6</td>
<td>t001</td>
<td>2</td>
<td>3</td>
<td>0.02</td>
</tr>
<tr>
<td>4</td>
<td>C1</td>
<td>21</td>
<td>t003</td>
<td>3</td>
<td>8</td>
<td>0.02</td>
</tr>
<tr>
<td>5a</td>
<td>F</td>
<td>31</td>
<td>t003</td>
<td>6</td>
<td>13</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>9</td>
<td>t002</td>
<td>2</td>
<td>3</td>
<td>0.036</td>
</tr>
<tr>
<td>7a</td>
<td>A</td>
<td>27</td>
<td>t001</td>
<td>11</td>
<td>37</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>8</td>
<td>G1</td>
<td>2</td>
<td>t032</td>
<td>2</td>
<td>5</td>
<td>0.022</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>9</td>
<td>t003</td>
<td>4</td>
<td>13</td>
<td>0.0003</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>5</td>
<td>t034</td>
<td>2</td>
<td>2</td>
<td>0.009</td>
</tr>
<tr>
<td>11</td>
<td>I</td>
<td>6</td>
<td>t243</td>
<td>2</td>
<td>2</td>
<td>0.011</td>
</tr>
<tr>
<td>12a</td>
<td>B</td>
<td>3</td>
<td>t032</td>
<td>2</td>
<td>6</td>
<td>0.047</td>
</tr>
<tr>
<td>13a</td>
<td>B2</td>
<td>2</td>
<td>t003</td>
<td>2</td>
<td>7</td>
<td>0.043</td>
</tr>
</tbody>
</table>

*aInvestigated by the ICP.
DOI: 10.1371/journal.pmed.0030033.t003

### Table 4. True and False Alerts/Clusters with the Various Methods Determined Using the 29 Alerts Associated with the 13 Significant Clusters in 2003

<table>
<thead>
<tr>
<th>Method</th>
<th>Alert</th>
<th>True-Positive</th>
<th>False-Positive</th>
<th>False-Negative</th>
<th>True-Negative</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>29</td>
<td>77</td>
<td>0</td>
<td>69</td>
<td>13/18/0</td>
<td></td>
</tr>
<tr>
<td>ICP</td>
<td>18</td>
<td>4</td>
<td>11</td>
<td>142</td>
<td>5/4/8</td>
<td></td>
</tr>
<tr>
<td>Clonal</td>
<td>29</td>
<td>7</td>
<td>0</td>
<td>139</td>
<td>13/7/0</td>
<td></td>
</tr>
</tbody>
</table>

*aNumbers of true-positive/false-positive/false-negative detections of significant clusters.
DOI: 10.1371/journal.pmed.0030033.t004
Automated Detection of MRSA Outbreaks

Table 5. Sensitivity, Specificity, PPV, and NPV of the Various Alert Methods (24% Pre-Test Probability) in 2003

<table>
<thead>
<tr>
<th>Method</th>
<th>% Sensitivity (95% CI)</th>
<th>% Specificity (95% CI)</th>
<th>PPV (%) (95% CI)</th>
<th>NPV (%) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>100 (NA)</td>
<td>47.3 (39.2–55.4)</td>
<td>27.4 (18.9–35.9)</td>
<td>100 (NA)</td>
</tr>
<tr>
<td>ICP</td>
<td>62.1 (44.4–79.7)</td>
<td>97.3 (94.6–99.9)</td>
<td>81.8 (65.7–97.9)</td>
<td>92.8 (88.7–96.9)</td>
</tr>
<tr>
<td>Clonal</td>
<td>100 (NA)</td>
<td>95.2 (91.7–98.7)</td>
<td>80.6 (76.7–93.5)</td>
<td>100 (NA)</td>
</tr>
</tbody>
</table>

NA, not applicable.
DOI: 10.1371/journal.pmed.0030033.t005

References

benefit analyses have demonstrated that the cost of MRSA infections far exceeds those costs involved in active surveillance and isolation procedures in a hospital [35,36]. Whether the expenditure required for spa typing is less than that for the labor-intensive manual review of patients’ charts and laboratory results needs to be determined.

Since the second half of 2004, when the study was finished and data analyzed, the Ridom StaphType software v. 1.5 beta that features an automated early warning based on clonal alerts came into routine use in our laboratory. A data-driven “re-admission alert” triggered by a hospital information system, which identifies the re-admission of any patient previously colonized or infected with MRSA, could enhance the accuracy of such a system [37].

All methods described above are based on underlying rules or seek predefined patterns. The advantages of these hypothesis-based methods are the high sensitivity and specificity achievable. However, a rapid method with a high discrimination involving gene typing is necessary to attain such a high specificity. Due to the predefined rules, unusual patterns of outbreaks might go undetected (e.g., retarded epidemics). A different approach is employed by data mining, such a high specificity. Due to the predefined rules, unusual patterns of outbreaks might go undetected (e.g., retarded epidemics). A different approach is employed by data mining, which identifies the re-admission of any patient previously colonized or infected with MRSA, could enhance the accuracy of such a system [37].

There are a number of limitations within the proposed method. There is no definitive proof available, with any method, to authenticate a MRSA transmission event. Furthermore, the “gold standard” used incorporates elements of a diagnostic test under study. Finally, if the epidemiological pressure of a certain clone changes rapidly, temporal clustering could fail and false-positive (e.g., as possible in the case of the “Rhine Hesse” MRSA clone) or false-negative clusters could be recorded.

Examining the occurrence of MRSA clusters in the way described here not only provides a useful early warning system, but can also be used to model infection dynamics and estimate important epidemiological parameters, e.g., cross-transmission rates [39,40]. Time series and typing data was used by Grundmann et al. with scan test statistics and risk factor analysis to show that the incidence of infection can be related to staffing levels [41]. In addition, molecular typing data can also be incorporated into geographical information systems combined with space-time scan statistics analysis on a regional and national level [42,43].

In conclusion, a surveillance method based on spa typing and automated alerts is useful as an early warning system in a hospital and is at least comparable to classical epidemiological approaches. We have shown that the combined use of medical informatics and molecular laboratory techniques makes intervention possible before limited clusters of preventable MRSA transmissions expand into outbreaks.

Acknowledgments
AM was funded by a grant of the Medical Faculty, University of Münster (ME 1–1–04–02). The funding organization was neither involved in the decision to submit the manuscript for publication nor influenced the organization and the research conducted in any way. We thank Ursula Keckevoet and Ralph Fischer for expert technical assistance.

References


Patient Summary

**Background.** Everyone carries many types of bacteria on or in their bodies; \textit{Staphylococcus aureus} is a normal bacteria for people to carry. About 25% to 30% of people have it, usually in the nose. It is usually harmless; however, this bacterium can also cause infections—especially in people who are otherwise unwell, or who have surgery. These infections need to be treated with antibiotics. Methicillin-resistant \textit{S. aureus} (MRSA) is an increasing problem in many of the developed world because, unlike other types of these bacteria, MRSA cannot be killed by most of the usual antibiotics that are used, such as methicillin. Without treatment, staphylococcal infection can become very severe.

**Why Was This Study Done?** MRSA is a particular problem in hospitals, where there is a need to be able to identify infected and colonized people quickly and isolate and treat them. These researchers wanted to test for the best way of identifying early clusters of MRSA outbreaks, which are more serious than just single cases and are an indication of hygiene deficiencies.

**What Did the Researchers Do and Find?** Between 1998 and 2003 the researchers analysed 557 MRSA strains from staff and patients admitted to one German university hospital. They collected information about the characteristics (in space and time) of these people, and genetically identified each of the strains. They then looked for the most efficient way to identify an outbreak, including assessment of the risk by specially trained hospital staff, with and without genetic analysis. They also assessed a specially designed computer programme (developed by some of the authors), which combined the genetic type of the MRSA as well as details about the outbreak, such as the characteristics of the patients involved; they found that this combination was more efficient and reliable method to identify outbreaks was to combine the genetic type of the MRSA with details about the outbreak, using the computer programme tested.

**What Do These Findings Mean?** The computer programme seems to be more efficient than other methods tested here in identifying when an outbreak is likely to occur. However, this is the first test of this method, and before being adopted more widely, further testing is needed in different settings and by other researchers.