Missed Opportunities to Diagnose Tuberculosis
Are Common Among Hospitalized Patients and
Patients Seen in Emergency Departments

Aaron C. Miller,1 Linnea A. Polgreen,2 Joseph E. Cavanaugh,3 Douglas B. Hornick,4 and Philip M. Polgreen4,5

1Department of Economics and Business, Cornell College, Mount Vernon, Iowa, Departments of 2Pharmacy Practice and Science, 3Biostatistics, 4Internal Medicine, and 5Epidemiology, University of Iowa, Iowa City

Background. Delayed diagnosis of tuberculosis (TB) may lead to worse outcomes and additional TB exposures.

Methods. To estimate the potential number of misdiagnosed TB cases, we linked all hospital and emergency
department (ED) visits in California’s Healthcare Cost and Utilization Project (HCUP) databases (2005–2011).
We defined a potential misdiagnosis as a visit with a new, primary diagnosis of TB preceded by a recent respiratory-related hospitalization or ED visit. Next, we calculated the prevalence of potential missed TB diagnoses for different
time windows. We also computed odds ratios (OR) comparing the likelihood of a previous respiratory diagnosis in
patients with and without a TB diagnosis, controlling for patient and hospital characteristics. Finally, we determined
the correlation between a hospital’s TB volume and the prevalence of potential TB misdiagnoses.

Results. Within 30 days before an initial TB diagnosis, 15.9% of patients (25.7% for 90 days) had a respiratory-related hospitalization or ED visit. Also, within 30 days, prior respiratory-related visits were more common in patients with TB than other patients (OR = 3.83; P < .01), controlling for patient and hospital characteristics. Respiratory diagnosis-related visits were increasingly common until approximately 90 days before the TB diagnosis. Finally, potential misdiagnoses were more common in hospitals with fewer TB cases (ρ = −0.845; P < .01).

Conclusions. Missed opportunities to diagnose TB are common and correlate inversely with the number of TB cases
diagnosed at a hospital. Thus, as TB becomes infrequent, delayed diagnoses may increase, initiating outbreaks in com-
munities and hospitals.

Keywords. missed diagnosis; transmission; tuberculosis.

In the past several years, the incidence of tuberculosis (TB) in the United States has decreased, and the major-
ity of cases occur among foreign-born persons [1]. Yet,
some reports indicate that the proportion of patients with advanced disease is increasing. This increase in
the number of cases with advanced disease may be a function of delayed diagnosis [2,3]. The delayed diagnosis of TB is
concerning for 2 reasons. First, delays are associated with
worse outcomes for patients [4–8]. Second, delays lead to
more exposures [9–11]. Delayed diagnosis in healthcare
settings is a major concern because it not only exposes
healthcare workers, but it also exposes other patients.
Several TB outbreaks have been reported in healthcare
settings [12–18] with many attributable to patients who
were initially undiagnosed [19–21].

Several reasons for delays have been identified and
have typically been attributed to either patients or the
health system [10, 22]. Patients not seeking medical
care in a timely fashion cause patient-related delays.
Health-system delays are primarily due to healthcare
professionals not considering TB at the point of care.
Reports have implicated inexperience diagnosing and
treating TB as a reason for missed opportunities to
diagnose TB [23, 24]. Unfortunately, there is currently
no standardized approach to investigate the frequency of
delayed or misdiagnosed TB that can be easily imple-
mented in a variety of geographic settings.
The purpose of this study is to propose a population-based approach for estimating the number of missed opportunities to diagnose TB, and then to use this approach for California, a state with a relatively high burden of TB [25].

METHODS

We created our cohort using the Healthcare Cost and Utilization Project (HCUP) state inpatient database (SID) and state emergency department database (SEDD) for California from 2005 through 2011. The SID contains records of all inpatient discharges for all non-federal hospitals in California. The SEDD contains records of emergency department visits that do not result in hospitalization, at hospital-affiliated emergency departments. In the state of California, patient records across hospitals and time can be linked between the SID and SEDD. Together, the SID and SEDD contain over 70 million linkable visits. These records represent over 21 million individual patients and cover 480 different hospitals. These data include measures of a patient’s principal and secondary diagnoses, procedures, patient demographics, length of stay, admission and discharge status, along with hospital charges and payment sources. Because of the absence of individual identifiers in the HCUP data, the University of Iowa’s Institutional Review Board views this as non-human-subjects research.

Inclusion Criteria

To estimate how often TB is potentially misdiagnosed, we used the administrative data in both the SID and the SEDD. We specifically defined a potential misdiagnosis as an episode of care in an emergency department or hospital that fulfilled 5 criteria: (1) a patient received a primary diagnosis of TB during either an emergency department visit or inpatient hospitalization; (2) a patient did not have a secondary diagnosis of TB; (3) the patient had a previous emergency department visit or inpatient hospitalization within a specified time window; (4) the patient did not have any TB diagnosis at the previous visit; and (5) the patient had a respiratory diagnosis at the previous visit.

We chose these criteria to increase the likelihood that the TB diagnosis (1) represented an initial diagnosis, (2) was not a follow-up indicator of a previous diagnosis, (3) was not an incorrectly recorded TB test, and (4) to confirm that the patient was in the study sample (ie, state of California) during the period when a potential misdiagnosis could have been identified. Observations were excluded for patients if their TB diagnosis occurred in a stay that was completely nested within another stay. The reason for excluding these cases is that we cannot determine the visit directly preceding the TB diagnosis, where a potential misdiagnosis may have occurred.

Cases of TB were identified using the International Classification of Disease, 9th Edition, Clinical Modification (ICD-9-CM) diagnosis codes beginning with 010, 011, 012, or 018. These codes correspond to primary, pulmonary, other and military TB, respectively. We only used the principal diagnosis code to identify cases of TB (ie, we excluded all secondary TB-related codes). To identify a respiratory diagnosis, we used both HCUP Clinical Classification Software diagnosis groupings and individual ICD-9 codes. Table 1 provides a description of the codes that were used to identify respiratory diagnosis.

We considered various time windows during which a potential misdiagnosis could have occurred. We performed 101 individual analyses on different time windows. These windows ranged from 5 to 365 days in increments of 1 to 5 days: the increments increased in size as the interval increased. For example, we performed individual analyses on time windows defined as 5–6 days, 5–7 days, 5–8 days, . . . , and 5–60 days. We used 5-day increments for 60 to 200 days: 5–65 days, 5–70 days, 5–75 days, . . . , 5–200 days. In addition, we used 10-day increments for 200–370 days: 5–210 days, 5–220 days, 5–230 days, . . . , and 5–370 days. We excluded discharges that occurred less than 5 days before the admission of the initial TB diagnosis to limit the possibility that the patient was tested for TB during the previous visit and returned to the hospital when the diagnosis was confirmed. We also excluded discharges that occurred more than 1 year before the TB diagnosis because these respiratory diagnoses are unlikely related to the TB diagnosis.

<table>
<thead>
<tr>
<th>Table 1. Diagnoses Used to Define a Respiratory-Associated Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
</tr>
<tr>
<td>Cancer of bronchus; lung</td>
</tr>
<tr>
<td>Cancer; other respiratory and intrathoracic</td>
</tr>
<tr>
<td>Pneumonia (except that caused by tuberculosis or sexually transmitted disease)</td>
</tr>
<tr>
<td>Influenza</td>
</tr>
<tr>
<td>Acute and chronic tonsillitis</td>
</tr>
<tr>
<td>Acute bronchitis</td>
</tr>
<tr>
<td>Other upper respiratory infections</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease and bronchiectasis</td>
</tr>
<tr>
<td>Asthma</td>
</tr>
<tr>
<td>Aspiration pneumonitis; food/vomitus</td>
</tr>
<tr>
<td>Pleurisy; pneumothorax; pulmonary collapse</td>
</tr>
<tr>
<td>Respiratory failure; insufficiency; arrest (adult)</td>
</tr>
<tr>
<td>Lung disease due to external agents</td>
</tr>
<tr>
<td>Other lower respiratory disease</td>
</tr>
<tr>
<td>Other upper respiratory disease</td>
</tr>
<tr>
<td>Respiratory distress syndrome</td>
</tr>
<tr>
<td>Foreign body in trachea bronchus and lung</td>
</tr>
<tr>
<td>Foreign body in pharynx and larynx</td>
</tr>
</tbody>
</table>

Abbreviations: CCS, Clinical Classification Software; ICD-9, International Classification of Disease, 9th Edition.
**Statistical Analysis**

Patients with a primary TB diagnosis who had a previous visit within a given window of time were compared with patients without a TB diagnosis who also had previous visit within the same window of time. Using multivariable logistic regression, odds ratios were then computed comparing the likelihood of a previous respiratory diagnosis in patients with and without a primary TB diagnosis. We controlled for a wide range of patient and hospital-stay characteristics, including patient age, gender, race, length of stay, number of procedures, payer type, admission type, discharge disposition, if the admission occurred on a weekend, patient income quartile, and whether or not a record contained a maternal-associated diagnosis. In addition, diagnostic conditions associated with a risk for TB, including human immunodeficiency virus, substance abuse, diabetes, kidney failure, head or neck cancer, and rheumatoid arthritis, were included in our model.

Finally, we calculated the prevalence of potential missed TB diagnoses. We determined the number of presumed missed cases for each window of time and divided that number by the total number of cases where TB was the primary diagnosis. To determine where these misdiagnoses were occurring, we calculated the prevalence of potential TB misdiagnoses at hospitals with different levels of TB admissions and emergency room (ER) visits. Finally, we used the Pearson correlation coefficient to analyze the association between hospitals’ volume of TB diagnoses rates and the prevalence of potential TB misdiagnosis.

**RESULTS**

We identified a total of 6707 nonnested cases of TB, and 5795 of these were the first time that the patient appeared in our data with a primary diagnosis of TB. Of these initial TB diagnoses, 3220 TB cases had at least 1 previous visit without a TB diagnosis. We also identified 11,781,328 different patients without TB who had a previous visit.

Figure 1A presents the number of TB cases who had a previous visit in a given time window. In this figure, there are 2 curves: one represents the number of primary TB cases with a previous respiratory diagnosis-related visit, and the other represents the number of TB cases with a previous nonrespiratory diagnosis-related visit. Tuberculosis cases are more likely to be preceded by a respiratory diagnosis-related visit than a visit with any other type of diagnosis. The marginal difference between those with a respiratory diagnosis-related visit and those without a respiratory diagnosis-related visit (represented by the vertical distance between the 2 curves) increases until approximately 90 days before a TB diagnosis and begins to steadily decrease after 90 days.

Figure 1B presents the same results as Figure 1A but for patients without TB. In contrast to Figure 1A, Figure 1B demonstrates that for patients without TB, respiratory diagnosis-related visits are much less common than nonrespiratory diagnosis-related visits, for any window of time. Moreover, for patients without TB, the marginal difference between the number of previous visits with and without a respiratory diagnosis-related visit increases continuously. Thus, for patients without TB, as the time between visits increases, patients are increasingly less likely to have a respiratory diagnosis in their previous visit. Together, Figures 1A and B demonstrate that patients with TB as a primary diagnosis are considerably more likely to experience a previous visit with a respiratory diagnosis than those without TB.

Figure 2 presents the odds ratios for our multivariable model. Even after controlling for the observable patient characteristics in our database, patients with a TB primary diagnosis are significantly more likely to experience a respiratory diagnosis-related visit preceding their TB diagnosis than are patients without TB. The odds ratios of a respiratory diagnosis preceding a TB diagnosis relative to a non-TB diagnosis are much greater than 1 for all time windows considered, but they decrease over time. Odds
ratios ranged from 5.85 for a 5- to 15-day window, 4.86 for 5–30 days, 3.83 for 5–90 days, down to 2.79 for 5–365 days.

Table 2 reports the prevalence of potential TB misdiagnoses, of all the patients with TB as a primary diagnosis included in the final analysis, for various potential misdiagnosis windows. These rates represent an estimate of the likelihood of a misdiagnosis occurring given windows where misdiagnoses may occur. The prevalence of a potentially missed TB diagnosis ranged from 15.9% using a 5- to 30-day window, 25.7% using a 5- to 90-day window, to over 33% using a 5- to 365-day window.

In addition, Table 3 categorizes hospitals into quintiles based on hospital TB volume, as measured by the number of TB diagnosis per 1000 patient discharges. Table 3 also reports the average potential misdiagnosis rate across the hospital quintiles, measured by the percentage of TB diagnoses that are potentially misdiagnosed using a 5- to 90-day window. Across all hospital years, TB diagnosis volume was significantly and negatively correlated with a hospital’s TB misdiagnosis rate ($\rho = -0.848; P = .0005$); as the number of primary TB diagnoses presenting to a hospital increases, the number of possible misdiagnoses decreases.

**DISCUSSION**

Our results demonstrate that a substantial proportion of patients, who were assigned TB as a primary diagnostic code, had previously presented to either a hospital or an ER and were diagnosed with a non-TB respiratory-related diagnosis. For example, in the 30 days before a newly recorded primary TB diagnosis, almost 16% had a previous visit with a respiratory-related diagnosis, and this number increases to 26% in the 90 days prior. Many of these prior respiratory-related visits were potential opportunities to diagnose TB that were likely missed. Without a TB diagnosis and treatment, undiagnosed patients undoubtedly put both members of the community, healthcare workers, and other patients at risk for contracting TB.

Without microbiologic data and medical charts to review, we cannot verify that cases assigned a TB diagnosis code actually had TB or an active case of TB. In addition, we cannot verify that preceding respiratory visits were missed cases of active TB. However, our cohort is population-based, and we examine the vast majority of all hospitalizations and all ER visits in California. Therefore, we can (1) make generalized, population-based estimates of potentially missed cases that would be difficult to replicate using TB registry data or (2) survey data focused only on patients with TB.

We realize that the respiratory-related visit preceding the TB cases we identified could have been completely coincidental and not related to TB. However, we controlled for a broad range of
both patient and hospital characteristics, and we found that respiratory-related visits were much more common before a TB diagnosis than before any other type of diagnosis (eg, odds ratio of 4.86 in the prior 30 days). Second, it is possible that patients diagnosed with TB are more likely to have respiratory problems than patients in the general population. Over time, however, the probability of a respiratory-related visit before a TB diagnosis is more likely than a nonrespiratory visit for all time windows considered. After 90 days, the probability decreases as other reasons for prior visits become more common. Because the association between TB and prior respiratory visits changes over time in an epidemiologically plausible fashion, we think that it is unlikely that our findings are due to an omitted variable bias. Finally, because instant point-of-care tests are not available for TB, we excluded respiratory-related visits 5 days before the initially recorded TB visit from our estimates of missed opportunities. Although we cannot independently verify the accuracy of our initial assumptions, our additional analyses make us more confident that we are detecting actual missed opportunities to diagnose TB. Future work should be focused on finding the profiles of patients with the highest risk for a missed diagnosis and the institutions at risk for a missed diagnosis.

Many prior reports investigate reasons for diagnostic delays for TB cases [21–23, 26]. In contrast to some other reports, we are unable to investigate patient-associated-diagnostic delays. However, we can investigate an important subset of TB cases, those who visited the ER or hospital before their diagnosis of TB. Given the administrative nature and scale of the data we used, we can generate our estimates for an entire state, and the estimate can be quickly updated without additional survey data, which is often subject to recall bias. A wide range of TB-related delays have been reported, ranging from 2 to 87 days [21]. Because we consider only the missed cases that present to the emergency department or are ultimately hospitalized, our results are not directly comparable to prior reports. Nevertheless, our results showing higher odds of missed opportunities (prior respiratory-related visits) for months prior seem reasonable.

Although TB cases have decreased in the United States in recent years, cases with a delayed diagnoses may be increasing [2, 3]. Physicians with less experience diagnosing TB may be less likely to accurately diagnose the disease in a timely fashion. Investigators in North Carolina used cavitary TB as a marker for TB cases with a delayed diagnosis, and they showed that such cases were more common where TB was less prevalent [3]. Indeed, our results show that possible misdiagnoses occur more commonly at hospitals where TB patients present less frequently. A similar result was reported when comparing private and public hospitals in the United States. The investigators also found that delays in TB diagnosis were negatively correlated with the frequency of TB cases [27]. The association between delayed diagnoses and clinical TB experience has also been observed outside the United States [28–33]. Thus, as TB becomes less common, we fear that delayed diagnoses may become an emerging problem, leading to outbreaks in community and healthcare settings.

Diagnostic delays attributable to the healthcare system are especially problematic. First, failure to diagnose and treat TB early puts members of the community at increased risk. Healthcare system delays also put healthcare workers and other patients at risk for acquiring TB. The complex contact patterns between healthcare workers and patients potentiate the spread of infectious diseases in healthcare settings. In such contact networks, a few infected patients can put many at risk [34–36]. A delay in placing patients with active TB in airborne infection isolation because the diagnosis is not considered at the time of admission, puts additional patients and healthcare workers at risk [5, 21, 37, 38]. If TB is considered early, chances for spread can be mitigated through the proper use of standard infection control measures [39]. The higher rate of latent TB among healthcare workers compared with the community is consistent with an increased risk for exposure to unsuspected cases of active TB cases [39–43].

To mitigate the threats posed by delayed diagnoses, more data-driven approaches are needed to help target resources and education towards areas where the potential for missed diagnoses occur more often. It is unfortunate that although it makes some sense to allocate TB resources according to TB burden, neglecting TB in lower-volume areas may exacerbate the spread of cases associated with unrecognized diagnoses, especially in healthcare settings, which exposes not only healthcare workers but also other patients to TB. A delayed diagnosis also places TB patients at greater risks for adverse outcomes.

Our study is subject to many limitations. First, as mentioned previously, we use administrative data exclusively and administrative codes without microbiologic data and pharmacy data, which may have a relatively low positive predictive value, compared with other ICD-9 codes for infectious diseases [44]. Thus, we excluded all patients with secondary diagnostic codes related to TB. Despite this limitation, we do know a great deal about the patients, and we can track patients over time, along with patients not diagnosed with TB. Second, our estimates of missed opportunities are probably underestimates because we have no records of clinic visits. Third, we do not have complete data on all patients. We do not have observations before 2005 nor do we have data on medical visits outside of California. Fourth, our analysis only considered missed opportunities in patient visits directly preceding a TB diagnosis. Patients might experience multiple missed diagnoses before TB. For example, using a 5- to 90-day window, we identified 826 patients that had a respiratory diagnosis in the visit directly preceding the initial TB diagnosis. Of these patients, 35 patients had multiple visits within 5 to 90 days prior that contained a respiratory diagnosis. This suggests that misdiagnoses may occur multiple times for some patients, and the likelihood of receiving a misdiagnosis may be greater than that derived from observing only the visit.
directly before a TB diagnosis. Finally, our work is based exclusively on the state of California and our results may not be generalizable to other regions.

CONCLUSIONS

Despite our many limitations, our results show that potential missed opportunities to diagnose TB are common. Some of these missed opportunities undoubtedly contribute to the spread of TB. Future work should use registry data to confirm these results. In addition to estimating the scale and scope of the problem of delayed diagnoses, our approach has the potential to identify areas for focused interventions, which may help reduce the rate of misdiagnoses. Because these data are collected at the state level, similar analysis can be used to generate results to help direct scarce public health resources more effectively. For example, one could target educational, prevention, and other resources toward institutions with relatively greater missed opportunities. We anticipate that such approaches will be more critical as TB becomes less common.

Acknowledgments

Financial support. This work funded by the National Heart Lung and Blood Institute at the National Institutes of Health (grant number K25HL122305; to L. A. P.) and the UI Health Care eHealth and eNovation Center (to P. M. P.).

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

References


