THE PERFORMANCE OF TUBERCULOSIS SPUTUM EXAMINATION AT DISTRICT HOSPITALS, KHON KAEN, THAILAND.

Keyword: tuberculosis, performance, sputum smear examination, district hospital

Somporn Sirsaenpong (1), MD.
Maarten Elink Schuurman (2), MD, MPH
Somdej Phitsontorn (1), MD, MPH.
Kriangsak Vaeeteewootacham (3), MD.
Samai Kungsaworn (4), MD
Aroon Chirawatkul (5), Mac.
Wchai Usawaphac (6), MD.

(1) Faculty of Medicine, Khon Kaen University
(2) Royal Tropical Institute, The Netherlands
(3) Office of Communicable Disease Control, Zone 6
(4) 6th Zonal Tuberculosis Control Centre
(5) Faculty of Public Health, Khon Kaen University
(6) Nampong District Hospital, Khon Kaen

พฤติกรรมการตรวจสแกนเพื่อหาเชื้อวัณโรคของโรงพยาบาลสุขุมวิท ณ จังหวัดขอนแก่น ประเทศไทย

สมพรท์ ศิรสาญปอง (1), พธ.
Maarten Elink Schuurman (2), MD, MPH.
สมเด็ด พิศมณี (1), พธ. ดอ.สม.
สมเกียรติ เวชกุลวิจารย์ (3), พธ.
สมัย กิจสาร (4), พธ.
ชูชน จิระกุลนาว (5), วชช.
วชัย อัศวภัคช์ (6), พธ.

(1) คณะแพทยศาสตร์, มหาวิทยาลัยขอนแก่น
(2) Royal Tropical Institute, The Netherlands
(3) สำนักงานควบคุมโรคติดเชื้อ 5
(4) สุขศึกษาโรงพยาบาล 6
(5) คณะสาธารณสุขศาสตร์, มหาวิทยาลัยขอนแก่น
(6) โรงพยาบาลสุขุมวิท จังหวัดขอนแก่น

บทคัดย่อ

องค์การ: ห้องปฏิบัติการโรคบกพร่องในโรงพยาบาลสุขุมวิท จำนวน 19 แห่ง และในศูนย์โรงเรียน จังหวัดขอนแก่น ประเทศไทย.

วัตถุประสงค์: เพื่อประเมินพฤติกรรมการตรวจสแกนเพื่อหาเชื้อวัณโรคโดยเจ้าหน้าที่ในห้องปฏิบัติการโรคบกพร่องใน โรงพยาบาลสุขุมวิท จังหวัดขอนแก่น และวิเคราะห์ผลให้มีการปรับปรุงการทำงาน.

วิธีการ: เป็นการศึกษาเพื่อสำรวจสภาพปัจจุบัน (Exploratory study) โดยวิเคราะห์ผลจากผลการตรวจสแกนของโรงพยาบาลสุขุมวิทจำนวน 19 แห่ง ที่มีผลผลิตสัมผัสกับการตรวจพบวัณโรค تفاصيلการตรวจและวิเคราะห์ผลการตอบแบบสอบถามที่มีค่าลัทธิสถิติการทดสอบของค่าสัมพันธ์ของผลผลิตสัมผัสกับการตอบแบบสอบถามของตัวแปรต่าง ๆ
สมบูรณ์

: ไม่ต้องมีผลต่อการจ่ายเงินจำนวน 1,057 แผน พบว่าการจ่ายแผนแต่ละกลุ่มระหว่างเรื่องปฏิบัติการรักษาสุขภาพของสุนัขในโรงพยาบาล ศัพท์เป็นสัตว์ 94 อย่างที่กล่าวมาจะร่าง 39 ของจำนวนแผนแต่ละกลุ่มมีคุณภาพของการบริการแผนแต่ละกลุ่มไม่ดี ในส่วนคุณภาพของแผน, การจ่ายแผนแต่ละ หรือการย้อมแผนแต่ละกลุ่ม ขึ้น ไม่คุ้มค่า แผนแต่ละกลุ่มที่มีคุณภาพสูงค่า แต่คุณภาพของการจ่ายแผนแต่ละกลุ่ม หรือการย้อมแผนแต่ละกลุ่มไม่ดี พบว่ามีผลต่อการจ่ายแผนแต่ละกลุ่ม เนื่องจากข้อมูลสำหรับการระดับ ขั้นตอนการจ่ายแผนแต่ละกลุ่ม คณะผู้รับผิดชอบแผนแต่ละกลุ่มจึงมีปฏิบัติการระดับ ขั้นตอนการจ่ายแผนแต่ละกลุ่ม และคุณภาพของแผน ทั้งที่มีผลต่อการสุนัขในโรงพยาบาล ศัพท์เป็นสัตว์ 94 อย่างที่กล่าวมาจะร่าง 39 ของจำนวนแผนแต่ละกลุ่มมีคุณภาพของการบริการแผนแต่ละกลุ่มไม่ดี

สรุป:

: ไม่ต้องมีผลต่อการจ่ายแผนแต่ละกลุ่มที่สั้นลง (องค์กร) เนื่องจากผลต่อขั้นตอน และผลต่อ แผนแต่ละกลุ่มที่มีคุณภาพสูงค่า แต่คุณภาพของการจ่ายแผนแต่ละกลุ่ม ไม่ดี ผลต่อการสุนัขในโรงพยาบาล ศัพท์เป็นสัตว์ 94 อย่างที่กล่าวมาจะร่าง 39 ของจำนวนแผนแต่ละกลุ่มมีคุณภาพของการบริการแผนแต่ละกลุ่ม ขึ้น ไม่คุ้มค่า แผนแต่ละกลุ่มที่มีคุณภาพสูงค่า แต่คุณภาพของการจ่ายแผนแต่ละกลุ่ม หรือการย้อมแผนแต่ละกลุ่มไม่ดี พบว่ามีผลต่อการจ่ายแผนแต่ละกลุ่ม และต้องคำนวณต้องคำนวณสำหรับการระดับ ขั้นตอนการจ่ายแผนแต่ละกลุ่ม คณะผู้รับผิดชอบแผนแต่ละกลุ่มจึงย้อมแผนแต่ละกลุ่ม ขึ้น ไม่คุ้มค่า แผนแต่ละกลุ่มที่มีคุณภาพสูงค่า แต่คุณภาพของการจ่ายแผนแต่ละกลุ่ม หรือการย้อมแผนแต่ละกลุ่มไม่ดี พบว่ามีผลต่อการจ่ายแผนแต่ละกลุ่ม เนื่องจากข้อมูลสำหรับการระดับ
Abstract:

Setting: Nineteen district hospital laboratories and a reference Zonal Tuberculosis control Centre laboratory in Khon Kaen, Thailand.

Objective: To assess the performance of sputum smear examinations by peripheral laboratory technicians in their own work setting.

Design: Exploratory study. 19 district hospital laboratories were asked to submit monthly a sample of sputum slides that were examined during 1990 and 1991. Those slides were reexamined by a reference laboratory technician. The reading results were compared and the influence of slide preparation on the reading result was analysed.

Results: In a sample of 1067 slides the agreement in reading results between the reference and peripheral laboratories was 94%. However, 39% of slides were not prepared well in sputum collection, smearing or staining. Among the slides with good sputum but with poor smearing or staining a significantly different proportion of positive slides were found compared with the well prepared slides when examined by the reference laboratory. The authors speculate that technicians in the district and reference laboratories make the same mistakes in reading these poorly prepared slides, so that a high agreement in results does not guarantee good performance.

Conclusion: Not only agreement in reading results but also the proportion well prepared slides is an important indicator for peripheral laboratory performance.

Keyword: tuberculosis, performance, sputum smear examination, district hospital

Request for reprints:
Assist. Prof. Dr. Somporn SRI-SIENPANG
Department of Community Medicine,
Faculty of Medicine, Khon Kaen University,
Khon Kaen, #0002, Thailand.

INTRODUCTION

The WHO strategy to improve case detection and cure rates in developing countries rests upon four cornerstones of national tuberculosis control programmes: a reliable sputum-smear microscopy service, registration of all tuberculosis patients, a secure drug supply, and supervision of programme activities at all levels (1). In developing countries tuberculosis is common but health care resources are limited and the major aim of tuberculosis control is to detect the infectious cases and render them rapidly non-infectious by modern short-course chemotherapy (2,3,4,5,6,7,8). In such a programme sputum microscopy is the mainstay of the diagnosis (9,10,11) since a skilled laboratory technician can detect the majority (80%) of infectious cases making use of this technique (9,12). The demonstration of the Mycobacterium tuberculosis in sputum by microscopy techniques is a clear indication of active tuberculosis (3,13). Sputum smear examination is not only vital for case detection and diagnosis but also for the follow-up of the response to chemotherapy (14,15,16). In areas where HIV infection is common it should be emphasized that the HIV epidemic does not change the strategy of tuberculosis control (17,18), and therefore the need for high quality microscopic sputum examination remains.

In Thailand the district hospital is the most peripheral health facility with a laboratory where sputum smear microscopy can be performed. A district hospital covers about 100,000 population. The quality control of these laboratories is carried out by the Regional Medical Science Centre. Four times per year two sputum test slides with a known bacterial load are sent to the peripheral laboratories.
for staining and reading. Written feedback of the test results is given to the laboratories. In this system the performance in staining and reading under routine field conditions (high workload, lack of motivation, not well maintained equipment) cannot be assessed and the performance in sputum collection and smearing is not assessed at all.

The MURLEP project (19) in Khon Kaen province was a multidisciplinary health systems research project aiming at the development of tools for tuberculosis and leprosy control programme managers to easily assess the quality of the programme and to identify appropriate interventions to improve performance. In the framework of this project we studied the performance of sputum smear examinations of peripheral laboratory technicians in their own work setting.

METHODS

During 12 months in 1990 and 1991 the 6th Zonal Tuberculosis Control Centre (TBZC), in its role as supervisory coordinator for 5 provinces in North-East Thailand, asked the district hospital laboratory technicians in Khon Kaen province to submit monthly the sputum slides they had prepared and examined from both TB suspects and patients registered on treatment, separately accompanied by the reading results. All 19 Khon Kaen Province district hospitals participated in this study. These slides (Ziehl-Neelsen stained) were reexamined by a senior laboratory technician of the TBZC, which is the reference laboratory for the Regional Medical Science Centre in sputum smear microscopy, within a month after the slide preparation at the district hospital. We regard the reading results of this laboratory technician as the gold standard. This technician was blinded about the reading results of the district hospital. He classified the quality of the sputum slide (sputum content, smearing and staining) according to the procedure described by Latini e.a. (20) : the sputum was classified as good, (mucous or mucopurulent) or poor (mainly saliva); the smearing was good or poor (too thin or too thick or uneven thickness); the staining was good or poor (precipitate, too red). The reading was done according to national standards which are the same as recommended by the International Union Against Tuberculosis and Lung Diseases (21). The slides were grouped in positive and negatives. A slide is negative when no AFB are found in 100 oil immersion fields. A Slide is positive when the reading result is 1+, 2+ or 3+. Doubtful slides are excluded. The results of the reference laboratory were compared with the district hospital laboratory results. The analysis consisted of determining agreement and the positive predictive value and the negative predictive value of the district hospital results. The findings were communicated to the technicians through the provincial control programme.

RESULTS

A total of 1057 slides were received from the 19 district hospitals and checked at the TBZC. Compared to the number of reported in the 6th Zonal Tuberculosis Control Centre Annual Report this is about 20% of the slides made during the year. In the beginning of the year the laboratory technicians were not yet accustomed to keep the slides, later they kept only the slides they thought had good quality. The selection bias in our slide population most likely leads to better results than in the whole population of slides made during the year.

The quality of slide preparation

The findings on quality of preparation are described in Table 1. About one-third were not good in one or more of the 3 stages of preparation. 15% of all slides contained only saliva, and are therefore completely unsuitable for finding bacilli. This may be
caused by the fact that also sputum specimens of patients on treatment, who may not easily cough up sputum any more, were taken. 647 (61%) slides were determined as "well prepared slides" which were good in all steps of sputum collection, smearing and staining. So 39% of all slides were poor at least in sputum collection, smearing or staining.

This table also shows the negative effect of a poorly conducted step in slide preparation on the following step. The effect of poor sputum collection on smearing is large: 96% of the poor sputum slides are poorly smeared, against 16% of the good sputum slides. The effect of poor staining on the following step, staining, is less but still considerable: 26% of the poorly smeared slides in poorly stained against 13% of the well smeared slides ($p<0.001$).

Quality of slide reading

Comparison of reading results between the district hospital laboratory and the TBZC (reference laboratory), shows a very good agreement, see table 2, section A.

In order to assess whether reader agreement was influenced by the quality of slide preparation, we tested the agreement in three different subgroups: B. Well prepared slides (all 3 qualities good, n=647), C. Slides with good sputum but poor smearing and/or staining (n=250) and D. Slides with poor sputum and poor smearing and/or staining (n=160), see table 2; section B,C and D.

For all subgroups the agreement remained about the same as for the group of all slides: 93%-94%, just as the negative predictive value of the district hospital results. However, the positive predictive value was smaller the slides with poor sputum and poor smearing and/or staining (group D).

It seems that if a slide is prepared badly the reference lab technician cannot read it much better than the DH lab technician. We therefore compared the percentage positive slides between the subgroups of slides as read by the reference laboratory, see figure 1.

The 40% positive slides in the group of well prepared slides (group B) served as the reference percentage for the subgroups of slides. Since smear preparation should not change the reading outcome this percentage should be found in all other groups. When we compare this percentage positive slides with that of group D it becomes clear that poor sputum collection is related to a very low percentage positive slides (significant difference, $p<0.001$). This is to be expected since few AFB will be present in saliva specimens. Although this low percentage positive smears may have come from follow-up patients who are almost cured, it may also indicate missed diagnoses when the laboratory technician is not keen on obtaining a good sputum specimen from a coughing patient.

When the slides contain good sputum but smearing and/or staining is poor (group C) we see a higher percentage positive slides ($p<0.05$). Further differentiation shows that poor staining (groups E and G) causes for more slides to be read as positive (51% and 61%, $p<0.05$) but that the combination poor smearing and good staining (group F) gives 40% positive slides, not different from the "true percentage"

**DISCUSSION**

In the 1057 slides the agreement in reading results between the reference and peripheral laboratories was 94%, but 15% of the slides did not contain good sputum (saliva) and 39% of the slides were poorly prepared by the district hospital laboratory technicians. Among the slides with good sputum but with poor smearing or staining a significantly different proportion of positive slides were found when compared with the well prepared slides when examined by the reference laboratory.
We concluded that technicians in the district and reference laboratories make the same mistakes in reading these poorly prepared slides, so that a high agreement in results does not guarantee good performance.

The sample of 1057 slides may not have been representative of all slides made during the year. The laboratory technicians most likely kept more positive slides and more slides that were well prepared. Since we aimed to look at the performance of the district hospital's technicians we assessed their performance in determining slides as positive separately from their performance in determining slides as negative in our sample of 1057 slides. In this way the positive bias in our sample did not influence the performance assessment outcome.

The saliva specimen might have been collected from patients in the follow-up phase of treatment, who often do not cough much any more and cannot easily produce sputum. From our study it is clear that it is rather meaningless to examine saliva specimens. Laboratory technicians should be trained and be supervised to collect good sputum specimens. Similar conclusions were made in Roelgard's, da Costa Santiago's and Pollak's study (22,23,24).

The effect of poor smearing on the reading result

We found poor smearing mostly among saliva specimens (Table 1), which is understandable: when there is no sputum one cannot smear it well. Among good sputum slides we found that poor smearing caused some more slides to be read as positive than among well prepared slides (Figure 1). So, poor smearing is associated with false positive reading.

The effect of poor staining on the reading result

We found that not only poor smearing but also poor staining causes more slides to be read as positive (Figure 1). So there are false positives among these slides. Apparently the poor staining causes more material to be stained red and falsely to be judged as acid fast bacteria by the laboratory technicians. This finding is similar to Latini's (20), who found an excess of false positive slides among poorly stained slides.

Agreement in reading results

The overall agreement between the district and reference laboratories was high (93-94%), whether the slides were well prepared or not (Table 2). The disagreement concerned mainly the PPV (87%) and not the NPV (99%). Latini (20) and Boujaibai (25) reported a similar reading agreements (98% and 95%). These relatively good agreements in reading results may be misleading because we have shown above that poorly prepared slides cause an excess of positive slides, also when examined by the reference laboratory technician. We conclude that poorly prepared slides cannot be read well neither by the peripheral laboratory technician, nor by the reference laboratory technician.

In group C (good sputum but poor smearing and/or staining) the district hospital laboratory diagnosed 51% of the slides positive (Table 2). This is 11% more than the "true" percentage of 40% positive among well prepared slides, as determined by the reference laboratory. This 11%, from 210 slides with good sputum but poor smearing and/or staining, translates into 28 probably false positive slides. Since it may be that two of these slides refer to one patient, we can speculate that among these patients at least 14 were wrongly diagnosed and put to treatment for tuberculosis. It is often argued "better safe than sorry". But over diagnosis obviously
causes problems: Toman's statement speaks for itself (26): "Man hours, drugs, films and finances are wasted on false positives, while the source of infection remains undetected and untreated. The effect on tuberculosis control is probably the same as underdiagnosis with the difference that such programmes cost more".

The method as presently used by the Regional Medical Science Centre in Thailand, which only checks staining and reading skills of laboratory technicians under "examination" circumstances does not seem adequate. Since it would cause a too heavy workload for the reference laboratory to reexamine all slides, we have developed a suitable slide sampling method per district hospital laboratory, which indicates whether performance is acceptable or not. This system is described in a separate paper.

Recommendations for quality control of peripheral laboratories are:
- Take regularly a sample of all slides kept for reexamination by the reference laboratory (27,28).
- Do not accept high agreement in slide reading result alone for good performance, since even experienced technicians cannot judge poorly prepared slides.
- The proportion well prepared slides is the first indicator for peripheral laboratory performance. The second indicators is the proportion correctly diagnosed slides, preferably among well prepared slides only.
- Supervisors should give clear feedback on the quality control results, stressing good slide preparation practices.

Acknowledgements

We are thankful to the inputs of other researchers and health service staff involved in the study: Dr. Sastra, Saewakontha, Dr. Pichet Leelapananetha, Dr. Kumron Chaisri, Dr. Pippop Sriapongrat, Mr. Udom Supunnawong and the staff of the TBZC. We thank Dr. M. Borgdorff and Dr. M. van Clift for their useful comments. The Netherlands Legality Relief Association is gratefully acknowledged for their financial support.

REFERENCES


Table 1. The distribution of 1057 district hospital laboratory sputum slides according to the 3 preparation qualities, as determined by the reference laboratory.

<table>
<thead>
<tr>
<th></th>
<th>all slides: 1057</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>good (85%)</td>
</tr>
<tr>
<td>Smearng</td>
<td>good 750 (84%)</td>
</tr>
<tr>
<td></td>
<td>good 7 (4%)</td>
</tr>
<tr>
<td>Staining</td>
<td>good 647 (87%)</td>
</tr>
<tr>
<td></td>
<td>good 109 (74%)</td>
</tr>
<tr>
<td></td>
<td>good 7 (100%)</td>
</tr>
<tr>
<td></td>
<td>good 84 (55%)</td>
</tr>
</tbody>
</table>

Table 2. Comparison of slide reading of all slides and different subgroups, between the district hospital laboratory and the reference laboratory (RL).

<table>
<thead>
<tr>
<th></th>
<th>Results of district hospital laboratory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>positive slides</td>
</tr>
<tr>
<td></td>
<td>No. confirmed</td>
</tr>
<tr>
<td></td>
<td>By RL</td>
</tr>
<tr>
<td>B. Good sputum, smearng and staining (well prepared slides)</td>
<td>292</td>
</tr>
<tr>
<td>C. Good sputum, poor smearng and/or staining</td>
<td>128</td>
</tr>
<tr>
<td>D. Poor sputum, poor smearng and/or staining</td>
<td>19</td>
</tr>
<tr>
<td>A. All slides</td>
<td>438</td>
</tr>
</tbody>
</table>

* PPV: Positive predictive value
** NPV: Negative predictive value
Figure 1. Percentages positive slides in different subgroups as determined by reference laboratory:

- A. all slides: 38% positive
- B. good: 40% positive
- C. 47% positive
- D. poor (saliva): 7% positive
- E. poor: 51% positive
- F. good: 43% positive
- G. poor: 61% positive