Why Are Clinicians Reluctant to Treat Smear-Negative Tuberculosis? An Inquiry about Treatment Thresholds in Rwanda

Paulin Basinga, Juan Moreira, Zeno Bisoffi, Bettina Bisig and Jef Van den Ende
Med Decis Making 2007; 27; 53
DOI: 10.1177/0272989X06297104

Why Are Clinicians Reluctant to Treat Smear-Negative Tuberculosis? An Inquiry about Treatment Thresholds in Rwanda

Paulin Basinga, MD, Juan Moreira, MD, Zeno Bisoffi, MD, Bettina Bisig, MD, Jef Van den Ende, PhD

Purpose. *The diagnosis of tuberculosis remains controversial between clinicians and public health officers. Public health officials fear to treat too many patients; clinicians fear that truly diseased will be denied treatment. We wondered whether an analysis of the treatment threshold could help making the often intuitive decision to treat smear-negative cases more evidence based.* ***Methods.*** *Eighteen clinicians and 10 public health specialists were asked for an intuitive estimate of their treatment threshold for tuberculosis and of key determinant factors for this threshold: the magnitude and subjective weight of mortality and morbidity due to both the disease and the treatment and risk and cost of the latter. With these factors, the authors calculated treatment thresholds and compared them to the intuitive thresholds of the interviewees. A prescriptive threshold was calculated based on literature data, omitting cost and subjective factors.* ***Results.*** *The median overall intuitive treatment threshold was 52.5%, the calculated 11.9%, and the prescriptive 2.7%. For 2 factors, public health officers provided significantly lower values than clinicians: cost of treatment (median = $20 v. $300;* $U = 2.5$*;* $P = 0.0002$*); cost of life (median = $500 v. $5000;* $U = 17.5$*;* $P = 0.009$*).* ***Conclusion.*** *These results suggest that clinicians and public health officers estimate wrongly the threshold even when using their own subjective estimate of influencing factors. Omitting treatment cost and subjective weight of provoked harm can result in a very low threshold. Sound training in threshold principles and providing tools to correctly assess data might help in making better decisions in tuberculosis in developing countries.* ***Key words:*** *threshold; tuberculosis; Rwanda; smear negative. (**Med Decis Making 2007;27:53-60**)*

Tuberculosis (TB) remains, despite the efforts of prevention and treatment, one of the most serious public health problems throughout the world.[1] In Rwanda, and most sub-Saharan African countries, this problem is amplified by the AIDS pandemic.[2]

In our fieldwork, over many years, we often found clear TB cases untreated, with frightening stories, both in developed and developing countries. When we questioned doctors, they told us the diagnosis was not certain, so they could not treat. When we explained briefly the threshold philosophy, they easily changed their minds, especially when they learned the real cost of 6 months of treatment ($11 international market price!).

After many years of formal clinical logic teaching (including thresholds) in various countries in 4 continents, we wondered what made all these clinicians put their threshold so high. If we could identify some particular beliefs or misbeliefs, we might focus on these influencing factors in our future teaching.[3]

Part of the problem arises from the well-known debate between clinicians and public health decision

Received 24 June 2005 from the School of Public Health, National University of Rwanda (PB); Institute of Tropical Medicine, Antwerp, Belgium (JM, BB, JVdE); Centro per le Malattie Tropicali, Negrar, Verona, Italy (ZB); Centro de Epidemiología Comunitaria y Medicina Tropical, Esmeraldas, Ecuador (JM); and Centre Hospitalier Universitaire de Kigali, Rwanda (JVdE). This work was previously presented at the Third European Congress on Tropical Medicine and International Health, Lisbon, Portugal, 8–11 September 2002, and at the NITD Symposium on Tuberculosis, Bagamoyo, Tanzania, 17–20 October 2005. Revision accepted for publication 3 April 2006.

Address correspondence to Jef Van den Ende, Clinical Sciences, Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerp, Belgium; e-mail: jvde@itg.be.

DOI: 10.1177/0272989X06297104

makers. Most of the latter are convinced that clinicians at the level of referral hospitals treat too many suspect patients without enough evidence. They are more concerned with resources and argue that clinicians make speculative rather than evidence-based decisions. In fact, the national program against tuberculosis in Rwanda suggests that not more than 15% of all TB treatments should be prescribed without bacteriological evidence.[4] Clinicians are often frustrated by these strict rules imposed on them by public health deciders. They work at the front lines of public health in the field but are confronted with a lack of a true gold standard to justify their decisions on tuberculosis treatment. In referral hospitals, they deal with a lot of smear-negative cases, and they argue that as the sensitivity of sputum examination is not higher than 60%, too many false-negative cases would be left untreated if they strictly adhered to the national guidelines.[5–7] (In pulmonary TB, patients can cough up secretions that contain Koch's bacilli, which can be detected by applying Ziehl's stain to a thin smear of these secretions and examining this under the microscope. A positive smear is considered as absolute proof, but it is not always present.)

This debate could be solved by tackling the problem from a threshold analysis perspective.[8] This approach, well developed in theory but rarely used in practice, would probably give a rational solution to this feeling of uncertainty.

There is a lot of confusion about thresholds. If we do not consider some disease-specific thresholds, as we have, for instance, in malaria,[9,10] the common meaning of *decision threshold* is the minimum posttest probability you need to treat the individual patient, also known as the probabilistic standard of proof.[11] This concept is different from the proportion of treated nondiseased patients/all treated patients a community can accept, often mistaken for the decision threshold (e.g., in Rwanda, some clinicians think they need 85% of posttest probability to treat TB because the national program accepts no more than 15% of all TB treatments without bacteriological evidence). The decision threshold is also sometimes confounded with the ratio of treated nondiseased to not-treated diseased, known as the Blackstone-like error ratio.[11]

For the decision threshold, 2 methodologies are described: The Pauker-Kassirer approach solves the utility equation for the final posttest probability, and the Zweig and Campbell approach solves the utility equation for sensitivity and specificity of a given test or a combination of tests in a score, with a fixed (and known) prevalence.[8,12] The threshold is defined as the (posttest) probability of disease at which the risk or cost of the diseased left untreated equals the risk or cost of treatment of diseased and nondiseased. Probabilities and values assigned both to the harm and the benefit of a treatment must be taken into account to estimate in a formal way the required level of certainty to justify a treatment. Pauker and Kassirer have further split their threshold in 2 (test threshold and test-treatment threshold), refining the concepts but creating a lot of confusion, as we observe with our students.[13] DeKay has shown that the use of the Blackstone-like error ratio does not lead to a defensible recommendation for policy.[11] Neither does a formal prescriptive algorithm exist for the proportion of treated nondiseased patients a community can accept.

We think clinicians do not have the time or perhaps the interest to perform complicated mathematical calculations. However, they probably use a kind of intuitive threshold they cannot make explicit but that is based on their knowledge and/or perception about the harm and benefit of a treatment. Neither these intuitive values, from which a required level of certainty is supposed to be derived, nor their relation to a formally calculated procedure have ever been explored in the field of pulmonary TB. We found only 1 single publication on the threshold value for TB treatment: Kopelmann and others computed an extremely low value of 0.5% for miliary TB in the United States.[14]

With our study of a representative group of clinicians and public health decision makers in Rwanda, we intend to explore 1) if the intuitive treatment threshold is high indeed; 2) which key determinant factors could be responsible for the high threshold; 3) if the problem lies in the estimation of the key factors or in the intuitive computing, that is, the taking into account; and 4) what a prescriptive threshold based on available literature data would look like. The results could also offer a valuable bridge between clinicians and public health decision makers and suggest key messages in training in decision analysis in Rwanda and in other developing countries.

MATERIAL AND METHODS

All medical professionals dealing with TB, both in 2 referral hospitals in Rwanda (Centre Hospitalier Universitaire de Kigali and Centre Hospitalier

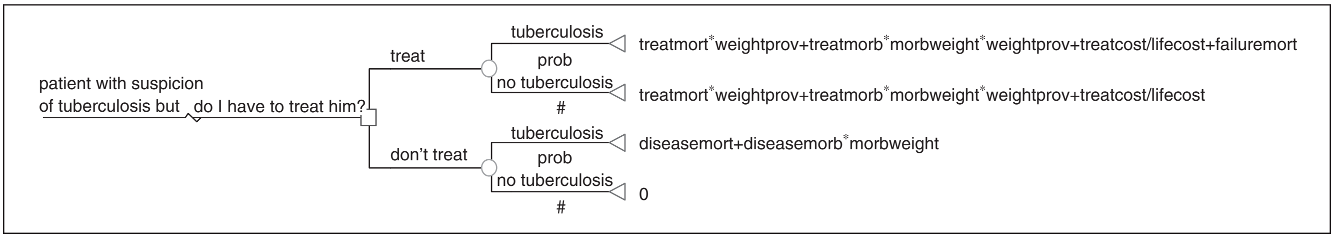

Figure 1 Threshold computation for pulmonary tuberculosis: the total harm per branch depends on a series of fixed values (for abbreviations, see the "Material and Methods" section) and on 1 unknown value, the probability of the disease. For a certain value of this probability, both treatment and no-treatment options will be in balance. This point is the threshold.

Universitaire de Butare) and in the National Program for Tuberculosis Control, as well as other public health professionals involved in the TB policy development in Rwanda, were asked to discuss their treatment thresholds. These 2 hospitals are the only reference centers in Rwanda admitting most of the TB cases characterized as difficult or dubious. Current treatment of smear-negative TB is decided by some of the respondents on a case-by-case basis, most often by consensus between several clinicians but without a formal protocol.

After an introductory explanation of the objectives of the study and the general concept of threshold, some key questions were asked by one of us (P.B.). As a surrogate for the intuitive treatment threshold, we applied the virtual cohort example of Pauker and Kassirer.[8] We asked the respondent how many patients without tuberculosis he or she would be ready to treat in order not to leave untreated a truly (but unknown) disease with a low posttest probability. Further questions focused on possible factors influencing the threshold estimate: mortality and morbidity of disease and treatment, efficacy, and cost of treatment. For comparison between mortality and morbidity, we asked how much a certain disease would subtract from the vital capacity, with full vital capacity being 1 and death being 0.[15] For the value of life estimated as a cost, we asked what the interviewee would be willing to pay to stop an imminent life threat against himself or herself.[16,17]

The last question concerned the number of people who would be infected by an untreated pulmonary TB patient. This item was not included in the following analysis as uncertain cases are smear negative by definition; therefore, their risk of spread is rather low: It is estimated at 11% (0.11), or as 100 smear-negative untreated patients will infect 11 persons.[18]

A decision tree to frame the problem, to integrate the probabilities and values, and to estimate the disutilities (risk and/or cost) was drawn (Figure 1). For a certain probability of disease,

disutility of treatment = disutility of withholding treatment.

If we take into account only mortality, the equation would be

$$\text{(All patients)} \times \text{(mortality by treatment)} + \text{(probability of disease)} \times \text{(mortality by treatment failure)} = \text{(probability of disease)} \times \text{(mortality by disease)}, \quad (1)$$

or

$$x = \frac{\mathit{treatmort}}{\mathit{diseasemort} - \mathit{failuremort}}, \quad (2)$$

where

- treatmort: probability of death due to the treatment,
- diseasemort: probability of death due to nontreated TB,
- failuremort: probability of death due to a treatment failure, and
- x: probability of disease, threshold.

Adding all factors, the following formula can be used to compute the calculated threshold:

$$x = \frac{\mathit{treatmort} \times \mathit{weightprov} + \mathit{treatmorb} \times \mathit{morbweight} \times \mathit{weightprov} + \mathit{treatcost} / \mathit{lifecost}}{\mathit{diseasemort} - \mathit{failuremort} + \mathit{diseasemorb} \times \mathit{morbweight}}, \quad (3)$$

where

- diseasemorb: probability of morbidity due to nontreated TB,
- lifecost: value of life estimated as a cost,
- morbweight: weighed value of morbidity with respect to mortality,
- treatcost: direct cost of a short course treatment,

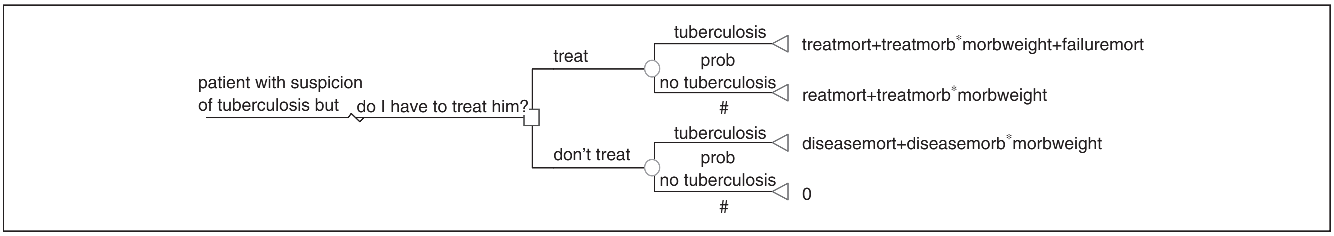

Figure 2 Threshold computation for pulmonary tuberculosis with restricted determinants. Cost and weight of provoked harm are withdrawn.

- treatmorb: probability of morbidity due to treatment, and
- weightprov: value of a provoked death, weighed relative to a natural death.

We also calculated a threshold based on data derived from international literature. Cost was withdrawn because treatment is entirely paid for by the government, and clinicians' subjective weight of harm was omitted because it should not interfere with a decision (patients' perspective; equation 4; Figure 2). Finally, this prescriptive threshold was compared with a restrictive threshold based on the interviewee's estimation but without cost and weight of provoked harm (omission bias).[19]

$$x = \frac{treatmort + treatmorb * morbweight}{diseasemort - failuremort + diseasemorb * morbweight}. \quad (4)$$

Data obtained from interviews were summarized with medians and quartiles according to professional group (public health background or not) and years of experience (junior, <5 years; senior, ≥5 years). The Mann-Whitney test was used to compare the difference of medians between independent groups.

An attempt to compute confidence intervals (95%) around the median of individual calculated thresholds showed extremely wide and nonsense values. Therefore, we generated a calculated threshold with a Monte Carlo simulation, sampling values randomly from table-type probability distributions, based on values between the 1st and 3rd quartile (q1–q3) of all estimated factors in a 1st simulation and of restricted factors in a 2nd simulation.

The analysis was performed using Epi Info 6.04 (CDC, WHO), SPSS for Windows 13.0.1, TreeAge Pro 2004 (TreeAge Software Inc.), and Microsoft Excel.

RESULTS

General

Twenty-eight of 38 professionals agreed to participate in the study. Respondents included 18 clinicians and 10 public health officers. The clinician group included 4 specialists and 8 residents in internal medicine and 4 specialists and 2 residents in pediatrics, none with a public health background. The public health officer group included 10 professionals with formal training in public health, who had several degrees of responsibility for the TB policy in Rwanda (decision makers) but were not involved in the individual care of patients. The 10 nonrespondents were either not available or did not want to participate.

Factors

The estimates of probabilities and values by respondents are summarized in Table 1. No significant difference was found for most factors between public health officers and clinicians, except for cost of treatment (median = $20 v. $300, U = 2.5, P = 0.0002) and cost of life (median = $500 v. $5000, U = 17.5, P = 0.009). Among clinicians, there was not a significant difference in the estimation of costs between their experienced and nonexperienced peers; however, a difference was found for the estimation of mortality due to a nontreated disease (median = 87.5% v. 60%, U = 14; P = 0.018), morbidity due to a nontreated disease (median = 12.5% v. 25%, U = 15.5, P = 0.026), and mortality due to the treatment (median = 1.5% v. 0.35%, U = 13, P = 0.014) . For literature data, we retrieved 11 relevant articles, 8 of which reported risk of treatment.[20–30] Except for

Table 1 Probabilities and Values of Influencing Factors

	Median of Estimations	Q1–Q3	Number of Respondents	Median of Literature Data	Range of Literature Data	Number of Retrieved Studies[a]
Disease mortality	75%	50%–85%	28	55%	49%–66%	3
Disease morbidity	22.5%	15%–25%	28	19%	18%–19%	3
Treatment morbidity	5%	2%–11.25%	28	4.9%	1.8%–8.6%	7
Treatment mortality	1%	0.28%–1%	28	0.1%	0.08%–0.12%	2
Mortality related to the failure of treatment	1.15%	1%–2%	10	10%	0.7%–15.6%	3
Cost of treatment (US$)	135	30–462.5	24	—	—	—
Weighted value of morbidity relative with mortality	0.25	0.1–0.4	28	—	—	—
Weighted value of provoked v. natural harm	3	2–5	28	—	—	—
Cost/value of life (US$)	5000	1000–10,000	23	—	—	—

a. For references, see text.

Table 2 Summary of Thresholds

Data Collection	Interviewees	Factors	Method	Threshold (%)
Interview	All		Median of intuitive thresholds	52.5
Interview	Public health		Median of intuitive thresholds	60
Interview	Clinicians		Median of intuitive thresholds	36.5
Interview	Senior clinicians		Median of intuitive thresholds	36.5
Interview	Junior clinicians		Median of intuitive thresholds	41.5
Interview	All	All	Calculated threshold on medians	11.9
Interview	All	Restricted	Calculated threshold on medians	2.8
Literature		Restricted	Calculated threshold on medians	2.7

treatment- and failure-related mortality, literature-derived data and estimates by interviewees were in the same order of magnitude (formal statistics do not make sense, as the method of collection is different in both arms).

Thresholds

The median intuitive threshold as estimated by the respondents as a whole group was 52.5% (Table 2). Although the median of the intuitive thresholds given by public health officers (60%) is higher than the one given by clinicians (36.5%), the difference was not significant ($U = 73.5$, $P = 0.426$). Nor was a significant difference found between experienced (median = 36.5%) and less experienced clinicians (median = 41.5%; $U = 34.0$, $P = 0.591$).

The threshold calculated with equation 3 based on the medians of all factors given by the interviewees was 11.9%; a Monte Carlo simulation resulted in a median of 15.2%, with a confidence interval of 4.7% to 65%. If only restricted factors were considered, the threshold was 2.8%, with a median by Monte Carlo simulation of 3.4% (confidence interval = 0.3%–25%).

Using data derived from the literature (medians and full range of considered [restricted] factors are depicted in Table 1), a prescriptive treatment threshold of 2.7% was calculated. A Monte Carlo simulation did not make sense, as data were sparse and scattered.

DISCUSSION

Our study found that the calculated treatment threshold for pulmonary TB in Rwanda, based on data from both interviewees and international literature, is quite low. This required certainty is probably similar in other developing countries. Thresholds intuitively estimated by the entire group were higher than formally calculated ones, and surprisingly, there was no statistically significant difference in intuitive thresholds between clinicians and public health officers, although this might be due to the small sample size. These findings should contribute to the revision of the clinical guidelines for smear-negative TB in Rwanda and other developing countries. When the intuitive estimate of cost and provoked harm was withdrawn, the calculated threshold came close to the prescriptive threshold. Destroying the myth of expensive antituberculous treatment and emphasizing the patient's perspective might be the most important messages our study suggests.

Methodology

Measuring values and incorporating them into the decision-making process ideally should be done from different points of view.[31,32] The perspective of the policy maker could be quite different from the perspective of the treating clinician and of the patient. We studied the difference between the perspective of the clinician and the patient.

One of the critical points was the estimate of an intuitive threshold. The difference between the threshold, the proportion of treated nondiseased patients/all treated patients, and the Blackstone-like error ratio is confusing: Some clinicians could be inclined to think that if a threshold is set at, for example, 20%, it would imply that 80% of patients would be erroneously treated or that for every 4 nondiseased treated patients, 1 diseased patient is not treated.[11,31] An explanation of the difference between the required probability (threshold, depending on utilities) and the obtained probability (posttest probability, conditional on prevalence and disease characteristics' discriminative value) is also of extreme importance to avoid confusion. All these definitions were thoroughly explained to the interviewees to obtain correct intuitive estimations.

Probabilities (e.g., probability of treatment side effects) were easier to define than values (e.g., the weight of provoked harm): Information about the former was available from the literature, whereas values depend on the perspective (of the clinician, the patient, the community).

Clinicians are more concerned by harm than by utility: With undesirable outcomes having a higher psychological impact, clinicians tend to recall more harm than success[33,34]; moreover, humans are more restrictive for harm provoked by action than by no action (omission bias).[19,35–37]

Two major values of harm were asked: the weight of a provoked versus a natural bad outcome and the cost of a life. The estimate of the former was quite consistent among the entire group: A provoked harm was 2 to 5 times more important than a natural bad outcome.

The cost of life was a hard topic to discuss. Its purpose was to estimate the equivalence between suffering and cost. We tried to solve the problem through a willingness-to-pay approach.[16,17] Although the estimated values of this factor showed a broad range, it was an easier way to ask such a difficult question. An alternative approach to estimate the cost of a life could have been the amount that an average Rwandese family might pay to save the life of a child. This approach would not be useful in Rwanda or other impoverished countries because most of the population could never afford the estimated cost of the treatment, resulting in an ever too high required certainty. The pertinence of these methods can be discussed, but it makes a bridge between mortality, morbidity, and cost of treatment.

We did not include the indirect cost. The amount spent on consultations, direct observation of therapy, organization of central purchase and distribution of drugs, and organization of a national reference laboratory is probably higher than the amount directly spent on drugs. On the other hand, threshold analysis concerns dubious cases. Fortunately, they represent only a fraction of all treated cases (J. Mugabekazi, unpublished), so we have to include only the marginal indirect cost, which is probably rather low.

Differences between factors as estimated by respondents versus those found in the literature were not formally compared, as this would require a meta-analysis, and as available, data are very heterogeneous. We relied only on ranges for difference estimation.

HIV infection was not included in the factor analysis, although we considered it in the first phase of the survey. After many hours of discussion, it was not clear in what direction HIV infection would alter the threshold: TB treatment is equally effective in HIV-positive and HIV-negative patients; excess mortality is mainly due to HIV infection itself and other opportunistic infections; TB treatment would postpone ARV treatment, if indicated; TB treatment would go with increased toxicity.

RESULTS

Threshold theory is not penetrating in medicine. The question arises as to why. First, the association of threshold theory with speculative or presumptive diagnosis creates a lot of confusion.[38] It is as if this theory gives clinicians a free hand, and nothing is more false. On the contrary, this theory provides a solid logical framework and strict rules for presumptive diagnosis. It also regulates the widely accepted therapeutic trial, a test with a combination of 4 antituberculous drugs for a few weeks, confirming the diagnosis if the symptoms disappear. It is challenged especially in TB because the drugs might cover a wide array of other infectious diseases. Second, a lot of clinicians are confused by the difference between required evidence and obtained evidence. A cause of this confusion might be that diagnostic criteria for certain diseases have traditionally been presented with sets of symptoms and signs and never with mathematical criteria.[39]

A recent review analyzed the relevant literature on the diagnosis of smear-negative pulmonary TB in low-income countries.[7] Surprisingly, none of the studies mentioned a threshold approach. One of them, carried out in the neighboring countries of Burundi and Tanzania, found a positive predictive value of 43% following a combination of any 2 among a given panel of clinical criteria.[40] The authors seem to think that it is too low and discuss alternative, more specific approaches. But the central question would be, Is a disease probability of at least 0.43 enough to undertake a treatment for TB or not? This is, precisely, the threshold notion.

When considering the entire group, intuitive thresholds were higher than thresholds computed from intuitive data. These findings could appear strange if we suppose that intuitively estimated data directly influence the final threshold estimate. Our study shows that at least there is a problem of computing or estimating by the human mind: Most of the estimated influencing factors are close to literature data, but intuitive calculation (estimation) by the interviewees does not correspond with the formally computed thresholds. A possible explanation for this systematic intuitive error could be that intuitive threshold levels are inflated in the context of the strong pressure by the Ministry of Health to limit the treatment of smear-negative cases to a small percentage.[41] The fact that some influencing factors are not included in the calculation might interfere also. These omissions may be due to the fact that these factors are not easily quantifiable (e.g., the stigmatization of the disease). The risk of spreading the infection (which would lower the threshold) and the risk of developing a drug resistance (which would inflate it) were not considered either, although clinicians and public health officers could intuitively take them into account.

Further Research

To explore the difference between the required (threshold) and the reached (posttest) probability, it would be interesting to seek real data from an investigation of a cohort of patients. In a developing country, this is not easy to do or to fund. This approach could also inform us about the more relevant factors (probabilities and values) to be taken into account for a TB treatment in a smear-negative patient and about the accuracy of a scoring system for estimating the posttest probability.

When looking for the relative weight of some factors, we did not specify the difference between a natural death and a death due to withholding a treatment in a truly diseased patient (omission bias).[19,35–37] We supposed that interviewees would understand natural death as withholding treatment, but it was not explained clearly. Moreover, we did not differentiate between deaths provoked by the treatment in nondiseased versus truly diseased patients. This might affect our figures. In future research, this question should be addressed.

The question arises as to whether the results of this study might be generalized. On one hand, 73% (28 of 38) of main deciders in TB issues in Rwanda were interviewed, adding to the representativeness of the study. On the other hand, the study assumes that all complicated cases of TB are transferred from health centers and district hospitals to the referral hospital where decisions regarding the treatment are made, as recommended by the national TB program. In reality, this is not the case because of difficulties of transferring a suspect but smear-negative patient to the referral hospital. In fact, doctors working in a district hospital make a lot of decisions, often without the consensus rule applied at the reference level. Future research should formally include the district level. Perhaps no inference of the results to other developing countries is possible. As it is an important issue to discuss, we need to keep in mind that these data coming from Rwanda could be considered as coming from a single case study or from an experiment. For this reason, an analytic generalization as used for case studies in social sciences could be a more relevant approach.[42] Nevertheless, the best way to solve this question could be a multicentric and/or multicountry approach.

REFERENCES

1. World Health Organization. Global Tuberculosis Control—Surveillance, Planning, Financing. Geneva (Switzerland): World Health Organization; 2002.

2. Colebunders R, Bastian I. A review of the diagnosis and treatment of smear-negative pulmonary tuberculosis. Int J Tuberc Lung Dis. 2000;4:97–107.

3. Pauker SG, Pauker SP, Wong JB. The utility of thresholds and the threshold utility. Med Decis Making. 1987;7:201–2.

4. Programme National Intégré de Lutte. Epidémiologie de la tuberculose au Rwanda. Kigali (Rwanda): Programme National Intégré de Lutte; 2000.

5. Aber VR, Allen BW, Mitchison DA, Ayuma P, Edwards EA, Keyes AB. Quality control in tuberculosis bacteriology. 1. Laboratory studies on isolated positive cultures and the efficiency of direct smear examination. Tubercle. 1980;61:123–33.

6. Levy H, Feldman C, Sacho H, van der MH, Kallenbach J, Koornhof H. A reevaluation of sputum microscopy and culture in the diagnosis of pulmonary tuberculosis. Chest. 1989;95:1193–7.

7. Siddiqi K, Lambert ML, Walley J. Clinical diagnosis of smear-negative pulmonary tuberculosis in low-income countries: the current evidence. Lancet Infect Dis. 2003;3:288–96.

8. Pauker SG, Kassirer JP. Therapeutic decision making: a cost-benefit analysis. N Engl J Med. 1975;293:229–34.

9. Van den Ende J, Lynen L. Diagnosis of malaria disease. In: Carosi G, Castelli F, editors. Handbook of Malaria Infection in the Tropics. Bologna (Italy): Organizzazione per la Cooperazione Sanitaria Internazionale; 1997. p 137–50.

10. Van den Ende J, Mukaminega M. Attributable fraction of malaria in Mali by Dicko et al. (2005). Trop Med Int Health. 2005; 10:1207.

11. Dekay ML. The difference between Blackstone-like error ratios and the probabilistics standards of proof. Law Soc Inq. 1996;21:95–132.

12. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem. 1993;39:561–77.

13. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med. 1980;302:1109–17.

14. Kopelman RI, Wong JB, Pauker SG. Clinical problem-solving: a little math helps the medicine go down. N Engl J Med. 1999; 341:435–9.

15. Ubel PA, Loewenstein G, Schwarz N, Smith D. Misimagining the unimaginable: the disability paradox and health care decision making. Health Psychol. 2005;24(4 suppl):S57–62.

16. Gafni A. Willingness-to-pay as a measure of benefits: relevant questions in the context of public decisionmaking about health care programs. Med Care. 1991;29:1246–52.

17. O'Brien B, Gafni A. When do the "dollars" make sense? Toward a conceptual framework for contingent valuation studies in health care. Med Decis Making. 1996;16:288–99.

18. Behr MA, Warren SA, Salamon H, et al. Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. Lancet. 1999;353:444–9.

19. Ritov I, Baron J. Protected values and omission bias. Organ Behav Hum Decis Process. 1999;79:79–94.

20. Springett VH. Ten-year results during the introduction of chemotherapy for tuberculosis. Tubercle. 1971;52:73–87.

21. Tuberculosis in a rural population of South India: a five-year epidemiological study. Bull World Health Organ. 1974;51:473–88.

22. Grzybowski S, Enarson D. Results in pulmonary tuberculosis patients under various treatment program conditions. Bull Int Union Tuberc. 1978;53:70–5.

23. Zierski M, Bek E. Side-effects of drug regimens used in short-course chemotherapy for pulmonary tuberculosis: a controlled clinical study. Tubercle. 1980;61:41–9.

24. Gonzalez Montaner LJ, Dambrosi A, Manassero M, Dambrosi VM. Adverse effects of antituberculosis drugs causing changes in treatment. Tubercle. 1982;63:291–4.

25. Combs DL, O'Brien RJ, Geiter LJ. USPHS Tuberculosis Short-Course Chemotherapy Trial 21: effectiveness, toxicity, and acceptability. The report of final results. Ann Intern Med. 1990;112: 397–406.

26. Steele MA, Burk RF, DesPrez RM. Toxic hepatitis with isoniazid and rifampin: a meta-analysis. Chest. 1991;99:465–71.

27. Ormerod LP, Horsfield N. Frequency and type of reactions to antituberculosis drugs: observations in routine treatment. Tuber Lung Dis. 1996;77:37–42.

28. Garner P, Holmes A. Tuberculosis. Clin Evid. 2002;7:747–56.

29. Yee D, Valiquette C, Pelletier M, Parisien I, Rocher I, Menzies D. Incidence of serious side effects from first-line antituberculosis drugs among patients treated for active tuberculosis. Am J Respir Crit Care Med. 2003;167:1472–7.

30. van Hest R, Baars H, Kik S, et al. Hepatotoxicity of rifampin-pyrazinamide and isoniazid preventive therapy and tuberculosis treatment. Clin Infect Dis. 2004;39:488–96.

31. Hunink M, Glasziou P, Siegel J, et al. Decision making in health and medicine: integrating evidence and values. Cambridge (UK): Cambridge University Press; 2001.

32. Holmes MM, Rovner DR, Rothert ML, et al. Women's and physicians' utilities for health outcomes in estrogen replacement therapy. *J Gen Intern Med.* 1987;2:178–82.

33. Feinstein AR. The "chagrin factor" and qualitative decision analysis. Arch Intern Med. 1985;145:1257–9.

34. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science. 1981;211:453–8.

35. Asch DA, Baron J, Hershey JC, et al. Omission bias and pertussis vaccination. Med Decis Making. 1994;14:118–23.

36. Cohen BJ, Pauker SG. How do physicians weigh iatrogenic complications? J Gen Intern Med. 1994;9:20–3.

37. Baron J, Spranca M. Protected values. Organ Behav Hum Decis Process. 1997;70:1–16.

38. Kassirer JP. Our stubborn quest for diagnostic certainty: a cause of excessive testing. N Engl J Med. 1989;320:1489–91.

39. Shulman ST. T. Duckett Jones and his criteria for the diagnosis of acute rheumatic fever. Pediatr Ann. 1999;28:9–12.

40. Samb B, Kamanfu G, Henzel D, et al. Diagnostic des tuberculoses à bacilloscopie négative en milieu hospitalier Est Africain. Int J Tuberc Dis. 1997;1:25–30.

41. Christensen-Szalanski JJ, Bushyhead JB. Physician's use of probabilistic information in a real clinical setting. J Exp Psychol Hum Percept Perform. 1981;7:928–35.

42. Yin R. Case Study Research: Design and Methods. 3rd ed. Thousand Oaks (CA): Sage; 2003.