A chain-binomial model for intra-household spread of *Mycobacterium tuberculosis* in a low socio-economic setting in Pakistan

S. AKHTAR[1,2]*, T. E. CARPENTER[3,4] AND S. K. RATHI[2]

[1] *Department of Community Medicine and Behavioural Sciences, Faculty of Medicine, Kuwait University, Safat, Kuwait*
[2] *Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan*
[3] *Department of Medicine and Epidemiology, University of California, Davis, CA, USA*
[4] *Center for Animal Disease Modeling and Surveillance, School of Veterinary Medicine, University of California, Davis, CA, USA*

SUMMARY

A simulation study using Greenwood's chain-binomial model was carried out to elucidate the spread and control of *Mycobacterium tuberculosis* among the household contacts of infectious pulmonary tuberculosis (TB) patients. Based on the observed data, the maximum-likelihood estimates ($\pm$ S.E.) of chain-binomial probabilities of intra-household *M. tuberculosis* transmission from an index case in 3-person and 4-person households were 0.313 ± 0.008 and 0.325 ± 0.009 respectively. The χ^2 goodness-of-fit test of observed and simulated mean expected frequencies of cases revealed good fit for 3-person ($P = 0.979$) and 4-person ($P = 0.546$) households. With the assumption of varying risk of *M. tuberculosis* transmission across the households under β-distribution, goodness-of-fit tests of observed and mean simulated expected frequencies revealed the inadequacy of Greenwood's chain-binomial model both for 3-person ($P = 0.0185$) and 4-person ($P < 0.001$) households. Simulated *M. tuberculosis* control strategy comprising efficient diagnosis, segregation and prompt antibiotic therapy of index pulmonary TB patients showed a substantial reduction of new cases among the household contacts in both household sizes. In conclusion, segregation coupled with prompt antibiotic therapy of the index case, chemoprophylaxis of *M. tuberculosis*-exposed household contacts, and the assessment of household environmental risks to devise and implement an educational programme may help reduce the TB burden in this and similar settings.

INTRODUCTION

Tuberculosis (TB) caused by *Mycobacterium tuberculosis* infection is one of the leading causes of mortality that accounts for 26% of preventable adult deaths in the developing world [1, 2]. Varying levels of endemicity of *M. tuberculosis* infection have been reported worldwide and South-East Asia including

Pakistan seems to be afflicted the most, where 44% of its population is reportedly infected by *M. tuberculosis* [3]. Screening of high-risk populations for *M. tuberculosis* infection to identify individuals eligible for chemoprophylaxis is one of the recommended strategies to combat TB [4, 5]. Household contacts of TB cases constitute a high-risk group for acquiring *M. tuberculosis* infection [6, 7]. It has been reported that $\sim 30\%$ of close contacts demonstrate evidence of *M. tuberculosis* infection and at least half of the *M. tuberculosis*-infected contacts exhibit progression to disease in the first 2 years [8]. However,

* Author for correspondence: S. Akhtar, Ph.D., Department of Community Medicine and Behavioural Sciences, Faculty of Medicine, Kuwait University, PO Box 24923, Safat 13110, Kuwait. (Email: saeed.akhtar@hsc.edu.kw)

consistent estimates of *M. tuberculosis* transmission rates in household contacts do not exist.

Infection with *M. tuberculosis* is dependent on nonlinear contact processes that are determined by population density as well as other factors. It is believed that occasional contacts with infectious TB patients rarely leads to transmission and that most secondary cases are the results of prolonged and sustained close contact with a primary case [9]. Classical TB epidemic models incorporate heterogeneity via proportional or weighted random mixing [10], or non-random mixing at the population level [11]. However, these extensions operate at a single mixing level and consequently their use in the study of TB dynamics neglects important features, namely the relative impact of close contacts. We report herein the results from a simulation study that used Greenwood's chain-binomial model to elucidate the spread and control of *M. tuberculosis* among household contacts of infectious TB patients.

MATERIAL AND METHODS

Epidemiological investigations of *M. tuberculosis* infection in household contacts

The study procedures including description of study design, setting and study population have been reported elsewhere [12], and are briefly outlined here. A cross-sectional study was conducted during August–September 1999 in Umerkot Taluka – an administrative sub-unit of the district of Umerkot, Sindh Province. It has population of ∼0·3 million. Agriculture is the main occupation in Umerkot. About 57% of the housing units are single-room houses with an average household size of five persons. The local health-care system in Umerkot Taluka is comprised of a government health centre, a non-governmental organization clinic, namely Umerkot Anti-tuberculosis Association (UATA) clinic, and about 20 private clinics. TB control is provided for most patients of this study area by the UATA clinic. The majority of patients are diagnosed and treated as outpatients. The index TB case was defined as the first member of a household, aged ⩾15 years, who presented and registered for TB treatment at the UATA clinic and was found positive on at least one acid fast bacilli (AFB) sputum-smear test [13, 14]. A list frame of 680 TB cases maintained at UATA clinic was obtained and used to select a simple random sample of 77 index cases based on the assumptions reported previously [12].

The household contact was defined as any person who was over 3 months of age, had lived in the same house as the index case for at least 3 months and had slept in the same house for an average of at least four nights per week for at least 3 months of co-habitation [15]. During the first visit to the house of a selected TB index patient, verbal informed consent was obtained from all persons aged ⩾12 years and from the parents of household contacts aged ⩽12 years who fulfilled contact definition. Household contacts of the index case were subjected to a tuberculin skin test (TST) using the Mantoux method in order to categorize them as infected or non-infected with *M. tuberculosis* [16, 17]. In total, 385 household contacts of 77 index patients were included in the study. The mean (±S.D.) and median number of contacts per case interviewed and tested for TST positivity were 5·4±3·1 and 4·0 respectively. The prevalence of *M. tuberculosis* infection among household contacts was 49·4% (95% CI 46·9–51·9). There was also a direct linear relationship between household size and the number of *M. tuberculosis*-infected contacts. The study protocol was reviewed and approved by the departmental ethics review committee of the Aga Khan University Hospital.

Greenwood's chain-binomial model

Greenwood's chain-binomial model was used to investigate the infectiousness of *M. tuberculosis* to the contacts of TB index cases only in household sizes of 3 and 4 [18], as epidemic patterns for these household sizes are more consistent with the Greenwood-type model, while epidemic patterns for the larger household sizes tend to be more compatible with a Reed–Frost type model [19]. We were unable to identify the links of the chain within households; therefore, the data on the total number of cases that occurred in each household were used. Greenwood's chain-binomial model takes the following form:

$$_aP_{nj} = \sum_{k=1}^{j-a} \binom{n-a}{k} p^k q^{n-a-k} {}_kP_{n-a,\,j-a},$$

where $_aP_{nj}$ = the probability that a group of total size n will have total j cases when there are a introductions (where $a \leqslant j \leqslant n$); k = number of cases and $n-a-k$ = uninfected susceptibles; $_kP_{n-a,j-a}$ = the probability of a further $J-a$ new cases (including k) to make j in all.

The model was programmed in spreadsheet software (Excel v. 7.0; Microsoft Corp., Redmond, WA, USA).

Model assumptions

Any disease qualifying for the application of Greenwood's chain-binomial model must have the following features:

(1) it must be highly infective;
(2) its transmission period must be short relative to the incubation period and where these are not known, there must be evidence of time-clustering of cases in an epidemic situation;
(3) it must produce lasting immunity after an infection;
(4) its incubation period must be constant.

The above criteria in general seem to be satisfied by the characteristics of *M. tuberculosis* infection. But for criterion (3), the duration of immune state following *M. tuberculosis* infection is quite variable; however, identified infectious cases undergo antibiotic therapy, and no longer remain infectious. For criterion (4), the incubation period of *M. tuberculosis* infection is somewhat variable, this assumption is, therefore, slightly violated but assuming that all the contacts were equally susceptible because of poor nutritional status and genetic homogeneity the mean incubation period could be considered as constant (3 months).

Parameter estimates

The chain-binomial probability (P) of *M. tuberculosis* transmission from infectious index case to household contact was estimated from the observed data using the maximum-likelihood method [18] separately for both 3- and 4-person households with spreadsheet add-in (Solver; Microsoft Corp.). The maximum-likelihood estimates of probabilities of *M. tuberculosis* transmission for the two household sizes were further used in model simulations. It was also assumed that the probability of acquisition of infection and household transmission of *M. tuberculosis* infection was homogeneous across all households. With an additional assumption that the average chance of infection from index case to susceptible may vary between different households according to a β-distribution, the maximum-likelihood estimates of P as a function of x and y (parameters of β-distribution) were obtained for both household sizes using spreadsheet add-in (Solver; Microsoft) and used in model simulations. The following function of the β-distribution was maximized to seek the estimates of P

$$\mathrm{d}F(p, x, y) = \frac{1}{B(x, y)} p^{x-1} q^{y-1} dp, \quad 0 \leqslant p \leqslant 1 \ [20],$$

where $\mathrm{d}F(p, x, y) =$ the probability density function as the derivative of cumulative distribution function; $B(x, y) = \beta$-function with parameters x and y, and $x > 0$ and $y > 0$; $p =$ probability of infection of susceptibles and is assumed to be same for all members of a given household and only varies between households; $q = 1 - p$.

Model simulation

A pre-constructed model in a commercially available spreadsheet (Solver; Microsoft) and simulation add-in for Monte Carlo sampling was used (@Risk, Palisades Corp., Newfield, NY, USA). For each of the two household sizes 1000 simulations were run by using the above calculated estimates of probabilities (P) of *M. tuberculosis* transmission from the index case. The proportions of simulations for the expected data for each household size were then used to compute the expected frequency distribution of *M. tuberculosis* transmission among contacts and the P value for goodness-of-fit χ^2 statistic in each case was obtained. Another set of 1000 simulations for each household size was run using estimates of x and y, the parameters of β-distribution and P values of goodness-of-fit χ^2 statistic were obtained as above.

To evaluate control strategies, including efficient diagnosis, segregation and prompt antibiotic therapy of the index case to prevent the spread of *M. tuberculosis* to household contacts, an improved chain-binomial probability of *M. tuberculosis* transmission was used (80% reduction in the estimated probabilities of transmission). The aforementioned proposed TB control strategy basically mimics the WHO recommended DOTS (directly observed therapy short course) strategy [21], which, if effectively employed is expected to substantially reduce *M. tuberculosis* transmission by achieving a corresponding cure rate of $\geqslant 80\%$. For each of the household sizes, 1000 simulations were run to obtain the expected distribution of cases. A χ^2 goodness-of-fit test was carried out to compare the observed and mean fitted frequency distributions of cases for each of the two household sizes using an in-built software routine to evaluate the effect of control strategies.

RESULTS

Maximum-likelihood estimates ($\pm$ S.E.) of the chain-binomial probability of *M. tuberculosis* transmission within household size of 3 ($n_1 = 4$) and household size

Table. *Observed and mean simulated (1000 simulations) expected frequencies of infected household contacts in a study of* Mycobacterium tuberculosis *spread from index cases under Greenwood's chain-binomial probability model*

| No. of *M. tuberculosis* infected contacts | Household size (number of susceptible household contacts plus the index infectious tuberculosis case) | | | | | |
| | Household size ($n=3$) | | | Household size ($n=4$) | | |
	O	E_{cb}	Prop	O	E_{cb}	Prop
(*a*) Assumption: uniform risk of *M. tuberculosis* transmission across households						
1	2	1·99	0·497	4	3·82	0·347
2	1	1·16	0·289	1	2·22	0·202
3	1	0·86	0·214	4	2·29	0·208
4	—	—	—	2	2·67	0·243
	P value = 0·978			*P* value = 0·546		
(*b*) Assumption: varying risk of *M. tuberculosis* transmission across households under β-distribution						
1	2	1·99	0·464	4	1·276	0·116
2	1	1·16	0·328	1	0·902	0·082
3	1	0·86	0·208	4	8·822	0·802
4	—	—	—	2	0·000	0·0
	P value = 0·0185			*P* value < 0·0001		

O, Observed number of *M. tuberculosis*-infected contacts; E_{cb}, expected number of *M. tuberculosis*-infected contacts under chain-binomial distribution; Prop, proportion of the 1000 simulations that yielded the E_{cb}; *P* value, *P* value associated with the χ^2 statistic computed for goodness-of-fit test of observed and expected case frequencies.

of 4 ($n_2 = 11$) were $0·313 \pm 0·008$ and $0·325 \pm 0·009$ respectively. Based on these maximum-likelihood estimates of *M. tuberculosis* transmission probabilities, simulated mean expected frequencies for 1, 2 or 3 cases or 1, 2, 3, or 4 cases in 3-person and 4-person households are given in the Table (part *a*). The χ^2 goodness-of-fit test of observed and expected frequencies revealed good fit both for 3-person ($P = 0·978$) and 4-person ($P = 0·546$) households.

The chain-binomial model was also evaluated under the assumption that the average probability of *M. tuberculosis* infection from the TB index case to household contacts might have been varied across the households under β-distribution. The mean simulated expected frequencies of cases under β-distribution with parameter estimates for 3-person ($x = 1·14$, $y = 0·27$) and 4-person ($x = 0·98$, $y = 0·28$) households, indicated an inadequacy of the chain-binomial model as revealed by χ^2 goodness-of-fit tests of observed and expected data both for 3-person ($P = 0·0185$) and 4-person ($P < 0·0001$) households (Table, part *b*). The results of simulated *M. tuberculosis* control strategy including, efficient diagnosis, segregation and prompt antibiotic therapy of the index case to prevent the spread of *M. tuberculosis* to household contacts showed significant ($P < 0·0001$) decrease in expected frequencies of *M. tuberculosis*-infected contacts in households of both sizes (Fig.).

DISCUSSION

The chain-binomial model is an appropriate model for an infection that spreads directly from one person to another within the household contacts of an index case. Furthermore, in comparison with a simple binomial model, the chain-binomial model fits better when the data are scanty or the household size is 3 or 4 persons [18]. One of the assumptions for a disease to be eligible for the application of a chain-binomial model is that the risk is uniform across all the households of a given size [18].

Maximum-likelihood estimates ($\pm$ S.E.) of chain-binomial probability of *M. tuberculosis* transmission within household sizes of 3 and 4 were $0·313 \pm 0·008$ and $0·325 \pm 0·009$ respectively. Relatively higher probability (0·46) of *M. tuberculosis* transmission has been reported among household contacts of infected patients in a high incidence area [22]. In contrast, substantially lower annual risk of *M. tuberculosis* transmission has been reported from South-East Asia (0·001–0·0023) and South and Central America (0·005–0·015) [23].
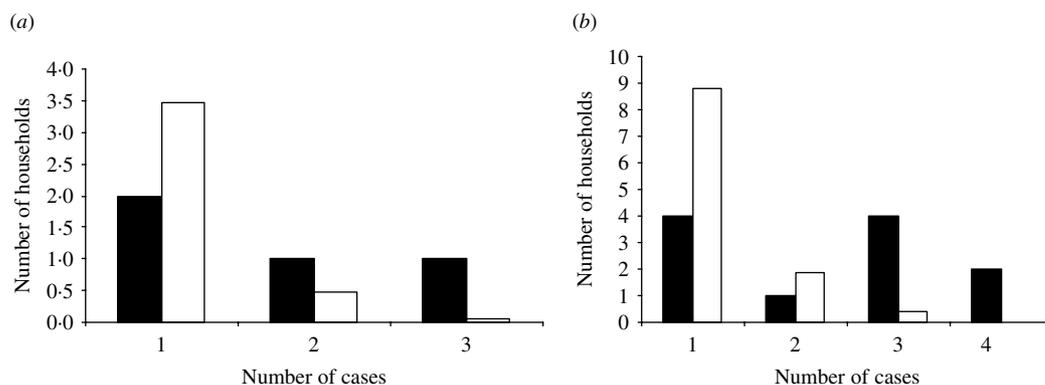
(a)

(b)

Fig. Comparison of observed (■) and simulated (□) (1000 simulations) case frequencies after instituting the control programme in (*a*) 3-person households (χ^2 *P* value $= 0.00019$) and (*b*) 4-person households (χ^2 *P* value < 0.0001).

We simulated the observed data under both assumptions, i.e. uniform risk of transmission across the households and variable risk of transmission between households under β-distribution. Our results showed excellent fit of observed and expected frequency distributions under the chain-binomial model with assumption of uniform risk of *M. tuberculosis* transmission, whereas, the chain-binomial model under the assumption of β-distribution exhibited poor fit. However, we cannot rule out several confounding effects which may influence the risk of *M. tuberculosis* transmission across households including age of contact, duration of cough, presence of pulmonary cavity and grade of sputum smear of index patient [24]. We have also previously demonstrated that in this particular population of household contacts there was a linear increase in *M. tuberculosis* positivity with increase in household size [12]. The chain-binomial model is also an appropriate model if there is an increase in subsequent attack rate with an increase in household size [18]. However, epidemic patterns in larger household sizes is aptly more compatible with the Reed–Frost type model [19].

The mechanism of person-to-person spread of *M. tuberculosis* infection is in-line with its known epidemiology. The fit of chain-binomial model-based simulated results with actual experience in the 3-person and 4-person households suggests that *M. tuberculosis* may be spread by a person from inside the family. It has been argued that in areas with high TB incidence, the risk of transmission from sources beyond the index case in the household may be much higher than in areas with low incidence [22, 25]. However, no extra-household risk factor for the acquisition of *M. tuberculosis* infection among these household contacts could be identified [12].

We had a sample of observed data for only four and 11 housing units for 3-person and 4-person household sizes and the maximum-likelihood estimates of chain-binomial probability of *M. tuberculosis* transmission from the TB index case to susceptible household contacts were 0·313 and 0·325 respectively. These estimates of *M. tuberculosis* transmission probabilities are quite high and perhaps reflect the high infectiousness (AFB smear positive) of index cases. Furthermore, the poor nutritional status of the susceptible population in the study area might have enhanced their vulnerability to *M. tuberculosis* infection. Therefore, these estimates of *M. tuberculosis* transmission probabilities should be taken into considerations in the design of any control programme.

The TB control strategy evaluated in this study was the efficient diagnosis, segregation and prompt antibiotic therapy of an index case, which closely follows the WHO recommended DOTS approach [21]. Assuming that strategy, when put in place, would reduce the estimated probability of *M. tuberculosis* transmission by 80%, the simulated results of the chain-binomial model predicted a substantial and significant reduction in the number of *M. tuberculosis*-infected household contacts of infectious TB index cases. We are not aware of any simulation study that has used a chain-binomial model to evaluate control strategies. However, it has been shown that early diagnosis and antibiotic therapy narrows down the time window of infectiousness of index cases and consequently reduces the risk of *M. tuberculosis* infection among household contacts [26]. In this simulation study, with the proposed control strategy, we targeted reduction in *M. tuberculosis* transmission by 80% in accordance with WHO recommendations. Furthermore, in India [15, 27, 28]

and other developing countries [29], the WHO recommended DOTS strategy [21] enabled the achievement of the desired cure rate of ⩾80% of TB cases with smear-positive sputum, resulting in a substantial reduction in annual risk of new infection.

In conclusion, *M. tuberculosis* infection is common among the household contacts of index cases in a low socio-economic setting in Pakistan as has been reported in similar settings in other parts of the world [30]. Simulation results of the chain-binomial model also predicted the increased risk of *M. tuberculosis* infection among household contacts. Therefore, improved diagnostic methods, segregation coupled with prompt antibiotic therapy of the index case, chemoprophylaxis of TST-positive household contacts of the index case, and assessment of household environmental risks to devise and implement an educational programme may reduce the TB burden in this and similar settings in the region.

ACKNOWLEDGEMENTS

DECLARATION OF INTEREST

None.

REFERENCES

1. **Small PM.** Tuberculosis research: balancing the portfolio. *Journal of the American Medical Association* 1996; **276**: 1512–1513.
2. **Rajeswari R, et al.** Socio-economic impact of tuberculosis on patients and family in India. *International Journal of Tuberculosis and Lung Disease* 1999; **3**: 869–877.
3. **Dye C, et al.** Global burden of tuberculosis-estimated incidence, prevalence, and mortality by country. *Journal of the American Medical Association* 1999; **282**: 677–686.
4. **Gourevitch MN, et al.** Lack of association of induration size with HIV infection among drug users reacting to tuberculin. *American Journal of Respiratory and Critical Care Medicine* 1996; **154**: 1029–1033.
5. **Behr MA, et al.** Predictive value of contact investigation for identifying recent transmission of *Mycobacterium tuberculosis*. *American Journal of Respiratory and Critical Care Medicine* 1998; **158**: 465–469.
6. **MacIntyre CR, Plant AJ.** Preventability of incident cases of tuberculosis in recently exposed contacts. *International Journal of Tuberculosis and Lung Disease* 1998; **2**: 56–61.
7. **Liippo KK, Kulmala K, Tala EOJ.** Focusing tuberculosis contact tracing by smear grading of index cases. *American Review of Respiratory Disease* 1993; **148**: 235–236.
8. **Binkin NJ, et al.** Tuberculosis prevention and control activities in the United States: an overview of the organization of tuberculosis services. *International Journal of Tuberculosis and Lung Disease* 1999; **3**: 663–674.
9. **Raffalii J, Sepkowitz KA, Armstrong D.** Community based outbreaks of tuberculosis. *Archives of International Medicine* 1996; **156**: 1053–1060.
10. **Feng Z, Castillo-Chavez C, Capurro AF.** A model for tuberculosis with exogenous re-infection. *Theoretical Population Biology* 2000; **57**: 235–247.
11. **Busenberg S, Castillo-Chavez C.** A general solution of the problem of mixing sub-population and its application to risk and age structure epidemic models for the spread of AIDS. *Journal of Mathematics Applied in Medicine and Biology* 1991; **8**: 1–29.
12. **Rathi SK, et al.** Prevalence and risk factors associated with tuberculin skin test positivity among household contacts of smear-positive pulmonary tuberculosis cases in Umerkot, Pakistan. *International Journal of Tuberculosis and Lung Disease* 2002; **6**: 851–857.
13. **Narain R, et al.** Distribution of tuberculosis infection and disease among households in a rural community. *Bulletin of the World Health Organization* 1966; **34**: 639–654.
14. **Datta M, et al.** Critical assessment of smear positive pulmonary tuberculosis programme. *Tubercle and Lung Disease* 1993; **74**: 180–186.
15. **Kamat SR, et al.** A controlled study of the influence of segregation of tuberculosis patients for year on the attack rate of tuberculosis in a 5 year period in close family contacts in South India. *Bulletin of the World Health Organization* 1998; **76**: 109–124.
16. **Alcaide J, et al.** Cigarette smoking as a risk factor for tuberculosis in young adults: a case-control study. *Tubercle and Lung Disease* 1996; **77**: 112–116.
17. **Al-Kassimi FA, et al.** The significance of positive Mantoux reactions in BCG-vaccinated children. *Tubercle* 1991; **72**: 101–104.
18. **Greenwood M.** On the statistical measure of infectiousness. *Journal of Hygiene* 1931; **31**: 336–351.
19. **Daley DJ, Gani J.** *Epidemic Modelling: An Introduction*. Cambridge: Cambridge University Press, 1999, 222 pp.
20. **Bailey NTJ.** The use of chain-binomials with variable chance of infection for the analysis of intra-households epidemics. *Biometrika* 1953; **40**: 279–286.
21. **Maher D, Mikulencak M.** What is DOTS? A guide to understanding the WHO-recommended TB control strategy known as DOTS. WHO/CDS/TB/99.270. Geneva: WHO, 1999.

22. **Verver S,** *et al.* Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet* 2004; **363**: 212–214.

23. **Murray CJ, Styblo K, Rouillon A.** Tuberculosis in developing countries: burden, intervention and cost. *Bulletin of the International Union against Tuberculosis* 1990; **65**: 6–26.

24. **Lienhardt C.** From exposure to disease: the role of environmental factors in susceptibility to and development of tuberculosis. *Epidemiological Review* 2001; **23**: 288–301.

25. **Lockman S,** *et al.* Molecular and conventional epidemiology of *Mycobacterium tuberculosis* in Botswana: a population-based prospective study of 301 pulmonary tuberculosis patients. *Journal of Clinical Microbiology* 2001; **39**: 1042–1047.

26. **Jaramillo E.** Encompassing treatment with prevention: the path of a lasting control of tuberculosis. *Social Science and Medicine* 1999; **49**: 393–404.

27. **Khatri GR, Frieden TR.** The status and prospects of tuberculosis control in India. *International Journal of Tuberculosis and Lung Disease* 2004; **4**: 193–200.

28. **Chadha VK,** *et al.* Annual risk of tuberculosis infection in the northern zone of India. *Bulletin of the World Health Organization* 2003; **81**: 573–580.

29. **Dye C,** *et al.* Evaluating the impact of tuberculosis control: number of deaths prevented by short course chemotherapy. *International Journal of Epidemiology* 2000; **29**: 558–564.

30. **Guwatudde D,** *et al.* Tuberculosis in household contacts of infectious cases Kampala, Uganda. *American Journal of Epidemiology* 2003; **158**: 887–898.