Diagnostic approaches for paediatric tuberculosis by use of different specimen types, culture methods, and PCR: a prospective case-control study

Richard A Oberhelman, Giselle Soto-Castellares, Robert H Gilman, Luz Caviedes, Maria E Castillo, Lenka Kolevic, Trinidad Del Pino, Mayuko Saito, Eduardo Salazar-Lindo, Eduardo Negron, Sonia Montenegro, V Alberto Laguna-Torres, David A J Moore, Carlton A Evans

Summary

Background The diagnosis of pulmonary tuberculosis presents challenges in children because symptoms are non-specific, specimens are difficult to obtain, and cultures and smears of *Mycobacterium tuberculosis* are often negative. We assessed new diagnostic approaches for tuberculosis in children in a resource-poor country.

Methods Children with symptoms suggestive of pulmonary tuberculosis (cases) were enrolled from August, 2002, to January, 2007, at two hospitals in Lima, Peru. Age-matched and sex-matched healthy controls were enrolled from a low-income shanty town community in south Lima. Cases were grouped into moderate-risk and high-risk categories by Stegen-Toledo score. Two specimens of each type (gastric-aspirate, nasopharyngeal-aspirate, and stool specimens) taken from each case were examined for *M tuberculosis* by auramine smear microscopy, broth culture by microscopic-observation drug-susceptibility (MODS) technique, standard culture on Lowenstein-Jensen medium, and heminested IS6110 PCR. Specimens from controls consisted of one nasopharyngeal-aspirate and two stool samples, examined with the same techniques. This study is registered with ClinicalTrials.gov, number NCT00054769.

Findings 218 cases and 238 controls were enrolled. 22 (10%) cases had at least one positive *M tuberculosis* culture (from gastric aspirate in 22 cases, nasopharyngeal aspirate in 12 cases, and stool in four cases). Laboratory confirmation of tuberculosis was more frequent in cases at high risk for tuberculosis (21 [14·1%] of 149 cases with complete specimen collection were culture positive) than in cases at moderate risk for tuberculosis (one [1·6%] of 61). MODS was more sensitive than Lowenstein-Jensen culture, diagnosing 20 (90·9%) of 22 patients compared with 13 (59·1%) of 22 patients (p=0·015), and *M tuberculosis* isolation by MODS was faster than by Lowenstein-Jensen culture (mean 10 days, IQR 8–11, *vs* 25 days, 20–30; p=0·0001). All 22 culture-confirmed cases had at least one culture-positive gastric-aspirate specimen. *M tuberculosis* was isolated from the first gastric-aspirate specimen obtained in 16 (72·7%) of 22 cases, whereas in six (27·3%), only the second gastric-aspirate specimen was culture positive (37% greater yield by adding a second specimen). In cases at high risk for tuberculosis, positive results from one or both gastric-aspirate PCRs identified a subgroup with a 50% chance of having a positive culture (13 of 26 cases).

Interpretation Collection of duplicate gastric-aspirate specimens from high-risk children for MODS culture was the best available diagnostic test for pulmonary tuberculosis. PCR was insufficiently sensitive or specific for routine diagnostic use, but in high-risk children, duplicate gastric-aspirate PCR provided same-day identification of half of all culture-positive cases.

Funding National Institutes of Health.

Published Online
July 26, 2010
DOI:10.1016/S1473-3099(10)70141-9

See Online/Reflection and Reaction
DOI:10.1016/S1473-3099(10)70145-6

Tulane School of Public Health and Tropical Medicine, New Orleans, LA, USA (Prof R A Oberhelman MD); **Asociación Benéfica Proyectos en Informatica, Salud, Medicina, y Agricultura (PRISMA), Lima, Peru** (G Soto-Castellares MD, Prof R H Gilman MD, M Saito MD, D A J Moore MD); **Department of Microbiology, Faculty of Sciences and Philosophy** (Prof R H Gilman, L Caviedes MS, M Saito, D A J Moore, Prof C A Evans MD), **and Department of Pediatrics, Faculty of Medicine** (Prof M E Castillo MD), **Universidad Peruana Cayetano Heredia, Lima, Peru; Infectious Diseases Service, Instituto Nacional de Salud del Niño; Lima, Peru** (Prof M E Castillo, L Kolevic MD, T Del Pino MD); **Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA** (Prof R H Gilman, M Saito, D A J Moore, Prof C A Evans); **Department of Pediatrics, Hospital Nacional Cayetano Heredia, Lima, Peru** (Prof E Salazar-Lindo MD, E Negron MD); **Universidad de Concepción; Concepción, Chile** (S Montenegro PhD); **US Naval Medical Research Center Detachment, Lima, Peru** (V A Laguna-Torres MD); **and Wellcome Trust Centre for Clinical Tropical Medicine and Deptartment of Infectious Diseases and Immunity, Imperial College, London, UK** (D A J Moore, Prof C A Evans)

Introduction

Although incidence of tuberculosis in adults has decreased in many high-burden countries, assessment of tuberculosis incidence in children remains difficult because of the absence of a dependable gold-standard test. Children account for an estimated 20% of the total tuberculosis caseload in high-incidence communities,[1–3] but most surveillance programmes only count acid-fast bacillus smear-positive cases, excluding more than 95% of presumptive tuberculosis cases in children younger than 12 years.[4] Adequate diagnosis of paediatric tuberculosis is difficult because of the lack of sputum production and paucity or absence of organisms in respiratory secretions, since tuberculosis bacilli typically remain confined to perihilar nodes that do not rupture into the bronchus.[5]

Technical and economic factors compound these diagnostic problems. The best available diagnostic tests are costly, whereas traditional methods are slow and insensitive. Even under optimum circumstances, *Mycobacterium tuberculosis* is isolated in fewer than 50% of children thought to have tuberculosis clinically.[6–10] Consequently, physicians often rely on poorly validated scoring systems.[11–13] Even in 2010, we still mainly depend on tools available since the 1950s to presumptively diagnose paediatric tuberculosis: purified protein derivative (PPD) skin test, chest radiography, history, and physical examination, usually without bacteriological confirmation.

The primary goal of this study was to assess new methods for diagnosis of pulmonary tuberculosis in

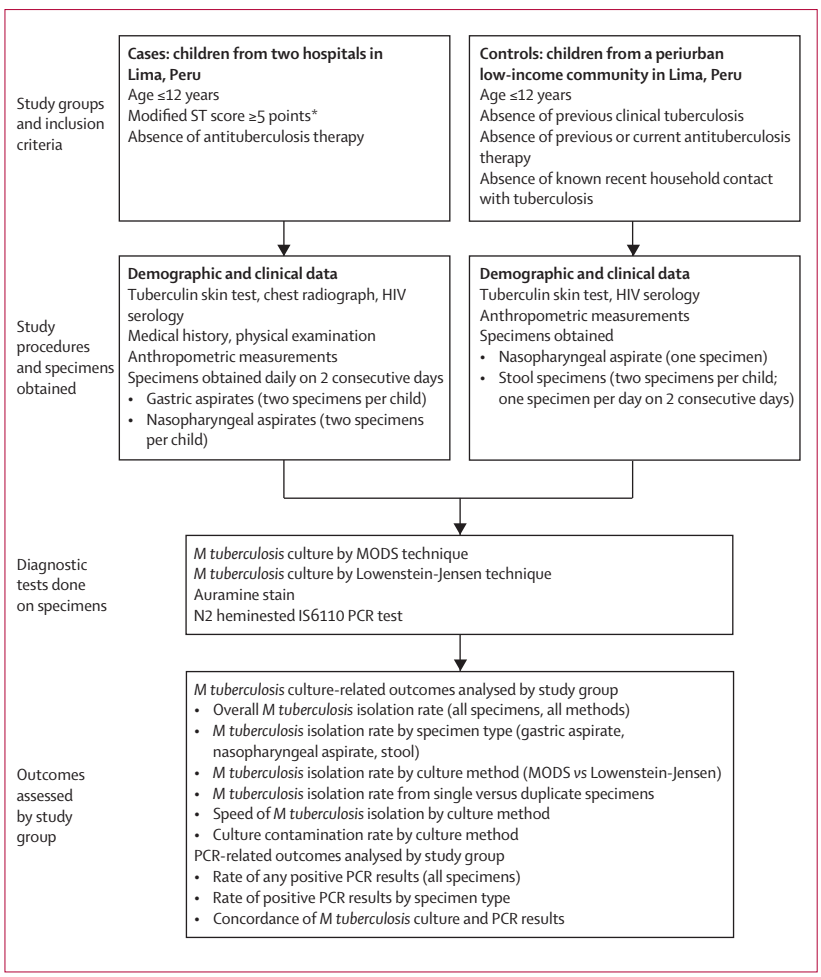

Figure 1: **Study design**
*Patients were classified according to the Stegen-Toledo (ST) score as moderate risk (score 5–6 points) and high risk (score ≥7 points). MODS=by microscopic-observation drug-susceptibility.

Correspondence to:
Prof Richard A Oberhelman, Department of Tropical Medicine, Tulane School of Public Health, 1440 Canal Street, SL 29, New Orleans, LA 70112, USA
oberhel@tulane.edu

children, including a heminested PCR assay (for potential rapid results), improved culture methods, and alternative clinical specimens. We analysed three outcomes in Peruvian children with suspected pulmonary tuberculosis. First, we examined *M tuberculosis* recovery from non-invasive specimens, such as nasopharyngeal aspirates and stool samples, compared with gastric aspirates, from single and duplicate specimens. Our previous data suggest that nasopharyngeal aspirates are a useful and less invasive alternative to gastric aspirates for diagnosis of tuberculosis in children,[14] and stool is potentially useful for tuberculosis PCR. Second, we analysed the speed and sensitivity of *M tuberculosis* isolation by microscopic-observation drug-susceptibility (MODS) culture compared with conventional Lowenstein-Jensen culture, and determined the added benefit of repeated testing. This study allowed us to expand our previous experience of the MODS technique in children[15] to establish the extent to which the enhanced microbiological sensitivity of this technique translates into clinical usefulness for diagnosis of tuberculosis in children. Third, we assessed the sensitivity, specificity, and predictive values of a heminested PCR assay compared with *M tuberculosis* culture. PCR technology might be available in reference laboratories in resource-poor countries, providing results within 6 h that can be used in initial clinical decision making. An age-matched and sex-matched control group was included to determine test specificity.

Methods

Cases and controls

The study design is summarised in figure 1. This study included cases with clinical evidence suggestive of pulmonary tuberculosis and healthy controls matched for age and sex. Children with evidence of HIV infection or AIDS were excluded. Cases were enrolled between April, 2002, and January, 2007, at the Instituto de Salud del Niño and the Hospital Nacional Cayetano Heredia in Lima, Peru. Paediatricians assessing children with respiratory illness in clinics at the two participating hospitals referred patients that they suspected of having pulmonary tuberculosis to a physician co-investigator or study nurse, who then assessed their eligibility for participation. Cases were classified according to the clinical criteria of Stegen and colleagues[16] (ie, Jones score in Africa) for diagnosis of paediatric tuberculosis as revised by Toledo and colleagues,[17] with modifications. This Stegen-Toledo scoring system is widely used throughout Latin America, and criteria used to determine Stegen-Toledo score are shown in table 1. Inclusion criteria for cases were age 12 years or younger, Stegen-Toledo score of 5 points or more, and absence of antituberculosis therapy. Although positive *M tuberculosis* culture is one of the Stegen-Toledo criteria, culture results were a primary outcome of this study and were not available at enrolment; therefore, this criterion was not included in our modified scoring system for participant eligibility. Cases were classified as either at moderate risk (Stegen-Toledo score 5–6 points) or high risk (score ≥7 points) for tuberculosis.

Empiric treatment for tuberculosis was given according to standard guidelines of WHO and the Peruvian Ministry of Health. Therapeutic decisions including inpatient versus outpatient management and treatment protocols used were established by local hospital physicians. The study procedures did not necessitate admission to hospital.

Age-matched and sex-matched controls were enrolled on a continuing basis from the Pampas de San Juan, a low-income shanty town community in south Lima that is within the area served by one of the hospitals from which cases were recruited. Inclusion criteria were an absence of chronic cough, fever, or evidence of pulmonary disease; absence of previous clinical tuberculosis; no previous or current treatment for tuberculosis; and absence of recent (past 2 months) household contacts with known or suspected pulmonary tuberculosis.

	High-risk cases (ST score ≥7; n=153)	Moderate-risk cases (ST score 5–6; n=65)	p value*	Controls (n=238)	p value†
Modified ST criteria‡					
Primary lesion by radiography (4 points)§	34/153 (22·2%)	0/65	<0·0001	NA	NA
Tuberculin skin test positive (>10 mm; 3 points)	144/152 (94·7%)	53/64 (82·8%)	0·0047	13/235 (5·5%)	<0·0001
Contact with patient with tuberculosis in past 2 years (2 points)	141/153 (92·2%)	42/65 (64·6%)	<0·0001	0/235	<0·0001
Suggestive radiograph (2 points)§	118/153 (77·1%)	26/65 (40·0%)	<0·0001	NA	NA
Cough >2 weeks (2 points)	80/153 (52·3%)	21/65 (32·3%)	0·0068	NA	NA
Admitted to hospital	13/153 (8·5%)	7/65 (10·8%)	0·5949	NA	NA
Income¶					
<400 Peruvian soles	68/141 (48·2%)	28/59 (47·5%)	..	190/238 (79·8%)	..
≥400 Peruvian soles	73/141 (51·8%)	31/59 (52·5%)	0·9209	48/238 (20·2%)	<0·0001
Mother's education					
Primary or less	34/153 (22·2%)	10/64 (15·6%)	..	82/238 (34·5%)	..
Secondary or more	119/153 (77·8%)	54/64 (84·4%)	0·2704	156/238 (65·5%)	0·0098
BCG vaccination					
Yes	124/150 (82·7%)	55/64 (85·9%)	..	216/235 (91·9%)	..
No	26/150 (17·3%)	9/64 (14·1%)	0·5537	19/235 (8·1%)	0·0059
ST score	7 (7–14)	5 (5–6)	<0·0001	NA	NA
Tuberculin skin test size (mm)	15 (12–20)	13 (10–18)	0·0301	0 (0–2)	<0·0001
Age (years)	3 (0–12)	4 (0–11)	0·7593	4 (0–12)	0·7134
Sex (male)	78/153 (51·0%)	34/65 (52·3%)	0·8577	128/238 (53·8%)	0·5882
Percentage weight-for-age	97% (88–106)	94% (85–104)	0·2244	94% (87–104)	0·0635
Percentage height-for-age	98% (95–101)	98% (95–103)	0·5555	96% (93–99)	<0·0001

Data are number (%) or median (range). For case groups, the number of cases positive for each of the factors used to determine the modified Stegen-Toledo (ST) criteria are reported. NA=not applicable. *Comparison of high-risk and moderate-risk groups. †Comparison of high-risk group and control group. ‡Standard interpretation for ST score is high risk for tuberculosis ≥7 points; moderate risk for tuberculosis 5–6 points; low risk for tuberculosis 3–4 points; unlikely tuberculosis 0–2 points. §In each case points were assigned for only one of these radiographic criteria. ¶400 Peruvian soles=US$142 or £93 (date of conversion July 14, 2010).

Table 1: **Demographic features and modified Stegen-Toledo criteria**[18,19]

Cases and controls were screened for HIV infection by duplicate commercial assays. HIV DNA PCR tests were done in children younger than 18 months old. Screening for HIV infection was requested, but not required if parents refused the test. Cases and controls who declined HIV testing were included and classified as HIV negative if they had no history of HIV exposure and no clinical evidence of HIV/AIDS.

Written informed consent was obtained from all cases and controls or their parents or guardians. This research adhered to human experimentation guidelines of the US Department of Health and Human Services. The protocol and consent forms were approved by the Institutional Review Boards of Tulane Medical Centre, Johns Hopkins Bloomberg School of Public Health, Asociación Benéfica PRISMA, the US Naval Medical Research Center Detachment (Lima, Peru), Hospital Nacional Cayetano Heredia, and the Instituto Nacional de Salud del Niño (Lima, Peru).

Data collection

Demographic and clinical data, including tuberculin skin test results, were obtained from cases and controls. Data collected from cases consisted of medical history, physical examination findings, and a chest radiograph read by a paediatric radiologist. Specimens obtained from cases for tuberculosis cultures and PCR were gastric aspirates, nasopharyngeal aspirates, and stool samples.

Gastric aspirates were obtained on two successive early mornings (0600–0700 h) by brief (<10 min) nasogastric intubation after an overnight fast. The volume of gastric aspirates was augmented as needed by injecting 5 mL sterile water and aspirating back. Nasopharyngeal aspirates were obtained daily for 2 days by inserting a soft flexible nasopharyngeal tube into the nasopharynx, lavaging with 5 mL saline solution, and aspirating with an electrical suction device or hand-held aspirator. The nasopharyngeal aspirate procedure induces a cough and sputum production, which is then aspirated from the nasopharynx. Stool specimens were collected daily for 2 days.

All specimens from cases were obtained in the hospital within 5 days of enrolment, before starting antituberculosis treatment. Blood samples were taken from cases for measurement of serum albumin, as an indicator of nutritional status.

Specimens taken from controls for tuberculosis cultures and PCR assays were one nasopharyngeal-aspirate and two stool samples. No gastric-aspirate specimens were obtained from controls because of the invasive nature of the test, and only one nasopharyngeal aspirate was taken

	High-risk cases (ST score ≥7)	Moderate-risk cases (ST score 5–6)	Adjusted odds ratio (95% CI)*	p value*	Controls	Adjusted odds ratio (95% CI)*	p value*
Culture							
Any specimen							
Positive	21/149 (14·1%)	1/61 (1·6%)	12·1 (1·5–97·1)	0·0189	0/238	..	<0·0001
Negative	128/149 (85·9%)	60/61 (98·4%)	Reference	..	238/238 (100·0%)	..	..
Nasopharyngeal aspirate							
Positive	11/152 (7·2%)	1/63 (1·6%)	6·2 (0·7–53·6)	0·0974	0/208	..	<0·0001
Negative	141/152 (92·8%)	62/63 (98·4%)	Reference	..	208/208 (100·0%)	..	..
Gastric aspirate							
Positive	21/152 (13·8%)	1/64 (1·6%)	13·1 (1·6–105·6)	0·0156	..	..	..
Negative	131/152 (86·2%)	63/64 (98·4%)	Reference	..	..	..	..
Stool sample							
Positive	3/148 (2·0%)	1/62 (1·6%)	1·3 (0·1–12·8)	0·8293	0/230	..	<0·0001
Negative	145/148 (98·0%)	61/62 (98·4%)	Reference	..	230/230 (100·0%)	..	..
PCR							
Any specimen							
Positive	40/152 (26·3%)	15/64 (23·4%)	1·2 (0·6–2·4)	0·6161	15/227 (6·6%)	5·0 (2·6–9·9)	<0·0001
Negative	112/152 (73·7%)	49/64 (76·6%)	Reference	..	212/227 (93·4%)	Reference	..
Nasopharyngeal aspirate							
Positive	22/153 (14·4%)	4/65 (6·2%)	2·8 (0·9–8·6	0·0763	5/228 (2·2%)	7·2 (2·5–20·2)	0·0002
Negative	131/153 (85·6%)	61/65 (93·8%)	Reference	..	223/228 (97·8%)	Reference	..
Gastric aspirate							
Positive	27/152 (17·8%)	8/65 (12·3%)	1·6 (0·7–3·8)	0·2622	..	..	..
Negative	125/152 (82·2%)	57/65 (87·7%)	Reference	..	..	..	..
Stool sample							
Positive	13/150 (8·7%)	4/64 (6·2%)	1·5 (0·5–5·1)	0·5067	10/237 (4·2%)	2·4 (1·0–5·9)	0·0578
Negative	137/150 (91·3%)	60/64 (93·8%)	Reference	..	227/237 (95·8%)	Reference	..

Data are number of children (%) with at least one positive *Mycobacterium tuberculosis* culture or PCR result in a clinical specimen, grouped by type of specimen. Analyses by specimen type limited to cases with complete data for specimens of the indicated type by the method used (culture or PCR). Any specimen indicates all specimen types combined (for cases, limited to those with two gastric-aspirate, nasopharyngeal-aspirate, and stool specimens; for controls, limited to those with one nasopharyngeal-aspirate and two stool specimens). *For the comparison between the high-risk group and the moderate-risk group, the odds ratios and p values were adjusted for confounding differences in patient status (inpatient *vs* outpatient). For the comparison between the high-risk group and the control group, the odds ratios and p values were adjusted for confounding differences in income and mother's education.

***Table 2*: PCR and culture results in high-risk and moderate-risk cases and controls**

from each control because this test is unpleasant for some children, and repeat nasopharyngeal-aspirate testing of controls was not acceptable to many parents.

Tuberculosis diagnostic techniques

Specimens were decontaminated with 0·5% N-acetyl-L-cysteine, 2% sodium hydroxide, and 1·45% sodium citrate, and the centrifuged pellet was resuspended in 2 mL 0·2% bovine serum albumin.[18] Before decontamination, stool samples were prepared by suspension of 0·1 g in 6 mL phosphate buffered saline, homogenisation, and settling for 10 min to separate, after which the supernatant was processed. This stool processing technique was shown in pilot work to provide the best possible compromise between detection sensitivity versus culture contamination. Specimens were tested by four methods. To reduce the risk of bias, samples were obtained from cases and randomly selected matched controls, and technicians were not aware of clinical characteristics, identity of participant groups, or the results of other tests.

In the first method, microscopic-observation drug-susceptibility (MODS),[19,20] 500 µL of each decontaminated specimen was inoculated into modified Middlebrook 7H9 media and separated into four 1·2 mL samples that were cultured in a sterile 24-well plate. Plates were placed in a plastic resealable bag, incubated at 37°C, and examined every other day without opening for up to 30 days by inverted light microscopy. Presumptive *M tuberculosis* isolates with cording morphology were reported as positive and all were subsequently confirmed by heminested IS6110 *M tuberculosis* PCR. Direct concurrent drug susceptibility testing was not done.

The second method was Lowenstein-Jensen agar culture. 250 µL of each decontaminated specimen was inoculated onto a Lowenstein-Jensen slant, incubated at 37°C, and examined two times a week from the first to eighth week after inoculation.

In the third test, auramine stain (smear test), two drops of each decontaminated specimen were dried on a microscope slide, stained with 0·1% auramine O (Sigma

Aldrich, St Louis, MO, USA), and examined at 100-times magnification. Positive tests had five or more bacilli counted in 300 fields.

N2 IS6110 PCR procedure (referred to as PCR testing) was the fourth method.[21] DNA was extracted from 500 µL of each decontaminated specimen, and two consecutive nested IS6110 PCRs were done with outer and inner primers. To reduce the risk of false-positive PCR reactions, separate rooms were used for DNA extraction, PCR mix preparation, amplification, and electrophoresis with protective clothing. Every heminested PCR assay included positive controls of genomic *M tuberculosis* DNA and negative controls of water added to the PCR reagents in place of the sample DNA.

Statistical analysis

We calculated a sample size to detect a 15% difference in sensitivity of detection of *M tuberculosis* between two tests done in the same individuals (eg, gastric aspirates *vs* nasopharyngeal aspirates) or between detection methods (eg, MODS *vs* Lowenstein-Jensen culture; culture *vs* PCR) or between cases and controls. On the basis of 80% power and an α error of p=0·05, 189 cases and 189 controls were required. To compensate for non-evaluable individuals (eg, second sample not provided), we increased the study sample by 10% to a target sample of 210–220 cases and an equal number of controls.

Demographic and clinical data, and culture and PCR results were compared in cases and controls by Stegen-Toledo risk group (moderate risk *vs* high risk). Culture and PCR results were compared by sample and by person. A negative culture result was deemed valid (ie, readable and contamination-free) if at least one valid Lowenstein-Jensen and one valid MODS result was available. A case with negative cultures was valid if all three types of specimen (gastric aspirate, nasopharyngeal aspirate, and stool sample) had valid results for duplicate specimens. For controls, nasopharyngeal-aspirate and stool cultures required valid results for inclusion in the analysis. χ^2 and McNemar's tests were used for categorical variables; two-tailed *t* test or Wilcoxon rank sum test were used for continuous variables. Multiple logistic regression was used to adjust for potentially confounding variables. Recovery rates for *M tuberculosis* were compared by culture method for cases and specimens, grouped by specimen type (gastric aspirate, nasopharyngeal aspirate, stool sample) and by auramine stain result. Cases with at least one clinical specimen that was culture positive for *M tuberculosis* by any method were included in the culture-positive group. Data analysis was done with STATA version 11 and EpiInfo version 6.

This study is registered with ClinicalTrials.gov, number NCT00054769.

Role of the funding source

The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study, had final responsibility for the decision to submit for publication, and takes responsibility for the data and the accuracy of the data analysis.

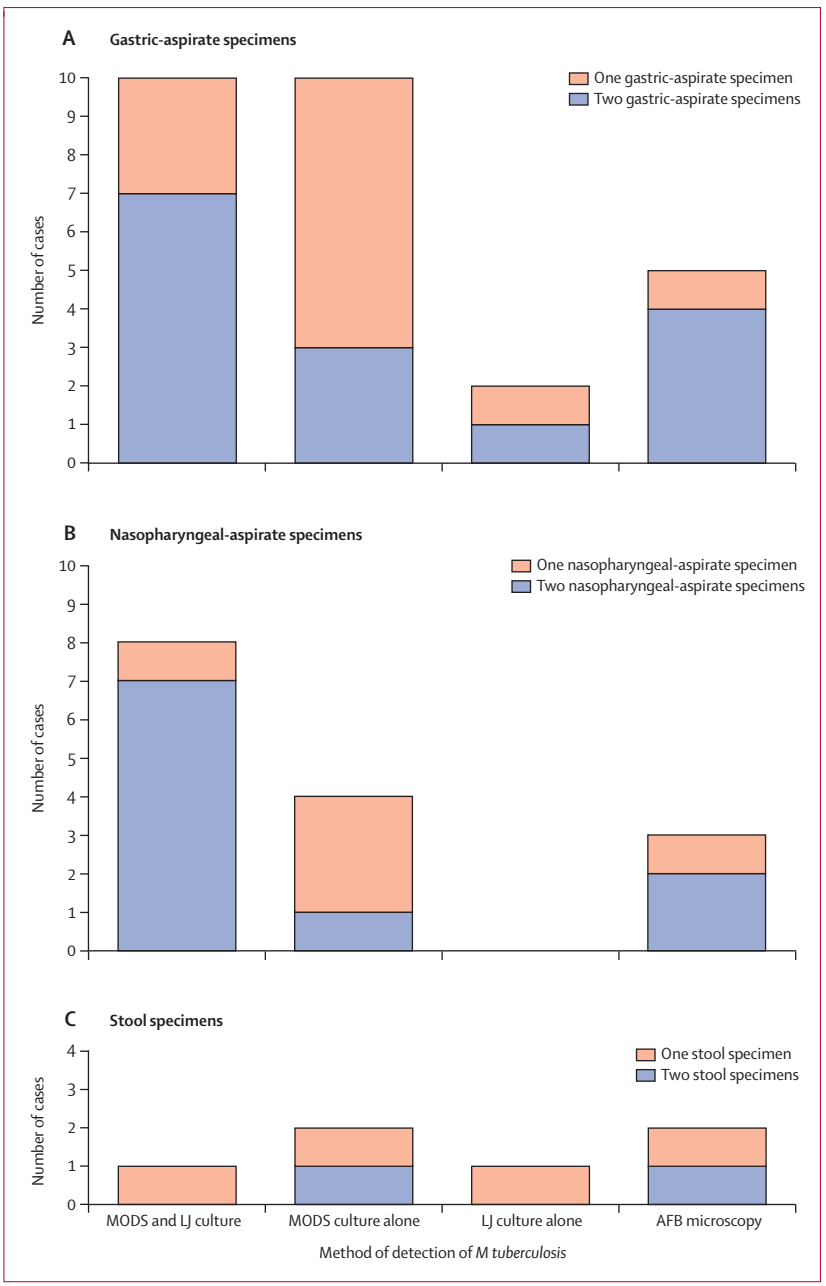

Figure 2: **Number of cases of tuberculosis detected by culture and microscopy, by specimen**
MODS=microscopic-observation drug-susceptibility. LJ= Lowenstein-Jensen agar. AFB=acid-fast bacilli.

Results

The study population consisted of 456 children with complete demographic information and, for most participants, complete laboratory data, with 218 presumptive tuberculosis cases (20 [9%] inpatients;

	Culture positive (n=21)	Culture negative (n=128)	Sensitivity	Specificity	Positive predictive value	Negative predictive value	κ
Any specimen							
PCR positive	13 (62%)	26 (20%)*	61·9%	79·7%	33·3	92·7	0·306
PCR negative	8 (38%)	102 (80%)	..	..	..	..	..
Nasopharyngeal aspirate							
PCR positive	9 (43%)	13 (10%)*	42·9%	89·8%	40·9	90·6	0·321
PCR negative	12 (57%)	115 (90%)	..	..	..	..	..
Gastric aspirate							
PCR positive	13 (62%)	13 (10%)*	61·9%	89·8%	50·0	93·5	0·471
PCR negative	8 (38%)	115 (90%)	..	..	..	..	..
Stool sample							
PCR positive	4 (20%)‡	9 (7%)†	20·0%	93·0%	30·8	88·1	0·152
PCR negative	16 (80%)‡	119 (93%)	..	..	..	..	..

*p≤0·0001 for comparison of culture-positive and culture-negative groups. †p=0·0567 for comparison of culture-positive and culture-negative groups. ‡Stool samples were unavailable from one culture-positive case.

Table 3: **Agreement of PCR and culture results, and predictive values of PCR for detection of children with at least one positive culture among high-risk cases**

198 [91%] outpatients) and 238 controls. 81 potential cases were not enrolled because participation was declined by their parents, and three controls without valid Lowenstein-Jensen results were excluded. These children were similar in age and sex distribution to enrolled participants.

Demographic characteristics of cases and controls and distribution of the characteristics used to determine the modified Stegen-Toledo criteria are shown in table 1. 153 (70%) cases had a Stegen-Toledo score of 7 points or more (high risk for pulmonary tuberculosis), and 65 (30%) cases had scores of 5–6 points (moderate risk). Most cases at high risk for tuberculosis had abnormal results on chest radiograph, and the most common findings were interstitial markings (83 [54%] cases) and hilar adenopathy (67 [44%]). A primary Ghon complex was present in 34 (22%) high-risk cases. Among moderate-risk cases, 26 (40%) had suggestive findings on chest radiograph (all interstitial markings or hilar adenopathy), and none had a primary Ghon complex.

Rate of BCG vaccination was lower in cases than in controls (p=0·006). Compared with high-risk cases, controls had lower levels of income and maternal education. HIV test results were negative in 179 (82%) cases and the test was declined in 39 (18%) cases. No cases or controls were HIV positive on the basis of screening tests.

22 (10%) cases had at least one positive *M tuberculosis* culture, on the basis of MODS and Lowenstein-Jensen results combined. *M tuberculosis* was isolated from gastric-aspirate specimens in all 22 cases, nasopharyngeal-aspirate specimens in 12 cases, and stool specimens in four cases. In all, 58 culture-positive specimens were obtained from cases, consisting of 33 gastric aspirates, 20 nasopharyngeal aspirates, and five stool specimens. 17 (29%) of these 58 specimens were positive on auramine microscopy.

There was a significant association between high-risk group and positive tuberculosis culture results (table 2). One (1·6%) of 61 moderate-risk cases with complete specimen collection was culture positive compared with 21 (14·1%) of 149 high-risk cases. The modified Stegen-Toledo scores in cases with a positive culture for *M tuberculosis* (median 9 points, range 6–13) were significantly greater than in culture-negative cases (7 points, 5–14; p<0·0001). 13 (8·5%) of 153 high-risk cases and seven (10·8%) of 65 moderate-risk cases were admitted to hospital, and hospital admission was associated with higher rates of positive *M tuberculosis* cultures (seven [36·8%] of 19 inpatient cases *vs* 15 [7·9%] of 191 outpatient cases; odds ratio 6·8, 95% CI 2·3–19·9, p<0·0001). None of the samples from the 238 controls were *M tuberculosis* culture positive.

Culture-positive cases were similar to culture-negative cases in terms of sex distribution and age. Median serum albumin concentrations (42 g/L [IQR 37–43] *vs* 45 g/L [43–47]) and prevalence of serum albumin lower than 45 g/L (17 [81%] of 21 *vs* 74 [40%] of 184) were both significantly different between culture-positive and culture-negative cases (p=0·0001 and p=0·0015, respectively).

Mean time from sample processing to isolation of *M tuberculosis* was significantly shorter for MODS cultures (mean 10 days, IQR 8–11) than for Lowenstein-Jensen cultures (25 days, 20–30; p=0·0001). MODS cultures were not interpretable because of bacterial or fungal contamination less often than were Lowenstein-Jensen cultures (14 [1·1%] of 1292 *vs* 102 [7·9%] of 1292, p<0·0001) and there were no significant differences in contamination rates between stool, nasopharyngeal-aspirate and gastric-aspirate specimens. MODS was more sensitive than was Lowenstein-Jensen culture, diagnosing 20 (90·9%) of 22 patients compared with 13 (59·1%) of 22 patients (p=0·015). Nine (41%) of the 22 culture-positive cases were detected by MODS culture only, two (9%) cases by Lowenstein-Jensen culture only, and in 11 (50%) cases at least one specimen was culture-positive by both methods (nine of 13, or 70% more cases detected by MODS that were missed by Lowenstein-Jensen culture).

As in our preliminary analysis,[15] the MODS technique was more sensitive than were the other tests. Here we extend previous results to establish how this test performance affected patient diagnosis. Figure 2 shows recovery of *M tuberculosis* from the 22 culture-positive cases by culture method and specimen type. Although all 22 cases were culture positive in at least one gastric-aspirate specimen, 11 of these 22 cases only had positive cultures from one of two gastric-aspirate specimens (figure 2). When results were analysed on the basis of the order of collection for duplicate specimens, *M tuberculosis* was isolated from the first gastric-aspirate specimen

obtained in 16 (72·7%) of 22 cases, whereas in six (27·3%), only the second gastric-aspirate specimen was culture positive (37% greater yield by adding a second specimen). Addition of a second nasopharyngeal-aspirate specimen increased yield by 50%.

Table 2 shows PCR results and culture results for high-risk and moderate-risk cases and controls. The proportions of patients with at least one positive PCR result were similar between high-risk and moderate-risk cases, and both groups had significantly higher proportions of positive PCR compared with the control group (p=0·0001).

15 of 238 controls had at least one positive PCR result (five from nasopharyngeal-aspirate specimens; ten from stool samples). No controls had more than one positive PCR result. 14 controls had a positive PPD skin test, but only one of these 14 children also had a positive PCR result (nasopharyngeal-aspirate specimen). All controls with a positive PPD were assessed by a physician, examined by chest radiography, assessed not to have tuberculous disease, and given preventive therapy. None of these controls with positive PPDs or PCRs developed further evidence of clinical tuberculosis during the year after enrolment. Controls had twice as many stool samples as nasopharyngeal aspirate samples per person, and since ten (4·2%) controls had a PCR-positive stool and five (2·1%) had a PCR-positive nasopharyngeal aspirate, this suggests that the rate of false-positive PCR results in any individual specimen (from these highly selected children with low risk of tuberculosis) was about 2%.

Table 3 compares PCR results in the 21 culture-positive cases compared with the 128 culture-negative cases with complete culture and PCR data for all specimens. Culture-positive cases were three times more likely to have at least one PCR-positive specimen than were culture-negative cases. Similar results were noted in analyses limited to each specimen type. To assess the potential for use of PCR to identify high-risk paediatric cases likely to be culture positive, PCR sensitivity, specificity, and predictive values were analysed with culture as the reference standard (table 3). In this high-risk group, a positive PCR from a gastric aspirate was associated with a positive predictive value (for a positive culture result) of 50%. Sensitivity of PCR in this subgroup was 62% for gastric aspirates and for any specimen, but false-positive PCR results (with culture as the reference standard) equalled or exceeded true positives for all specimen types. However, both specificity and positive predictive values must be reduced by the limited sensitivity of culture as a reference standard.

Discussion

This study presents a direct comparison of culture methods and PCR, with different specimen types for both approaches, for diagnosis of paediatric tuberculosis in a resource-poor area. In these Peruvian children, MODS roughly doubled the diagnostic sensitivity of culture and halved the time required compared with traditional Lowenstein-Jensen culture. 70% more cases were diagnosed as having culture-positive tuberculosis by MODS compared with Lowenstein-Jensen culture, and all specimens from these additional cases were auramine smear negative. The speed and sensitivity of the MODS technique results from the use of culture broth with microscopic detection of positive cultures before they are large enough to be visible to the naked eye and the capacity of this technique to culture a large volume of each clinical specimen.[20] This increased diagnostic yield in children with culture-proven tuberculosis by addition of MODS greatly exceeds the 10–20% additional yield seen in high-risk Peruvian adults,[20] as would be expected for diagnostic samples from children that usually contain few mycobacteria. Our results lend support to WHO recommendations that suggest that liquid tuberculosis culture techniques are superior to agar-based techniques.[18] MODS performed better than did Lowenstein-Jensen culture for auramine-negative specimens, which include most culture-positive paediatric specimens.

M tuberculosis recovery by culture from gastric aspirates (22 of 22 cases) was clearly superior to recovery from stool samples (four of 22 cases) and from nasopharyngeal aspirates (12 of 22 cases). The 37% additional yield of culture-positive patients identified by a second gastric aspirate greatly exceeded the 5–10% additional yield from a second sputum culture in Peruvian adults.[20] One explanation of this major incremental benefit from a second specimen in children is that adults tend to be already smear positive and with high bacillary loads at detection, whereas children with paucibacillary disease benefit much more from a sensitive method like MODS. In almost half the patients studied, gastric aspirates were the only specimens that were culture positive, and in five of these cases only one specimen was culture positive, showing the importance of testing at least duplicate specimens.

We noted that high-risk clinical scores were associated with higher likelihood of a positive *M tuberculosis* culture. Larger numbers of moderate-risk cases would be useful to substantiate this finding, since only one moderate-risk culture-positive case was identified in this study. Although this result suggests that the Stegen-Toledo score might correlate with rates of isolation of *M tuberculosis* in children, it also emphasises the fact that, even with intensive culture methods, 85% of high-risk cases were still culture negative. Determining how many of these children have culture-negative pulmonary tuberculosis is difficult, but data from cases with positive cultures (eg, the predominance of smear-negative specimens, the high proportion of cases confirmed by a single MODS culture-positive gastric aspirate) suggest that many of these cases are tuberculosis. Other new

techniques, such as induced sputum with increased recovery of *M tuberculosis* by culture in African children,[22,23] might help to confirm these suspicions.

The N2 IS6110 PCR assay was positive more often in subgroups with greater clinical suspicion of tuberculosis or positive cultures than in those at low risk or with negative cultures. Our findings corroborate those of a previous retrospective study of IS6110 stool PCR,[24] and this larger prospective study also shows that sensitivity of PCR to detect culture-positive children was greater for both nasopharyngeal aspirates and gastric aspirates than with stool specimens. However, several healthy controls had positive PCR results despite having no evidence of tuberculosis exposure or disease, implying that these were false positives, occurring in about 2% of specimens despite extensive precautions to prevent PCR contamination. These results raise the possibility that the PCR assay is detecting the acquisition or presence of latent tuberculosis or early, asymptomatic disease. Our 1-year follow-up of controls for signs of clinical tuberculosis did not suggest this, but additional studies might be needed to better assess this possibility.

Despite these non-specific positive results, in some cases we noted that PCR might be useful for identification of children who are likely to have culture-positive pulmonary tuberculosis. Among high-risk children, a positive PCR result from a gastric-aspirate specimen was associated with a positive predictive value (for a positive culture result) of 50%—ie, it identified a subgroup of high-risk patients with a 50% rate of positive cultures with a test result that can be available within hours. By contrast with these encouraging results, eight (38%) of 21 children in the high-risk, culture-positive group were PCR negative in all specimens. Thus, our results do not support the routine use of PCR for diagnosis of paediatric tuberculosis, but suggest that the same-day gastric-aspirate PCR test might be useful as a screening test for high-risk children because it rapidly identifies those children in the high-risk group who are three times more likely to have culture-proven tuberculosis.

The results reported in this study are subject to several limitations. First, cases were enrolled on the basis of the Stegen-Toledo score, which although widely used is poorly validated. The association between Stegen-Toledo risk group and proportion of culture-confirmed cases was notable, somewhat diminishing this limitation. Second, potential participants who did not enrol might have been different from those who did enrol in the study. Most parents of the potential cases and controls who declined to participate were unwilling or unable to return to the clinic for subsequent sample collection, because participation required several days' attendance. This situation was usually the result of parental work demands or distance from home to the clinic. These children were not different from enrolled participants in terms of age or sex, but there could have been differences in other variables such as income and level of education. Third, we chose to include cases and controls who declined HIV testing and lacked evidence of HIV/AIDS. Although we cannot exclude occult HIV infection, there is substantial cultural stigma associated with HIV testing and a very low HIV seroprevalence in Lima (about 0·5% in adults and <0·1% in children younger than 12 years),[25] so including these children in the analysis seemed appropriate. Finally, controls were only assessed with a single nasopharyngeal-aspirate specimen and no gastric-aspirate specimens, whereas two nasopharyngeal aspirates and two gastric aspirates were obtained from cases. The extended sampling strategy for cases is justified from a standard of care perspective, since there is potential direct benefit to the patient from extended sampling. We did not believe that the discomfort of gastric aspiration was justified for a well child. Nasopharyngeal aspiration is unpleasant for some, and many parents were unwilling to agree to repeated sampling.

Although our sampling strategy therefore differed between cases and controls, the study provided the most extensive data so far on specificity of nasopharyngeal-aspirate culture for tuberculosis in children, providing samples from more than 200 children in a population in which *M tuberculosis* infection is common. Among cases with positive nasopharyngeal-aspirate cultures, around two-thirds were culture positive in both nasopharyngeal-aspirate specimens collected, suggesting that examination of a single nasopharyngeal-aspirate specimen was reasonable in view of the logistical and ethical considerations.

Thus, MODS culture increased sensitivity and speed of diagnosis of pulmonary tuberculosis compared with conventional Lowenstein-Jensen culture. Gastric-aspirate cultures improved case detection compared with nasopharyngeal-aspirate cultures. Although most children treated for presumed pulmonary tuberculosis were culture-negative in all specimens, MODS culture of duplicate gastric aspirates substantially improved the yield of laboratory confirmation. PCR was insufficiently sensitive and specific for routine diagnostic use, but in high-risk children duplicate gastric-aspirate PCR rapidly identified a subgroup with a 50% rate of positive cultures. Collection of duplicate gastric-aspirate specimens from high-risk children for MODS culture increased microbiological diagnosis of tuberculosis by more than a third.

Contributors

GS-C participated in enrolment of participants, collection of study data, and quality control. MEC, LK, ES-L, and EN contributed to participant referral and medical support. TDP contributed to interpretation of radiographs. SM, LC, VAL-T, DAJM, and RHG contributed to laboratory support and diagnostic testing. MS undertook data analysis. DAJM, CAE, and RHG participated in study design and data interpretation. The principal investigator (RAO) was involved in study design, quality control, data analysis, data interpretation, and reporting.

Acknowledgments

This study was supported in part by the following grants: National Institutes of Health (NIH) 1 RO1 AI-49139, USAID—Tuberculosis Award HRN-5986-A-00-6006-00, NIH ITREID grant 5D43-TW00910, the Fogarty-NIH AIDS training programme 3T22-TW00016-05S3, and the

National Institute of Allergy and Infectious Diseases tutorial training grant 5T35-AI07646-02. Some co-investigators were supported by IFHAD and FIND (CAE) and the Wellcome Trust (CAE, DAJM). Naval Medical Research Center Detachment work was supported by Work Unit Number (WUN) 62787A.873.H.B000. The research study described was supported entirely by the first grant listed, whereas the other grants supported personnel and expenses of the tuberculosis diagnostic laboratory used in the study. VAL-T is employee of the US Government. This work was prepared as part of their official duties. The views expressed in this report are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the US Government.

Conflicts of interest

We declare that we have no conflicts of interest.

References

1 Eamranond P, Jaramillo E. Tuberculosis in children: reassessing the need for improved diagnosis in global control strategies. *Int J Tuberc Lung Dis* 2001; **5:** 594–603.

2 Donald PR, Maher D, Qazi S. A research agenda to promote the management of childhood tuberculosis within national tuberculosis programmes. *Int J Tuberc Lung Dis* 2007; **11:** 370–80.

3 Mwinga A. Challenges and hope for the diagnosis of tuberculosis in infants and young children. *Lancet* 2005; **365:** 97–98.

4 WHO. Global tuberculosis control: surveillance, planning, financing. 2004. http://www.who.int/tb/publications/global_report/2004/en/ (accessed March 10, 2010).

5 Kelynack TN. Tuberculosis in infancy and children. New York: William Wood and Company, 1908; 80–85.

6 Abadco D, Steiner P. Gastric lavage is better than broncho-alveolar lavage for isolation of *Mycobacterium tuberculosis* in childhood pulmonary tuberculosis. *Pediatr Infect Dis J* 1992; **11:** 735.

7 Hesseling AC, Schaaf HS, Gie RP, Starke JR, Beyers N. A critical review of diagnostic approaches used in the diagnosis of childhood tuberculosis. *Int J Tuberc Lung Dis* 2002; **6:** 1038–45.

8 Starke JR. Pediatric tuberculosis: time for a new approach. *Tuberculosis* 2003; **83:** 208–12.

9 Marais BJ, Gie RP, Schaaf S, Beyers N, Donald PR, Starke JR. Childhood pulmonary tuberculosis: old wisdom and new challenges. *Am J Respir Crit Care Med* 2006; **173:** 1078–90.

10 Marais BJ, Hesseling AC, Gie RP, Schaaf HS, Enarson DA, Beyers N. The bacteriologic yield in children with intrathoracic tuberculosis. *Clin Infect Dis* 2006; **42:** e69–71.

11 Marais B, Gie RP, Schaaf HS, Hesseling AC, Enarson DA, Beyers N. The spectrum of disease in children treated for tuberculosis in a highly endemic area. *Int J Tuberc Lung Dis* 2006; **10:** 732–38.

12 Marais BJ, Graham SM, Cotton MF, Beyers N. Diagnostic and management challenges for childhood tuberculosis in the era of HIV. *J Infect Dis* 2007; **196:** S76–85.

13 Salazar GE, Schmitz TL, Cama R, et al, and the Working Group on TB in Peru. Pulmonary tuberculosis in children in a developing country. *Pediatrics* 2001; **108:** 448–53.

14 Franchi LM, Cama RI, Gilman RH, Montenegro-James S, Sheen P. Detection of *Mycobacterium tuberculosis* in nasopharyngeal aspirate samples in children. *Lancet* 1998; **352:** 1681–82.

15 Oberhelman RA, Soto-Castellares G, Caviedes L, et al. Improved recovery of *Mycobacterium tuberculosis* from children using the microscopic observation drug susceptibility (MODS) method. *Pediatrics* 2006; **118:** e100–06.

16 Stegen G, Jones K, Kaplan P. Criteria for guidance in the diagnosis of tuberculosis. *Pediatrics* 1969; **43:** 260–63.

17 Toledo A, Katz F, Montiel J, Rico FG. Diagnostic criteria for pediatric tuberculosis. *Rev Mex Pediatr* 1979; **46:** 239–43 (in Spanish).

18 American Thoracic Society. Diagnostic standards and classification of tuberculosis in adults and children. *Am J Respir Crit Care Med* 2000; **161:** 1376–95.

19 Caviedes L, Lee TS, Gilman RH, et al. Rapid, efficient detection and drug susceptibility testing of *Mycobacterium tuberculosis* in sputum by microscopic observation of broth cultures. *J Clin Microbiol* 2000; **38:** 1203–08.

20 Moore DAJ, Evans CAW, Gilman RH, et al. Microscopic-observation drug-susceptibility assay for the diagnosis of TB. *N Engl J Med* 2006; **355:** 1539–50.

21 Montenegro SH, Gilman RH, Sheen P, et al. Improved detection of *Mycobacterium tuberculosis* (MTB) in Peruvian children using a nested IS6110-PCR assay. *Clin Infect Dis* 2003; **36:** 16–23.

22 Zar HJ, Hanslo D, Apolles P, Swingler G, Hussey G. Induced sputum versus gastric lavage for microbiological confirmation of pulmonary tuberculosis in infants and young children: a prospective study. *Lancet* 2005; **365:** 130–34.

23 Iriso R, Mudido PM, Karamagi C, Whalen C. The diagnosis of childhood tuberculosis in an HIV-endemic setting and the use of induced sputum. *Int J Tuberc Lung Dis* 2005; **9:** 716–26.

24 Wolf H, Mendez M, Gilman RH, et al. Diagnosis of pediatric pulmonary tuberculosis by stool PCR. *Am J Trop Med Hyg* 2008; **79:** 893–98.

25 Joint United Nations Programme on HIV/AIDS. Epidemiological fact sheet on HIV and AIDS. Core data on epidemiology and response: Peru. 2008. http://apps.who.int/globalatlas/predefinedReports/EFS2008/full/EFS2008_PE.pdf (accessed March 10, 2010).