JOURNAL OF CLINICAL MICROBIOLOGY, Oct. 1999, p. 3229–3232 Vol. 37, No. 10
0095-1137/99/$04.00+0

Evaluation of the Tuberculin Gamma Interferon Assay: Potential To Replace the Mantoux Skin Test

SUDHA POTTUMARTHY,[1] ARTHUR J. MORRIS,[1]* ADRIAN C. HARRISON,[2] AND VIRGINIA C. WELLS[1]

Departments of Microbiology[1] and Respiratory Medicine,[2] Green Lane Hospital, Auckland, New Zealand

Received 19 January 1999/Returned for modification 6 March 1999/Accepted 10 July 1999

We evaluated an in vitro test of cell-mediated immunity, the tuberculin gamma interferon assay, *Quanti*FERON-TB (QIFN), in 455 individuals from three groups: group I, 237 immigrants from high-risk countries; group II, 127 health care workers undergoing Mantoux testing; group III, 91 patients being investigated for possible active tuberculosis (79 patients) or *Mycobacterium avium-Mycobacterium intracellulare* complex infection (12 patients). The QIFN results were compared either to those of the Mantoux test or to microbiological and clinical diagnosis, as appropriate. In each group the correlation between the diameter of induration for the skin test and the magnitude of QIFN response was significant and of moderate strength (Spearman's rank correlation coefficient; $\rho = 0.59$ to 0.61; $P < 0.001$). For group I, the agreement between QIFN and Mantoux results was 89% for Mantoux-negative and 64% for Mantoux-positive individuals. For group II, when ≥10-mm-diameter induration was taken as positive, the agreement was 81% for Mantoux-negative and 67% for Mantoux-positive individuals. For group III, agreement was 81% for Mantoux-negative and 86% for Mantoux-positive patients. For patients being evaluated for active tuberculosis, the performance of the Mantoux test was not statistically different from that of the QIFN assay. In patients with active tuberculosis, the assay had a sensitivity of 77%, not significantly higher for extrapulmonary than pulmonary cases (83% versus 74%). QIFN sensitivity was not significantly different for smear-negative or smear-positive cases (80% versus 71%). The QIFN assay is a potential replacement for the Mantoux test. The acceptability of these performance values and those of similar evaluations will determine the place this test will have in detecting evidence of mycobacterial infection.

Intradermal injection of tuberculin, the Mantoux test (tuberculin skin test), is used worldwide to determine whether an individual has immunological reactivity to mycobacterial antigens. While the Mantoux test is a useful aid in identifying tuberculous infection, it has a number of drawbacks, including the need for a return visit to allow reading, problems in interpretation due to cross-reactivity with other mycobacterial species, the booster effect, and false-negative results because of intercurrent immunosuppression, as well as the variability inherent in its application and reading (3). The tuberculin gamma interferon (IFN-γ) assay, *Quanti*FERON-TB (QIFN), has recently been developed by CSL Ltd Australia (5). Principally, this test involves detection and quantitation of the cytokine IFN-γ produced by T lymphocytes stimulated with tuberculin purified protein derivatives (PPDs) obtained from either *Mycobacterium tuberculosis* (human), *Mycobacterium avium* (avian), or *Mycobacterium bovis* (bovine). Due to the QIFN assay's ability to quantitate the differential response to the tuberculin PPDs, e.g., human versus avian, it has the potential to discriminate between *M. tuberculosis* and *M. avium-Mycobacterium intracellulare* complex (MAC) infections. The QIFN assay is an adaptation of the BOVIGAM test, which is licensed as an official test for diagnosing bovine tuberculosis in both Australia and New Zealand (13). QIFN overcomes some of the shortcomings of the Mantoux test, namely, the need for return visits and reader variability. Also, depending on specimen transport time and laboratory efficiency, it has the potential to provide an earlier result. Like the Mantoux test, the QIFN assay could have a role in screening for *M. tuberculosis* infection, contact tracing, and diagnosing tuberculosis, particularly when the acid-fast smear is negative. We evaluated the performance of the QIFN assay with three groups of individuals, immigrants, health care workers (HCWs), and patients, and compared the results to either the Mantoux test results or the microbiological and clinical diagnoses.

* Corresponding author. Mailing address: Microbiology Laboratory, Green Lane Hospital, Green Lane West, Auckland 1003, New Zealand. Phone: (649) 638-9909. Fax: (649) 630-9785. E-mail: arthurm@ahsl.co.nz.

MATERIALS AND METHODS

Study population. Over 16 months, November 1996 to February 1998, 455 individuals were evaluated. All were human immunodeficiency virus negative except one patient in whom active tuberculosis was excluded. Group I consisted of 237 immigrants from countries with a high prevalence of tuberculosis. This group consisted of 191 New Zealand quota refugees and 46 asylum seekers undergoing screening for infectious diseases. They had a median age of 28 years (range, 1 to 72), and 151 were males. Group II consisted of 93 HCWs undergoing employment screening at the occupational health clinics of Auckland (AKH) and Green Lane (GLH) hospitals and 34 microbiology laboratory staff at these two hospitals and a community laboratory in Auckland. This group had a median age of 35 years (range, 20 to 56), and 16 were males. No HCW was from a high-risk or high-prevalence group for tuberculosis, was a close contact of a person with an active case of tuberculosis, or had radiological evidence of tuberculosis. In groups I and II QIFN assay results were compared to those of the Mantoux test. The Centers for Disease Control and Prevention (CDC) interpretive criteria for positive Mantoux test results were followed (2). The chest X rays of individuals in group I with discrepant Mantoux and QIFN results were reviewed by a respiratory physician (A.C.H.) for radiological evidence of infection. Group III consisted of 91 patients being evaluated for either tuberculosis ($n = 79$) or MAC disease ($n = 12$). This group had a median age of 44 years (range, 13 to 89), and 49 were males. Among the 79 cases investigated for tuberculosis, active disease was excluded in 19. Sixty cases of active tuberculosis were detected, 42 (70%) pulmonary and 18 (30%) extrapulmonary. The extrapulmonary cases included lymphadenitis (eight), pleural (three), miliary (two), and one each of disseminated, renal, skeletal, pericardial, and adrenal gland infection. Forty-eight cases (80%) were culture proven, 28 smear positive and 20 smear negative. The 12 culture-negative patients were diagnosed on the basis of symptoms, exposure history, radiology and histology findings, and response to antituberculous treatment. Mantoux results were available for 51 of 79 (65%) patients evaluated for tuberculosis, including 38 of 60 (63%) with active disease, 26 pulmonary and 12 extrapulmonary. In this group the QIFN assay results were compared to those of the Mantoux test as well as to microbiological and clinical diagnoses. Twelve

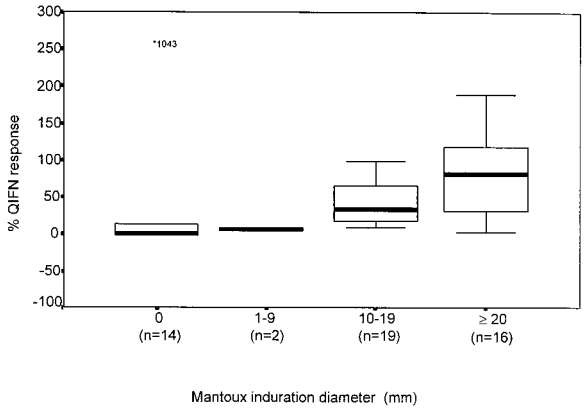

FIG. 1. Correlation between percent QIFN response and Mantoux induration diameter in patients being evaluated for active tuberculosis. The percent QIFN response versus the Mantoux induration diameter is stratified as either 0, 1 to 9, 10 to 19, or ≥20 mm of induration. The solid bars represent medians, the boxes indicate the interquantile ranges, and the lines show the most extreme observations within 1.5 times the interquantile range. There was one extreme outlier with a value of 1,043% QIFN response in the 0-mm Mantoux group. The correlation between the QIFN result and Mantoux induration diameter was significant and of moderate strength (Spearman's rank correlation coefficient = 0.61; $P < 0.001$).

patients were investigated for active MAC infection: eight had recent positive cultures for MAC and were clinically thought to represent either colonization ($n = 7$) or MAC disease ($n = 1$). Four patients had been treated for MAC disease in the previous 4 years and did not have evidence of recurrence during the period of study. In this group the results of the QIFN assay were compared to microbiological and clinical diagnoses.

QIFN assay. The QIFN assay was performed in accordance with the manufacturer's instructions. In brief, testing was conducted in two stages, overnight culture of blood with stimulation antigens and the subsequent quantification of IFN-γ production by enzyme immunoassay (EIA). In the first stage, 1-ml aliquots of heparinized whole blood were incubated in tissue culture wells with different tuberculin PPDs (human, avian, and bovine), sterile phosphate-buffered saline (no-antigen control), and phytohemagglutinin (positive mitogen control). Following 18 h of incubation (range, 16 to 24 h) at 37°C in a humidified atmosphere, the supernatant plasma was harvested. The IFN-γ in the plasma supernatant was subsequently quantified by EIA.

The results were calculated and interpreted according to the manufacturer's instructions: *M. tuberculosis* infection was indicated by a percent human response (human PPD/mitogen response) of >15% and a percent avian difference (human PPD − avian PPD/human PPD response) of >−10%. MAC infection was indicated by a percent avian response (avian PPD/mitogen response) of >20% and a percent avian difference of <−10%.

Statistical analysis. The correlation between the degree of QIFN response and the Mantoux induration diameter was assessed by Spearman's rank correlation test. The kappa statistic was used to measure the strength of agreement between the Mantoux and QIFN results, with a kappa statistic value of >0.75 representing excellent agreement, 0.40 to 0.75 representing good to fair agreement, and <0.40 representing poor agreement. Differences in the performance of the test(s) were analyzed by the χ^2 test.

RESULTS

Agreement between QIFN and Mantoux tests. In all three groups the correlation between the diameter of induration for the Mantoux test and the magnitude of QIFN response was significant and of moderate strength (Spearman's rank coefficient; $\rho = 0.59$ to 0.61; $P < 0.001$). For group III, the correlation between the median percentage of QIFN response and the Mantoux result, stratified as 0, 1 to 9, 10 to 19, and ≥20 mm, is shown in Fig. 1. Similar results were obtained for groups I and II (data not shown).

For the 237 immigrants in group I, for whom a positive Mantoux reaction was defined as a ≥10-mm-diameter induration (2), the agreement was 89 and 64% for Mantoux-negative and -positive individuals, respectively, with a kappa statistic of 0.55 (Table 1). None of the 47 immigrants with discrepant Mantoux and QIFN results, i.e., Mantoux positive and QIFN negative ($n = 31$) or Mantoux negative and QIFN positive ($n = 16$), had radiological evidence of active tuberculosis. Only one immigrant who was Mantoux positive but QIFN negative had minor radiological evidence of old tuberculosis.

For the 127 HCWs in group II, if ≥10-mm-diameter induration was defined as a positive Mantoux test result, the agreement was 81 and 67% for Mantoux-negative and -positive individuals, respectively, with a kappa statistic of 0.48 (Table 1). When the CDC guidelines for the interpretation of Mantoux reactions were followed, i.e., ≥15-mm-diameter induration was considered positive (2), the agreement was 70 and 68% in the Mantoux-negative and -positive individuals, respectively, with a kappa statistic of 0.26.

For the 51 patients in group III for whom a Mantoux test result was obtained, the agreement between the two tests was 81 and 86% for Mantoux-negative and -positive individuals, respectively, with a kappa statistic of 0.65 (Table 1).

Agreement between QIFN assay and clinical diagnosis. The sensitivity of the QIFN assay to detect *M. tuberculosis* disease is shown in Table 2. The QIFN results were not statistically different for pulmonary versus extrapulmonary disease, for culture-positive versus clinically diagnosed cases, or for smear-positive versus smear-negative cases. None of the patients with active tuberculosis tested positive for MAC by the QIFN assay.

For the 38 patients with evaluable Mantoux test results, the sensitivities of the Mantoux and QIFN tests to detect *M. tu-*

TABLE 1. Agreement between QIFN and Mantoux tests

Group no.	Group members	No.	Mantoux test result	QIFN test result (no.)		Agreement (%)	Kappa statistic[a]
				Negative	Positive		
I	Immigrants	237	Negative (<10 mm)	135	16	89	0.55
			Positive (≥10 mm)	31	55	64	
II	HCWs	127	Negative (<10 mm)	64	15	81	0.48
			Positive (≥10 mm)	16	32	67	
			Negative (<15 mm)	73	32	70	0.26
			Positive (≥15 mm[b])	7	15	68	
III	Patients	51	Negative (<10 or <15 mm[c])	13	3	81	0.65
			Positive (≥10 or ≥15 mm[c])	5	30	86	

[a] A kappa statistic of ≥0.75 represents excellent agreement, 0.40 to 0.75 represents good to fair agreement, and <0.40 represents poor agreement.
[b] The CDC criterion of ≥15-mm-diameter induration was taken as a positive Mantoux result because no HCW was from a high-risk or high-prevalence population (2).
[c] Interpretive criteria varied depending on whether the patient was from a group with a high prevalence of tuberculosis (2).

TABLE 2. Performance of QIFN test in active tuberculosis

Group	No.	No. QIFN positive	Sensitivity (%)
Total	60	46	77
Pulmonary	42	31	74[a]
Extrapulmonary	18	15	83[a]
Culture positive	48	36	75[b]
Clinical diagnosis	12	10	83[b]
Smear positive	28	20	71[c]
Smear negative	20	16	80[c]

[a] $\chi^2 = 0.639$; $P > 0.1$.
[b] $\chi^2 = 0.373$; $P > 0.1$.
[c] $\chi^2 = 0.457$; $P > 0.1$.

berculosis disease are shown in Table 3. Mantoux test results were not different from those of the QIFN assay. Similarly, there was no significant difference between the sensitivities of the assays when the patients were stratified by pulmonary or extrapulmonary disease (Table 3).

The QIFN assay was positive for MAC in three of seven cases of MAC colonization. One patient with active MAC disease, who was clinically suspected to have concurrent MTB disease and had a positive blistering Mantoux test result, had a positive *M. tuberculosis* response in the QIFN assay. Two of the four cases of previously treated MAC disease had a positive *M. tuberculosis* QIFN response; one of these had been treated for tuberculosis in the past.

DISCUSSION

The resurgence of tuberculosis and the increase in multidrug-resistant strains has intensified the need for rapid and accurate screening and case finding. Despite the widely acknowledged limitations of the Mantoux test, due to the lack of an effective alternative, it remains the "gold standard" for identifying tuberculous infection (9). The sensitivity of the Mantoux test in patients with culture-positive tuberculosis has been estimated to be 87%, with a specificity of 80% in a nonvaccinated healthy population and a 13 to 15% variability in reading of the skin test (3). In comparison, the sensitivity of the QIFN assay was recently determined to be 83% in individuals with active tuberculosis and 90% in individuals with tuberculous infection (Mantoux positive but no disease) (12). Its specificity was estimated to be 98% in nonvaccinated Australian-born military recruits with low tuberculosis exposure (12).

The Prophit Survey showed a direct relationship between a strong reaction on Mantoux conversion and development of tuberculosis: 11% of the strong converters developed active disease versus 3% of regular converters (10). Thus, the ability to quantitate the reaction to tuberculin may help estimate the risk of developing active disease. Our findings (Fig. 1) support those of Converse et al., which also showed a correlation between the Mantoux induration diameter and the magnitude of the QIFN response, i.e., the assay is able to be quantified in a manner similar to the Mantoux test (4). This is desirable, as it could allow for different cutoff values to be applied to different groups, based on the risk of tuberculosis, just as different induration diameters are used in Mantoux testing (2). This possibility requires further evaluation.

In the present study the agreement between the Mantoux test and the QIFN assay was good to fair, with kappa statistics of 0.55 and 0.65 for group I (immigrants) and group III (patients), respectively. None of the 47 immigrants with discrepant test results had radiological evidence of active disease. It may be speculated that the discrepancies could be accounted for in Mantoux-positive and QIFN-negative participants by the cross-reactivity of the Mantoux test due to exposure to nontuberculous mycobacteria, and in Mantoux-negative and QIFN-positive participants it could be accounted for by a greater sensitivity of the QIFN assay in detecting *M. tuberculosis* infection compared to that of the Mantoux test, as suggested by others (4, 12). One method of resolving this issue would be long-term follow-up of those persons with discrepant results to determine if there is any difference in the rates of developing active disease according to either the Mantoux or QIFN result.

For group II (HCWs), the agreement between the tests was good to fair, with a kappa statistic of 0.48 when a positive Mantoux test result was defined as a ≥10-mm-diameter induration. The kappa statistic decreased to 0.26, indicating poor agreement, when the CDC interpretive criterion of ≥15-mm-diameter was used. Due to the antigens shared across mycobacterial species and the lack of specificity of the tuberculin used, both the Mantoux test and the QIFN assay can suffer from cross-reactivity in *M. bovis* BCG-vaccinated individuals. For the Mantoux test this is compensated for in the New Zealand guidelines by increasing the cutoff from ≥10-mm to ≥15-mm diameter for BCG-vaccinated people, to increase its specificity (8). A corresponding modification in interpretation of QIFN assay results is not recommended by the manufacturer for BCG-vaccinated individuals. However, newer antigens, such as ESAT-6 and MTP-64, may increase the specificity of the QIFN assay in these situations. ESAT-6, a low-molecular-weight secreted antigen, and the gene encoding it, *esat-6*, are absent in BCG strains (7). Preliminary studies indicate that immunological response to ESAT-6 is not influenced by previous BCG vaccination, in that 11 of 54 (20%) individuals tested QIFN positive 5 months after BCG vaccination with tuberculin PPD (human), but none were positive when ESAT-6 was used in its place in the QIFN assay (5a).

A positive acid-fast smear remains the most rapid diagnostic test for tuberculosis. However, smear-positive cases only comprise about 50% of patients with pulmonary disease and are less frequent in cases of extrapulmonary disease and pediatric tuberculosis (1). If the QIFN assay was equivalent to or better than the Mantoux test at detecting active tuberculosis, it would be useful in cases where conventional laboratory means of diagnosis fail. Our results show that the sensitivity of the QIFN assay in detecting various forms of *M. tuberculosis* disease was 70 to 80%. This difference from the performance of the Mantoux test was not statistically significant, and therefore the assay could potentially replace the Mantoux test in clinical situations where tuberculosis is suspected.

By virtue of its ability to quantitate a differential response to the human and avian PPDs, the QIFN assay has the potential to discriminate between *M. tuberculosis* and MAC disease. The QIFN assay correctly differentiated 50 culture-confirmed *M. tuberculosis* cases from 10 cases of MAC cervical lymphadenitis

TABLE 3. Performance of QIFN and Mantoux tests in active tuberculosis

Group (*n*)	Test	No. positive	Sensitivity (%)
Total (38)	Mantoux	33	87[a]
	QIFN	30	79[a]
Pulmonary (26)	Mantoux	23	88[b]
	QIFN	20	77[b]
Extrapulmonary (12)	Mantoux	10	83
	QIFN	10	83

[a] $\chi^2 = 0.835$; $P > 0.1$.
[b] $\chi^2 = 1.209$; $P > 0.1$.

in children (6). The present study showed that all cases of active tuberculosis detected by the QIFN assay had a predominant response to human PPD and were identified as *M. tuberculosis* infection. None had a predominant MAC response. The assay also accurately detected three of the seven cases of MAC colonization. Too few cases of MAC disease were tested to draw any conclusions about the discriminatory power of the QIFN assay for adult patients with respiratory MAC disease or colonization. More data are required for patients with sputum cultures containing MAC to determine if the QIFN assay can help differentiate colonization from disease in this difficult to evaluate patient group.

QIFN assay works on the same principles as the Mantoux test but, being an in vitro test, has several inherent advantages: it only requires a single patient visit, it lacks a booster effect, and its interpretation is more objective. The inclusion of a phytohemagglutin as a positive mitogen control allows the identification of those who are negative because they cannot mount an in vitro response. This is not possible in routine Mantoux testing, where a negative result due to immunosuppression is not detected. Furthermore, it has been suggested that by showing a reduction in human PPD response the QIFN assay could also be useful in monitoring the response to treatment (11). Sodhi et al. showed that reduced IFN-γ production by peripheral blood mononuclear cells stimulated by heat-killed MTB is a marker of severe tuberculosis in both HIV-negative and -positive patients with tuberculosis (11). Whether the QIFN assay could be used similarly to assess the severity of tuberculosis requires investigation.

Conclusion. We evaluated the QIFN assay in three different settings, immigrant and HCW screening and evaluation of cases of *M. tuberculosis* and MAC disease. The QIFN results showed a correlation with Mantoux test results. When CDC interpretive guidelines were used, the agreement between the QIFN assay and Mantoux test was good to fair for immigrants and patients but poor for the HCW group. The sensitivity of the assay was not significantly different from that of the Mantoux test in cases of active tuberculosis, and it detected three of the seven cases of MAC colonization. The QIFN assay seems to be a promising new approach to detect evidence of mycobacterial infection. In some situations it may have the potential to replace the Mantoux skin test. The assay does, however, require laboratory facilities to stimulate viable lymphocytes and EIA to quantify IFN-γ. More experience is needed from long-term studies to determine whether the management of contacts of individuals with infectious tuberculosis, based in part on Mantoux results, can be extrapolated to the QIFN assay results. The acceptability of the performance values observed in this study and those of similar evaluations will determine the place this test will have in detecting evidence of mycobacterial infection, either as a screening test or for those suspected of having active disease.

ACKNOWLEDGMENTS

We thank Lester Calder, Martin Reeves, Alison McCleod, Tony Wansborough, Chris Walls, the staff of the tuberculosis ward (GLH) and the occupational health clinics (GLH & AKH), and the staff of the Mangere Refugee Resettlement Centre for their help with the study and the immigrants, the Microbiology staff at GLH, AKH, and Diagnostic laboratories, and the staff attending occupational health clinics at GLH and AKH for their participation. We also thank Jim Rothel, Gavin Horrigan, and Judy Woodard of CSL Biosciences for their technical advice and assistance with the statistical analysis.

REFERENCES

1. **Bothamley, G. H.** 1995. Serological diagnosis of tuberculosis. Eur. Respir. J. **8**(Suppl. 20)**:**676–688.
2. **Centers for Disease Control and Prevention.** 1994. Guidelines for preventing the transmission of *Mycobacterium tuberculosis* in healthcare facilities. Morbid. Mortal. Weekly Rep. **43**(RR-13)**:**59–64.
3. **Chaparas, S. D., H. M. Vandiviere, I. Melvin, G. Koch, and C. Becker.** 1985. Tuberculin test. Variability with the Mantoux procedure. Am. Rev. Respir. Dis. **132:**175–177.
4. **Converse, P. J., S. L. Jones, J. Astemborski, D. Vlahov, and N. M. H. Graham.** 1997. Comparison of a tuberculin interferon-γ assay with the tuberculin skin test in high-risk adults: effect of human immunodeficiency virus infection. J. Infect. Dis. **176:**144–150.
5. **Desem, N., and S. L. Jones.** 1998. Development of a human gamma interferon enzyme immunoassay and comparison with tuberculin skin testing for detection of *Mycobacterium tuberculosis* infection. Clin. Diagn. Lab. Immunol. **5:**531–536.

5a. **Johnson, P. D. R.** Personal communication.

6. **Mukherjee, S., R. Lumb, P. Robinson, and R. Stapledon.** 1995. Evaluation of gamma interferon (IFN-γ) assay in human mycobacterial infection. Aust. N. Z. J. Med. **25:**436.
7. **Pollock, J. M., and P. Anderson.** 1997. The potential of the ESAT-6 antigen secreted by virulent mycobacteria for specific diagnosis of tuberculosis. J. Infect. Dis. **175:**1251–1254.
8. **Public Health Group.** 1996. Guidelines for tuberculosis control in New Zealand, p. 18–22. Ministry of Health, Wellington, New Zealand.
9. **Reichman, L. B.** 1998. A scandalous incompetence . . . continued. Chest **113:** 1153–1154.
10. **Sepkowitz, K. A.** 1996. Tuberculin skin testing and the health care worker: lessons of the Prophit Survey. Tuber. Lung Dis. **77:**81–85.
11. **Sodhi, A., J. Gong, C. Silva, D. Qian, and P. F. Barnes.** 1997. Clinical correlates of Interferon γ production in patients with tuberculosis. Clin. Infect. Dis. **25:**617–620.
12. **Streeton, J. A., N. Desem, and S. L. Jones.** 1998. Sensitivity and specificity of a gamma interferon blood test for tuberculosis infection. Int. J. Tuber. Lung Dis. **2:**443–450.
13. **Wood, P. R., L. A. Corner, J. S. Rothel, C. Baldock, S. L. Jones, D. B. Cousins, B. S. McCormick, B. R. Francis, J. Creeper, and N. E. Tweddle.** 1991. Field comparison of the interferon-gamma assay and the intradermal tuberculin test for the diagnosis of bovine tuberculosis. Aust. Vet. J. **68:**286–290.