AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 46:453–462 (2004)

The Beryllium Lymphocyte Proliferation Test: Relevant Issues in Beryllium Health Surveillance

Arthur W. Stange, PhD,* F. Joseph Furman, MD, DSc, and Duane E. Hilmas, DVM, PhD

Background *The beryllium lymphocyte proliferation test (Be-LPT) measures beryllium-specific cellular immune response, and is useful in medical surveillance of beryllium sensitivity and chronic beryllium disease (CBD).*

Methods *Current and former employees (n = 12,194) of 18 United States Department of Energy (DOE) sites were tested for beryllium sensitization at four laboratories with Be-LPT expertise. Beryllium sensitized individuals were offered evaluations for CBD. The sensitivity, specificity, and positive predictive value (PPV) of the Be-LPT were determined, as was inter- and intra-laboratory agreement.*

Results *False positives were calculated to be 1.09%, with a laboratory range of 0.00–3.35% for the 10-year investigation. Be-LPTs performed on inter-laboratory split blood specimens from sensitized individuals showed a false negative rate of 31.7%. The intra-laboratory repeatability of abnormal Be-LPT results ranged from 80.4–91.9%. The sensitivity of the Be-LPT was determined to be 0.683, with a specificity of 0.969. The PPV of one abnormal Be-LPT was 0.253.*

Conclusions *The Be-LPT is efficacious in medical surveillance of beryllium-exposed individuals. The PPV of the Be-LPT is comparable to other widely accepted medical tests. Confirmation of an abnormal result is recommended to assure appropriate referral for CBD medical evaluation.* Am. J. Ind. Med. 46:453–462, 2004. *Published 2004 Wiley-Liss, Inc.*†

KEY WORDS: beryllium; lymphocyte proliferation test; sensitivity; chronic beryllium disease; health surveillance; occupational

INTRODUCTION

The beryllium lymphocyte proliferation test (Be-LPT) is an in vitro measure of beryllium antigen-specific, cell-mediated immune response that has been shown to be useful in identifying people at elevated risk of developing chronic beryllium disease (CBD) [Williams and Williams, 1983; Rossman et al., 1988; Mroz et al., 1991, Newman et al., 1996; Newman, 1996; Stange et al., 1996b]. The Be-LPT can identify individuals who, upon exposure to sufficient airborne beryllium, initiate an immune response, herein referred to as beryllium sensitivity that plays a central role in the immunopathogenesis of CBD [Newman et al., 1989; Newman, 1994, 2000; Fontenot et al., 1999]. The electronics, aerospace, defense and nuclear weapons industries have been large users of beryllium and beryllium alloys.

The Be-LPT measures the response of T-cell lymphocytes isolated from heparinized peripheral venous blood or lung lavage cells. The harvested T-cells are cultured in a medium containing the DNA precursor, thymidine, which has been labeled with tritium (^{3}H), a radioactive isotope of hydrogen. Cell uptake of the tritiated thymidine reflects the immune response-driven cell proliferation of T-cells

Center for Epidemiologic Research, Oak Ridge Institute for Science and Education, Arvada, Colorado

Disclaimer: The opinions and observations contained in this manuscript are based upon the authors' review and analysis of the medical surveillance program information reported herein. The opinions and observations are solely those of the authors and are not to be considered as reflecting any approval of these opinions or observations by the United States Department of Energy, or by any past or present management and operating or integrating contractors or subcontractors of the Rocky Flats Environmental Technology Site.

Contract grant sponsor: Oak Ridge Institute for Science and Education (ORISE); Contract grant number: DE-AC05-00OR22750.

*Correspondence to: Arthur W. Stange, Center for Epidemiologic Research, Oak Ridge Institute for Science and Education, 9950 West 80th Avenue, Suite 17, Arvada, Colorado 80005. E-mail: stangea@orau.gov

Accepted 8 July 2004
DOI 10.1002/ajim.20082. Published online in Wiley InterScience (www.interscience.wiley.com)

stimulated with 1, 10, and 100 μM beryllium salts. Between 1992 and 2001, all of the Be-LPT laboratories used beryllium sulfate. One of the laboratories also used beryllium fluoride in performing the Be-LPT.

A stimulation index is the ratio of the radioactivity (counts per minute) of beryllium-exposed cell cultures to the count rate of unstimulated cultures. The stimulation indices for tested individuals are compared to those of a non-exposed control group that define the normal range for the Be-LPT results [Rossman et al., 1988; Newman, 1996]. Each test has six possible stimulation index values resulting from three beryllium concentrations and two incubation periods. Two or more stimulation index values of six possible values that exceed the normal range index by at least two standard deviations is the definition of an abnormal test. One of six stimulation index values that exceed the normal range index by at least two standard deviations is a borderline-abnormal test. From 1992 to 2000, two of the four laboratories used a standard ratio of 3.0 (stimulated to unstimulated) as the cut-off for an abnormal result, and the other two laboratories used a statistically calculated lab-specific abnormal cut-off. Beginning in 2001, a third laboratory began using a statistically calculated lab-specific abnormal cut-off.

In previous studies, Be-LPT results were nearly always abnormal in persons with CBD, indicating that the Be-LPT might be an effective surveillance test for both beryllium sensitivity and CBD [Williams and Williams, 1983; Rossman et al., 1988; Morz et al., 1991; Newman et al., 1996; Newman, 1996; Stange et al., 1996b]. However, questions remained regarding the test sensitivity and specificity and the rates of false positives and false negatives when using the Be-LPT [Deubner et al., 2001].

A very limited amount of information has been published concerning the sensitivity and specificity of the Be-LPT [Bobka et al., 1997; Newman, 2000; Deubner et al., 2001]. Test sensitivity and specificity cannot be easily defined for beryllium sensitivity because the Be-LPT is the only practical means to determine beryllium sensitivity and there is no other standard test with which to compare.

Beryllium sensitivity and CBD cases have been reported previously for current and former employees of the United States Department of Energy (DOE) Rocky Flats Environmental Technology Site (RFETS) [Kreiss et al., 1989, 1993; Stange et al., 1996a,b, 2001]. This investigation significantly increased the size of the RFETS study population, and provided additional information from 17 other DOE sites. The authors estimate the sensitivity and specificity of the Be-LPT, the false positive and negative rates, and the PPV using RFETS Be-LPT data collected from 1992 to 2001. Of the 18 DOE sites represented in this investigation, the RFETS site provided the largest cohort with the longest history of beryllium sensitivity testing. The authors also estimate the background rates of abnormal Be-LPT results found in RFETS individuals with no known beryllium exposure. Inter- and intra-laboratory agreement were analyzed for RFETS Be-LPT results and Be-LPT data from 17 other DOE sites where sensitivity testing was performed during the 1999 to 2001 time period.

MATERIALS AND METHODS

Be-LPT results (n = 25,643) obtained from 1992 to 2001 for current and former employees from 18 DOE and predecessor sites were analyzed to evaluate the efficacy of the Be-LPT for determining beryllium sensitivity and CBD. These 18 sites represent a variety of DOE missions (national laboratories, production, and support) and accordingly a variety of beryllium exposure frequency and exposure level scenarios. Beryllium exposure at 17 of the sites included either beryllium or beryllium oxide, while exposure at one site was limited to a beryllium-copper alloy. In January 2000, DOE sites implemented 10 CFR 850, the Chronic Beryllium Disease Prevention Program, to reduce exposure to beryllium and establish medical surveillance using the Be-LPT (10 CFR 850, 1999). Additionally, in April 2001, the DOE published Specification-1142-2001, Beryllium Lymphocyte Proliferation Testing (Be-LPT), which provides the technical specifications for laboratories performing the test [United States Department of Energy, 2001]. By this specification, an abnormal Be-LPT is defined as two or more of six possible stimulation index values above a calculated cut-off value specific for each laboratory.

Four laboratories with demonstrated expertise in performing Be-LPTs processed the submitted blood specimens. Be-LPT results from the largest cohort, RFETS, were used to determine the sensitivity and specificity of the test, the false positive and false negative rates for the test, and the usefulness of serial Be-LPTs. A subcohort of the RFETS population and beryllium non-exposed new hires was tested to determine background levels of beryllium sensitivity and CBD. Inter- and intra-laboratory split-specimen agreement included Be-LPT results from all 18 DOE sites where testing occurred between 1992 and 2001.

An Institutional Review Board (IRB) reviewed all procedures and materials used by the surveillance program on an annual basis. An informed consent approved by the IRB was obtained from each participant prior to any testing associated with the program. Participants were asked to complete a self-administered questionnaire regarding their medical and occupational history and demographics. An interviewer with subject matter expertise reviewed the questionnaire for completeness with each participant. All participants were offered testing for beryllium sensitivity using the Be-LPT.

Beryllium Sensitivity Surveillance

To determine if individuals not exposed to known sources of beryllium could be identified as having beryllium sensitivity, 458 participants with no known beryllium

exposure were tested. The potential for exposure to beryllium was based on the results of a work and exposure history questionnaire and interviews. The non-exposed cohort included 291 newly hired RFETS employees who had no known exposure to beryllium in prior employment and 167 current RFETS employees who had no known exposure to beryllium prior to their RFETS employment and had not entered RFETS beryllium buildings or work areas. Based on industrial hygiene monitoring (1959–1998) and hazardous materials documentation, beryllium characterization for the RFETS site was available. The interviewer used this information to confirm that current RFETS non-exposed employees had not entered buildings, work areas, or storage locations where beryllium was used or maintained. As part of this non-exposed cohort testing, new hires received an initial Be-LPT prior to their entering the site. The non-beryllium-exposed current employees received initial, 3- and 6-year follow-up Be-LPTs. The non-exposed current employees originally tested completed a follow-up exposure and work history questionnaire and an interview to determine their potential for beryllium exposure since their initial test, prior to receiving the Be-LPT. Employees who reported entering a beryllium work area were removed from the non-exposed cohort.

Serial Be-LPT testing was offered at approximate 3-year intervals to current and former employees who had normal Be-LPT results and who were no longer occupationally exposed. Current employees with the potential for ongoing exposure to beryllium were tested annually. Participants (no longer exposed to beryllium) who had an abnormal Be-LPT result that had not been confirmed by a second Be-LPT received serial retesting at 6- to 12-months, decreasing in frequency to every 3 years. Serial testing allowed for the identification of Be-LPT normal-to-abnormal conversions and of initial false negative Be-LPT results. Individuals identified as sensitized did not receive additional peripheral blood Be-LPTs as part of the surveillance program, but were offered ongoing tests as part of their CBD medical evaluation.

Blood specimens consisting of 30 ml of peripheral venous heparinized blood were collected from program participants and sent via overnight carrier to predetermined laboratories. Specimens for inter-laboratory comparisons were selected from individuals on a random basis (5–8% of individuals on a weekly basis) and from those with other than normal Be-LPT results.

In recognition of the possibility of inter- and intra-laboratory variability in the test results, a person was not considered beryllium sensitized unless an initial abnormal Be-LPT result was confirmed by one or more concurrent or subsequent Be-LPTs.

CBD Medical Evaluations

Participants sensitized to beryllium were offered clinical evaluations on an ongoing basis to monitor for the development of CBD. Participants diagnosed with CBD were offered periodic clinical evaluations to monitor the progression of the disease and to provide treatment where appropriate. The frequency of beryllium sensitivity and CBD medical evaluations were at the discretion of the pulmonologist based on individuals' health status and ranged from 3 months to 3 years.

All participants with beryllium sensitivity were offered a diagnostic clinical evaluation at one of several major medical facilities in the United States having expertise in the diagnosis of CBD. Pulmonologists at different medical centers did not use the same medical evaluation protocol. A ''complete'' clinical evaluation generally included a work history, physical examination, blood chemistry (including CBC with differential), peripheral blood Be-LPT, exercise physiology/pulmonary function, posterior-anterior chest X-ray with B-reader interpretation [International Labor Organization (ILO), 1980], CT of the lungs, bronchoalveolar lavage (BAL) with biopsy, and a BAL Be-LPT. A diagnosis of CBD resulted from a history of beryllium exposure, histopathologic evidence on biopsy of non-caseating granulomas or mononuclear cell infiltrates in the lung, a confirmed abnormal peripheral blood Be-LPT, and an abnormal BAL Be-LPT. CBD was also diagnosed in participants with normal BAL Be-LPTs, confirmed abnormal peripheral blood Be-LPTs and with histologic evidence on biopsy or CT evidence of pulmonary granulomatous disease. Because of the difficulty in obtaining adequate biopsy specimens, a CBD diagnosis was made with immunologic evidence of beryllium sensitivity (peripheral blood Be-LPT or BAL Be-LPT) and other compelling evidence of pulmonary disease such as decreased diffusing capacity, CT evidence of granulomatous disease, or increased level of mononuclear cell infiltrates consistent with CBD. CBD was rarely diagnosed in individuals with only abnormal peripheral blood and BAL Be-LPTs. Participants who were identified as sensitized to beryllium or who were diagnosed with CBD were notified of the potential hazards related to further exposure to airborne beryllium.

Analysis of Be-LPT Results

Sensitivity and specificity of the Be-LPT were calculated using the more extensive RFETS Be-LPT data set containing results from laboratories performing a high volume of tests each year.

False positives and false negatives were calculated using the following definitions. If an abnormal test result could not be confirmed by additional (minimum of two) retest Be-LPTs (first retest usually performed within 2 months of the abnormal), then the original abnormal result was considered a false positive. False negative results were identified only among those participants found sensitized based on two or more abnormal results. Among this group, a false negative

result was defined as a normal test result concurrent with or following within 2 years of an abnormal result.

If a normal result occurred 2 or more years after the last of two or more abnormal results, it was considered a true negative. This is because we assume that immune response might have diminished with time in some individuals. Weakly sensitized individuals can show both normal and abnormal Be-LPT results. In these individuals, no effort was made to eliminate the normal results from the calculation of false negatives; however, the normal results would not be a failure of the test, but a manifestation of the inability of weakly sensitized individuals to mount a consistently abnormal result.

The positive predictive values (PPV) (true positives/(true positives + false positives)) for the first abnormal Be-LPT result being confirmed as beryllium sensitivity (2 abnormal Be-LPT results), for the first abnormal Be-LPT resulting in a CBD diagnosis, and for beryllium sensitivity resulting in a CBD diagnosis were calculated.

Inter- and intra-laboratory agreement were analyzed for the entire 10-year investigation period by comparing Be-LPT results for split blood specimens sent to different laboratories and sequential specimens sent to the same laboratory. For inter-laboratory comparisons, the percent of split-specimen pairs that resulted in an abnormal-abnormal agreement of results was calculated. These percents were tabulated both for all split specimens and for only those split specimens taken from individuals with confirmed beryllium sensitivity. For intra-laboratory comparisons, performed for sensitized individuals only, we calculated the percent of abnormal results obtained for sequential specimens.

Using split-specimen results that were obtained from the RFETS sensitized participants, the authors estimate the percent of false negatives according to the following equation: $x = (r/(2 + r))*100$, where: x = fraction of false negatives, r = (number of normal-abnormal pairs)/(number of abnormal-abnormal pairs), 2 = constant reflecting the two possibilities for the result (normal and abnormal), and 100 = conversion from fraction to percentage.

The second approach in examining inter- and intra-laboratory agreement was the calculation of the weighted kappa statistic. Kappa measures the level of agreement that exceeds the agreement that would occur by chance alone, and can range from −1.0 to +1.0 [Cohen, 1960]. Using a weighted kappa allows comparisons among sample pairs of three Be-LPT result types (abnormal, borderline-abnormal, and normal), with weighting factors used to reflect the level or relative seriousness of disagreement among the types [Spitzer et al., 1967]. Other possible Be-LPT outcomes, including unsatisfactory and uninterpretable results, were excluded from the analysis. The weighting factors for the Be-LPT comparisons were calculated using the method of Fleiss and Cohen [1973]. The ranges of kappa, <0.21, 0.21–0.40, 0.41–0.60, 0.61–0.80, and 0.81–1.00, are considered to represent poor, fair, moderate, good, and excellent agreement, respectively [Landis and Koch, 1977].

For the inter-laboratory kappa calculations, the authors used data from all 18 DOE sites for split-specimen Be-LPTs performed between 1992 and 2001 (n = 7,394 paired-tests). Likewise, for intra-laboratory comparisons, normal, borderline-abnormal, and abnormal results from sequential sampling were used in calculating kappas.

RESULTS

Identification of Beryllium Sensitivity and CBD

A total of 12,194 current and former employees of 18 DOE sites, who had known potential for exposure to beryllium, were tested for beryllium sensitivity. Eighty-one percent of those tested were former employees and had no potential for ongoing exposure to beryllium. A total of 25,643 Be-LPTs were performed on the 12,194 participants. The 1,902 Be-LPT results that were borderline-abnormal, unsatisfactory, or uninterpretable were excluded from the analyses except for the inter- and intra-laboratory agreement and kappa. Unsatisfactory and uninterpretable Be-LPT rates, 1992 to 2001, averaged 4.12% (range 2.55–6.13%). No correlation was found between the annual number of Be-LPTs performed by a laboratory and the unsatisfactory and uninterpretable Be-LPT rates.

The authors examined 20,275 RFETS Be-LPT results (1,427 borderline-abnormal, unsatisfactory, or uninterpretable results were excluded) to determine the test sensitivity and specificity of the Be-LPT. The prevalence of beryllium sensitivity without CBD in the subject population is 2.35% (184/7,820), the prevalence of CBD is 1.50% (117/7,820), and the combined beryllium sensitivity and CBD prevalence is 3.85% (301/7,820). Ultimately, 38.9% (117/301) of the sensitized participants were diagnosed as having CBD.

The data in Table I track the progression of diagnosis from sensitivity to CBD. Of the 117 diagnosed with CBD, 75.2% were diagnosed on their initial CBD clinical evaluation, 90.5% by their second evaluation, and 96.6% by the third. Participants may not have received the same diagnostic tests, due to variations in protocol at different medical centers or participant-specific medical contraindications. Ninety-four percent of beryllium-sensitized participants diagnosed with CBD had a ''complete'' clinical evaluation, and 37% of participants diagnosed as beryllium-sensitized without CBD had a ''complete'' evaluation. Because of variation in medical evaluation diagnostics and individual preference, some sensitized individuals may not have been offered all of the tests necessary to determine whether they had CBD. As a result, the PPV of the Be-LPT in determining CBD may be underestimated.

TABLE I. Progression of Sensitivity to Chronic Beryllium Disease (CBD)

Number of clinical evaluations	Number of CBD cases diagnosed	Mean number of years*: sensitivity to CBD	Percent of total	Cumulative percent of total
1	88	0.34	75.2	75.2
2	18	2.73	15.4	90.5
3	7	3.65	6.0	96.6
4	3	4.83	2.6	99.1
5	1	4.48	0.9	100.0
Totals	117		100.0	100.0

*Denotes time from the identification of beryllium sensitivity to the diagnosis of CBD.

Non-Exposed Population

Based on interviews and completed exposure and work history questionnaires, 291 RFETS new hires and 167 current RFETS employees identified as having no known exposure to beryllium, received Be-LPTs. As shown in Table II, of the 458 receiving an initial Be-LPT, 7 had an abnormal response (3 new hires, 4 current employees). Upon retesting at the laboratory that had identified the abnormal response and a second quality control laboratory, none of these abnormal results were repeatable. As a result of completed questionnaires and interviews with current employees regarding the potential for beryllium exposure, 149 of the 167 who received an initial Be-LPT received a 3-year retest. Of the 149 current employees who received a 3-year serial retest, 2 individuals with previously normal results were identified with abnormal results, but were not confirmed upon retesting. Based on interviews and completed questionnaires, 136 of the 149 who received a 3-year retest were offered and completed a 6-year retest. None of the 6-year retests were identified as having an abnormal response.

False Positives

The annual false positive rates for the individual laboratories range from 0.00% to 3.35%, with an average for the four laboratories over the entire 10-year period of 1.09%. Among the 7,820 RFETS participants tested during the 10-year period, abnormal results for 218 (2.79%) could not be confirmed with repeat testing. The Be-LPT laboratory average false positive rate of 1.09% is lower than the 2.79% of RFETS participants with unconfirmed abnormal results because RFETS participants were serially tested for 10 years, providing an increased opportunity of having a false positive result.

False Negatives

When only normal and abnormal results were considered, there were 153 normal-abnormal pairs and 165 abnormal-abnormal pairs. The percent of false negatives was calculated as $((153/165)/(2+(153/165)))*100\% = 31.7\%$. When pairs with borderline-abnormal results were included and the borderline-abnormal results counted as abnormal results, there were 164 normal-abnormal pairs and 214 abnormal-abnormal pairs. With borderline-abnormal results included, the percent of false negatives was $((164/214)/(2+(164/214)))*100\% = 27.7\%$. The two laboratories with the most consistent agreement between their results have a 24.6% false negative rate, based on 43 normal-abnormal pairs and 66 abnormal-abnormal pairs. When borderline-abnormal results were counted as abnormal results, their rate of false negatives was 21.0% (44 normal-abnormal pairs and 83 abnormal-abnormal pairs).

Table III lists initial Be-LPT results for the 301 RFETS participants who had a known potential for beryllium exposure and who ultimately were determined to be sensitized. The rates of false negatives calculated above using split-specimen pairs can be compared to the values listed in Table III for the Percent of Total Cases for the Normal Results on Initial Tests: 24.7% for CBD; 24.5% for Sensitized Only; and 24.6% for the 2 categories combined (Total). A total of 519

TABLE II. Be-LPT Results for Beryllium Non-Exposed RFETS Employees*

	Employees tested	Abnormal Be-LPTs	Abnormals confirmed	Number with 3-year retests	Abnormal Be-LPTs	Abnormals confirmed	Number with 6-year retests	Abnormal Be-LPTs
New hire (RFETS)	291	3	0	n/a	n/a	n/a	n/a	n/a
Current (RFETS)	167	4	0	149	2	0	136	0
Total	458	7	0	149	2	0	136	0

*Rocky Mountain Flats Environmental Technology Services.

TABLE III. Initial Be-LPT Results for RFETS Participants with a Known Potential for Exposure and Who Ultimately Were Determined to be Sensitized

	Be-LPT results-initial test		
	Normal (%)	Abnormal (%)	Borderline abnormal (%)
CBD (n = 117)	29 (24.7)	78 (66.7)	10 (8.5)
Sensitized, no CBD (n = 184)	45 (24.5)	124 (67.4)	15 (8.2)
Total (n = 301)	74 (24.6)	202 (67.1)	25 (8.3)

confirmational Be-LPT tests were performed for the 301 individuals who ultimately were found to be sensitized. False negative Be-LPTs occurred in 138 of the 519 (26.6%) tests performed for these participants.

To determine the number of serial tests after which no further testing is informative or necessary, the authors looked at how many normal tests preceded the first abnormal test in all of those participants ultimately confirmed to be sensitized. The distribution of these results is listed in Table IV. When the serial Be-LPT results for the 301 sensitized participants were examined, it was observed that 13 of these participants had three or more normal Be-LPT results prior to the first abnormal result. Using a single-test false negative rate of 30%, the combined false negative rate for four tests in series is 0.81% (30%*30%*30%*30%); the combined false negative rate for five tests in series is 0.24%. The rate of negatives represented by the 13 cases exceeds what would be expected from false negatives alone, and at least some of these 13 sensitized participants may have expressed true immune system conversion. However, 3 of the 13 had some minimal potential for exposure to beryllium between their serial Be-LPTs because they continued to work at RFETS, although they were not assigned to buildings in which beryllium was used. The majority (69.1%) of sensitized participants were found to have an abnormal Be-LPT result on their initial test.

Inter- and Intra-Laboratory Agreement

Tables VA and VB summarize the comparisons of inter-laboratory split-specimen quality control testing data for the Be-LPT laboratories. Test results other than abnormal, borderline-abnormal, and normal were excluded from the analysis. These data include split specimens from the 18 DOE sites where Be-LPTs were performed during the 1992–2001 time period. As shown in Table VA, the inter-laboratory agreement for abnormal results ranged from a low of 26.2% to a high of 61.8%. In those instances where borderline-abnormal results were received from both laboratories (n = 15), the results were counted as an abnormal match. In Table VB the inter-laboratory agreement for abnormal results was analyzed for the sensitized population only. The sensitized population was selected in order to eliminate the potential for false positive Be-LPT results influencing the analysis. Table VC summarizes the intra-laboratory agreement for abnormal results among sequential specimens analyzed for the sensitized population from the 18 DOE sites. This agreement ranges from 80.4% to 91.9%.

Using kappa statistics, the best agreement occurred between laboratories 1 and 4, with a kappa of 0.61 indicating good agreement. Moderate agreement was demonstrated between laboratories 1 and 3 (kappa = 0.48) and laboratories 2 and 3 (kappa = 0.42). The agreement for sequential specimens tested within the same laboratory ranged from moderate to fair using weighted kappa, but, as noted above, the agreement rate for abnormal results of sequential specimens ranged from 80.4% to 91.9%. Laboratory 1 with 1,967 sequential specimen tests (86.2% agreement among abnormal results, kappa = 0.51) and laboratory 3 with 1,851 sequential specimen tests (91.9%, kappa = 0.50) showed moderate agreement.

For each laboratory, sensitivity and CBD outcomes for individuals whose serial Be-LPT results went from borderline-abnormal to normal, borderline-to-borderline and borderline-abnormal to abnormal also were examined. Only the

TABLE IV. Frequency of Normal Be-LPTs in Serially Tested RFETS Participants Ultimately Determined to be Sensitized

	Number of normal Be-LPT results preceding the first abnormal result					
	0	1	2	3	4	5
CBD	85	17	7	0	1	0
Sensitized only	123	38	18	9	1	2
Total	208	55	25	9	2	2
Cumulative total	208	263	288	297	299	301
Cumulative % of total sensitized	69.1	87.4	95.7	98.7	99.3	100
Participants tested	7,820	7,261	5,119	2,783	1,543	674
% Abnormal	2.66	0.76	0.49	0.32	0.13	0.30

TABLE VA. Inter-Laboratory Split-Specimen Agreement, 1992–2001

Laboratories	Number of test pairs	Number of abnormal-abnormal pairs	% Agreement for abnormal results
1 and 2	1,716	126	26.2
1 and 3	2,604	272	39.7
1 and 4	106	34	61.8
2 and 3	2,496	210	32.4
2 and 4	17	0	n/a
3 and 4	455	22	27.3

TABLE VC. Intra-Laboratory Split-Specimen Agreement, Sensitized Cases Only, 1992–2001

Laboratory	Number of tests	Number of abnormal tests	% Agreement for abnormal results
1	427	368	86.2
2	285	229	80.4
3	345	317	91.9
4	72	64	88.9

first borderline-abnormal result and the subsequent confirmational retest result for each individual was used in the analysis. Of the 19,396 normal, borderline-abnormal and abnormal Be-LPT results received by 7,820 RFETS participants tested between 1992 and 2001, 355 participants exhibited borderline-abnormal results. Borderline-abnormal to normal results (271/355) occurred significantly ($P < 0.01$) more often than borderline-to-borderline results (37/355) or borderline-abnormal to abnormal results (47/355). Individuals, who received a borderline-abnormal result followed by a normal result at the same laboratory, eventually were found to have beryllium sensitivity 9.23% of the time and CBD 3.32% of the time. In individuals who received a borderline-abnormal result followed by another borderline-abnormal result at the same laboratory, a significant ($P < 0.01$) increase was found in the percentage of individuals found to have beryllium sensitivity (35.16%) and CBD (8.11%) over the borderline-abnormal to normal cohort. Similarly, in individuals who received a borderline-abnormal result followed by an abnormal result, a significant ($P < 0.01$) increase was found in the percentage of individuals found to have beryllium sensitivity (51.06%) and CBD (23.40%) versus the borderline-abnormal to normal cohort.

Statistical Parameters

Table VI lists calculated statistical parameters for the Be-LPT. The calculations are based on the data collected from 1992 to 2001 for the RFETS population, as this population provides the largest homogeneous set of beryllium sensitivity and CBD data. The test sensitivity (0.683), true positives/(true positives + false negatives) is the probability that a patient with CBD will have an abnormal Be-LPT result. The test specificity (0.969), true negatives/(true negatives + false positives) is the proportion of normal tests in all patients who do not have CBD. The PPV for the first abnormal Be-LPT result (true positives/(true positives + false positives)) for CBD was 0.253. The PPV for the first abnormal Be-LPT for beryllium sensitivity was 0.580.

TABLE VB. Inter-Laboratory Split-Specimen Agreement, Sensitized Cases Only, 1992–2001

Laboratories	Number of test pairs	Number of abnormal-abnormal pairs	% Agreement for abnormal results
1 and 2	93	34	36.6
1 and 3	218	107	49.1
1 and 4	34	22	64.7
2 and 3	159	65	40.9
2 and 4	0	n/a	n/a
3 and 4	9	5	55.6
Total	513	233	

DISCUSSION

Analysis of data from the RFETS beryllium medical surveillance program demonstrates the efficacy of the Be-LPT in determining beryllium sensitivity. The calculated values for test sensitivity (0.683) and specificity (0.969) of the Be-LPT for a 10-year period in a serially tested population of 7,820 current and former employees illustrate the usefulness of the Be-LPT in a properly designed beryllium medical surveillance program. The calculated PPV indicates that 25.3% of participants with one abnormal Be-LPT result and 38.9% of those found sensitized ultimately were diagnosed with CBD. The PPV for CBD may be underestimated by factors such as the thoroughness of the clinical evaluation, and the number of repeat evaluations a sensitized individual had before diagnosis of CBD. This predictive capability is

TABLE VI. Calculated Parameters for RFETS Be-LPT Results (n = 19,396)

CBD prevalence	1.50%
Sensitized only prevalence	2.35%
Background beryllium sensitivity prevalence	0.00%
False positive rate	1.09%
False negative rate	
Borderline-abnormal results included	27.7%
Borderline-abnormal results excluded	31.7%
Be-LPT sensitivity	0.683
Be-LPT specificity	0.969
Be-LPT (1st abnormal) PPV for sensitivity	0.580
Be-LPT (1st abnormal) PPV for CBD	0.253

similar to that of other widely accepted medical screening tests, such as the prostate-specific antigen, Pap smear, tuberculin skin test, and mammography [Nash and Douglass, 1980; Kerlikowske et al., 1993; Brown et al., 1995; Fahey et al., 1995; Gann et al., 1995; Marshall, 1995; Bass, 2001; Schroeder et al., 2001; Solomon et al., 2001; Kulasingam et al., 2002].

The PPV of the first abnormal Be-LPT for beryllium sensitivity was 0.580. Using the Be-LPT to accurately identify beryllium sensitization is extremely important as it portends a future risk of developing CBD. It may also assist health and safety personnel in determining the effectiveness of changes in workplace practices that are implemented to minimize the risk of beryllium sensitization and CBD.

To assess the potential for environmentally induced beryllium sensitivity, the authors measured beryllium sensitization in a population of RFETS non-beryllium exposed individuals. Normal results from the initial Be-LPTs of 291 non-beryllium exposed RFETS newly hired employees, and from 167 non-beryllium exposed current employees supports the idea that beryllium sensitivity, as the result of non-occupational exposure, is rare. In addition to these initial Be-LPTs, 149 non-beryllium exposed current employees received a second serial test and 136 received a third serial retest without the identification of beryllium sensitivity.

The relatively high rate of false negatives seen in this study demonstrates the importance of serially testing beryllium exposed individuals. False negatives range from about 25% to 38%, depending upon which laboratories and DOE sites are included in the data set and whether borderline-abnormal test results are included and treated as abnormal results. The range of false negatives might be the result of factors that influence a laboratory's ability to repeat an abnormal test. Such factors may include the frequency and level of exposure, the latency of testing, and the type of beryllium exposure at each of the tested sites. False negative results can occur in individuals who are in the process of becoming sensitized to beryllium. As the immune system response grows, the likelihood of a repeatable abnormal test increases while the likelihood of false negative results decreases. This is demonstrated by the inspection of serial Be-LPT results for a beryllium-exposed population.

A small percentage of individuals tested with normal Be-LPT results may, as the result of false negative tests, be sensitized, have CBD and not be offered a CBD medical evaluation. This investigation offered serial Be-LPTs and used a multiple laboratory testing protocol to minimize the probability of missing an individual as the result of a false negative Be-LPT result.

It should be noted that because abnormal tests were generally repeated in the original laboratory, plus one other laboratory, the calculated false negative rates might be artificially increased as the result of our retesting scheme and the fact that a different serum is used in each of the laboratories. Serum differences play a role in laboratory agreement as shown by the differences between intra-laboratory and inter-laboratory results. Additional factors that can affect a laboratory's ability to repeat a positive test include the number of Be-LPTs that the laboratory regularly performs, the proficiency and experience of the laboratory personnel, and the statistics used to analyze the Be-LPT data.

Of the 301 RFETS participants diagnosed with beryllium sensitivity or CBD, 98.7% had abnormal results by the fourth serial Be-LPT. The average period between serial Be-LPTs in the RFETS population was 3.2 years. Four former employees had three or more normal Be-LPT results prior to the first abnormal Be-LPT result, and another 17 had two normal results prior to an initial abnormal result. Based on interviews and questionnaire data, these 21 former employees were found to have no identifiable opportunity for beryllium exposure for at least 9 years (range 9–23 years). These findings indicate the potential for the development or recognition of beryllium sensitivity long after exposure to beryllium has ceased. For individuals who no longer have the opportunity for exposure to beryllium and no B-reader chest X-ray findings or symptoms suggestive of CBD, our results indicate that the value of the Be-LPT is diminished after the third or fourth normal test result, and that testing can be stopped if serial Be-LPTs are normal. However, periodic surveillance testing using the Be-LPT is recommended, if the potential for exposure to beryllium continues. Individuals with known exposure to beryllium and unexplained pulmonary findings consistent with CBD should be considered for further clinical evaluation even with repeatedly normal Be-LPTs.

These results aid those planning long-term beryllium medical surveillance efforts by providing an empirical basis for determining the length of follow-up required to determine the presence or absence of beryllium sensitivity or CBD, in individuals no longer exposed to beryllium. Individuals in beryllium surveillance testing programs should be instructed that if unexplained respiratory symptoms arise that are suggestive of CBD, they should contact their beryllium medical surveillance provider for further guidance and potential evaluation.

The authors previously reported a 21% to 38% agreement among three Be-LPT laboratories for split-specimen tests that followed previous abnormal results [Stange et al., 2001]. Other investigations have indicated inter-laboratory variability in Be-LPT results [Kreiss et al., 1997; Deubner et al., 2001]. These previous studies reflect data that is nearly 10 years old. Since that data was reported, the increase in inter-laboratory agreement for abnormal results may be attributed to several factors including improved laboratory technique (the standardization of Be-LPT protocols) and the laboratories increased experience in performing the Be-LPT. We believe the data and analysis of this current investigation is more reliable. The data set used in this investigation to

examine laboratory agreement for split specimens was expanded to include data from 17 other DOE sites. Agreement between some laboratories is markedly different than between other laboratories, for example, Lab 1 has better agreement with Lab 4 (61.8%) than with the other laboratories (Lab 2: 26.2%; Lab 3: 39.7%). These inter-laboratory differences may be the result of several factors including, the statistical interpretation of the Be-LPT data, the differences or similarities in the human serum pools used by each of the laboratories, the protocols and equipment used by the laboratories, and the amount of experience of the laboratories technicians.

This study found that the overall inter-laboratory agreement (1992–2001) for the four Be-LPT laboratories for all Be-LPT results ranged from 93.8% to 98.1%, and the range of agreement for abnormal results ranged from 26.2% to 61.8% (kappa 0.42–0.61). The kappa values may be underestimated as the result of factors intrinsic to sequential testing, such as immune system variability over time, and the differences in environmental conditions during specimen shipment.

The better intra-laboratory agreement suggests that much of the poorer inter-laboratory agreement may be the result of test serum differences. Consideration has been given to having all of the Be-LPT laboratories use the same human serum in their test. However, if the difference seen among the laboratories is due to the serum, using a single human serum might decrease the ability to identify sensitized individuals. In addition, coordinating the use of a single serum is not practical because of several factors including highly variable levels of testing between laboratories and the limited size of available serum lots.

In this investigation if a participant was found through serial testing to have an initial borderline-abnormal result followed by an abnormal result from the same laboratory, the majority (87.1%) of these participants eventually developed either beryllium sensitivity or CBD. However, if an initial borderline-abnormal result was followed by a normal result from the same laboratory, only a small percentage (9.9%) of these participants have been identified with beryllium sensitivity or CBD. Based on these observations, the authors recommend that individuals with borderline-abnormal results should be offered two repeat Be-LPTs, using the initial as well as a second testing laboratory.

Inter- and intra-laboratory agreement varied among the four Be-LPT laboratories that were used to perform the testing. Careful examination of Be-LPT results and confirmation of abnormal results is recommended to assure appropriate referral for additional clinical evaluation. The Be-LPT is a highly effective tool for the identification of beryllium-sensitized individuals who may be at risk for the development of CBD. However, in cases where beryllium sensitivity is not clear, proper assessment and disposition of these cases requires careful examination of serial Be-LPT results and knowledge of medical and exposure history through questionnaires and trained participant interviewers. The Be-LPT can be used as a means to identify areas of excess beryllium exposure where remediation is warranted, and to identify employee populations at risk.

The RFETS data indicate that sensitized individuals with no evidence of CBD or other lung disease (i.e., pulmonary function tests are normal; there are no B-reader chest X-ray findings possibly related to CBD; and there are no symptoms of CBD) can be clinically evaluated every 2 years. Of the 117 participants diagnosed with CBD, 90.5% were identified by their second clinical evaluation and 99.1% by their fourth, with the mean time to these evaluations of 2.73 years and 4.83 years, respectively, although diagnoses occurred up to 9.77 years following the identification of sensitivity. The results suggest that the frequency and the comprehensiveness of CBD clinical evaluations for sensitized individuals could be progressively reduced if CBD is not diagnosed by the fourth evaluation, and if the individual's pulmonary condition remains stable.

ACKNOWLEDGMENTS

This work was performed under contract DE-AC05-00OR22750 between the United States Department of Energy (DOE) and Oak Ridge Associated Universities (ORAU) through the Oak Ridge Institute for Science and Education (ORISE). ORAU is a university consortium leveraging the scientific strength of 86 major research institutions to advance science and education by partnering with national laboratories, government agencies, and private industry. ORAU manages ORISE, a DOE facility that focuses on scientific initiatives to research health risks from occupational hazards, assess environmental cleanup, respond to radiation medical emergencies, support national security and emergency preparedness, and educate the next generation of scientists. This work was authored by a contractor of the United States Government. Accordingly, the United States Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this work or to allow others to do so for United States Government purposes. The authors thank Nancy Daugherty, Center for Epidemiologic Research, Oak Ridge Institute for Science and Education, for her review of this manuscript.

REFERENCES

Bass JB, Jr. 2001. The tuberculin skin test. In: Field MJ, editor. Tuberculosis in the workplace. Washington, DC: National Academy Press. 340 p.

Bobka CA, Stewart LA, Engelken GJ, Golitz LE, Newman LS. 1997. Comparison of in vivo and in vitro measures of beryllium sensitization. J Occup Envi Med 39(6):540–547.

Brown ML, Houn F, Sickles EA, Kessler LG. 1995. Screening mammography in community practice: Positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. Am J Roentg 165(6):1373–1377.

Code of Federal Regulations. 1999. Title 10, Volume 64, Part 850. Part III, Department of Energy, Chronic Beryllium Disease Prevention Program. Washington, D.C: U.S. Department of Energy.

Cohen J. 1960. Coefficient of agreement for nominal scales. Educ Psychol Meas 20:37–46.

Deubner DC, Goodman M, Iannuzzi J. 2001. Variability predictive value, and uses of the beryllium blood lymphocyte proliferation test (BLPT): Preliminary analysis of the ongoing workforce survey. Appl Occup Environ Hyg 16(5):521–526.

Fahey MT, Irwig L, Macaskill P. 1995. Meta-analysis of pap test accuracy. Am J Epidemiol 141(7):680–689.

Fleiss JL, Cohen J. 1973. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 33:613–619.

Fontenot AP, Falta MT, Freed BM, Newman LS, Kotzin BL. 1999. Identification of pathogenic t-cells in patients with beryllium-induced lung disease. J Immun 163:1019–1026.

Gann PH, Hennekens CH, Stampfer MJ. 1995. A prospective evaluation of plasma prostate-specific antigen for detection of prostatic cancer. JAMA 273(4):289–294.

International Labor Organization (ILO). 1980. Guidelines for the use of ILO international classification of radiographs of pneumoconioses. Occupational Safety and Health Series, No. 22 (Rev.), Geneva Switzerland: International Labor Office.

Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernster V. 1993. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 270(2):2444–2450.

Kreiss K, Newman LS, Mroz MM, Campbell PA. 1989. Screening blood test identifies subclinical beryllium disease. J Occup Med 31(7):603–608.

Kreiss K, Mroz MM, Zhen B, Martyny JW, Newman LS. 1993. Epidemiology of beryllium sensitization and disease in nuclear workers. Am Rev Respir Dis 148:985–991.

Kreiss K, Mroz MM, Zhen B, Wiedemann H, Barna B. 1997. Risks of beryllium disease related to work processes at a metal alloy and oxide production plant. Occup Environ Med 54:605–612.

Kulasingam SL, Hughes JP, Kiviat NB, Mao C, Weiss NS, Kuypers JM, Koutsky LA. 2002. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities. JAMA 288(14):1749–1757.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics 33(1):159–174.

Marshall KG. 1995. The utility of prostate-specific antigen for detecting prostate cancer. JAMA 274(8):608–609.

Mroz MM, Kreiss K, Lezotte DC, Campbell PA, Newman LS. 1991. Re-examination of the blood lymphocyte transformation test in the diagnosis of chronic beryllium disease. J Allergy Clin Immunol 88:54–60.

Nash DR, Douglass JE. 1980. Anergy in active pulmonary tuberculosis: A comparison between positive and negative reactors and an evaluation of 5 tu and 250 tu skin test doses. Chest 77:32–37.

Newman LS. 1994. Beryllium lung disease: The role of cell-mediated immunity in pathogenesis. In: Dean JH, Luster MI, Munson AE, Kimber I, editors. Immunotoxicology and immunopharmacology. New York: Raven Press. p 377–394.

Newman LS. 1996. Significance of the blood beryllium lymphocyte proliferation test. Environ Health Persp 104(Suppl 5):953–956.

Newman LS. 2000. Immunologic mechanisms in granulomatous lung disease. Immunopharm 48:329–331.

Newman LS, Kreiss K, King TE, Jr., Seay S, Campbell PA. 1989. Pathologic and immunologic alterations in early stages of beryllium disease: Re-examination of disease definition and natural history. Am Rev Respir Dis 139:1479–1486.

Newman LS, Lloyd J, Daniloff E. 1996. The natural history of beryllium sensitization and chronic beryllium disease. Environ Health Persp 104(Suppl 5):937–943.

Rossman MD, Kern JA, Elias JA, Cullen MR, Epstein PE, Preuss OP, Markham TN, Daniele RP. 1988. Proliferative response of bronchoalveolar lymphocytes to beryllium: A test for chronic beryllium disease. Ann Intern Med 108:687–693.

Schroeder FH, Roobol-Bouts M, Vis AN, van der Kwast T, Kranse R. 2001. Prostate-specific antigen-based early detection of prostate cancer-validation of screening without rectal examination. Urology 57(1): 83–90.

Solomon D, Schiffman M, Tarone R. 2001. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: Baseline results from a randomized trial. J Natl Cancer Inst 93(4):293–299.

Spitzer RL, Cohen J, Fleiss JL, Endicott MS. 1967. Quantification of agreement in psychiatric diagnosis: A new approach. Arch Gen Psych 17:83–87.

Stange AW, Hilmas DE, Furman FJ. 1996a. Possible health risks from low-level exposure to beryllium. Toxicol 111:213–224.

Stange AW, Hilmas DE, Furman FJ. 1996b. Rocky Flats Beryllium Health Surveillance. Environ Health Persp 104(Suppl 5):981–986.

Stange AW, Hilmas DE, Furman FJ, Gatliffe TR. 2001. Beryllium sensitization and chronic beryllium disease at a former nuclear weapons facility. Appl Occup Environ Hyg 16(3):5405–5417.

United States Department of Energy. 2001. In: Wambach PW, editor. DOE-SPEC-1142-2001: Beryllium Lymphocyte Proliferation Testing. Washington, DC: US Department of Energy.

Williams WJ, Williams WR. 1983. Value of beryllium lymphocyte transformation tests in chronic beryllium disease and in potentially exposed workers. Thorax 38:41–44.