BLIND TESTING OF CERTIFIED ALPHA TRACK AND CHARCOAL LABORATORIES BY THE COMMONWEALTH OF PENNSYLVANIA

Robert K. Lewis
DEP/BRP, Radon Division
Harrisburg, PA

INTRODUCTION

In an attempt to assess the quality of radon measurement results being provided to Commonwealth residents by certified alpha track and charcoal laboratories, blind, spiked samples were introduced into the normal processing routine of three alpha track (ATD) laboratories and 18 charcoal laboratories.

Radon test kits were purchased anonymously from the 21 laboratories (22 devices) by Radon Division staff. The test kits were exposed at the Wilkes University Radon Chamber as per laboratory instructions. The test kits were then returned to the lab for analysis. The results were received by Radon Division staff and some relatives, compared to a pass/fail criteria, and then letters were sent to the laboratory confirming there successful results and acknowledging their good work or informing them of unsuccessful results and asking for investigation, corrective action, and retesting.

Eight of the 22 devices failed (36%) the initial blind test criteria. This group of eight failures included two ATD labs and six charcoal labs. Two of the six charcoal labs also failed to meet our criteria during retesting. They did pass on the third attempt. The two labs that failed the retesting were both suspended from providing any more measurement results to Commonwealth residents until they could demonstrate that their measurement process was "back in control."

SCOPE OF STUDY

This initial study consisted of all of the state-certified ATD and charcoal labs as of November 1995, which was the start of this study. This included three ATD labs and 18 charcoal labs. The three ATD labs all reside out of state and six of the charcoal labs reside out of state. The ATD and charcoal labs together comprise 55% of all radon measurement results reported to Commonwealth residents. The three ATD labs tested were Landauer, Radon Environmental Monitoring (REM), and Radiation Safety Services Inc. (RSSI). The 18 charcoal labs tested were Alpha Energy, REM, TCS Industries (uses two different devices), Key Technology, Radon Testing Corporation of America (RTCA), ABE Labs, UST Labs, Air Check, Best Inspect, Microbac, Pennsylvania Home Inspection, Prosser Laboratories, Franklin Environmental Analysis,

Low Level Radiation Monitoring Laboratory, RAdata Inc., Radiation Data, Airtech Radon Testing, and Total Home Technology.

During the second phase of this study, yet to be started, we will attempt to expose devices from E-PERM and continuous radon monitor users. This study will most likely not be done using the blind approach, due to the fact that most E-PERMs and continuous monitors are placed directly by the certified tester. They are not sent through the mail where we could purchase them anonymously.

EXPOSURE FACILITY

The Wilkes University Radon Chamber (chamber) was used as the exposure facility. This chamber is a 50 cubic meter chamber in the basement of the physics building on the Wilkes campus. The Radon-222 (radon) is supplied by two Pylon flow through Radium-226 sources. The chamber environmental parameters are computer controlled via feedback from temperature and relative humidity sensors located within the chamber. The radon in the chamber is continuously monitored by a Pylon AB-5/Passive Radon Detector located inside the chamber and by an Eberline RGM-3 continuous radon monitor located outside the chamber. The Eberline is used as the primary instrument.

Wilkes University participates in radon intercomparison programs with both the DOE Environmental Measurements Laboratory and the EPA, National Air and Radiation Environmental Laboratory in Montgomery, Alabama.

During all blind exposures the Radon Division carried out simultaneously, radon sampling in the chamber using an Eberline RGM-3. This unit is intercompared at the DOE EML radon chamber in New York City. Additional, simultaneous sampling was also carried out at most, but not all exposures, using E-PERMs and grab scintillation cells.

CHAMBER CONDITIONS

During all exposures, chamber conditions were held constant. The following chamber parameters were held within the listed ranges.

Parameter	Range
Radon Concentration	4 - 20 pCi/L
Temperature	55 - 80 degrees Fahrenheit
Relative Humidity	20 - 75%
Air Velocity	Negligible
Particle Concentration	Not Measured
Progeny Concentration	Not Measured
Equilibrium Ratio	Not Measured

EXPOSURE CONDITIONS

1. All chamber parameters as listed above were stabilized prior to and during all exposures.

2. Three test devices were exposed to a low radon concentration of between 4 and 10 pCi/L. One test device was left unexposed as a blank.

3. Three test devices were exposed to a high radon concentration of between 11 and 20 pCi/L. One test device was left unexposed as a blank.

4. Exposure duration was as stated on each test device instruction sheet. If a range for exposure duration is given, the exposure will be that which is most convenient for placement and retrieval, however, still within the stated range.

5. All exposed devices were returned to the laboratory the same day they were removed from the chamber using first class mail.

6. All test devices were placed and retrieved from the chamber by Radon Division staff.

7. Pictures were taken of all test devices in the chamber to document exposure and proper placement.

DEVICE ANALYSIS

After both the high and low exposure is completed for a laboratory, the six results are compared to the chamber concentration, as obtained from the Wilkes Eberline RGM-3. The Individual Relative Error (IRE) is calculated for each measurement result. The IRE is the reported value minus the target value, divided by the target value, and that quantity multiplied by 100.

The participants outcome is one of the following.

Pass: The IRE for each individual device, from both the high and low exposure groups are all less than or equal to +/-25%.

Fail: The IRE for any one device from either the high or low exposure group is greater than +/- 25%.

A letter was sent to all participants who failed the initial blind test, requesting that they investigate the reason(s) for the unsatisfactory results, as well as provide solutions to any problems found. After corrective actions have been taken, a second round of non-blind exposures will be conducted. A failure during this second round of testing will result in immediate

suspension of all radon testing activities for that particular company, for the device in question. Control of laboratory analysis will have to be demonstrated, through additional exposures, before the suspension can be lifted.

QUALITY ASSURANCE (QA)

All purchased test kits are inspected for any damage and inconsistencies with other similar devices. The devices are secured in a locked cabinet on the 13th floor of the Rachel Carson State Office Building until they are taken to the chamber for exposure. The annual average radon concentration on the 13th floor is 0.5 pCi/L.

Upon arrival at the chamber, four scintillation cells are connected in series and the chamber atmosphere is sampled from outside the chamber. The average of the four cells is compared to the display on the Eberline RGM-3 during sampling and to the average RGM-3 value for the entire sampling period of the exposed devices. There were 19 separate exposure runs performed and the Mean Absolute Relative Error (MARE) of the average of the four cells from all exposure runs compared to the chamber RGM-3 was 4.4%.

All test kits were placed in the chamber by Radon Section personnel, primarily one individual. After the test kits were opened, pictures were taken to document placement. Along with the laboratory test kits, Radon Section short-term E-PERMs and a RGM-3 were also placed in the chamber. These devices were both used to confirm chamber radon concentrations and also used as backup instrumentation in case of chamber instrumentation failure. The MARE for the Radon Section RGM-3 compared to the chamber RGM-3 for all 19 exposure runs was 4.1%, and the MARE for the E-PERMs compared to the chamber RGM-3 for all 19 exposure runs was 5.8%.

The Radon Section RGM-3 and scintillation cells were both intercompared with the DOE Environmental Measurements Laboratory (EML) radon chamber prior to the start of this blind study. Both measurement methods were within +/- 1% of EML chamber values. The RGM-3 was again intercompared in May 1997 after the completion of the blind study and the unit was now 6.5% high.

In addition to our QA requirements Wilkes University Radon Chamber was also responsible for maintaining their own QA program throughout the blind study.

RESULTS AND DISCUSSION

For comparison purposes Table 1 lists the Mean Absolute Relative Error (MARE) for each laboratory for both the high and low exposure groups. The MARE is the absolute value of the mean from the three individual IREs for the high exposure group and the low exposure group. The Coefficient of Variation (COV) is also listed for both high and low exposure groups. The

COV is the standard deviation divided by the mean times 100 to give percent. The below table should not be used to determine whether a laboratory has passed or failed this study, since that criteria is based on the IRE and not the MARE.

Table 1. Results Summary Table

Laboratory	Test Method	Study ID #	Low Exp. MARE	COV	High Exp. MARE	COV
ATD1	Alpha Track	AT001	14.2%	15.5%	6.8%	3.6%
ATD2	Alpha Track	AT002	33.4%	5.5%	9.5%	2.8%
ATD3	Alpha Track	AT003	41.8%	9.4%	4.9%	5.2%
Charcoal 1	Charcoal OF	AC004	13.2%	3.7%	40.7%	4.1%
Charcoal 2	Charcoal LS	AC005	15.7%	19.4%	13.2%	18.6%
Charcoal 3	Charcoal OF	AC006	6.0%	6.9%	48.0%	26.5%
Charcoal 3	Charcoal DB	AC007	4.2%	2.9%	15.1%	1.4%
Charcoal 4	Charcoal OF	AC008	5.6%	8.6%	5.6%	5.4%
Charcoal 5	Charcoal DB	AC009	13.6%	5.2%	13.6%	1.4%
Charcoal 6	Charcoal DB	AC010	16.0%	1.9%	12.3%	1.8%
Charcoal 7	Charcoal DB	AC011	5.6%	6.3%	5.6%	0.5%
Charcoal 8	Charcoal DB	AC012	8.1%	10.1%	3.0%	4.4%
Charcoal 9	Charcoal DB	AC013	13.9%	19.3%	15.6%	6.4%
Charcoal 10	Charcoal OF	AC014	21.9%	8.1%	33.1%	8.0%
Charcoal 11	Charcoal OF	AC015	2.5%	3.4%	15.4%	2.5%
Charcoal 12	Charcoal DB	AC016	14.4%	17.0%	7.9%	13.0%
Charcoal 13	Charcoal DB	AC017	9.6%	1.7%	17.9%	11.3%
Charcoal 14	Charcoal DB	AC018	14.7%	19.0%	3.6%	4.4%
Charcoal 15	Charcoal OF	AC019	15.9%	7.1%	13.3%	12.5%
Charcoal 16	Charcoal OF	AC020	9.4%	13.6%	17.1%	14.8%
Charcoal 17	Charcoal DB	AC021	4.4%	6.9%	11.4%	4.4%
Charcoal 18	Charcoal OF	AC022	0.5%	0.8%	3.3%	3.8%

LS = Liquid Scintillation, DB = Diffusion Barrier, OF = Open Face

Table 2. Summary Statistics

Alpha Track Detectors (3 Labs)

Average Low Exp. MARE = 29.7%
Average Low Exp. COV = 10.1%
Average High Exp. MARE = 7.0%
Average High Exp. COV = 3.8%

Charcoal Detectors (19 labs)

Average Low Exp. MARE = 10.3%
Average Low Exp. COV = 8.5%
Average High Exp. MARE = 15.5%
Average High Exp. COV = 7.6%

The average MARE for the ATDs at 29.7% does not speak well for the performance of these devices at the low exposure range. The total integrated exposure, at 256 pCi/L-day was well above the lower limit of detection for the ATDs in question, which ranged from 15-30 pCi/L-day. As would be expected at the higher exposures (515 pCi/L-day) the average MARE was much improved at 7.0%. The average MAREs for both the low and high charcoal exposures were both quite reasonable at 10.3% and 15.5% respectively. However, with the charcoals, the better performance was observed at the lower concentrations compared to the higher concentrations, just the opposite to the ATDs and also not what would be expected due to counting statistics.

Those laboratories that failed the initial study included two ATD labs, ATD 2 and ATD 3, and six charcoal labs, Char. 1, Char. 2, Char. 3, Char. 10, Char. 13 and Char. 16. The individual IREs are presented in Table 3 to give some indication of the degree of results for each laboratory.

Table 3. Failed Laboratory Results

Laboratory	Test Method	Low Exp. IREs	High Exp. IREs
ATD 2	Alpha Track	**+41.6%,+ 27.8%,+ 30.6%**	+5.9%, +11.2%, +11.2%
ATD 3	Alpha Track	**+53.0%, +45.1%, +27.0%**	-2.8%, +7.7%, +4.2%
Charcoal 1	Charcoal OF	-10.2%, -16.6%, -12.8%	**-39.2%, -39.2%, -43.4%**
Charcoal 2	Charcoal LS	+10.2%, +7.6%, **+29.4%**	+8.3%, -23.2%, +8.3%
Charcoal 3	Charcoal OF	-5.1%, -12.8%, 0.0%	**-32.1%, -55.3%, -56.5%**
Charcoal 10	Charcoal OF	**+32.9%**, +13.9%, +19.0%	**+25.4%, +28.6%, +45.2%**
Charcoal 13	Charcoal DB	-11.3%, -8.2%, -9.3%	-12.1%, **-28.6%**, -13.1%
Charcoal 16	Charcoal OF	-3.3%, +23.9%, +1.1%	+17.4%, **+30.3%**, -3.5%

As can be seen from the above table, five of the laboratories convincingly failed to meet the +/-25% IRE criteria and three of the laboratories just marginally missed meeting the pass criteria with only one device failing for each lab. Each laboratory that failed the initial blind study was asked to investigate the reason(s) for the failure and provide a written explanation and any corrective actions that may have been taken. The following are the reasons provided for the laboratory failures.

- Excessive background from Cosmic Ray exposure due to detectors being ~ six years old.
- A "jitter" in the image analysis system caused some tracks to be double counted.
- Calibration error.
- Found nothing abnormal from analysis records.
- New personnel lead to operator error.
- New batch of devices. Old calibration curves not satisfactory for new batch of devices.
- Operator error. Operator failed to use correct absorption factor in calculation.
- Beyond possibility of unusual statistical fluctuation, could not detect any problems with laboratory system.

After corrective actions were taken based on the above postulated problems, a second round of non-blind retesting was carried out to confirm the effects of the corrective actions. Table 4 below shows the results from this retesting program.

Table 4. Non-blind Retest Results

Laboratory	Initial Blind Low Exp. MARE	Non-Blind Retest Low Exp. MARE	Initial Blind High Exp. MARE	Non-Blind Retest High Exp. MARE
ATD 2	33.4%	1.5%	9.5%	8.5%
ATD 3	41.8%	3.5%	4.9%	7.7%
Charcoal 1	13.2%	15.8%	40.7%	8.0%
Charcoal 2	15.7%	25.3%	13.2%	23.5%
Charcoal 3 (OF)	6.0%	13.2%	48.0%	17.0%
Charcoal 10	21.9%	10.9%	33.1%	5.7%
Charcoal 13	9.6%	4.4%	17.9%	9.7%
Charcoal 16	9.4%	29.9%	17.1%	13.7%

The corrective actions and/or the added attention given the retest devices produced improved results in all but two of the laboratories. The one charcoal laboratory (Char. 2) had much poorer performance during the retesting than during the initial blind testing. During the initial blind testing, Char. 2 had one result out of six greater than +/- 25%. During the retesting, Char. 2 had three out of six results greater than the +/- 25% criteria. Charcoal 16's initial and retest results were similar. They had one out of six results greater than +/- 25% during both initial testing and retesting. Additional test devices were again requested and the two laboratories were asked to continue to investigate the reason(s) for discrepancies between the target and measured values. These investigations by the two laboratories (# 2 and 16) produced the following two conclusions.

- Problems in charcoal transfer process to the LS cocktail.
- Hardware problem. Loose connection caused excessive counts to be dumped into region of interest.

Additional test devices were again exposed non-blind. Eight Char. 2 charcoals were exposed to a high concentration producing a MARE of 12.8%, and 7 Char. 2 charcoals were exposed to a low concentration producing a MARE of 8.8%. Twenty Charcoal 16 charcoals were exposed to only a high concentration, however, they were exposed at varying time periods of 2, 3, 4, 5, and 6 days. The MARE for the 20 charcoals was 5.5%. Based on these values it was thought that both laboratories were "in control."

COMPARISON STUDIES

In 1989 the Radon Division conducted its first blind study of the certified testing community. This study differed in a number of ways from this present blind study. Table 5 lists the major differences between the studies.

Table 5. 1989 Vs. 1997 Blind Study Comparison

	1989	1996/1997
Radon Chambers:	Radon QC, Palmer, PA	Wilkes Univ., Wilkes-Barre, PA
Pass/Fail Criteria:	MARE >25%	IRE >25%
Detectors Exp.:	2 at low conc. 2 at medium conc. 1 blank	3 at low conc. plus 1 blank 3 at medium conc. plus 1 blank
Participants:	47 laboratories	21 laboratories
Methods:	charcoal, ATD, E-PERM	charcoal and ATD
Concentration:	9 to 70 pCi/l	4 to 20 pCi/l

Fifteen of the 44 laboratories that participated in the 1989 study were still in business and participated in the 1996/1997 study. In the 1989 study 6 out of 47 labs (12.7%) failed during the initial testing, in the 1996/1997 study 8 out of 22 devices (36%) failed during the initial testing. During the 1989 study the overall MARE for all devices (charcoal, E-PERM and ATD) was 14.1%. During the 1996/1997 study the overall MARE for all devices (ATD and charcoal) was 13.6%. During the 1989 study there were 5 laboratories who passed using the MARE criteria who would not have passed the 1996/1997 study using the IRE criteria. These 5 labs plus the other 6 who failed would bring the total to 11 out of 47 (23.4%) who would have failed the 1989 study. This is still almost 13 percentage points below our 1996/1997 fail rate of 36%.

Of the 8 labs who failed the 1996/1997 study, 6 of them (ATD 3, Char. 1, Char. 3, Char. 10, Char. 13, and Char. 16) also participated in the 1989 study. Of these 6, three of them, ATD 3, Char. 1, and Char. 16 also failed the 1989 study.

The July 1995 Consumer Reports magazine provided ratings for 10 different radon test kits. Their analysis found six out of seven short-term test kits as acceptable and two out of three long- term test kits were acceptable. Four of the six (Air Check, Key Technology, First Alert (Alpha Energy), and RTCA) short-term test kits in the Consumer Reports study also participated in our 1996/1997 study. Only one of the long-term test kits was included in our 1996/1997, Landauer's RadTrak.

The EPA through the Radon Measurement Proficiency Program (RMP) conducts non-blind testing (proficiency testing) on most all companies and laboratories providing radon testing services within the country. .All types of test methods go through the RMP program. Additionally, companies and labs know they are being tested and since RMP listing is very

important to the companies and labs continued existence, it can be expected that "best effort" will be given to all analyses. However, RMP data can still be used to provide some comparison examples of pass/fail rates for different test methods. For comparison purposes for this report, activated charcoal had an initial fail rate of 34.9% and alpha track detectors had a fail rate of 58%. The 34.9% fail rate for activated charcoal is close to our 1996/1997 study fail rate for charcoal labs of 32% and the 58% fail rate for alpha track detectors is close to our 1996/1997 study results of 66% failure for alpha track detectors. These similar fail rates would tend to imply that giving "best effort" does not appreciably improve test result quality.

In an attempt to try and rank the quality of results provided by each laboratory, a graph was produced that plotted accuracy verse precision. Accuracy was plotted on the X axis, as the MARE for the high and low exposures combined. Precision was plotted on the Y axis, as the COV for the high and low exposures combined. Performance Graph 1 plots each laboratory's initial performance prior to any corrective actions.

On each graph are three rectangles. The rectangles are produced by joining the 15% MARE on the X axis with the 15% COV on the Y axis, and so on for 10% and 5%. These rectangles are referred to as windows; the 15/15 window, 10/10 window, and the 5/5 window. There is also a line running at a 45 degree angle from the origin. Points below this line would correspond to a greater degree of inaccuracy, points above this line would correspond to a greater degree of imprecision.

After all laboratories were tested, it can be seen that six labs fall outside of the 15/15 window. Five of those labs show a greater degree of inaccuracy and one lab shows a greater degree of imprecision. The six labs that fall outside of the 15/15 window all failed the initial blind test. Additionally, there were two labs, numbers 17 and 20, that also failed the initial blind test but were inside the 15/15 window.

All eight labs that failed the initial blind test took corrective actions. This data was then plotted on Performance Graph 2. The points now show a much tighter grouping. Two of the points fall on the 15/15 window margin and all other points now fall within the 15/15 window. The majority of labs still show greater degrees of inaccuracy compared to imprecision.

One should not place too much emphasis on the laboratory ratings seen in Performance Graph 2, at least at this point in time. However, a laboratory consistently falling within the 5/5 or 10/10 window over several years worth of blind test results would more convincingly point to that laboratory's commitment to sound QA practices and the production of highly dependable results.

Based upon Performance Graph 2 and DEP/EML radon intercomparison charcoal results, it would appear that most "good" labs should consistently perform within the 10/10 window or better.

CONCLUSIONS

Whereas the 36% initial fail rate for labs seems high, it does not appear to be out of line with some other studies. Certainly the more stringent pass/fail criteria of the 1996/1997 study contributed to the higher fail rate compared to the 1989 study. The higher 1996/1997 fail rate may also be due to some QA/QC complacency over time. As Alfred Edward Perlman said, "After you've done a thing the same way for two years, look it over carefully. After five years, look at it with suspicion. And after ten years, throw it away and start all over." This blind study has at least caused laboratories to "look it over carefully," and this should certainly benefit those receiving test results.

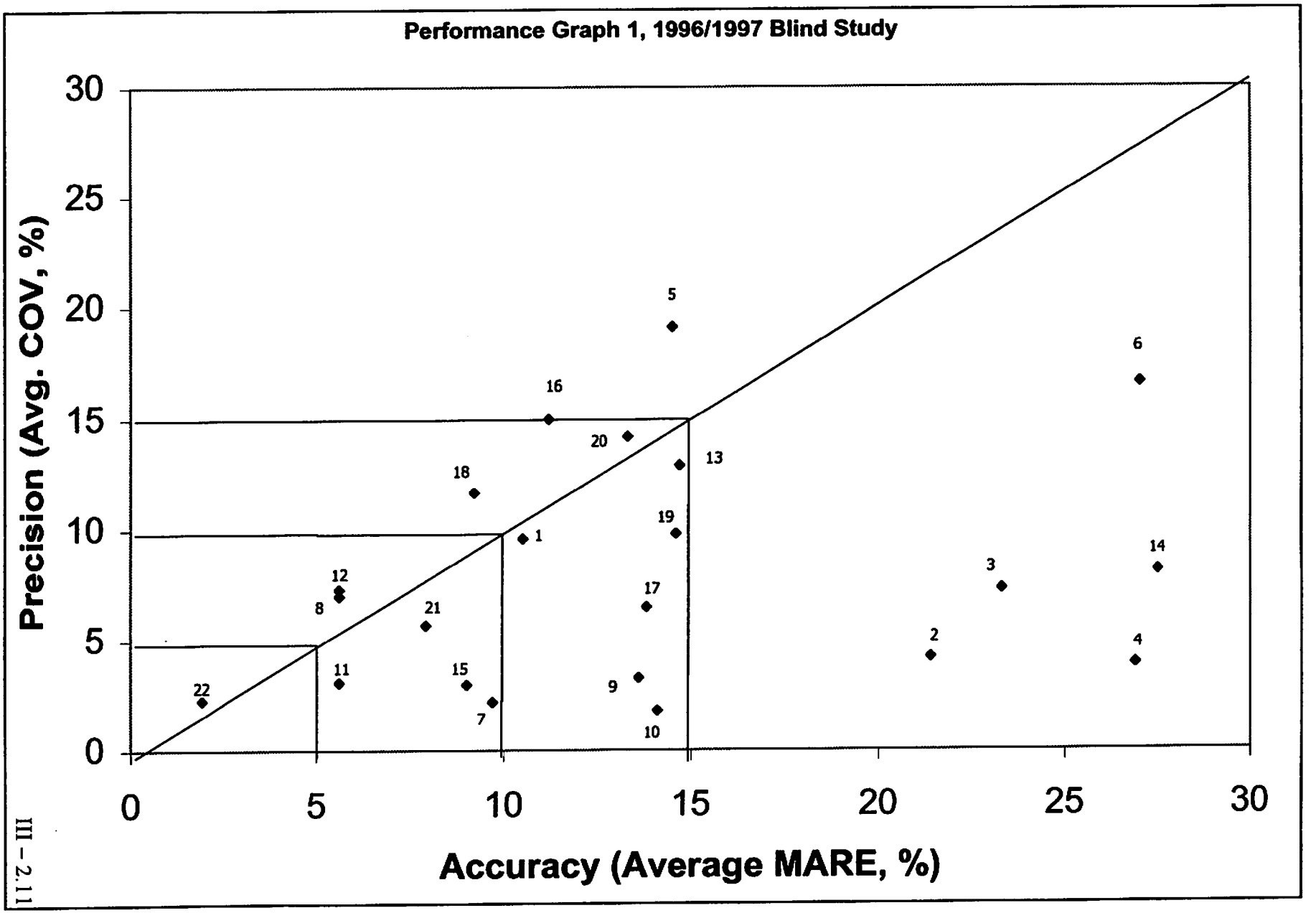
REFERENCES

Consumers Report, July 1995, pg. 465. A publication of Consumers Union.

Granlund, C. ; Kaufman, M., 1989 Double Blind Study. PA Dept. of Environmental Resources, Bureau of Radiation Protection, Radon Division, Unpublished report, 1989.

Larkin, K, ; Hornesky, R, ; Maxwell, R. Wilkes University Radon Chamber Quality Assurance Plan. Draft Copy, 1995.

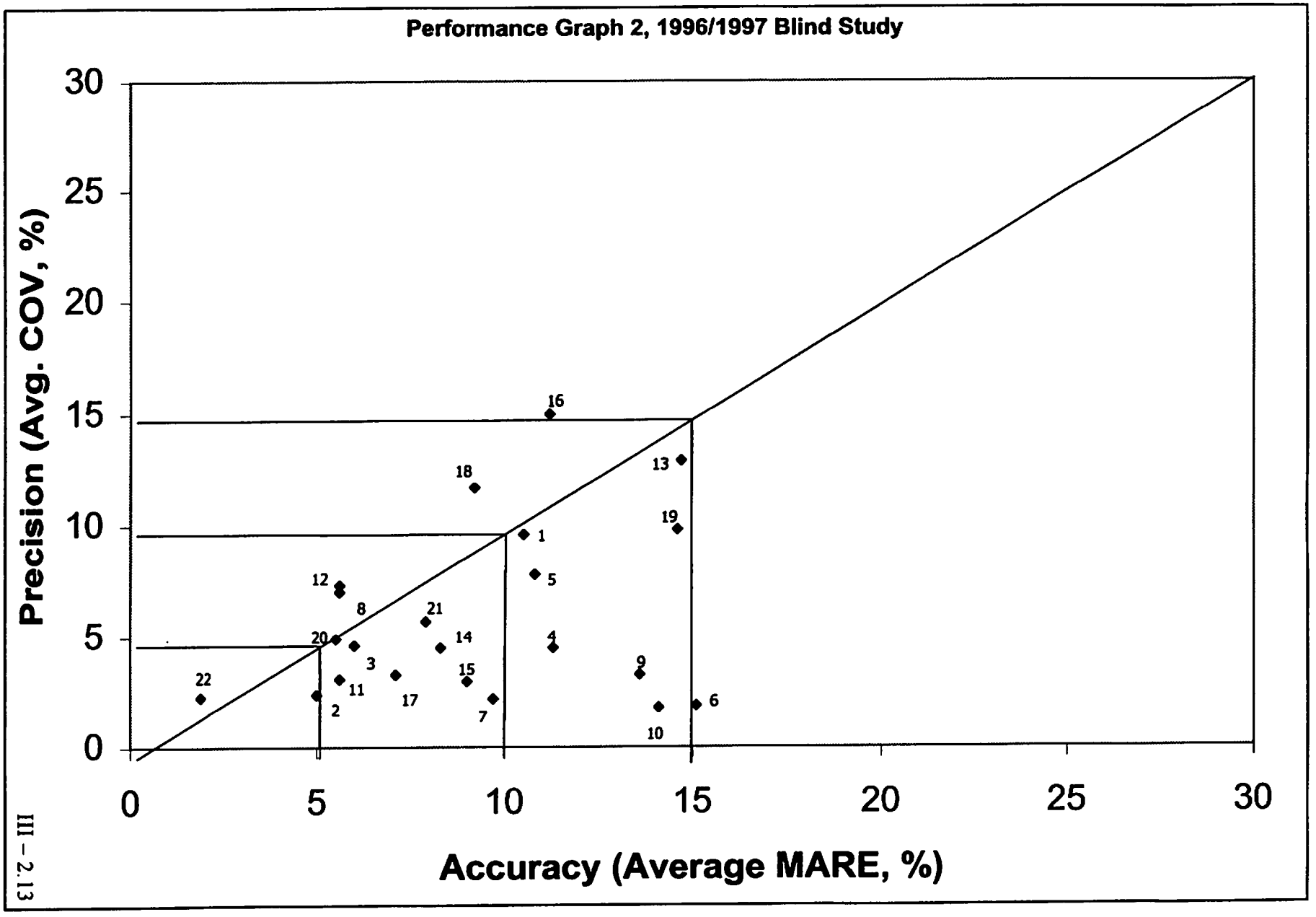
U.S. Environmental Protection Agency. National Radon Proficiency Program Cumulative Data, 01-January-1991 to 8-July-1997. EPA 402-F-93-003-I.


Performance Graph 1, 1996/1997 Blind Study
Precision (Avg. COV, %)
Accuracy (Average MARE, %)
0
5
10
15
20
25
30
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Performance Graph 1 Data

Graph ID # (Study ID #)	Laboratory	Average MARE	Average COV
1	ATD1	10.5%	9.6%
2	ATD2	21.4%	4.2%
3	ATD3	23.3%	7.3%
4	Charcoal 1	26.9%	3.9%
5	Charcoal 2	14.5%	19.1%
6	Charcoal 3 (OF)	27.0%	16.6%
7	Charcoal 3 (DB)	9.7%	2.2%
8	Charcoal 4	5.6%	7.0%
9	Charcoal 5	13.6%	3.3%
10	Charcoal 6	14.1%	1.8%
11	Charcoal 7	5.6%	3.1%
12	Charcoal 8	5.6%	7.3%
13	Charcoal 9	14.7%	12.9%
14	Charcoal 10	27.5%	8.1%
15	Charcoal 11	9.0%	3.0%
16	Charcoal 12	11.2%	15.0%
17	Charcoal 13	13.8%	6.5%
18	Charcoal 14	9.2%	11.7%
19	Charcoal 15	14.6%	9.8%
20	Charcoal 16	13.3%	14.2%
21	Charcoal 17	7.9%	5.7%
22	Charcoal 18	1.9%	2.3%

The above table is for results prior to any corrective action.


Performance Graph 2, 1996/1997 Blind Study
Precision (Avg. COV, %)
Accuracy (Average MARE, %)
0
5
10
15
20
25
30
16
18
13
19
1
12
5
8
21
20
14
4
3
9
15
22
11
17
2
7
6
10

Performance Graph 2 Data

Graph ID # (Study ID #)	Laboratory	Average MARE	Average COV
1	ATD1	10.5%	9.6%
2	ATD2	5.0%	2.4%
3	ATD3	6.0%	4.6%
4	Charcoal 1	11.3%	4.5%
5	Charcoal 2	10.8%	7.8%
6	Charcoal 3 (OF)	15.1%	1.9%
7	Charcoal 3 (DB)	9.7%	2.2%
8	Charcoal 4	5.6%	7.0%
9	Charcoal 5	13.6%	3.3%
10	Charcoal 6	14.1%	1.8%
11	Charcoal 7	5.6%	3.1%
12	Charcoal 8	5.6%	7.3%
13	Charcoal 9	14.7%	12.9%
14	Charcoal 10	8.3%	4.5%
15	Charcoal 11	9.0%	3.0%
16	Charcoal 12	11.2%	15.0%
17	Charcoal 13	7.1%	3.3%
18	Charcoal 14	9.2%	11.7%
19	Charcoal 15	14.6%	9.8%
20	Charcoal 16	5.5%	4.9%
21	Charcoal 17	7.9%	5.7%
22	Charcoal 18	1.9%	2.3%

The above table is for results after the eight labs made their final corrections.