

Evaluating the Performance of Residential CO Alarms

Final Report

Prepared by:

Paul K. Clifford
Mosaic Industries, Inc.
5437 Central Ave., Suite 1
Newark, CA 94560
510-790-8222

For:

The Gas Research Institute
1700 S. Mount Prospect Rd.
Des Plaines, IL 60618-1804

GRI Contract No. 5091-254-2309
Report No. GRI-02/0112

June, 2002

Acknowledgments

The authors and GRI would like to express their appreciation to the experts in the CO alarm industry and research community who contributed to this report, and particularly to thank the reviewers who made many valuable suggestions. We are indebted to them for the valuable information they provided. A special debt of gratitude is owed to Steve J. Wiersma of GRI, for his support, encouragement and insight.

Disclaimer

LEGAL NOTICE: This report was prepared by Mosaic Industries, Inc. of Newark, California, as an account of work sponsored by the Gas Research Institute (GRI). Neither Mosaic Industries, Inc., GRI, members of GRI, nor any person acting on behalf of Mosaic Industries, Inc., GRI, or its members:

- a. makes any warranty or representation, express or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or
- b. assumes any liability with respect to the use of, or for any and all damages resulting from the use of, any information, apparatus, method, or process disclosed in this report.

References to trade names, brand names, or specific commercial products, commodities, or services in this report do not represent nor constitute an endorsement, recommendation, criticism, condemnation, approval, disapproval, disapprobation, favoring or disfavoring by GRI or Mosaic Industries, Inc. of the commercial product, commodity or service.

Report Documentation Page

1. Title and Subtitle Evaluating the Performance of Residential CO Alarms:	2. Sponsoring Agency Report No. GRI-02/0112
4. Author(s) Paul K. Clifford	3. Report Date June, 2002
6. Performing Organization's Name and Address Mosaic Industries, Inc. 5437 Central Ave., Ste. 1 Newark, CA 94560 Contact: Paul Clifford, Tel: 510-790-8222, paul@mosaic-industries.com	5. Report Type and Period Covered Final Feb. 1999 – Dec. 2001
8. Sponsoring Agency's Name and Address Gas Technology Institute 1700 S. Mount Prospect Rd. Des Plaines, IL 60618-1804 Contact: Vipul Srivastava, Tel: 847-768-0539	7. Contract No. GRI 5091-254-2309
9. Supplementary Notes	
10. Abstract <p>This study assessed the performance of commercially available carbon monoxide (CO) alarms, several preproduction prototypes not available commercially, and several alarms removed from residences by first responders. Alarm performance was assessed in terms of the Underwriters Laboratories UL 2034 design specification for sensitivity to CO, their immunity to interfering gases, the accuracy of their digital displays, and their ability to detect CO in ambient air of various relative humidity.</p> <p>Alarms of three of ten commercially available brands alarmed within the UL 2034 specification. They performed consistently, without false alarms, and always alarmed when presented with dangerous CO levels. Alarms of six of ten commercially available brands behaved inconsistently, with one half of the devices failing to alarm in at least one test at dangerous CO levels and moderate humidity, and four out of five failing to alarm at low humidity. Three of seven commercially available brands with digital displays were accurate to within $\pm 30\%$, with one brand being accurate to within $\pm 10\%$. The other four brands were highly inaccurate with a significant portion of alarm units of two brands reading 0 ppm when exposed to CO concentrations as great as 100 ppm.</p> <p>The questionable performance of commercially available residential CO alarms may be attributable in part to shortcomings of the UL 2034 standard, particularly its continued omission of the requirements for quality assurance testing recommended by the Consumer Product Safety Commission (CPSC) in October 1996. These recommendations have been incorporated into the CSA 6.19-01 Standard for Residential CO Alarms, administered by CSA International of Cleveland, Ohio.</p> <p>This study points out the importance of time-of-manufacture and in-service testing as a primary means of assuring the reliability of CO alarms sold to the public. Gas utilities, appliance manufacturers and other consumers and resellers of large numbers of alarms are urged to scrutinize their performance carefully, through a program of technical review and laboratory testing. It is recommended that whenever possible these parties should require certification to the CSA 6.19-01 standard.</p>	
11. Subject Terms Carbon monoxide, carbon monoxide alarm, CO detector, CO alarm, carboxyhemoglobin, Gas Research Institute, UL 2034, reliability of carbon monoxide alarms, CO poisoning.	
12. Availability Statement Unlimited, unclassified.	13. Number of Pages 172

Contents

Research Summary	ix
Introduction	1
Background.....	1
Objective and Scope	3
Technical Approach	5
Brands Tested	5
Test Apparatus.....	6
Dynamics of Gas Presentation.....	6
Calibration	7
Data Collection.....	7
Specific Tests.....	9
Test Sequence.....	19
Test Results	21
Analysis and Recommendations	86
Overall Reliability – Types and Numbers of Failures Observed.....	86
Insufficient Sensitivity to CO at Lower Relative Humidity	91
Poor Integration of CO Exposure	98
Digital Display Inaccuracy	107
Sensitivity to Interference Gases	110
Conclusions	111
Summary of Technical Recommendations.....	112
Available Standards for CO Alarms	113
References	117
Appendix A: The Use of the Coburn Equation in the Design of CO Alarms	123
Appendix B: Guidelines for Assuring the Time-of-Manufacture In-Service Reliability of CO Alarms	137

Research Summary

Title *Evaluating the Performance of Residential CO Alarms*

Contractor Mosaic Industries, Inc. under GRI Contract Number: 5091-254-2309

Principal Investigator Paul K. Clifford

Report Type Final

Report Period February 1999 - December 2001

Objective This study was aimed at assessing the performance of newly purchased commercially available (“retail”) carbon monoxide (CO) alarms certified to the most recent revision of the Underwriters Laboratories Standard for their performance, UL 2034¹. In addition to testing devices of ten retail brands, we also tested uncertified preproduction prototypes of several alarm brands under development as well as several certified devices withdrawn from a residence by first responders. Performance was assessed in terms of conformance to the UL 2034 specification for sensitivity to CO, their immunity to interfering gases, and their ability to detect CO at various relative humidity. This study was not intended to duplicate UL tests or to rank models of alarms; rather, it was undertaken to assess their performance, identify the causes of any deficiencies, and suggest remedies for those deficiencies.

Technical Perspective A prior study⁷ by this laboratory of retail-off-the-shelf CO alarms found that a majority of alarm brands available to consumers in 1997 and 1998 suffered from one or more deficiencies in performance, including: out-of-the-box failures, poor control of sensitivity to CO, lack of compliance to the UL 2034 specification, false alarms on exposure to interference gases, excessive supervised failures, late alarms and false negatives (times when the alarm should sound, but does not), and inaccurate digital displays.

Other field and laboratory studies^{4-6,8,9,61-66} have shown that despite CO alarms’ certification to UL 2034, they have serious performance problems, including false alarms and failures to alarm at life-threatening CO concentrations, particularly at low ambient humidity. One of these studies⁶⁴ resulted in a recall⁷⁶ of approximately one million UL-certified CO alarms. Further, a recent UK government study¹⁶ of 27 models of residential CO alarms found that only one functioned fully in both laboratory tests and throughout a one year field trial.

Similar deficiencies in performance were also noted in 1999 in a UL survey of retail alarms¹⁵. Presumably, since that time manufacturers have worked to remedy the sources of alarm failure and to bring them into compliance with the latest revision of UL 2034, effective October 1998.

In light of the prior poor performance of commercial CO alarms, the current battery of tests were undertaken to answer the following continuing questions about their reliability and performance:

- Has the performance of alarms offered to the public in the past been improved by newer brand offerings certified to the October 1998 revision to UL 2034?
- Are CO alarms particularly susceptible to failures to alarm when exposed to low relative humidity?
- How accurate are the digital displays of CO alarms?

Technical Approach All alarm brands widely available through retail outlets in the Chicago, IL and San Francisco, CA areas were purchased for test. Additionally, preproduction units of several newly developed alarm models not commercially available were tested, and several other alarm units were retrieved from consumers after involvement in CO poisoning events. All alarms were installed in a dynamic-flow gas system and subjected to a battery of tests of their sensitivity to CO. Their performance was judged against the Underwriters Laboratories UL 2034 sensitivity specification and against a less stringent requirement that they alarm at a cumulative dose greater than 2.5% and less than 10% carboxyhemoglobin (%COHb).

Results The commercially available brands certified to the October 1998 revision of UL 2034 were highly varied in their performance. Some behaved as specified by the revised standard: they complied with the specification for sensitivity to CO, they always alarmed at or before cumulative exposures of 10% COHb, they were immune to the interference gases tested, and their digital displays accurately conveyed the CO concentrations. Others missed the mark on one or another of these performance benchmarks, while some brands performed poorly.

All alarms of the three best performing of ten retail brands consistently actuated within the UL 2034 specification. As the following summary table shows, they never exhibited a false alarm, and none failed to alarm at dangerous CO doses. Their one drawback was a high level of supervised failure (in which the alarm reports an error on its display), a nuisance which does not compromise the safety of the devices.

In contrast to these, alarms of the six most poorly performing retail brands often failed to actuate at dangerous CO levels.

	Alarm Failure Rates	
	Three Well-Performing Brands	Six Poorly Performing Brands
Supervised Failures	15%	0%
False Alarms	0%	8%
Failure to Alarm at 5% RH	0%	79%
Failure to Alarm at 50% RH	0%	47%
Alarm with Interference Gas	0%	30%

The well performing retail brands comprised 33 units distributed among three brands while the poorly performing retail brands comprised 38 units distributed among six brands. One retail brand showed performance intermediate between

these two groups. All retail brands except one of the well performing brands had been certified to UL 2034.

Among the significant failings observed were the following:

False Alarms – Of the commercially available alarm brands tested, 4% of units false alarmed in clean air. It should be noted that this study did not investigate alarms' behavior at high humidity, which is associated with high false alarm rates in the field, particularly false alarms resulting from morning fog and rain.⁶¹⁻⁶³

Failure to Alarm – A significant portion of alarms of seven of ten retail brands failed to alarm at hazardous CO levels. Thirty-five percent of retail alarms tested failed to alarm when presented with one or more of the CO concentrations of the UL 2034 sensitivity test or a progressive concentration test at moderate ambient humidity, and 37% failed to alarm at low humidity.

Poor Sensitivity at Low Humidity – When tested at an absolute humidity of 1320 ppm (two times greater than the lowest absolute humidity required by the UL 2034 standard) many retail alarms were insufficiently sensitive to CO, with 37% failing to actuate. In general, those retail brands which failed to alarm at lower relative humidity were also those most likely to fail at moderate humidity. While many brands lost sensitivity at low relative humidity, one was adequately sensitive at low humidity but inadequately sensitive at moderate humidity.

Inaccurate Digital Displays – Three of seven retail brands with digital displays were accurate to $\pm 30\%$, with one of the three accurate to within $\pm 10\%$ of the true concentration. In contrast, the other four brands were not adequately accurate, with the inaccuracy of the displays reflecting an overall inaccuracy sufficient to prevent the alarm from meeting the basic sensitivity specification of the UL Standard. A significant portion of alarms of two retail brands read 0 ppm when exposed to CO concentrations as great as 100 ppm, even at moderate humidity.

Poor Response to Varying CO Levels – Even when some alarms reliably actuate at the UL test concentrations they do not necessarily activate an alarm on exposure to other life-threatening CO concentrations. Of the fourteen brands tested, seven show decreased sensitivity to steadily increasing CO concentrations in comparison to fixed concentrations. The alarms may be classified into two groups, based on their response to CO. The first group comprises several brands which do consistently integrate CO exposures and activate their alarms at a repeatable threshold for cumulative exposure, whatever the time course of CO concentration. Units of the second group activate at the fixed concentrations of the UL test, but often do not respond to 10% COHb when presented with the increasing CO concentrations typical of CO poisoning events.

This study sheds light on the physical and algorithmic determinants of the poor performance of CO alarms, and proposes solutions for some of their most serious technical deficiencies. Specifically:

- A method of temperature and humidity compensation is proposed that should prevent failures to alarm at low humidity for semiconductor sensor-based alarms.
- A detection algorithm is developed that ensures that alarms respond “biomimetically”, integrating CO concentrations as an exposed person would, and always alarming at or before 10% COHb. The physiological and medical basis for this detection algorithm is provided in Appendix A, along with examples of its practical implementation.
- Guidelines are provided in Appendix B delineating quality assurance methods for sampling and testing CO alarms to assure their reliability.

Conclusions The results of this study provide the following answers to its motivating questions:

- Has the performance of alarms been improved by the October 1998 revision to UL 2034?

These findings parallel the results of earlier studies by ourselves⁷, other researchers^{8-9,61-66}, a British field trial¹⁶, and UL¹⁵ – the failure rates of the retail alarms certified to the revised UL standard continue to be comparable to those of alarms certified to prior revisions. We conclude that the overall performance of alarms offered to the public has not improved with the new revision. Many UL-certified alarms continue to fail the basic sensitivity requirement of the UL standard.

- Are CO alarms particularly susceptible to failures to alarm when exposed to low relative humidity?

The results of this study confirm and expand the observations of prior studies^{7-9,61-64} that the sensitivity of retail CO alarms depends strongly on ambient humidity. In particular, several brands of CO alarms fail to alarm at life-threatening CO concentrations when exposed to humidity typical of the indoor environment in the wintertime. While these brands are particularly susceptible to failure at low humidity, operation at moderate humidity still does not assure safe operation. Those brands most likely to fail at low humidity are also those most unreliable at moderate humidity.

- How accurate are the digital displays of CO alarms?

While several retail alarms show good or adequate display accuracy, many do not. The inaccuracy of several brands is so great as to indicate that the alarms do not meet their basic sensitivity specification, and will not alarm at dangerous CO levels.

Recommendations This study reiterates the Consumer Product Safety Commission’s (CPSC) October 1996 recommendations¹⁹ to UL directed at improving the reliability and safety of CO alarms provided to the public. The incorporation of a UL 2034 standard requirement for time-of-manufacture and in-service testing will go a long way to assuring that CO alarms fulfill their basic function.

Unlike UL 2034, the recently developed Canadian Standards Association CSA 6.19-01 Standard for Residential CO Alarms has incorporated key CPSC and GRI recommendations for reliability, including specifications for alarm longevity and time-of-manufacture testing. Improving the reliability of CO

alarms in the US will be best facilitated if the CPSC and local regulators recommend installation of alarms certified to standards that include requirements for quality assurance, lifetime and reliability, such as CSA 6.19-01.

Project Implications When CO alarms were first widely installed in 1994 it quickly became apparent that they suffered from a number of drawbacks including numerous false alarms. During the next several years laboratory tests revealed failures to alarm also, particularly at lower relative humidity.^{7-9, 62-3} After two years of CPSC sponsored workshops and hearings, in October, 1996 CPSC staff made recommendations to UL for extensive modifications to the UL 2034 standard that they believed “if adopted, will provide greater protection for consumers through a more reliable and effective safety device.” GRI offered similar, detailed recommendations in UL 2034 and CSA 6.19 technical advisory panel meetings from 1996 to 2001. The CPSC and GRI recommendations included requirements for quality assurance, time-of-manufacture and lifetime testing of alarms, and an alarm threshold of 10% COHb for any time course of CO exposure. These recommendations for reliability and quality assurance were not adopted by UL 2034. Subsequent tests, including those reported here, reveal a continuation of the technical deficiencies found in CO alarms since their inception.

These tests reveal that a significant portion of some widely installed alarm brands did not alarm in CO poisoning incidents with potentially lethal consequences. In particular, they did not adequately protect at low relative humidity, or for the CO profiles of likely poisoning scenarios.

While these tests reveal that three of ten retail brands work well, consumers have no way of choosing the better brands – most brands carry the UL mark (although, the best performing brand of this study was not UL certified).

The technical deficiencies revealed in these tests have persisted since their widespread recognition in 1994, even though technical and regulatory solutions have been available. That several brands of alarms perform well is testament to the industry’s ability to produce well functioning devices at reasonable cost.

Subsequent to the testing reported here, two of the three “well-performing” brands, including the best performing brand with the most accurate digital display, have become unavailable in the market and are no longer available to consumers. Consequently, the average performance of the remaining retail brands may be similar to that of the six “poorly performing” brands in the table above. Further, this study focused on newly purchased alarms. Owing to degradation with age and drift in sensitivity devices installed in residences are not expected to function as well as newly purchased units.

Gas utilities, appliance manufacturers and other users of large numbers of alarms are urged to scrutinize the performance of alarms carefully, through a program of technical review and laboratory testing. It is recommended that whenever possible these parties should require certification to the CSA 6.19-01 standard.

GTI Project Manager Vipul Srivastava, Gas Technology Institute

Introduction

Background

The Poor Field Performance of CO Alarms

From 1993 to 1998 the installed base of residential wall-mount CO alarms grew rapidly in the United States. Several factors were responsible for their popularity: the new availability of low cost CO alarms and their nationwide marketing; the adoption in 1992 of an Underwriters Laboratories Standard for their performance¹; publicity campaigns by the American Lung Association; and the advocacy of the US Consumer Products Safety Commission (CPSC).² In addition, there has been an increased public and media awareness of the dangers of CO. By 1998 approximately 20 million units had been shipped and an estimated 8-15% of homes had at least one alarm installed.³

As CO alarm installation rates rose it quickly became apparent that many alarms were not behaving properly. An inordinate number of false alarms created a noisy and expensive nuisance for many consumers and first responders. At the same time, a silent but insidious problem, that of false negatives in which an alarm should sound but does not, needlessly endangered others.

A multi-year field survey of CO alarm responders found that more than 70% of CO alarm activations occurred at harmless CO levels, and that the incidence of these false alarms varied greatly from brand to brand.^{4,5} Another detailed examination of over 4000 alarm activations found that alarms were not correlated with the incidence of malfunctioning appliances, but they depended instead on conditions of ambient temperature, humidity and outdoor CO levels.⁶¹ Further, alarm incidence was geographically distributed, following patterns of high levels of outdoor pollution.⁶¹ Other GRI-sponsored studies⁶ have found alarm activations to depend more on alarm brand than on the level of CO present. Laboratory studies of commercially available CO alarms have shown that despite their certification to UL 2034, several widely installed CO alarm brands have serious performance problems, including both false alarms and failure to alarm at life-threatening CO concentrations.^{7,8,9,62,63}

False negatives may be a greater problem than false alarms even though they incur no immediate response cost. The documented failures-to-alarm of some UL-certified brands have led to the recall of one million alarms.^{64,76} Because consumers cannot check the sensitivity of their alarms there are undoubtedly many who are not protected while being given a false sense of security. Further, while UL 2034 is a design standard, it does not specify the degree to which alarms must maintain their performance during field installation, and as a result, they have become notorious for their poor reliability as they age.

In field studies⁶²⁻⁶³, CO alarms installed at Southern California Air Quality Management District monitoring sites alarmed in response to ambient CO and moisture levels. The study identified both false alarms and false negatives, concluding,

*... the detector's response is a direct function of CO and moisture. Increasing ambient levels of CO and moisture increase the probability that a detector will alarm. Very low ambient moisture decreases the detector's ability to respond to high levels of CO. Therefore, two types of errors may occur. False positive alarms can occur when the CO is well below hazardous levels but when the ambient moisture is high. False negatives (no alarms) may occur when the CO is above hazardous levels but when the ambient moisture is very low.*⁶²⁻⁶³

These studies⁶²⁻⁶³, as well as laboratory tests⁸, identified the particular liability of CO alarms to fail to activate at low humidity. But even at moderate and high humidity failures to alarm have been observed. Tests by the Southern California Gas Company found that the majority of units of certain models failed to alarm at deadly CO concentrations (greater than 50% COHb) at moderate humidity⁶⁴, and led to recalls of several brands.⁷⁶

The costs of alarms' poor performance are significant.¹⁰ Most calls to responders involve false alarms^{11,12}, and the resulting investigations have been disruptive to consumers and costly to service providers. These costs are not borne by the alarm manufacturers but are instead imposed on consumers and the various responders to the alarm including fire departments, utilities, ambulance service providers, and hospitals. The cumulative costs of these false alarms account for the greater portion of the total response cost and greatly decreases the cost-effectiveness of CO alarm installation.¹³

Beyond the direct costs, false alarms also diminish consumers' confidence in CO alarms. After a false alarm some consumers may intentionally disable their alarms; or worse, tired of an alarm that "cried wolf" too many times, they may ignore future alarms.¹⁰ For smoke alarms the consequences of such alarm fatigue significantly limit their effectiveness, as a CPSC study finds. Owing to their high false alarm rate, and the inability of consumers to determine the legitimacy of an alarm, an even more pronounced effect may be occurring with respect to CO alarms.¹⁰ A comprehensive study of consumer attitudes toward CO alarms found that consumers place little trust in CO alarms; 62% of consumers responded to an alarm activation by ignoring it.¹⁴

CO alarms' reputation for poor performance appears to have slowed installation rates. Shortly after the introduction of CO alarms into the marketplace several cities (e.g., Albany, Chicago, and Toronto) passed closely watched ordinances requiring CO alarms in residential housing. Their poor experience with mandatory installation has discouraged other municipalities from following suit, and even in ordinance cities installation rates have stalled at low levels.

There have been efforts to quantify the performance of CO alarms. Chamber tests⁸ of CO alarms found that many were inadequately sensitive at the relative humidity typical of northern states in the winter, corroborating field reports of failures to alarm in cold weather. Another study of retail-off-the-shelf CO alarms focused on their performance, identified their deficiencies, and determined the degree of their conformance to the UL 2034 standard for sensitivity to CO.⁷ Intended to provide gas utilities with the basis for assessing the comparative capabilities of different alarm models, the study found that a majority of alarm brands available to consumers in 1997 and 1998 suffered from one or another serious deficiency in performance. These deficiencies included numerous out-of-the-box failures, poor control of sensitivity to CO, lack of compliance to the UL 2034 design specification, false alarms on exposure to interference gases, excessive supervised failures, late alarms and false negatives, and inaccurate digital displays. Many of these deficiencies in performance were confirmed in 1999 by tests at Underwriters Laboratories.¹⁵

Further, a recent British Government study¹⁶ found that when newly purchased only 13 of 27 brands fully complied with the sensitivity requirements of the British CO Alarm Standard. Of those thirteen brands, six brands remained commercially available sufficiently long enough for a field trial to be conducted. After one year in the field, only one of those six brands continued to fully comply with the Standard.

Efforts to Improve Alarm Reliability

Consumers may assume that certification to the UL standard assures the full functioning of purchased CO alarms. However, certification does not appear to guarantee functionality in the field. Over one million nonfunctional CO alarms have been recalled from the market as a result of incidental testing by the

Southern California Gas Company.^{64, 76} Despite their inability to alarm at deadly CO levels, these alarms had been certified by Underwriters Laboratories to the UL 2034 standard.

The poor field performance of the installed base of CO alarms points out the need for greater reliability and more stringent testing. There have been concurrent efforts to address these needs. In the summer of 1994, after the field problems had become apparent in areas of widespread installation (i.e., Chicago and Los Angeles), the CPSC sponsored a workshop on their field reliability.¹⁷ The workshop set up continuing task forces charged with developing recommendations for improving CO alarm reliability. In 1995 and 1996 the CPSC continued to hold public hearings on the deficiencies of CO alarms, at which a number of technical recommendations were suggested.¹⁸ In October, 1996 the CPSC issued numerous recommendations to UL for changes to UL 2034¹⁹. Subsequently, UL released a revision in October 1997, which became effective for alarms manufactured after October 1998. This revision requires more rigorous testing; in particular, it requires that individual alarms be subject to multiple tests and that sensitivity testing and selectivity testing be performed using the same alarm.

However, UL did not incorporate CPSC's key recommendations to specify quality assurance and reliability testing at an alarm's time-of-manufacture or over its lifetime. That is, the UL standard does not require the quality control and assurance measures that are standard in many other industries – that a statistically significant sample be tested to assure that alarms do not have inordinately great failure rates. Such tests have been proposed by the CPSC, GRI, and others as essential to assuring alarms' reliability.^{19,20,60}

While the need for more comprehensive testing of CO alarms has been obvious, how to implement these tests has not been as clear. A recent GRI study suggests performance tests which, if implemented as part of a comprehensive program of alarm certification and manufacturing quality assurance, may improve the reliability of alarms installed in the future.²⁰ That study provides a framework for understanding the technical determinants of the poor reliability of currently installed alarms, and specifies the performance parameters to be tested. It points out that a properly designed testing program should evaluate a sufficient number of alarms to develop a statistically valid assessment of their reliability, and provides methods for determining the number of units to be tested.

To address this need for reliability testing, in June 1998 the International Approval Services issued a supplemental requirement to be used as an addenda to UL 2034, designated IAS No. 6-96. This supplemental requirement specifies CO alarms' manufacturing quality and lifetime failure rates.²¹ In brief, it requires that at their time of manufacture at least 99% of alarms meet the UL Standard's tests for sensitivity to CO, and, that throughout a presumed three year lifetime, no more than 14.6% fail to continue to conform to that test, and no more than 33% fail with a warning signal. While these failure rates may seem great for a life-safety device, they are substantially less than current failure rates. These requirements have subsequently been incorporated in a comprehensive standard developed by manufacturers and public agencies under the aegis of the Canadian Standards Association, designated CSA 6.19-01, which became effective in 2001.²² Similar requirements for reliability have not yet been incorporated into the UL 2034 Standard. Because UL 2034 is the primary standard in use in the United States, the improvements incorporated in the CSA standard can not be expected to improve the reliability of CO alarms installed here unless they are also adopted by UL.

Objective and Scope

The purpose of the present study is diagnostic and remedial: to identify deficiencies in CO alarms, to ascertain the physical determinants of their poor performance, to determine the adequacy of prior changes in CO alarm standards in light of the actual performance of alarms, and to recommend improvements to

the standards that are directed to the root causes of poor alarm performance. In particular, this study sought to evaluate the performance of the cohort of alarms marketed after the most recent revisions of the UL Standard to determine the effectiveness of those revisions. These revisions, which became effective in October 1998, do not include the reliability requirements of IAS 6-96 or CSA 6.19-01.

The tests reported here concerned only the basic CO sensitivity of the devices. No efforts were made to reproduce UL tests related to electrical or mechanical reliability, even though there is considerable evidence of poor electrical reliability among installed alarms. For example, laboratory tests⁶⁵⁻⁶⁶ that were initiated by consumer complaints have demonstrated the pronounced susceptibility of some units to electromagnetic interference (for example, from cell phones), despite their certification to UL 2034, which contains tests intended to prevent the certification of susceptible devices. Our own tests have raised questions of the ability of alarms to withstand the electrical surges of UL 2034 Section 48, but we have not yet undertaken a systematic study of the problem.

This study was also not intended to determine the full conformance of commercially available CO alarms to their UL Standard, or to attempt any *a posteriori* remedy for inadequacies in UL's testing or certification process. Even so, the sensitivity test of UL 2034 was used as a primary benchmark for assessing alarms' functioning. Further, this study was not intended to provide a direct comparison of model brands or to rank brands of alarms by performance or suitability to the marketplace; but it was intended as an investigative study of the design and environmental factors which influence the field performance of installed alarms. For this reason, while brands were all initially subjected to a battery of identical tests, for diagnostic purposes additional tests were devised to further explore anomalies of behavior exhibited by particular brands.

In order to gain the broadest perspective on alarm performance, alarms from several sources were studied. As well as testing alarms readily available to the public, several alarms withdrawn from the field and several brands under development, but not yet available commercially, were tested. Specifically, tests were undertaken to answer the following questions about the reliability and performance of CO alarms:

- Has the past performance of alarms been improved in newer brand offerings?
- Do CO alarms function adequately at the lower relative humidity typical of winters in the northeastern and midwestern United States?
- How accurate are the digital displays of CO alarms?

Technical Approach

Brands Tested

While this study focused on testing commercially available alarms, several alarms withdrawn from service and new alarms under development by manufacturers were also assessed.

Nine different brands of commercially available residential CO alarms were purchased predominantly from retail outlets in Chicago and San Francisco, but one brand was purchased directly from the manufacturer. All newly purchased brands but Brand A15 were manufactured in 1999; Brand A15 was manufactured in 1997. These brands are marked “Retail” in Table 1.

Table 1 CO Alarm Brands Tested							
Brand ID	Status	Qty Fully Tested	Sensor Type	Plug-In	Plug-In w/Battery Back-up	Battery Operated	Digital Display
A4	Retail Field	3 1	Semiconductor	--	Yes	--	Yes
A11	Field	1	Semiconductor	Yes	--	--	Yes
A12	Research	6	Electrochemical	Yes	--	--	--
A13 ¹	Retail	10	Electrochemical	--	--	Yes	Yes
A14	Research	7	Electrochemical	--	--	Yes	--
A15	Retail	9	Semiconductor	Yes	--	--	Yes
A16	Research	5	Electrochemical	Yes	--	--	--
A17	Retail	10	Electrochemical	--	--	Yes	--
A18	Retail	10	Electrochemical	Yes	--	--	Yes
A19 ^{1,2}	Retail	13	Electrochemical	--	--	Yes	Yes
A20	Retail	10	Semiconductor	Yes	--	--	Yes
A21	Retail	8	Colorimetric	--	--	Yes	--
A22	Retail	6	Semiconductor	Yes	--	--	--
A23	Research	10	Electrochemical	--	--	Yes	--

¹ As of the writing of this report these brands A13 and A19 are no longer commercially available. It should be noted that these are two of the three “well-performing” brands of this study.

² All retail brands except A19 were certified to UL 2034.

Two alarm units, one each of Brands A4 and A11, designated “Field” in Table 1, were withdrawn from service by a fire department responding to a CO poisoning call. These alarms had not activated, but fire department personnel thought they should have. While the units had been mounted in a conditioned residential area, the incident occurred in early January in a mid-western state after a week of cold dry weather, raising the suspicion that low relative humidity had prevented alarm functioning. These alarms had been manufactured in February and October 1998, and were withdrawn from service shortly after installation in early 1999.

In addition to the field unit of Brand A4, three additional retail units of the same brand were tested at the same time. Brands A4 and A11 appeared to contain identical circuitry except for the battery backup

feature of Brand A4; consequently, they were tested together so that their responses could be directly compared.

Four other models tested, marked "Research" in Table 1, comprised pre-production prototypes, manufactured in 1999 and 2000, but not widely available commercially.

Brands included alarms based on colorimetric, semiconductor, and electrochemical sensor technologies as shown in Table 1. Some brands were wall mounted plug-in units, some required an AC adapter, and some were battery operated. Seven of the brands featured digital displays of CO concentration. The properties of each brand tested are summarized in Table 1.

Test Apparatus

The tests made use of state-of-the-art dynamic gas delivery and response characterization facilities in the laboratories of Mosaic Industries, Inc. This automated gas delivery system complies with a draft ASTM standard for testing systems for measuring the responses of CO alarms.²³ It accurately delivers mixtures of CO and interference gases and vapors over wide concentration ranges in a pure air carrier stream with controlled relative humidity. These facilities comprise a pure air system, a computer-controlled humidification apparatus capable of delivering relative humidity from under 1% to over 80%, multi-channel gas and solvent delivery modules, an 80 liter test chamber, and an operator console that controls the gas delivery system. An on-line gas chromatograph (GC) facilitates verification of gas delivery during testing.

Unlike the "static" gas systems typically used to implement alarm testing, the computer-controlled dynamic gas delivery equipment continuously mixes the target gases and vapors to achieve well-controlled concentrations, high gas purity, and excellent experimental repeatability.

This test apparatus is described fully elsewhere.⁷

Dynamics of Gas Presentation

Each presentation of gas to the CO alarms requires a short time to establish a steady state concentration. The UL Specification for CO Alarms, UL 2034, suggests that CO concentrations should be established within 3 minutes of sealing a test chamber. In our test station the chamber is kept sealed and the CO concentration is established by continuous flow. For the first five minutes of a presentation, gas is supplied in sufficient concentration to charge the test chamber's volume to the desired concentration; thereafter the steady state concentration is maintained by a steady flow of the target concentration. If there were no delays in the gas delivery system the gas concentration would linearly increase throughout the five-minute charge period to its final value. However, transit times of up to 15 seconds combined with the mixing times of 20 seconds and 1 minute in the gas delivery module smooth the transitions in concentration as shown in Figure 1.

The time from the onset of a gas presentation to 50% of the final value is 3.5 minutes, and the gas is considered stable to within 2% of its final value at 7 minutes. Alarms that integrate a response to CO experience a response equivalent to that which would occur if the gas system provided an ideal step in concentration 3.5 minutes after the onset of the gas presentation. For this reason, when comparing alarm times to those required by the UL 2034 sensitivity criteria, the times to alarm activation are measured from the 50% concentration level, a time 3.5 minutes after the onset of the gas presentation.

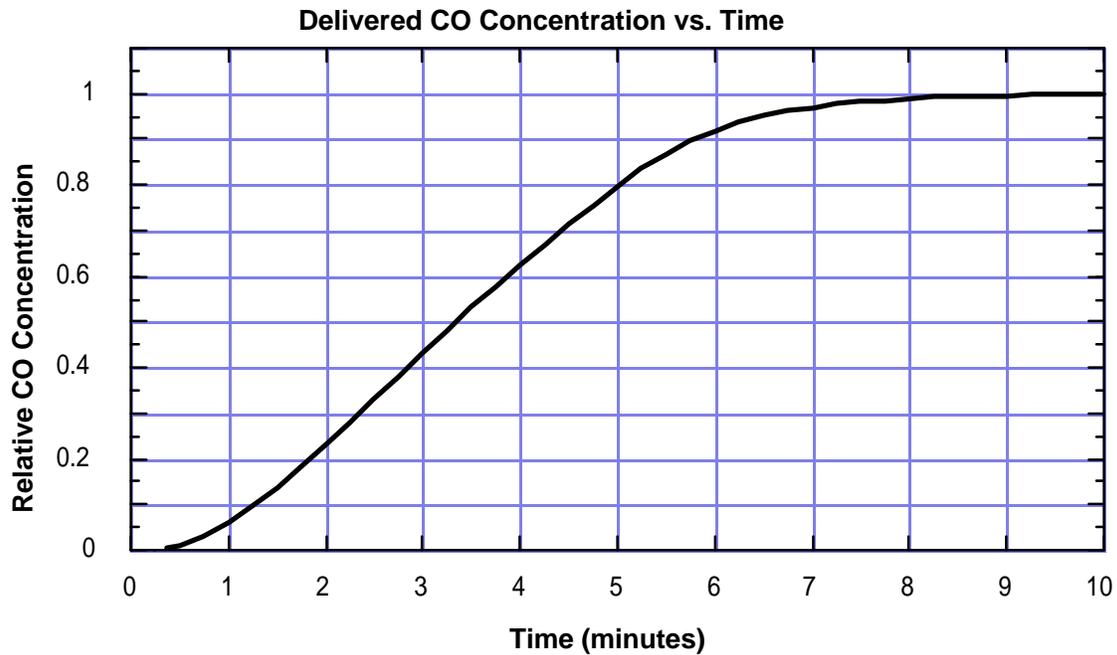


Figure 1: CO concentration as a function of time after the onset of a gas presentation. Concentrations attain 50% of their final value within 3.5 minutes and are effectively stable after 7 minutes.

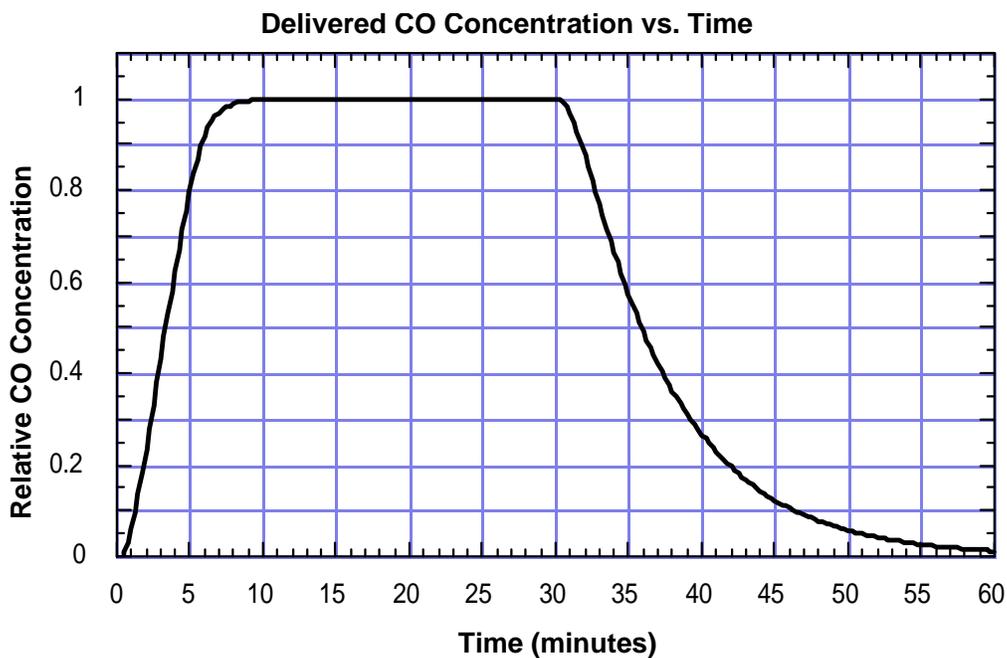


Figure 2: CO concentration as a function of time throughout a gas presentation. At the termination of a presentation the concentration decays with an exponential dilution time of 6.5 minutes.

Data is acquired from the alarms with a temporal resolution of one minute, commencing at the onset of the presentation. Consequently, when comparing measured times to alarm activation to the UL sensitivity criteria, an additional minute is appended on each side of the UL spec, effectively broadening the range of allowed times, to accommodate the maximum quantization error in time measurement.

At the termination of a gas presentation the gas concentration decays with an exponential dilution time of approximately 6.5 minutes as shown in Figure 2.

Calibration

A gas chromatograph (GC) was calibrated against purchased NIST-traceable gas standards and used to calibrate each range of the gas mixing and solvent modules. Thereafter, an electrochemical-cell-based CO meter which was frequently recalibrated against the gas standards was used to monitor the CO levels produced in the test chamber. A more detailed description of gas verification can be found elsewhere.⁷

Data Collection

Wherever possible data acquisition from the CO alarms was automated. Acquired data include times of alarm activations, displayed CO concentrations for those alarms with digital displays, analog voltage outputs from some alarms, and chamber measurements of actual CO concentrations. Data acquisition of each of these parameters is discussed in turn.

Alarm Activations

A microcomputer based data acquisition system was designed to record alarm activations. Microphones were attached to each alarm directly over their piezoelectric sounders and connected to specially designed sound level threshold detectors. A transient alarm of only a few cycles duration (a few cycles of their 2 kHz output), or several milliseconds, was sufficient to be detected. The threshold detectors were extensively tested to assure the reliable detection of an alarm activation without crosstalk among detectors. Once actuated, these threshold detectors provide a logic level output of several seconds duration. The outputs of the threshold detectors were sampled once per second by a microcomputer (a QED-ICS[™] manufactured by Mosaic Industries), which also logged the alarm activation times into nonvolatile memory with a one-minute resolution. If an activation occurred anytime within a one-minute period the CO alarm was considered to be activated for the duration of that minute. The recorded data was then forwarded over a multi-drop serial communications link to the Macintosh control computer where it was combined with the gas delivery information. Alarm activations from up to sixteen CO alarms could be simultaneously recorded with this system.

Digital Display Indications

During each of the sensitivity tests the alarms' digital displays were manually read and recorded.

Actual CO Measurement

Measurements of the CO in the test chamber were manually recorded periodically throughout the tests and each time a digital display was recorded. The CO meter was periodically recalibrated against tank concentrations of 10 and 400 ppm CO to verify its accuracy.

Specific Tests

Testing consisted of alternating exposures to CO, to assess alarm sensitivity, with exposures to a variety of interference gases, to determine the alarms' immunity to false alarms. Exposures to CO were presented at various relative humidity to assess alarms dependence on humidity. Alarms could not all be placed in the chamber at once, so testing was performed in batches of ten to fifteen alarms each.

These tests provide data to evaluate the short-term performance of residential CO alarms, establish their basic functionality, assess their dependence on relative humidity, and determine the degree of their compliance with UL 2034 standards for sensitivity to CO.

All tests provided gas concentrations that stabilized to within 2% of their final concentrations in 7 minutes and decayed at the end of their presentation with a characteristic exponential decay time of 6.5 minutes.

Tests of any particular brand comprised various combinations of the following fundamental test sequences:

Sensitivity Test A – 50 and 70 ppm CO

Sensitivity Test A provided a 10.5-hour assessment of sensitivity to 50 and 70 ppm CO, with a 3-hour recovery between the two presentations. The 50 ppm presentation was required in an early version of UL 2034 and the 70 ppm concentration is a test concentration in the current (post Oct. 1998) version.

Throughout the test, the times to alarm for all devices were measured with one-minute resolution. Digital display readings were recorded 25, 55, and 85 minutes into the 50 ppm presentation, and 30, 90, 150, and 210 minutes into the 70 ppm presentation.

Sensitivity Test A provided the following sequence of CO concentrations:

Table 2. Sensitivity Test A	
Duration	Exposure
30 min	clean air
90 min	50 ppm CO
3 hours	clean air
5 hours	70 ppm CO
30 min	clean air

Figure 3 shows the time course of CO concentration and simulated carboxyhemoglobin throughout Sensitivity Test A. Carboxyhemoglobin was determined by computer simulations of a COHb model similar to that cited in UL 2034, and described in detail in Appendix A. The UL 2034¹ model provides COHb as a function of time given initial COHb and external CO concentration as,

$$\%COHb_t = \%COHb_o[e^{-(t/2398B)}] + 218 [1 - e^{-(t/2398B)}] [C_o + (\text{ppm CO} / 1316)] \quad \text{Eqn. 1}$$

where %COHb_t is the percentage of blood saturation with COHb at time t, %COHb_o is the initial percentage of COHb, t is the time in minutes, and the parameter B pertains primarily to ventilation rate consequent to physiological work level. UL 2034 uses B = 0.0404, signifying a heavy work effort.

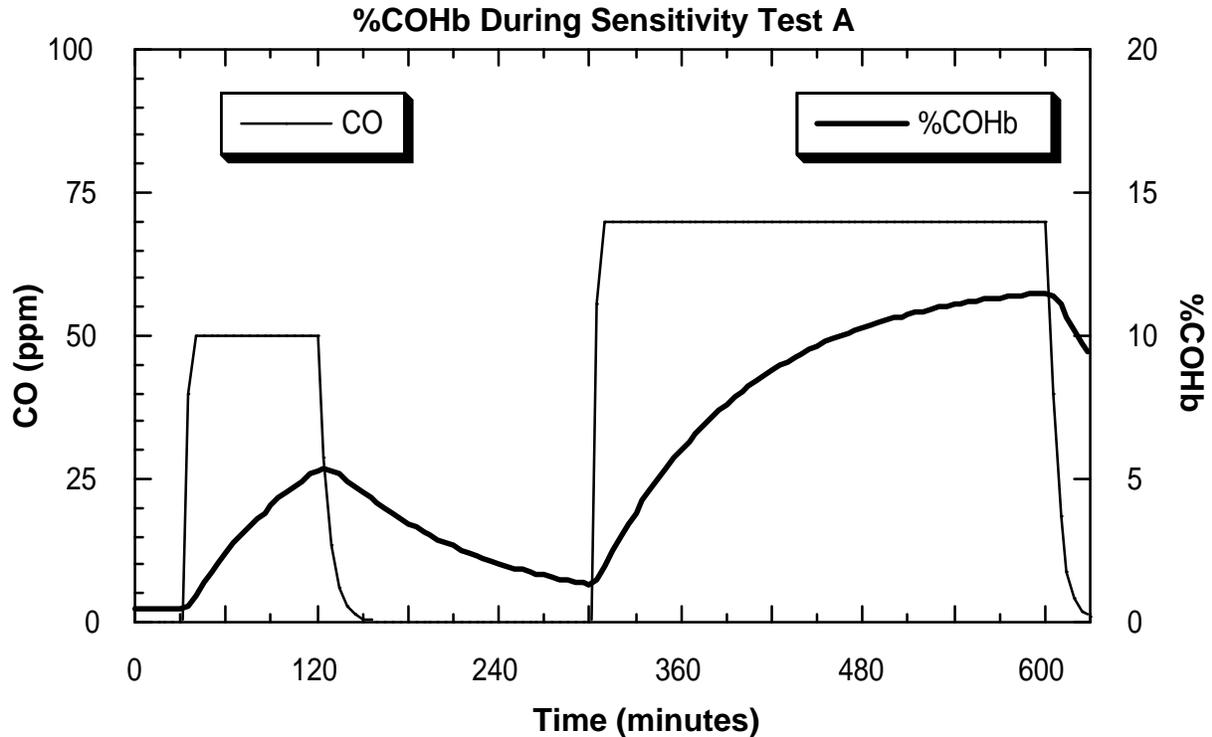


Figure 3. CO concentration and simulated carboxyhemoglobin (%COHb) as a function of time throughout Sensitivity Test A.

The parameter C_o in Eqn. 1 represents endogenous COHb production resulting from normal metabolism. The equation cited in UL 2034 uses $C_o = 0.0003$, which is a level of endogenous CO production equivalent to breathing 0.4 ppm CO in the absence of endogenous production, and that results in a steady state %COHb of only 0.06%. This level is far less than the normal physiological levels of 0.3 – 0.8 %COHb cited by many sources^{24, 25, 26, 27, 28, 29} for nonsmokers, suggesting that the value of C_o used in UL 2034 is mistaken. This use of an essentially zero percent background COHb level was criticized by the Consumer Product Safety Commission in their recommendations for changes in UL 2034.¹⁹ An accurate value should be approximately an order of magnitude greater in order to represent a realistic physiological level.

For this reason we use a value different than that quoted by UL 2034. Throughout this report simulated carboxyhemoglobin levels are computed using a value of $C_o = 0.002$. This endogenous rate of CO production is equivalent to that produced by continuously breathing 2.6 ppm CO in the absence of endogenous production, and results in a steady state %COHb of 0.44%.

Modifications of Eqn. 1, as explained in Appendix A, were used to emulate COHb production on a minute-to-minute basis, considering the concentrations of CO presented throughout the sensitivity tests and the dynamics of gas delivery and decay of Figure 3. The initial %COHb used in the simulations was the endogenous steady state level of 0.44%.

During the 50-ppm exposure of Sensitivity Test A, the simulated %COHb rose to more than 5%, and during the 70 ppm exposure it exceeded 10%. After the recovery period between the two presentations it had fallen to less than 1.5%.

Sensitivity Test B – 150 and 400 ppm CO

Sensitivity Test B followed the format of Sensitivity Test A but used CO concentrations of 150 and 400 ppm with a recovery period of 5 hours between, and lasted 9.5 hours. These concentrations are taken from the UL 2034 sensitivity test. Sensitivity Test B provided the following sequence of CO concentrations:

Table 3 Sensitivity Test B	
Duration	Exposure
30 min	clean air
2 hours	150 ppm CO
5 hours	clean air
1 hour	400 ppm CO
1 hour	clean air

As during Sensitivity Test A, CO was provided in a background of clean air. Times to alarm for all devices were measured with one-minute resolution. Digital display readings were recorded 25, 55, 85, and 115 minutes into the 150 ppm presentation, and 25, 40, and 55 minutes into the 400 ppm presentation.

Figure 4 shows the time course of CO concentration and simulated carboxyhemoglobin throughout Sensitivity Test B. Carboxyhemoglobin was determined as described above for Sensitivity Test A. Gas presentations were of a sufficient duration so that during the 150 ppm exposure the simulated %COHb rose to nearly 20%, and during the 400 ppm exposure it exceeded 30%. After the recovery period between the two presentations it had fallen to less than 1.5%.

Sensitivity Test C – Progressively Increasing CO Concentration

The prior sensitivity tests are modeled on the tests of the UL standard in which CO concentrations are abruptly applied to the devices under test. However, realistic CO poisoning events do not involve the abrupt appearance of a fixed concentration of CO; instead, they involve steadily increasing concentration. Both theoretical models and experimental measurements in test houses show that in the most likely acute poisoning scenarios CO concentrations rise steadily and nearly linearly.^{68, 69, 71, 72, 73, 74}

Detailed measurements in test houses using a variety of surrogate and actual appliances reveal typical rates of CO increase of 50 ppm/hr extending over a period of approximately four hours.^{68,69,71,73,74} Sensitivity Test C is designed to mimic these most likely poisoning scenarios by providing a step-wise concentration increase with a 30 minute resolution and an average rate of 50 ppm/hr.

Sensitivity Test C was also used to determine the accuracy of alarms' digital displays and the level, in terms of %COHb, at which the alarms actuated. It provided progressively graded concentrations, stepping from zero to 175 ppm of CO in 25 ppm increments, over a span of 4.5 hours. This test provided each concentration for a one half-hour period. We found that this time was sufficient for the digital

displays of all brands tested to stabilize. Digital displays were recorded 25 to 30 minutes after the start of each presentation.

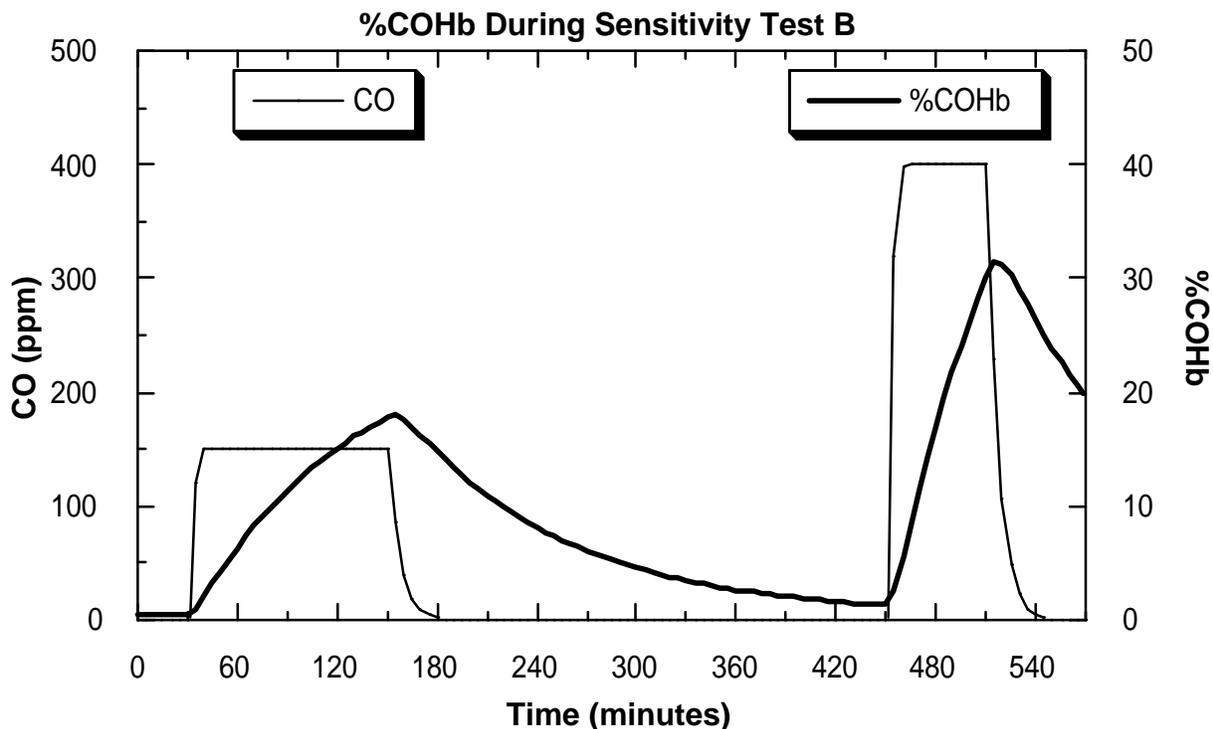


Figure 4 CO concentration and simulated carboxyhemoglobin (%COHb) as a function of time throughout Sensitivity Test B.

The sequence of CO concentrations used in Sensitivity Test C is shown in Table 4.

Table 4 Sensitivity Test C	
Duration	Exposure
30 min	clean air
30 min	25 ppm CO
30 min	50 ppm CO
30 min	75 ppm CO
30 min	100 ppm CO
30 min	125 ppm CO
30 min	150 ppm CO
30 min	175 ppm CO
30 min	clean air

Again, times to alarm were measured with one-minute resolution. Additionally, for alarms with digital displays the displays were read 25 to 30 minutes after the start of each gas presentation.

Figure 5 shows the time course of CO concentration and simulated carboxyhemoglobin throughout Sensitivity Test C. Carboxyhemoglobin was determined as described above for Sensitivity Test A. COHb rose continuously throughout the gas presentations. The times to alarm were compared to the times to attain 2.5% and 10% COHb.

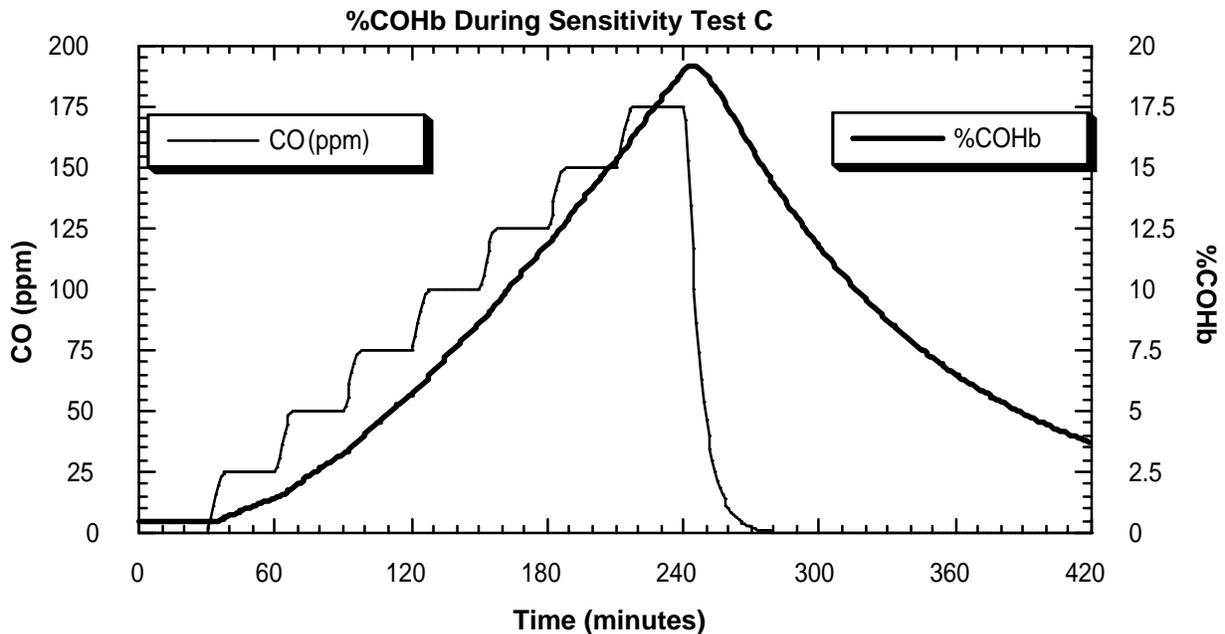


Figure 5 CO concentration and simulated carboxyhemoglobin (%COHb) as a function of time throughout the graded Sensitivity Test C.

Sensitivity Test H – 70, 150, and 400 ppm CO

It became apparent throughout the course of testing that some models of alarms de-integrated CO concentrations with a greater time constant than they integrated. Consequently, they required a greater recovery time than Sensitivity Tests A and B provided. Sensitivity Test H was devised to provide a much greater recovery time between presentations of the UL sensitivity test concentrations. The time course of gas presentations is summarized in Table 5 and Figure 6, and detailed in Figure 7, Figure 8, and Figure 9. This sensitivity test was the one most frequently used.

Table 5 Sensitivity Test H	
Duration	Exposure
30 min	clean air
5 hours	70 ppm
7 hours	clean air
2 hours	150 ppm CO
15 hours	clean air
1 hour	400 ppm CO
1 hour	clean air

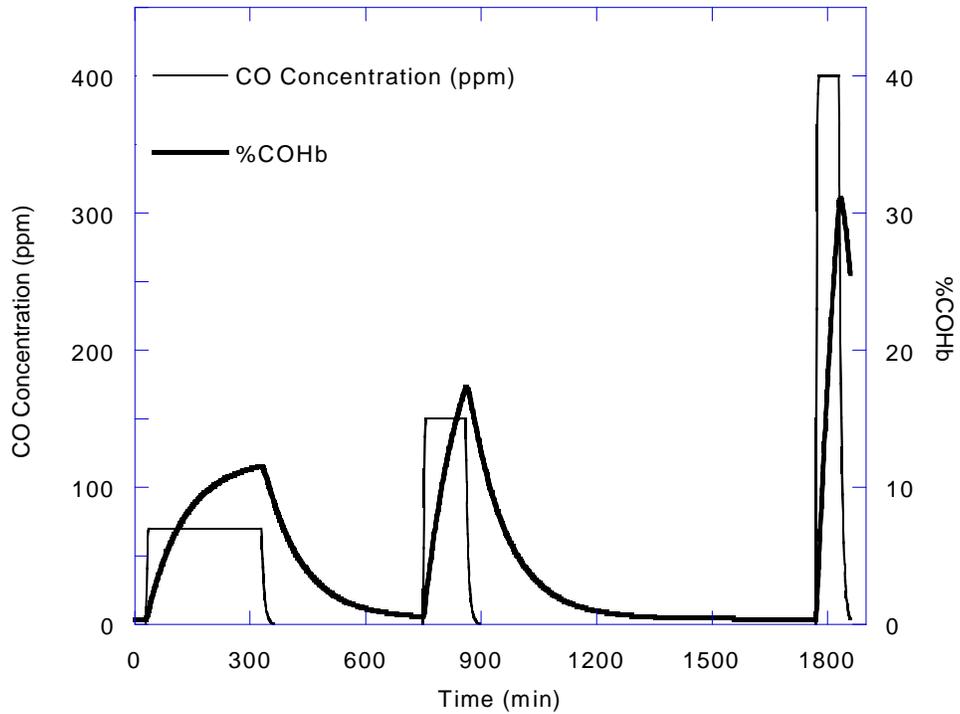


Figure 6 CO concentration and simulated carboxyhemoglobin (%COHb) as a function of time throughout the Sensitivity Test H to 70, 150, and 400 ppm CO.

Sensitivity Test H, 70 ppm presentation:

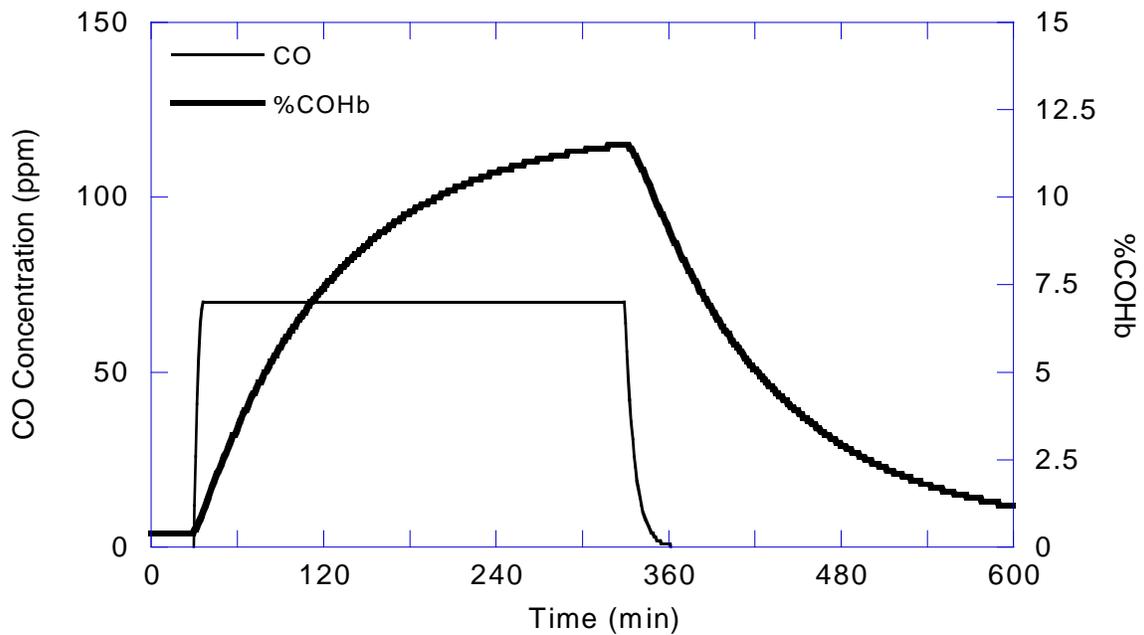


Figure 7 CO and COHb concentrations during the 70 ppm presentation of Sensitivity Test H. COHb exceeds 12% by the end of the presentation.

Sensitivity Test H, 150 ppm presentation:

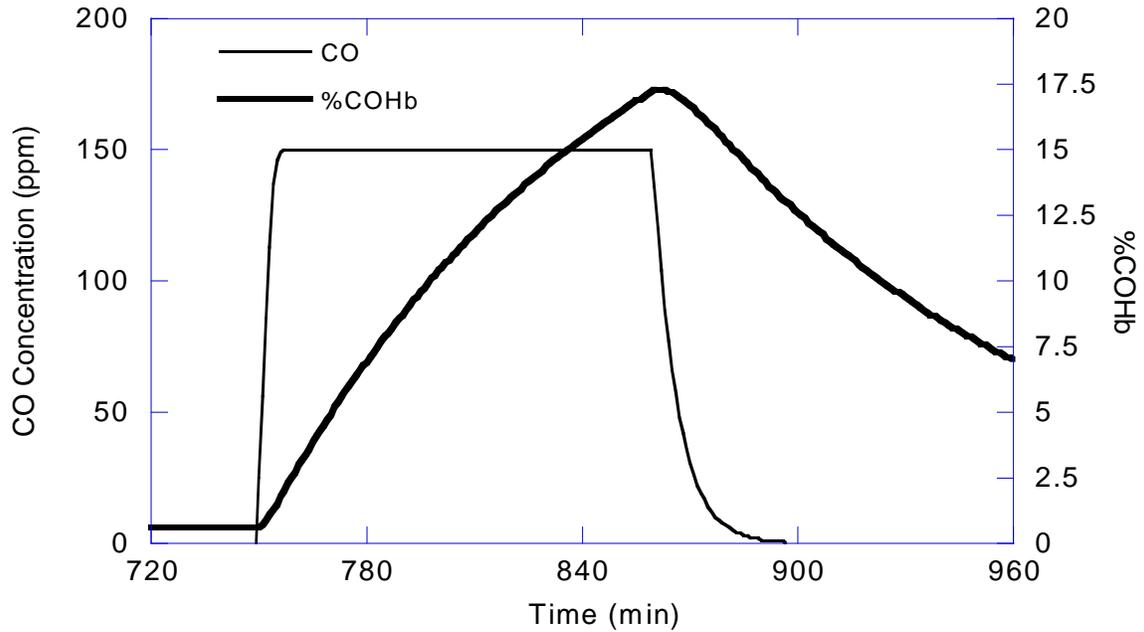


Figure 8 CO and COHb concentrations during the 150 ppm presentation of Sensitivity Test H. COHb attains 17.5% by the end of the presentation.

Sensitivity Test H, 400 ppm presentation:

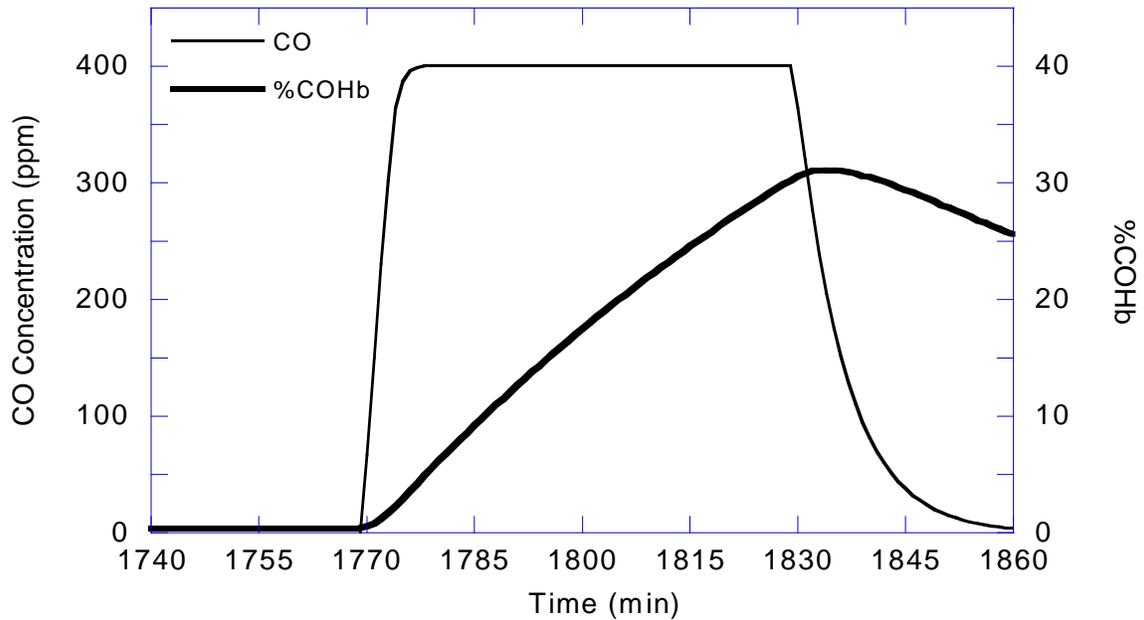


Figure 9 CO and COHb concentrations during the 400 ppm presentation of Sensitivity Test H. COHb exceeds 30% by the end of the presentation.

Sensitivity Test D – 30 ppm CO

Sensitivity Test D is a 24 hour test for immunity to 30 ppm CO, with 30 minute setup and recovery times. Throughout the 24-hour test the times to alarm for all devices are measured with a resolution of one minute. Digital display readings are recorded 30, 90, 150, and 210 minutes into the presentation.

Figure 10 shows the time course of CO concentration and simulated carboxyhemoglobin throughout Sensitivity Test D. Carboxyhemoglobin was determined as described above for Sensitivity Test A. Times to alarm were tested against the UL 2034 specification, which requires that the units do not alarm for 24 hours, and against the time needed to attain 2.5% COHb.

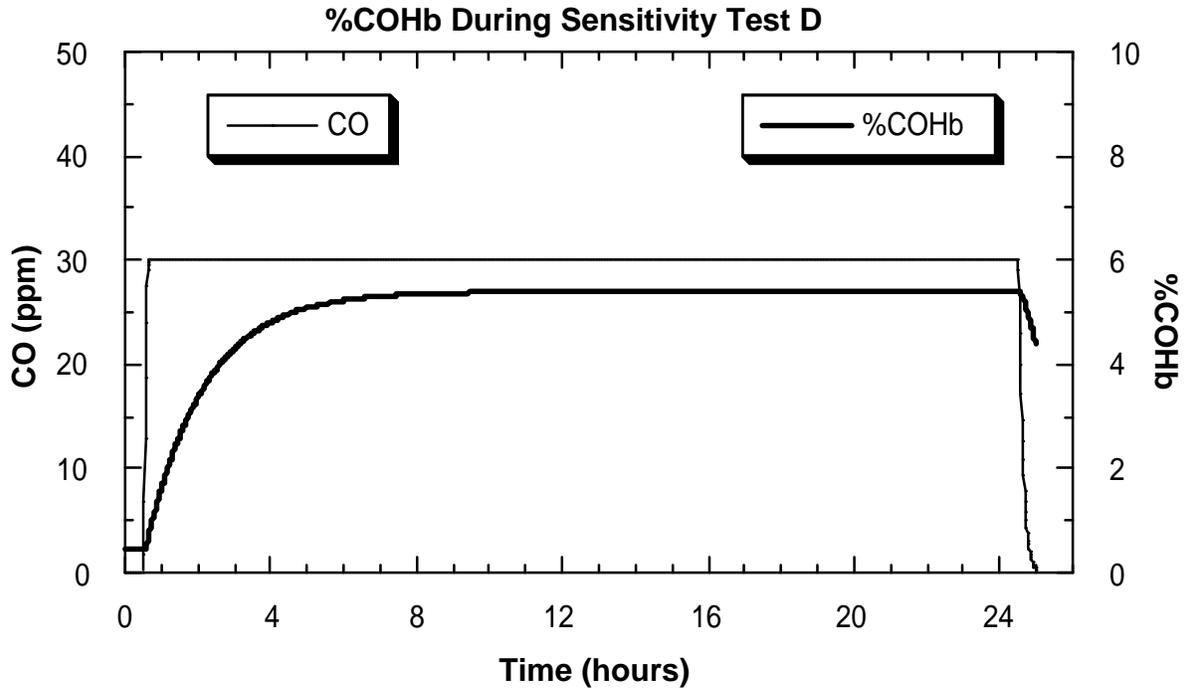


Figure 10 CO concentration and simulated carboxyhemoglobin (%COHb) as a function of time throughout Sensitivity Test D

Sensitivity Tests E and F – Varying Relative Humidity

Table 6 Sensitivity Test E		
CO	RH	Time
200 ppm	50%	30 minutes
200 ppm	80%	90 minutes
200 ppm	50%	90 minutes
200 ppm	20%	90 minutes
200 ppm	10%	90 minutes
200 ppm	5%	90 minutes

Table 7 Sensitivity Test F		
CO	RH	Time
200 ppm	5%	90 minutes
200 ppm	10%	90 minutes
200 ppm	20%	90 minutes
200 ppm	50%	90 minutes
200 ppm	80%	90 minutes
200 ppm	50%	30 minutes

In order to directly observe the influence of relative humidity on the sensitivity of alarms with digital displays a test was devised that subjects the alarms to a constant CO concentration and varying relative humidity. After a warm-up period of 30 minutes at 50% RH and 200 ppm CO, Sensitivity Test E maintained a continuous exposure to 200 ppm CO with decreasing relative humidity as per the following table. While Sensitivity Test E implemented a stepped decrease in relative humidity, Sensitivity Test F provided a stepped increase. These tests were implemented in succession, so that the alarms were exposed to a stepwise decrease in humidity followed by a stepwise increase.

CENELEC Sensitivity Test

Several alarm brands were tested to the European Union CO Standard (the CENELEC standard). This test sequentially exposed the alarms to concentrations of 33, 55, 110, 330, 5000, and 55 ppm in order, with only short recovery times in between.

Interference Test A

Interference Test A implemented the UL 2034 requirement to test some interference gases while expanding the number of gases tested. Over a duration of 28 hours it provided two hour exposures to each of the gases in Table 8.

The shaded rows of Table 8 represent gases that are required by the interference tests of UL 2034. These are i-butane, methane, i-propanol, ethyl-acetate, and n-heptane. UL 2034 also calls for exposure to carbon dioxide at a concentration of 5000 ppm, which was not tested in this project. The additional gases of Interference Test A are some that are typically found in household residences in the form of cleansers, glues, and solvents. Ethanol, toluene, and acetone were recommended by the Consumer Product Safety Commission for testing CO alarms³⁰; in addition to these, ammonia and methanol have been recommended as indoor pollutants.²⁰

Table 8 Interference Test A		
Gas or Vapor	Concentration (ppm)	Duration (minutes)
i-Butane	300	120
Methane	500	120
Ammonia	50	120
Methanol	200	120
Ethanol	200	120
Acetone	200	120
i-Propanol	200	120
Ethyl-acetate	200	120
n-Heptane	500	120
Toluene	200	120
Trichloroethylene	200	120

Interference Test A included a 30-minute setup time and 30-minute recovery times after each gas presentation. Throughout the 28-hour test the times to alarm for all devices were measured with a resolution of one minute. Digital display readings were not recorded during this test.

Our interference test is more extensive than the UL test, the applicable revision of which tests to methane, butane, heptane, ethyl-acetate, i-propanol, and carbon dioxide. UL has however indicated that it will add some of these other gases to a standard revision to take effect in 2003. That revision will include ammonia, ethylene, ethanol, toluene, trichloroethane, and acetone.

Table 9 Interference Test C		
Gas or Vapor	Concentration (ppm)	Duration (minutes)
Clean Air		30
i-Butane	300	120
Clean Air		120
Methane	500	120
Clean Air		120
Ammonia	50	120
Clean Air		120
Methanol	200	120
Clean Air		120
Ethanol	200	120
Clean Air		120

Table 10 Interference Test D		
Gas or Vapor	Concentration (ppm)	Duration (minutes)
Clean Air		30
Acetone	200	120
Clean Air		120
i-Propanol	200	120
Clean Air		120
Ethyl-acetate	200	120
Clean Air		120
n-Heptane	500	120
Clean Air		120
Toluene	200	120
Clean Air		120

Interference Tests C and D

During the course of testing, as will be seen, many alarms responded to the interference gases in such a way as to suggest that their protective carbon filters were becoming fully loaded. In an effort to reduce this effect Interference Gas Tests C and D were devised. These provide a full two hour recovery period between gas presentations, so that the entire set of gas presentations takes place over a two day period.

Test Sequence

Generally, the tests comprised the following sequence of events:

Inspection and Conditioning: On receipt of the alarms they are inspected for any obvious out-of-the-box failures. They are then conditioned under power for three days at a nominal 50% relative humidity (or more precisely, at 13,200 ppm water vapor concentration) in clean air at room temperature (20-24 °C).

UL Sensitivity Tests: The alarms are then tested for their sensitivity to CO at 70, 150, and 400 ppm in a background of clean air containing 50% relative humidity, using Sensitivity Test H.

Progressive CO Dose Test: The alarms are tested to a progressive increase in CO concentration to determine the effective COHb at which the alarms actuate, at 50% RH, using Sensitivity Test C. Responses of alarms with digital displays are recorded at each of the concentrations.

Interference Test: The alarms are exposed sequentially to concentrations of isobutane, methane, ammonia, methanol, ethanol, acetone, isopropanol, ethyl acetate, heptane, toluene, and trichloroethylene, for two hours each, using Interference Tests A, C, or D. During this time the relative humidity is maintained at 50%.

Exposure to Low Relative Humidity: The sensitivity and progressive dose tests are then repeated at a nominal 5% relative humidity (or more precisely, at 1,320 ppm water vapor concentration) at room temperature after a conditioning period at 5% RH. We have used conditioning periods at low relative humidity of from 2 hours to 2 weeks before repeating the CO sensitivity tests and have found, surprisingly, that the period at lower relative humidity has little influence on the outcome of the test.

UL Sensitivity Tests at 5% RH: The alarms are retested fully for their sensitivity to CO at 70, 150, and 400 ppm in a background of clean air containing 5% relative humidity.

Progressive Exposure Test at 5% RH: Finally the alarms are retested to a progressive increase in CO concentration to determine the effective COHb at which the alarms actuate, now at 5% RH. Responses of alarms with digital displays are recorded at each of the concentrations to determine their sensitivity at 5% RH.

Ideally the performance of the alarms would be fully repeatable so that exposures to CO do not influence their subsequent behavior. However, in the event that some alarms are not repeatable, the above sequence is chosen to minimize the possibility of adversely affecting the sensitivity of the alarms before that sensitivity can be measured. For this reason the first exposure of the alarms to gas is to the UL test concentrations (70, 150 and 400 ppm).

Most of the brands were subjected to the following typical time course of tests:

<u>Action</u>	<u>Duration</u>
1. Inspection for out-of-box failures	
2. Alarm conditioning under power at 50% relative humidity in clean air	3 days
3. Sensitivity Test H at 50% RH	2 days
4. Sensitivity Test C at 50% RH	1 day
5. Interference Test C at 50% RH	1 day
6. Interference Test D at 50% RH	1 day
7. Alarm conditioning under power at 5% relative humidity in clean air	3 days
8. Sensitivity Test H at 5% RH	2 days
9. Sensitivity Test C at 5% RH	1 day

While the alarms were often conditioned for three days at each relative humidity, we found that this conditioning is unnecessary. Exploratory tests of many brands revealed that alarms' behavior was insensitive to the conditioning time, for conditioning times ranging from two hours to two weeks.

For some brands there were variations from this test sequence in order to test to standards other than UL 2034 (for example the CENELEC standard), or to further explore alarms' humidity response. Whenever other test sequences were used the sequence implemented is described in conjunction with the test results in the brand-by-brand summaries.

Test Results

A simple tally of alarm activations does not adequately represent the complexity of the responses of residential CO alarms. So that readers of this report may gain a fuller appreciation of that complexity, all test results are presented here in a series of graphs for each alarm brand, in brand-by-brand order. These graphs representing the raw test results are generally of two types. In the first, alarm activations are represented as a horizontal trace for each device above a time axis. For most tests, several traces, representing the alarm actuations of several devices simultaneously exposed, are shown on a single graph. A flat line represents a device that does not alarm, while an abrupt step in the trace represents an alarm actuation. The trace remains high as long as an alarm signal is detected any time within a sixty second window. Whenever the alarm signal is not present for a duration greater than one minute the trace returns to a low position.

For alarms with a digital display a second type of graph shows the display readings as a function of the actual CO concentration, as they were displayed during the progressive CO concentration test, Sensitivity Test C.

For readability and consistency in presentation the results are placed in a standard sequence (sensitivity to CO at low humidity, followed by sensitivity to CO at high humidity, followed by digital display accuracy) even though the tests may not have been performed in that order.

This section presents the raw test results accompanied where appropriate by notes about brand-specific behavior. The next section of this document, "Analysis", presents summaries across alarm brands.

Brands A4 and A11

Brands A4 and A11 were tested at the request of a Fire Department that had retrieved two devices from the site of a CO poisoning event. In early January, 1999 a vehicle had been inadvertently left running in an attached garage and discovered three and a half hours later only after residents had become ill. On arrival, fire department personnel measured 800 ppm CO in the garage. Occupants of an attached apartment unit that did not contain CO alarms were transported to the hospital.

A second attached apartment unit contained the alarms, which had not sounded. Both alarms had been installed in November, 1998, one on the first floor and one on the second. The fire department measured CO levels of 110 ppm in the apartment unit containing the alarms. Both devices have digital displays and two buttons, one for test/reset and one for displaying a peak value. The upstairs device (which had been manufactured in February, 1998, and which we designate A11-B10-0) registered a peak level of 68 ppm, while the downstairs device (manufactured in October, 1998 and designated A4-B10-1) registered a peak level of 86 ppm. These peak recorded levels were both significantly less than the levels measured by the fire department, although the fire department report does not specify what the display readings were when the 110 ppm was measured. As the fire department's 110 ppm reading was taken well after the poisoning incident and some time after ventilation of the dwelling fire department personnel surmised that actual CO levels were likely to have been much greater.

Because the alarms had not sounded, and their peak recorded readings were significantly less than the CO levels measured by the fire department, there was a suspicion that they were not functioning adequately. Consequently, they were removed for testing. In particular, the fire department questioned whether the low ambient relative humidity conditions prevailing at the time (resulting from cold outdoor conditions) might have influenced the alarms.

These two devices, one Brand A4 and one Brand A11, are nearly identical in design, except that Brand A4 is AC wall adapter powered with a 9 V backup battery, while Brand A11 was AC powered alone.

Simultaneously with testing these two devices, three other brand A4 devices were tested for comparison. Testing consisted of alternating exposures to CO, to assess alarm sensitivity, with exposures to varied relative humidity. Testing comprised these steps:

1. The alarms were first inspected on receipt for obvious failures.
2. Initial Sensitivity Tests: The devices were powered for two days in a clean air ambient nominally containing 50% relative humidity (RH) (or more precisely, a controlled 13,200 ppm H₂O) and then tested for their sensitivity to CO at 150 and 400 ppm, while maintaining 50% RH. The alarms were then tested to a progressive increase in CO concentration to determine the effective COHb at which the alarms actuated, at 50% RH. During these tests gas presentations were computer-automated and alarm responses and activations were recorded. Times to alarm activation were compared to those required by the UL 2034 specification and to times to attain 2.5% to 10% carboxyhemoglobin (%COHb) levels.
3. Exposure to 5% Relative Humidity: The devices were then powered for two weeks in a clean air ambient containing 5% RH (or more precisely, 1320 ppm H₂O).
4. Sensitivity Tests at 5% RH: Maintaining the alarms at 5% RH, they were subjected to a battery of tests of their CO sensitivity including tests at 30, 50, 70, 150, and 400 ppm. They were also subjected to a test in which the CO concentration is progressively increased.
5. Exposure to 50% Relative Humidity: The devices were then powered for one week in a clean air ambient containing 50% RH.
6. Sensitivity Tests at 50% RH: Maintaining the alarms at 50% RH, they were subjected to a battery of tests of their CO sensitivity including tests at 30, 50, 70, 150, and 400 ppm. They were also subjected to a test in which the CO concentration is progressively increased.
7. Sensitivity to CO at Varying Relative Humidity: The devices were then tested at a fixed concentration of CO (200 ppm) and their digital display readings recorded as relative humidity was varied from 80% RH to 50%, 20%, 10% and 5% RH and back. The duration of each humidity step was 90 minutes.
8. Repeated Tests without Conditioning: The devices were then retested as in steps 2-6 above, but without the two week and one week conditioning periods.

The results of these tests showed the following:

- The inter-device variation of sensitivity is great – from unit to unit display readings range over a factor of three to one at a fixed humidity, and up to six to one over the humidity range. See for example Figure 21 and Figure 22
- Sensitivity to CO is strongly dependent on relative humidity, with sensitivity dropping off profoundly at low relative humidity. Depending on their initial calibration (some devices were strongly over-sensitive while some were under-sensitive), some devices would become inadequately sensitive to CO anywhere between 50% RH and 5% RH. At 5% RH all devices were inadequately sensitive to CO.

- Alarms’ dependence on relative humidity is a prompt effect, tracking variations in humidity with a time constant of 30 minutes or less. Although in our first tests we provided a two-week conditioning period at low relative humidity, we found in subsequent tests that the same drop off in sensitivity at low relative humidity is seen without any conditioning time.
- Alarms’ decline in sensitivity to CO is a smooth and gradual function of decreasing humidity. A greater and greater proportion of devices become inadequately sensitive to CO at progressively lower and lower humidity, but there is no threshold humidity at which sensitivity to CO abruptly changes. This gradual decline in sensitivity as humidity is lowered is depicted in Figure 27 and Figure 28.
- The two “defective” devices withdrawn by fire department personnel were not defective in the sense that their performance differed from other units of the same brand. The following raw test results show that their performance was typical of other units of the same brand.

As regards the two devices withdrawn from the field, these test results suggest that their inadequate sensitivity is likely to have resulted from the prevailing condition of low ambient relative humidity. The downstairs unit (Brand A4) read a peak of 86 ppm and upstairs unit (Brand A11) read a peak of 68 ppm. If the actual relative humidity were as low as 5%, the actual CO exposures would have been as great as 230 ppm (A4) and 225 ppm (A11). (Note that when corrected for their measured calibration curves the alarms agree rather well, with 230 vs 225 ppm.) If the actual relative humidity were greater than 5% the alarms would have been more sensitive and their peak readings during the incident closer to the actual CO concentration. Unfortunately, owing to the relative humidity dependence of the sensitivity we can not know the actual concentration present in the residence, only that it was between 68 and 230 ppm, if the humidity was between 50% and 5%, as shown in Table 11. In any case, as a result of their inadequate sensitivity, these two devices did not protect their users against harmful CO levels.

Table 11. Peak Reading vs Range of Likely Actual CO Concentrations

	Peak Reading	Actual CO at 50% RH	Actual CO at 5%
Upstairs unit (Brand A11)	68 ppm	65 ppm	225 ppm
Downstairs Unit (Brand A4)	86 ppm	70 ppm	230 ppm

The following graphs illustrate the raw data collected during the tests. In each graph the alarms are designated with a serial number, for example “A4-B5-13” in which the “A4” identifies the brand and the “B5-13” uniquely identifies the individual instrument. (The funny format “B5-13” is just an artifact of our accounting system for alarms.) The bottom horizontal trace represents the times of presentation of the CO gas, while the remaining horizontal traces represent the alarm actuations.

Those graphs marked “Sequence A” represent tests performed on initial receipt of the devices and after a two week period of 5% relative humidity, while “Sequence B” tests were done without conditioning periods at high and low relative humidity. There were no significant differences between the two sets of results indicating that the decrease in sensitivity observed at lower relative humidity did not result from long term exposure to dry conditions, but was rather a prompt response to lower humidity.

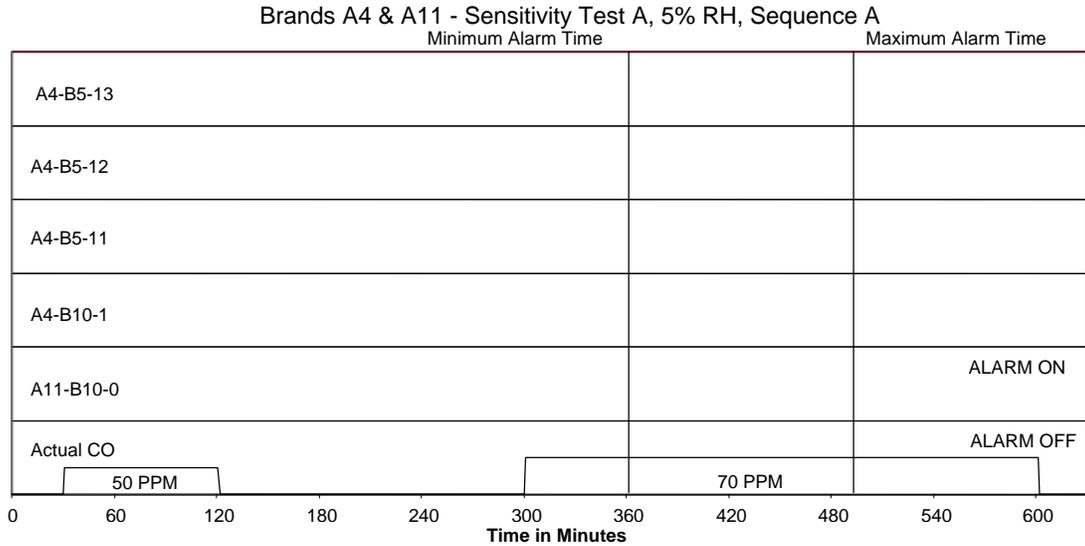


Figure 11 The bottom trace represents the times of presentation of 50 and 70 ppm CO and the two vertical lines represent the times at which 2.5% and 10% COHb would be achieved in a hypothetical exposed person. The two lower alarm traces, labeled A11-B10-0 and A4-B10-1 pertain to the two units removed from service, and the remaining Brand A4 units were tested for comparison. At 5% RH all units tested of Brand A4 and A11 were insufficiently sensitive to alarm at 70 ppm; they never alarmed.

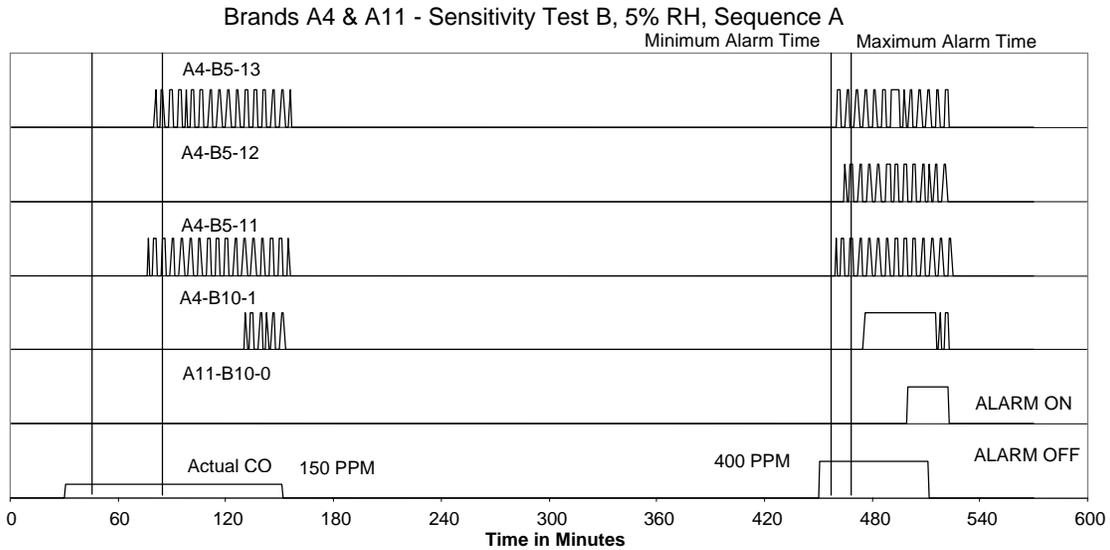


Figure 12 The bottom trace represents the times of presentation of 150 and 400 ppm CO and the two vertical lines represent the times at which 2.5% and 10% COHb would be achieved in a hypothetical exposed person. Both of the units removed from service, A11-B10-0 and A4-B10-1, failed to alarm at less than 10% COHb at 5% RH. The comparison units either failed to alarm or alarmed intermittently.

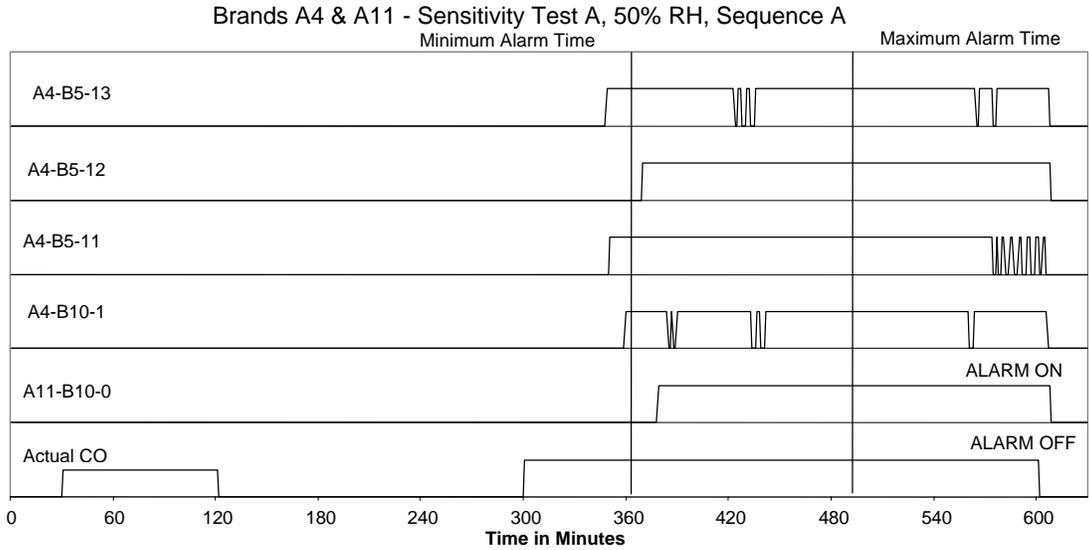


Figure 13 The bottom trace represents the times of presentation of first 50 and then 70 ppm CO and the two vertical lines represent the times at which 2.5% and 10% COHb would be achieved in a hypothetical exposed person. The two lower alarm traces, labeled A11-B10-0 and A4-B10-1 pertain to the two units removed from service, and the remaining A4 units were tested for comparison. At 50% RH all A4 and A11 units tested alarmed at less than 10% COHb, and some at less than 2.5% COHb.

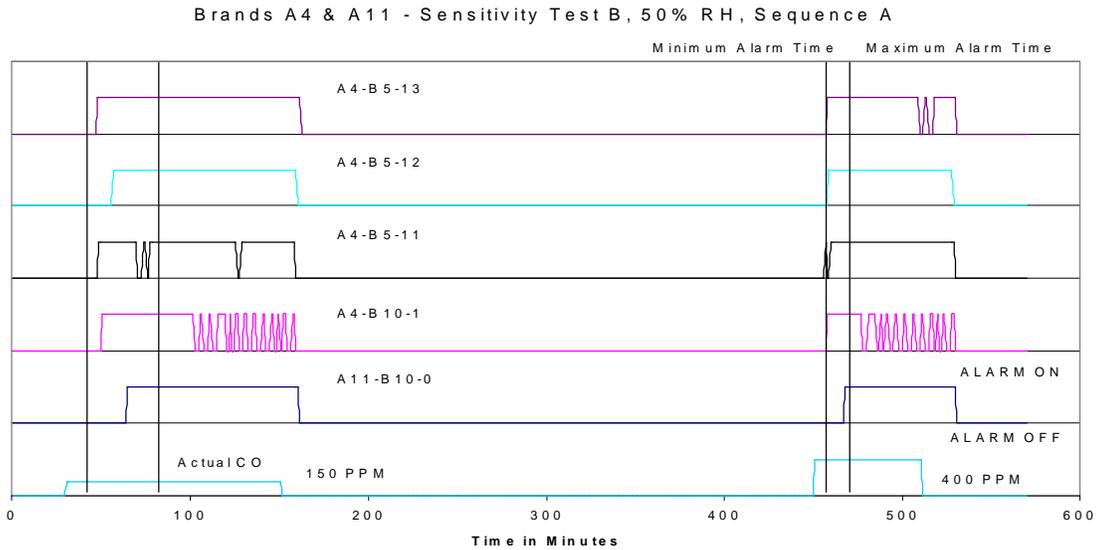


Figure 14 The bottom trace represents the times of presentation of 150 and 400 ppm CO and the two vertical lines represent the times at which 2.5% and 10% COHb would be achieved in a hypothetical exposed person. At 50% RH the units removed from service, A11-B10-0 and A4-B10-1, alarmed within the 2.5% - 10% limits.

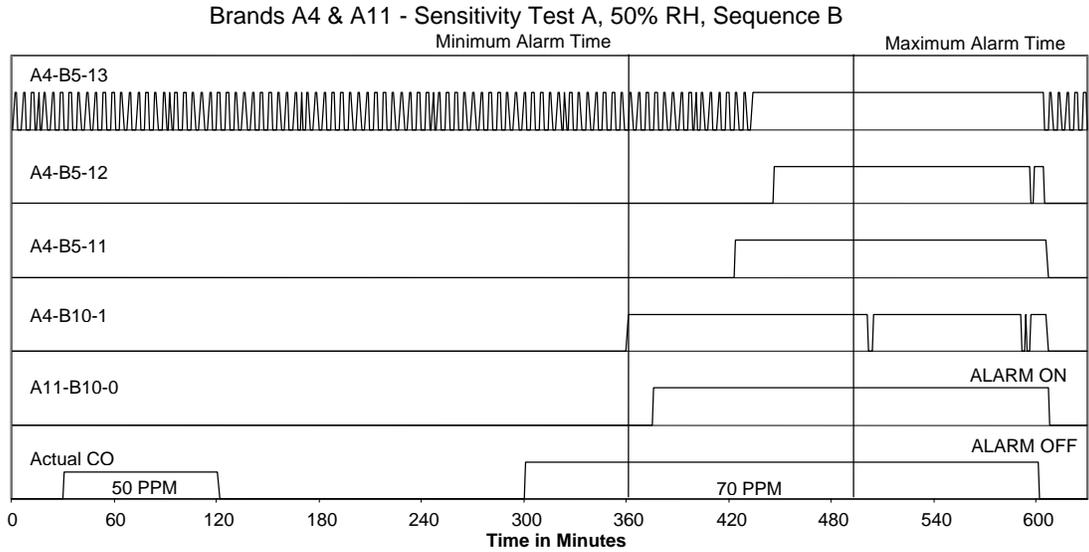


Figure 17 When presented with 70 ppm CO at 50% RH all units responded at times corresponding to between 2.5% and 10% COHb. The bottom trace represents the times of presentation of 50 and 70 ppm CO.

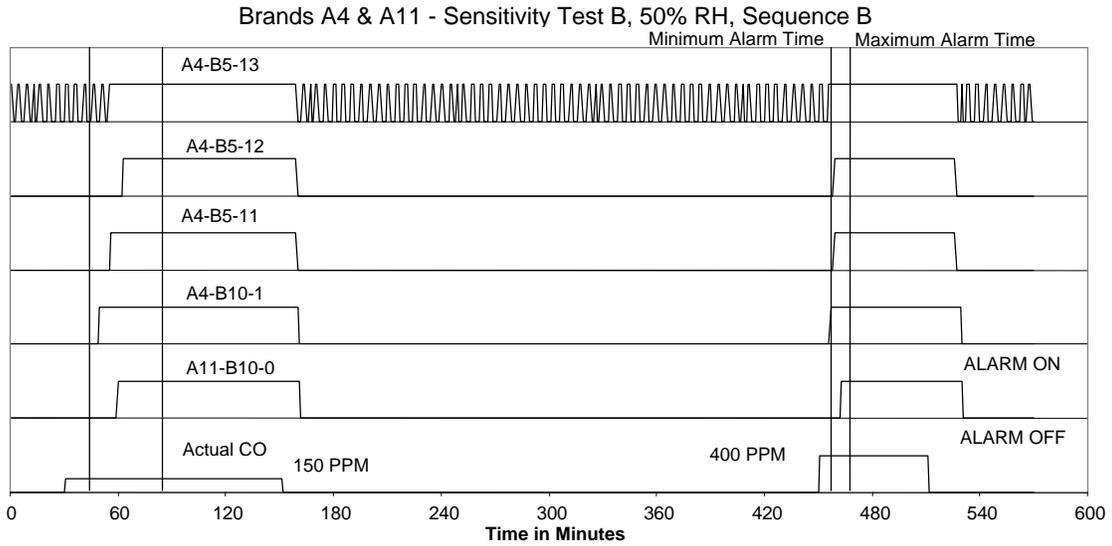


Figure 18 In contrast to their behavior at 5% RH, at 50% RH all units responded at less than 10% COHb.

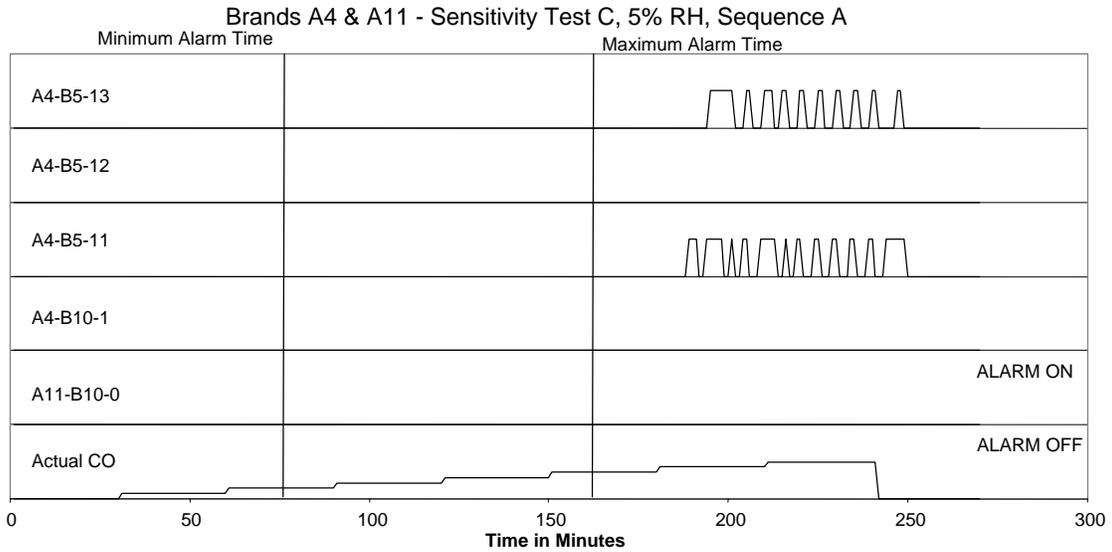


Figure 19 In response to a stepwise increase in CO concentration at 5% RH the units removed from service, A11-B10-0 and A4-B10-1, did not alarm, even at up to approximately 20% RH. Two of the comparison units did alarm, at 12-15% COHb, while two did not. The bottom trace shows concentrations increasing from 25 to 175 ppm in 25 ppm steps.

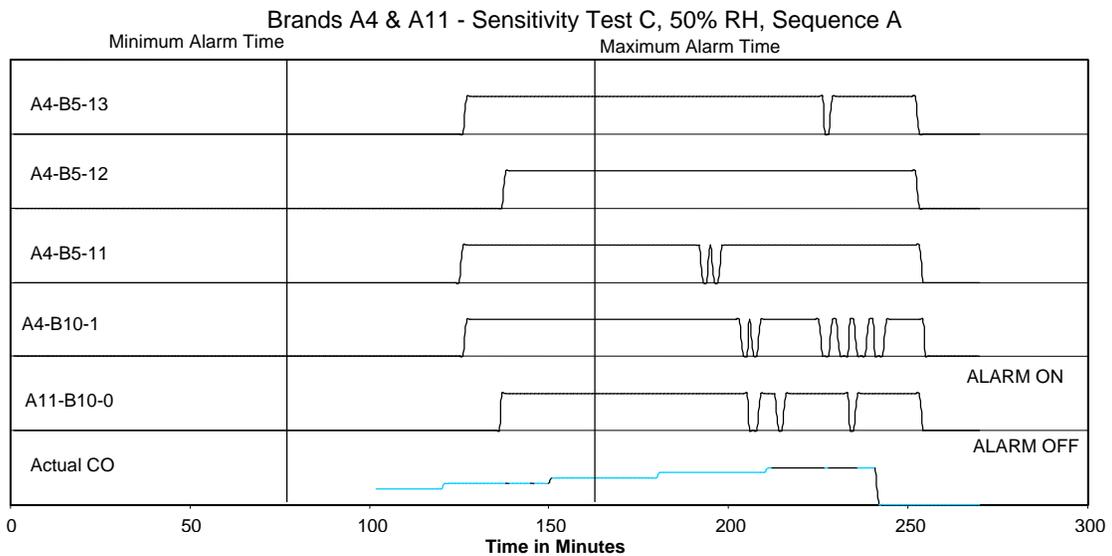


Figure 20 In response to a stepwise increase in CO concentration at 50% RH all did alarm at less than 10% COHB, indicated by the vertical line near 160 minutes. The bottom trace shows concentrations increasing from 25 to 175 ppm in 25 ppm steps.

Brands A4 & A11 - Sensitivity Test C, 5% RH, Sequence A

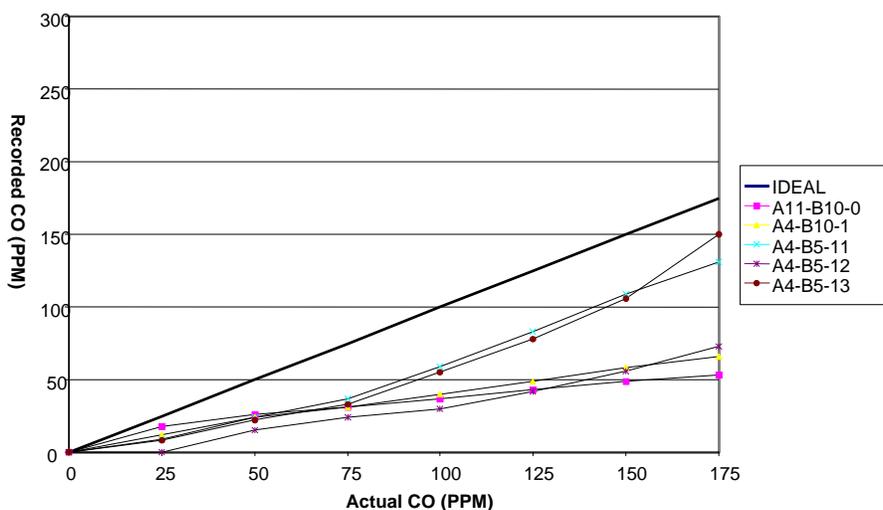


Figure 21 The displayed digital readouts showed the alarms to be insufficiently sensitive at 5% RH, with some having only one third of the required sensitivity.

Brands A4 & A11 - Sensitivity Test C, 50% RH, Sequence A

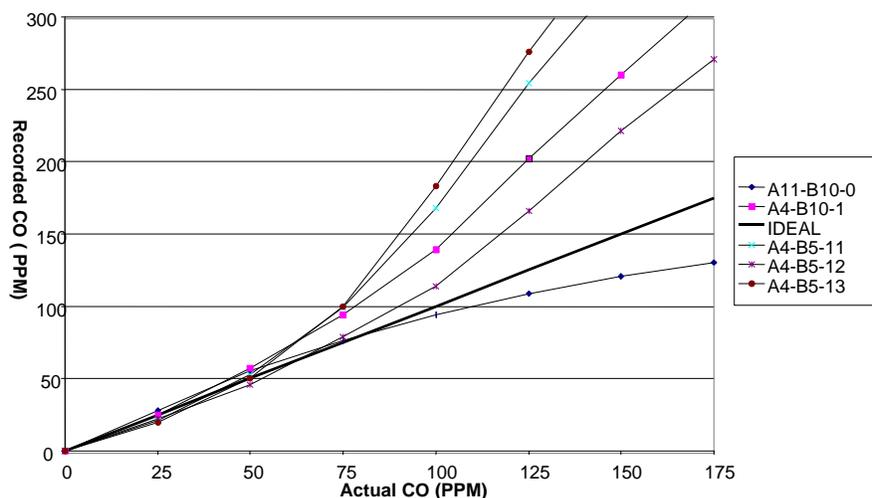


Figure 22 At 50% RH, the digital displays read predominantly high, by more than 100%.

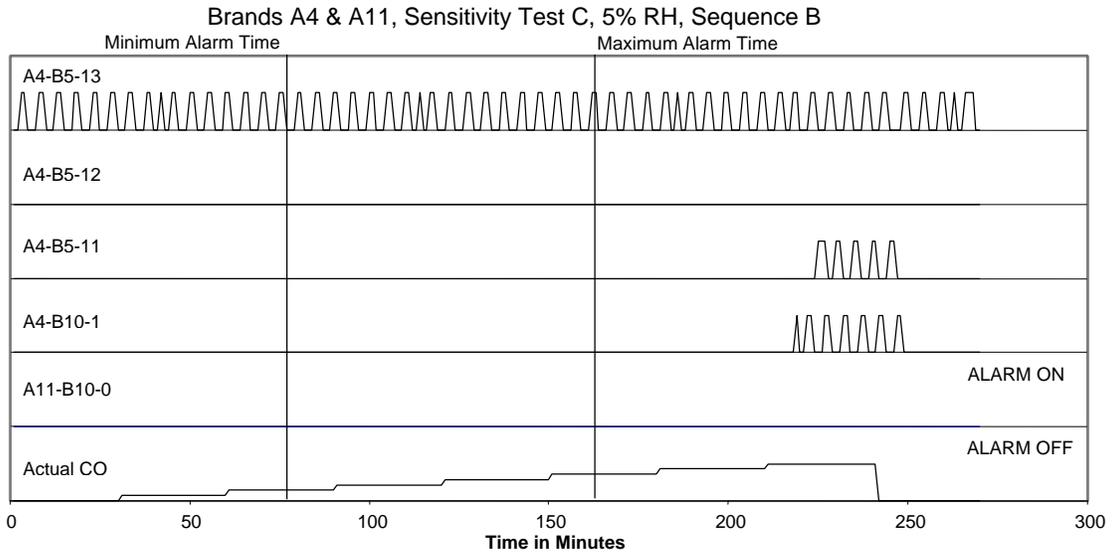


Figure 23 In the second sequence of tests, after a one week exposure to 5% RH, all units were insufficiently sensitive to increasing CO concentration. The prompt effect of low relative humidity is similar, indicating that this insensitivity did not result from the extended low humidity, but was rather an immediate effect of low humidity. The bottom trace shows concentrations increasing from 25 to 175 ppm in 25 ppm steps.

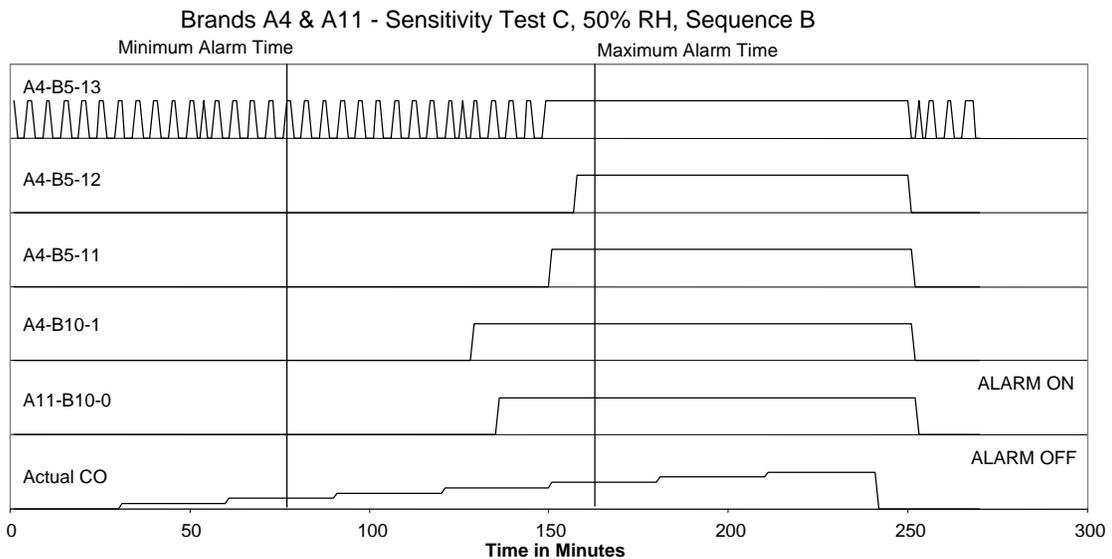


Figure 24 At 50% RH all units were sufficiently sensitive to increasing CO concentrations. The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. One unit exhibited a persistent false alarm.

Brands A4 & A11 - Sensitivity Test C, 5% RH, Sequence B

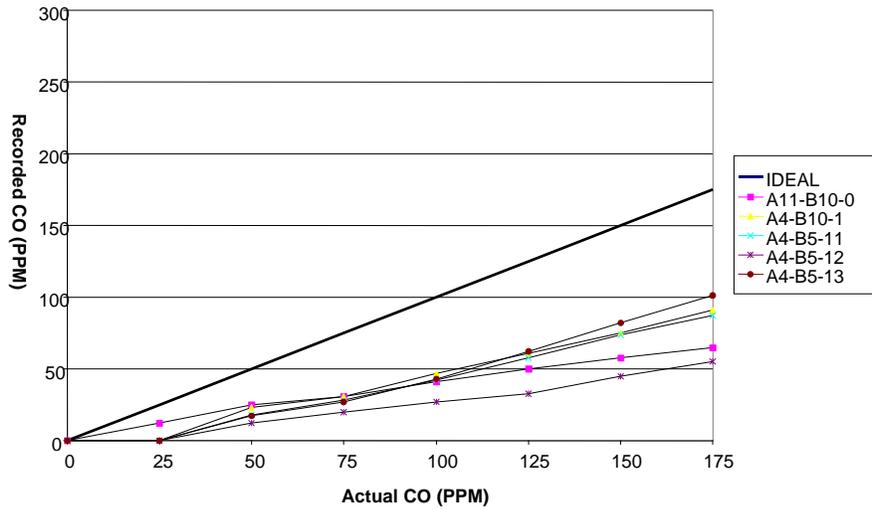


Figure 25 Digital display readings taken during Sensitivity Test C are compared to the ideal. The displayed digital readouts showed the alarms to be insufficiently sensitive at 5% RH, with a sensitivity less than one third of that required for adequate detection.

Brands A4 & A11 - Sensitivity Test C, 50% RH, Sequence B

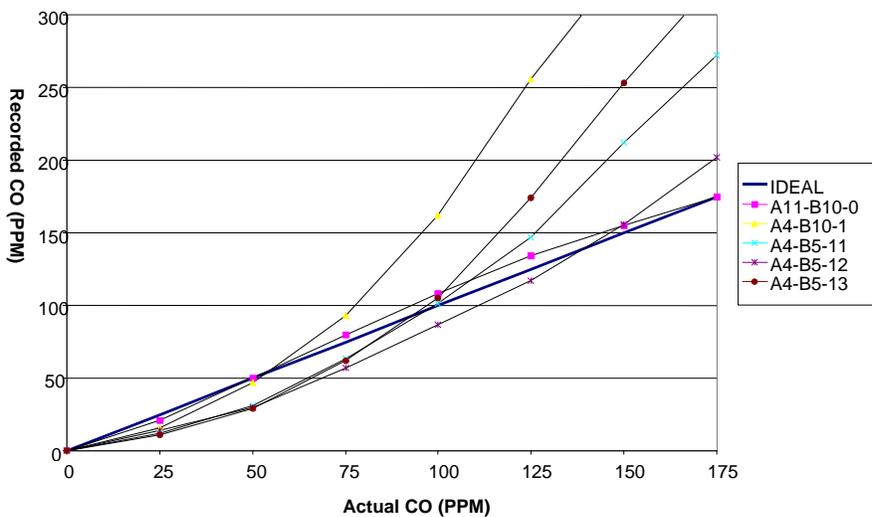


Figure 26 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. At 50% RH the alarms are overly sensitive, some by more than 100%.

Brands A4 & A11 - Sensitivity Test E

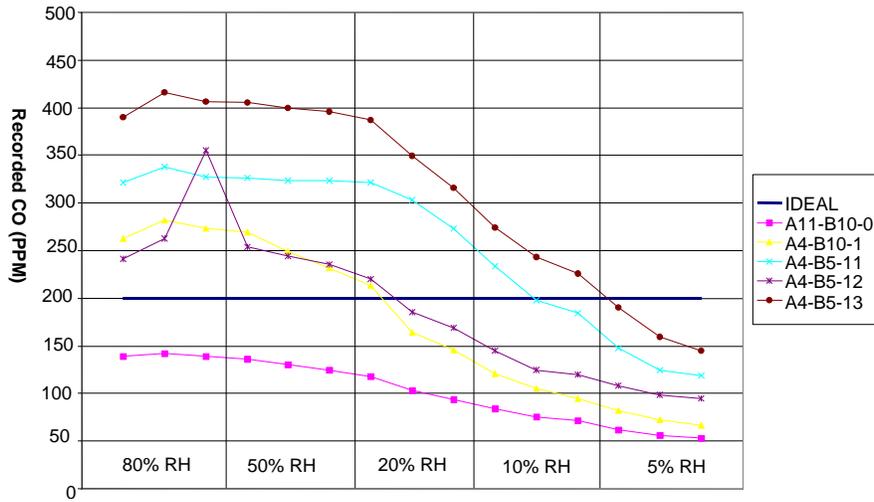


Figure 27 After conditioning at 50% RH alarm digital display readings in response to 200 ppm CO were measured at relative humidity of 80%, 50%, 20%, 10% and 5%. Each relative humidity was presented for 90 minutes and displays were read every 30 minutes. The alarm sensitivity is strongly dependent on relative humidity, with readings varying over a three to one range.

Brands A4 & A11 - Sensitivity Test F

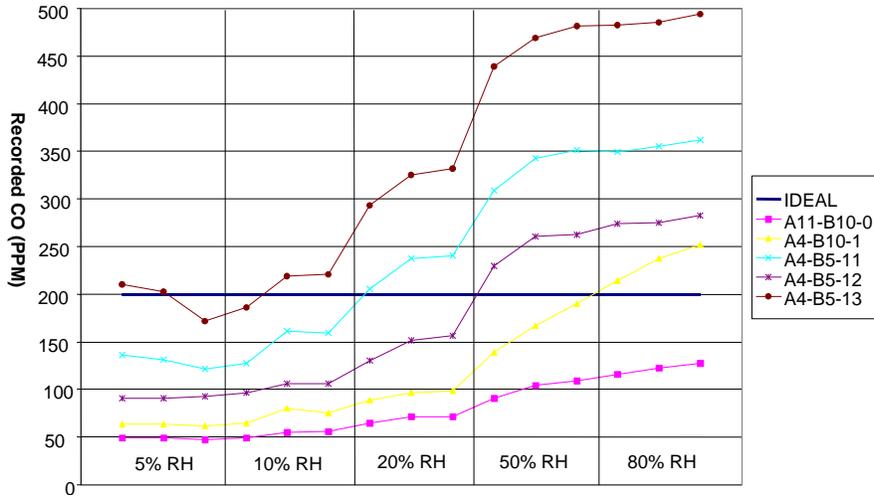


Figure 28 After conditioning at 5% RH alarm digital display readings in response to 200 ppm CO were measured at relative humidity of 5%, 10%, 20%, 50% and 80%. Each relative humidity was presented for 90 minutes and displays were read every 30 minutes. The alarm sensitivity is strongly dependent on relative humidity, with readings varying over a three to one range.

Brand A12

At the request of the manufacturer, Brand A12 was tested to the sensitivity requirements of the CENELEC (European Union) Standard. Tests of Brand A12 exhibited the following notable response characteristics:

- The CENELEC Standard requires a rigorous recovery test – alarms must sound in response to 55 ppm CO just a short time after a 5000 ppm exposure. These alarms did not meet this test, as shown in the following figures.
- Alarms responded within their sensitivity specification at 55 ,110, 330, and 5000 ppm.
- Devices of this brand showed no significant dependence of CO sensitivity on humidity.
- Alarm activations during the interference test (Figure 31) likely resulted from filter breakthrough resulting from the long exposure to the solvent gases. Subsequent tests used a recovery period comprising two hours of exposure to clean air.

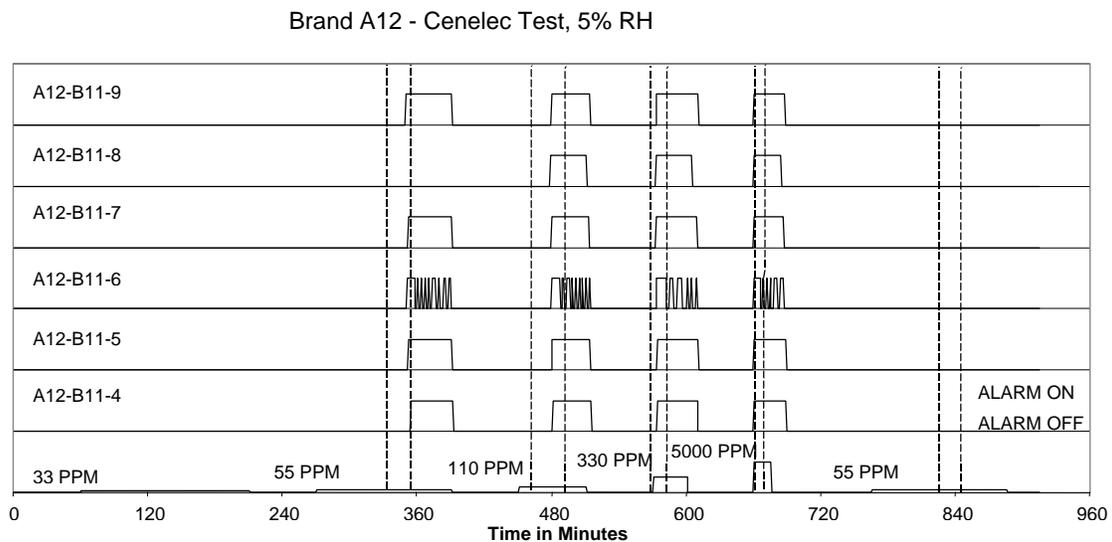


Figure 29 Response to CENELEC alarm concentrations of 55, 110, 330, and 5000 ppm at 5% RH. One of six alarms failed to alarm at 55 ppm, and they all did not meet the CENELEC recovery test – they failed to alarm at 55 ppm shortly after the 5000 ppm exposure.

Brand A12 - Cenelec Test, 50% RH

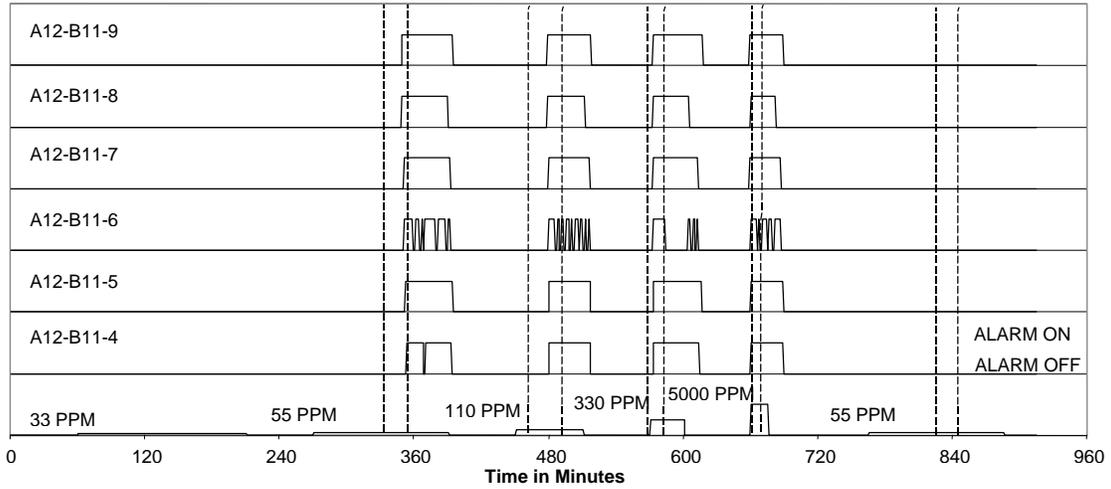


Figure 30 Response to CENELEC alarm concentrations of 55, 110, 330, and 5000 ppm at 50% RH. All alarms sounded reliably from 55 ppm to 5000 ppm, but they did not meet the CENELEC recovery test – they failed to sound at 55 ppm shortly after the 5000 ppm exposure.

Brand A12 - Interference Test A, 50% RH

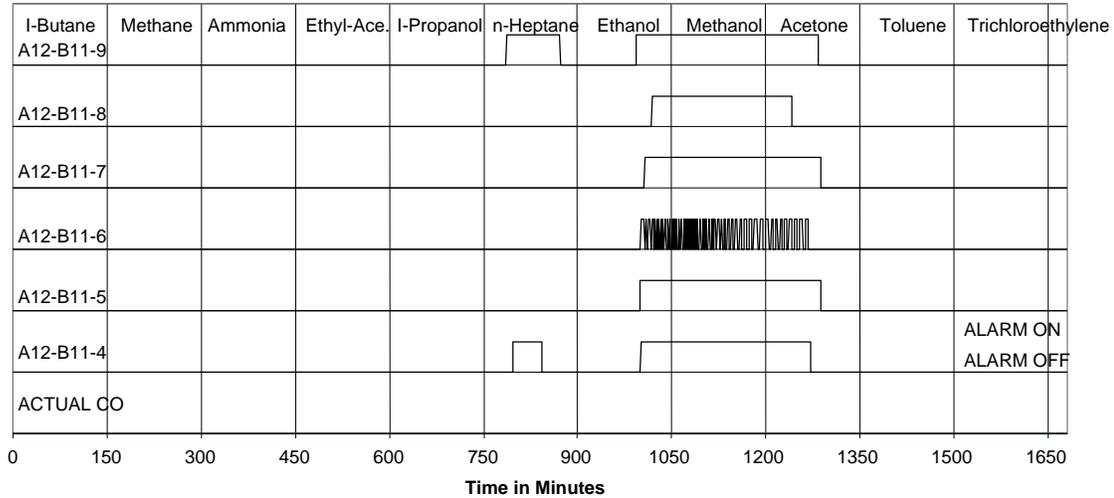


Figure 31 Response to 2 hour exposures to interference gases. It appears that there is some sensitivity to n-heptane and that filter breakthrough occurred during the alcohol exposures.

Brand A13

Aside from mediocre, but servicable, digital display accuracy and the failure of one LCD display, devices of Brand A13 performed well. Tests of Brand A13 exhibited the following notable response characteristics:

- Alarms sounded well within the sensitivity requirements of UL 2034 for fixed CO concentrations, and within the 2.5% and 10% COHb bounds for steadily increasing CO concentration.
- Devices of this brand showed no significant dependence of CO sensitivity on humidity.
- Devices showed no sensitivity to each of the interference gases tested.
- Digital displays were adequately accurate, reading between 100 and 140 ppm when presented with 100 ppm CO.
- The digital display of one of ten alarms failed by flickering intermittently.

There were no hardware failures other than a single intermittent LCD display.

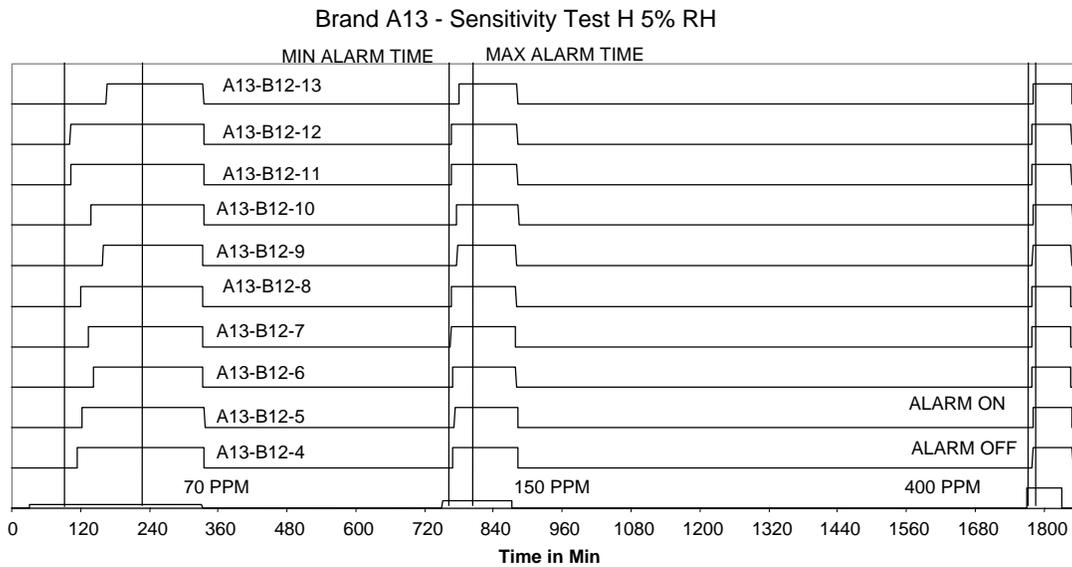


Figure 32 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms sounded within the 2.5 - 10% alarm limits, even at low relative humidity.

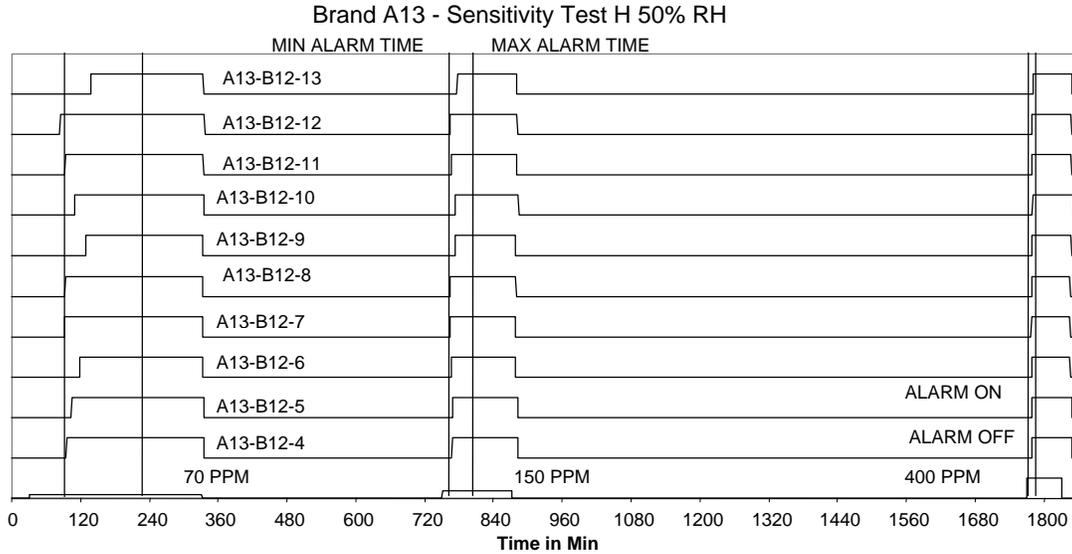


Figure 33 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms sounded within the 2.5 - 10% alarm limits.

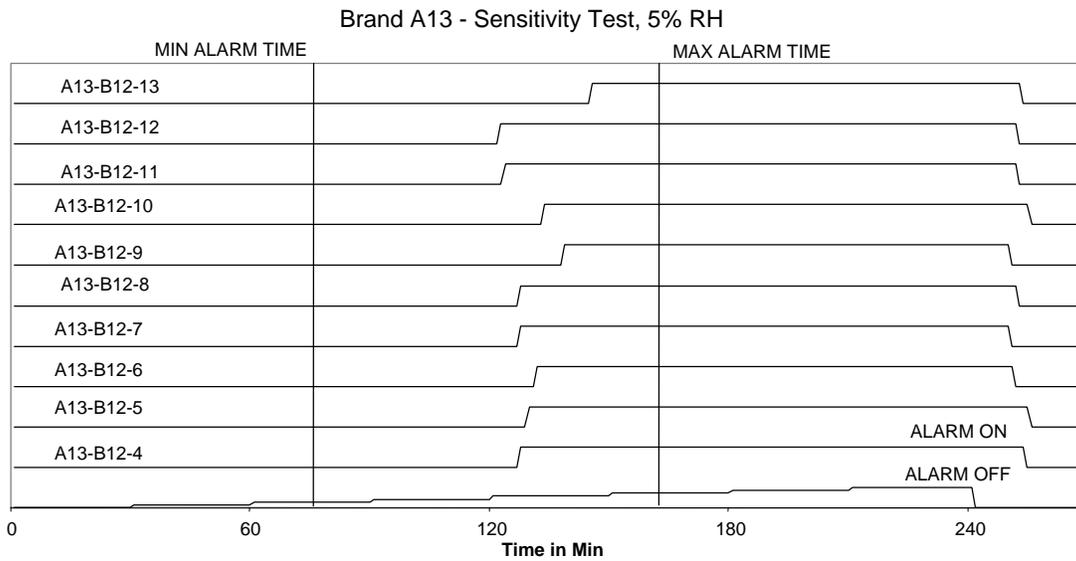


Figure 34 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms sounded well before 10% COHb was attained.

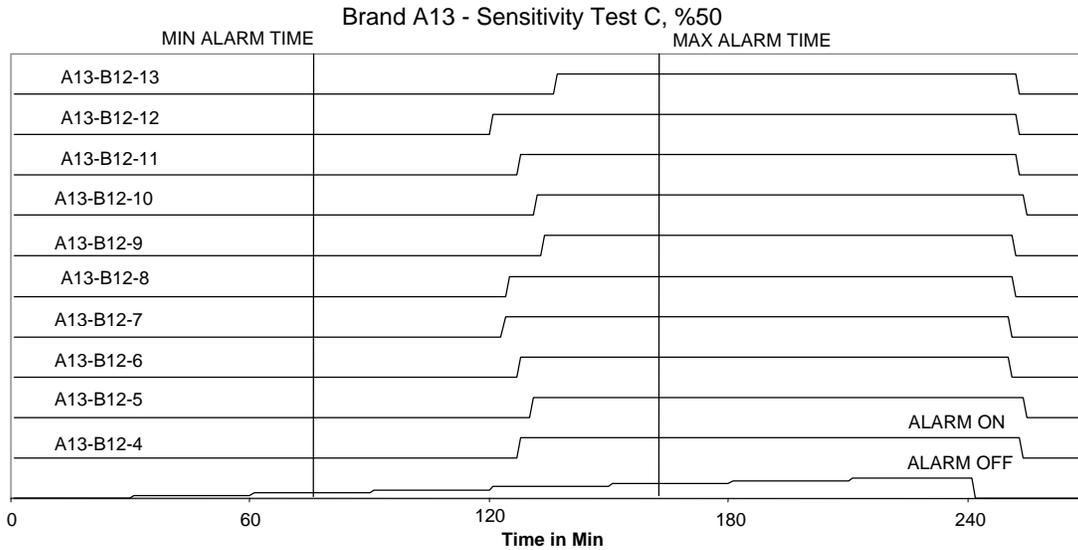


Figure 35 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb were attained. The peak COHb attained was approximately 20% COHb. All alarms sounded well before 10% COHb is attained.

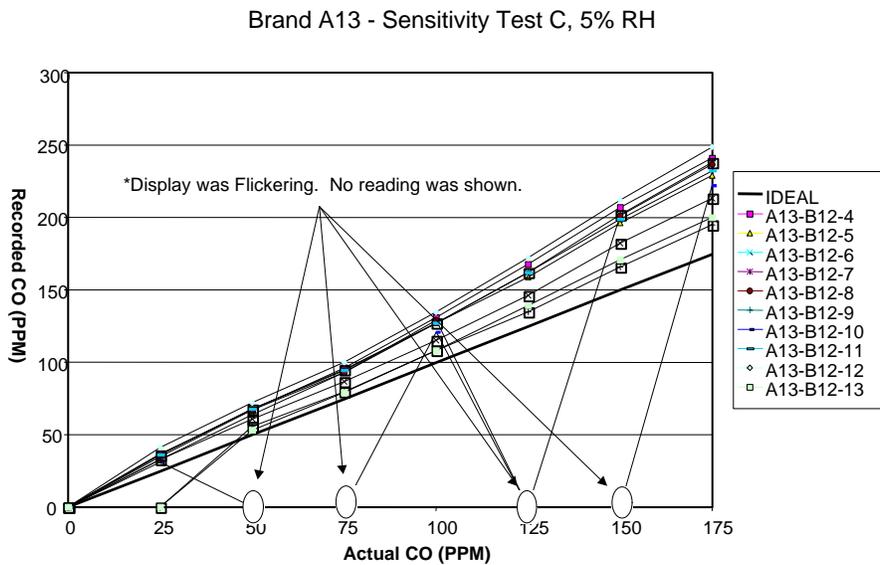


Figure 36 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. Some alarms read high by up to 40%. One alarm's display flickered intermittently.

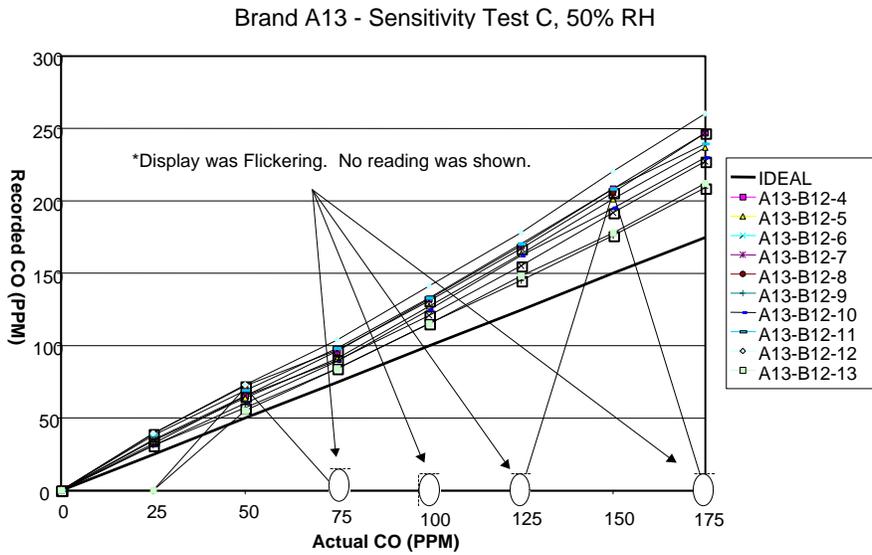


Figure 37 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. Some alarms read high by up to 40%. One alarm's display flickered intermittently.

Brand A13 - Interference Test A, 50% RH

	I-Butane	Methane	Ammonia	Ethyl-Ace.	I-Propanol	n-Heptane	Ethanol	Methanol	Acetone	Toluene	Trichloroethylene
A13-B12-13											
A13-B12-12											
A13-B12-11											
A13-B12-10											
A13-B12-9											
A13-B12-8											
A13-B12-7											
A13-B12-6											
A13-B12-5											
A13-B12-4											

Time in Minutes

Figure 38 Brand A13 showed no sensitivity to any of the interference gases tested.

Brand A14

The performance of alarms of Brand A14 was representative of electrochemical cell based devices using counting/timing algorithms.

- At fixed CO concentrations, alarm activations were repeatable, and occurred well within the sensitivity requirements of UL 2034. However for a steadily increasing CO concentration activation times were more variable, with one of seven alarms failing to sound at less than 10% COHb.
- Typical of electrochemical cell alarms, this brand showed no significant dependence of CO sensitivity on humidity.
- Alarms activated in response to several interference gases. These activations likely resulted from filter breakthrough after becoming loaded with alcohol vapors (see Figure 43).
- After removal of the test gases during the fixed concentration tests alarms uniformly exhibited spurious activations. We assume these activations resulted from logical errors in resetting counters or timers in the instruments' detection algorithm. Even though under the conditions of our tests these spurious activations only occurred after exposure to CO gas, their occurrence raises questions about the validity and reliability of the instrument's embedded software.

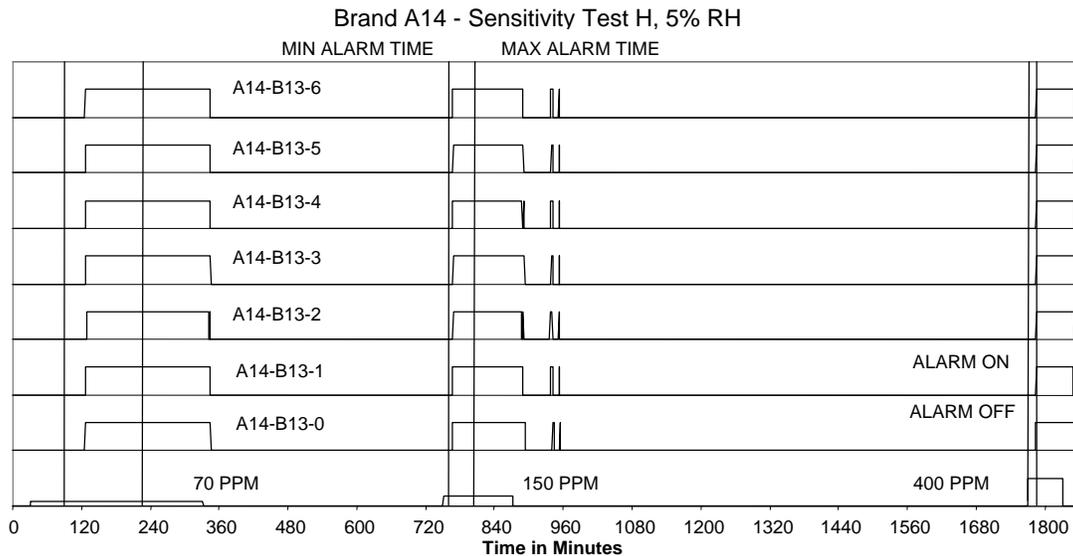


Figure 39 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. The spurious alarm activations after the 150 ppm presentation is removed likely result from an error in the alarm's timer/counter algorithm.

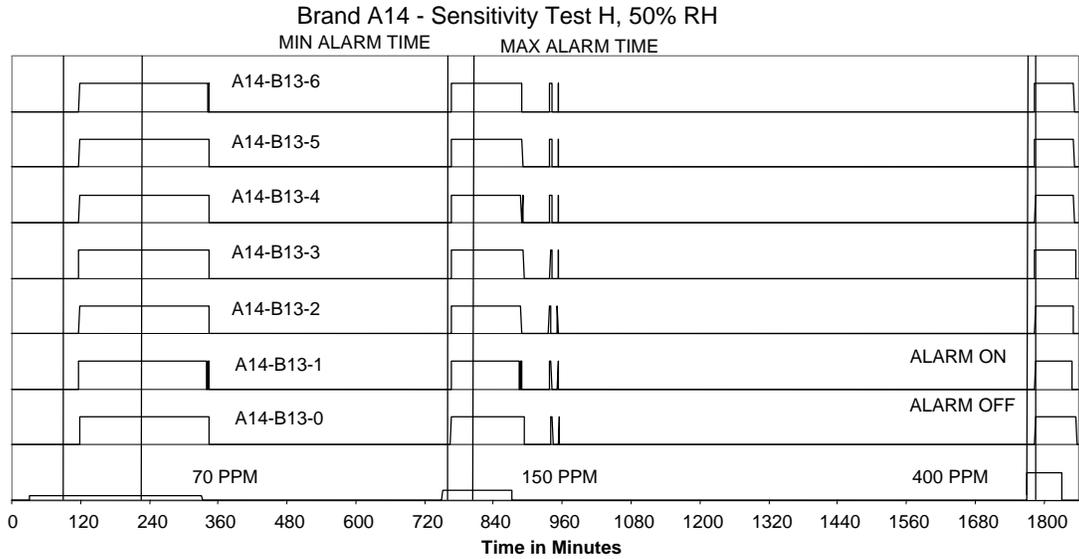


Figure 40 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. The spurious alarm activations after the 150 ppm presentation is removed likely result from an error in the alarm's timer/counter algorithm.

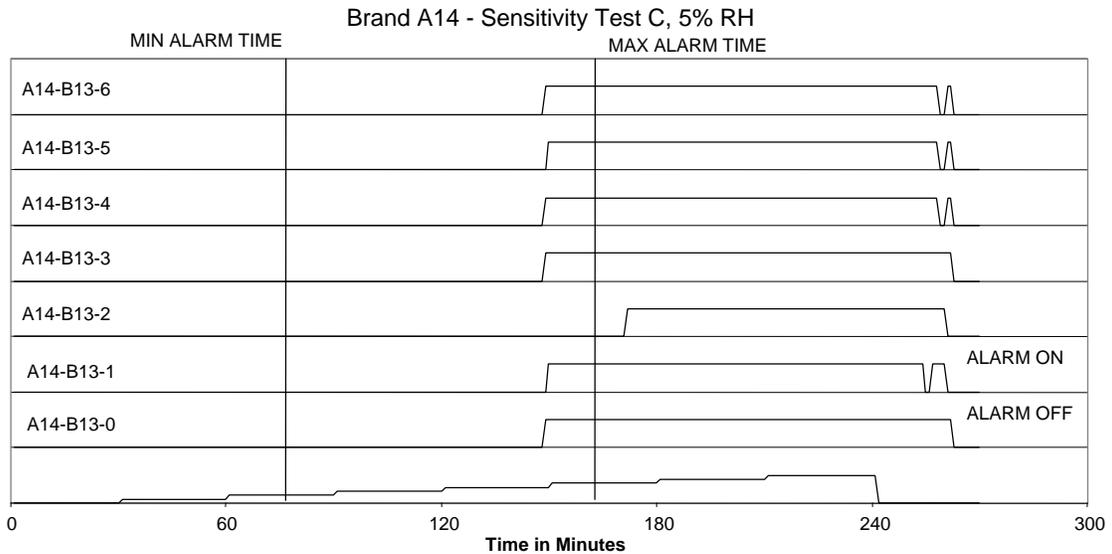


Figure 41 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. Although alarm A14-B13-2 activated readily on exposure to fixed CO concentrations, it activated late on exposure to a steadily increasing concentration. This behavior seems typical of alarms that do not integrate concentration.

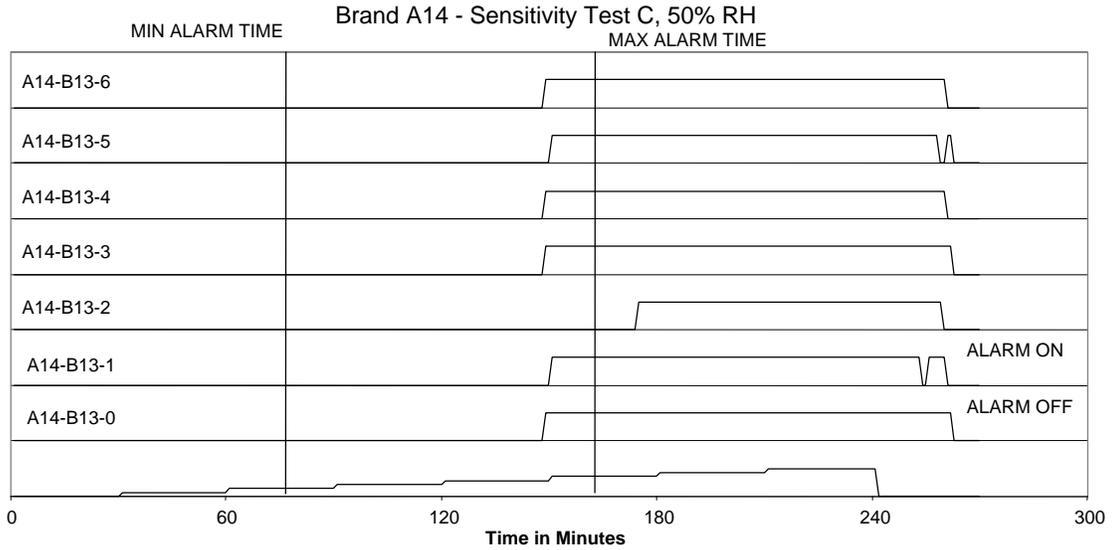


Figure 42 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. Again, alarm A14-B13-2 activated late on exposure to a steadily increasing concentration. This behavior seems typical of alarms that do not integrate concentration.

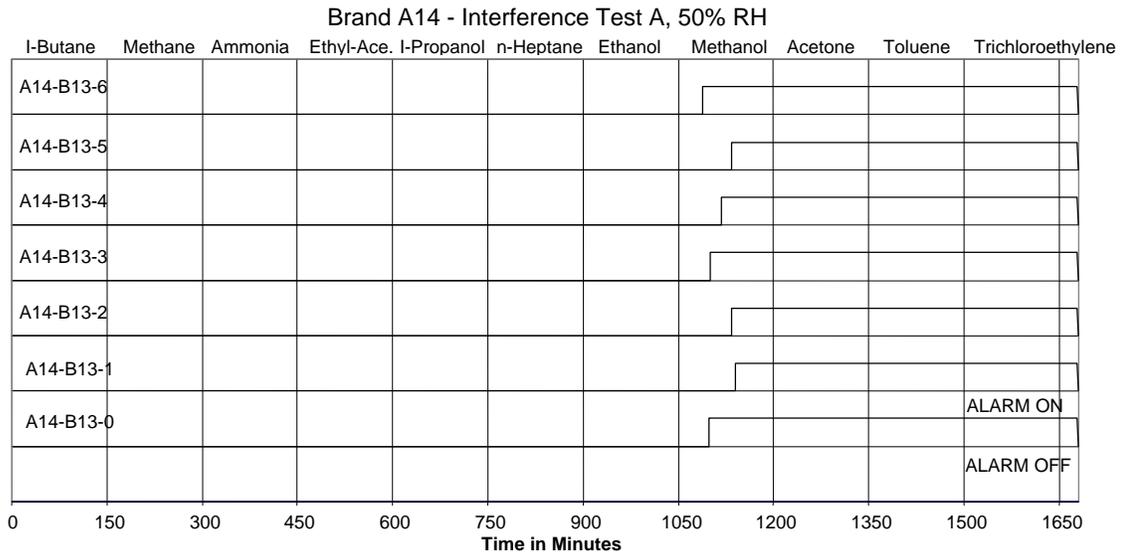


Figure 43 The alarms were immune to the first six interference gases tested but it appears that filter breakthrough occurred during the second alcohol exposure, and persisted throughout the remaining gas presentations.

Brand A15

Devices of Brand A15 had highly variable sensitivity. They predominantly alarmed prematurely at 150 and 400 ppm CO while often failing to respond to 70 ppm. While they showed no consistent dependence on humidity, during Sensitivity Test C one third of the alarms failed to sound in time at only the greater humidity (50% RH). Their response showed the following features:

- Although these alarms were not sensitive to relative humidity, at the greater of the two humidities tested (50% RH) all devices failed to sound at 70 ppm CO.
- In response to fixed CO concentrations alarms predominantly activated early.
- Alarm activation times were highly variable during tests with steadily increasing concentration, and several alarms failed to activate at less than 10% COHb. This high variability in alarm times coupled with the tighter variance when tested at fixed concentrations indicates the use of a counter/timer algorithm that inadequately integrates CO exposure.
- Digital displays were highly inaccurate, with three of nine alarms failing to indicate any concentration when exposed to 100 ppm at 50% RH, while other displays indicated up to approximately 400 ppm (see Figure 48 and Figure 49).
- The alarms were sensitive to many of the interference gases tested. The pattern of response was not typical of a filter whose capacity had been overwhelmed, as had been seen with other brands, but rather of a filter that either inadequately captured particular gases, or was short circuited by unintended flow paths.

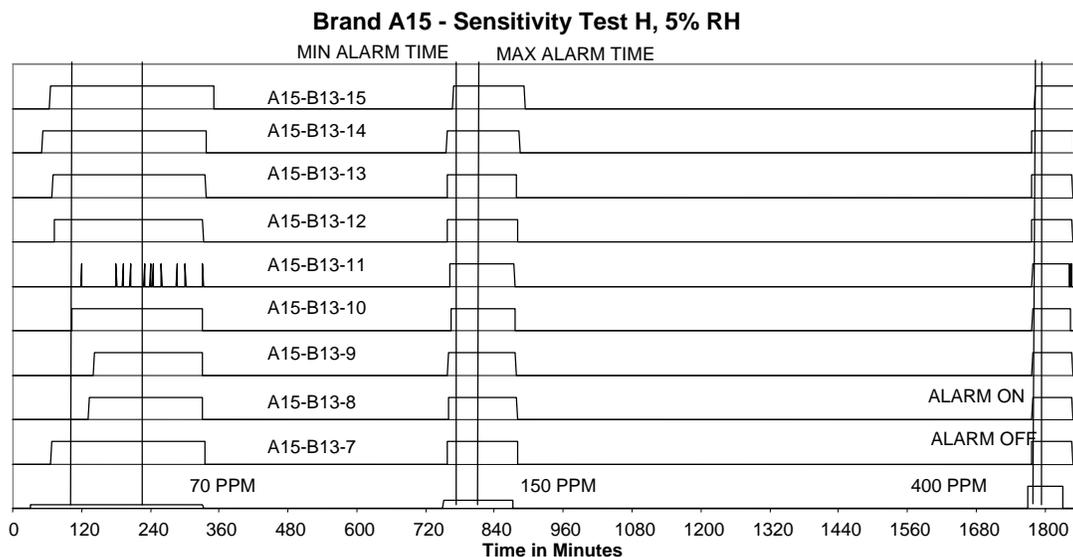


Figure 44 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. These alarms are predominantly oversensitive, even at 5% RH, often activating at less than 2.5% COHb. Alarm A15-B13-11 activated only intermittently at 70 ppm.

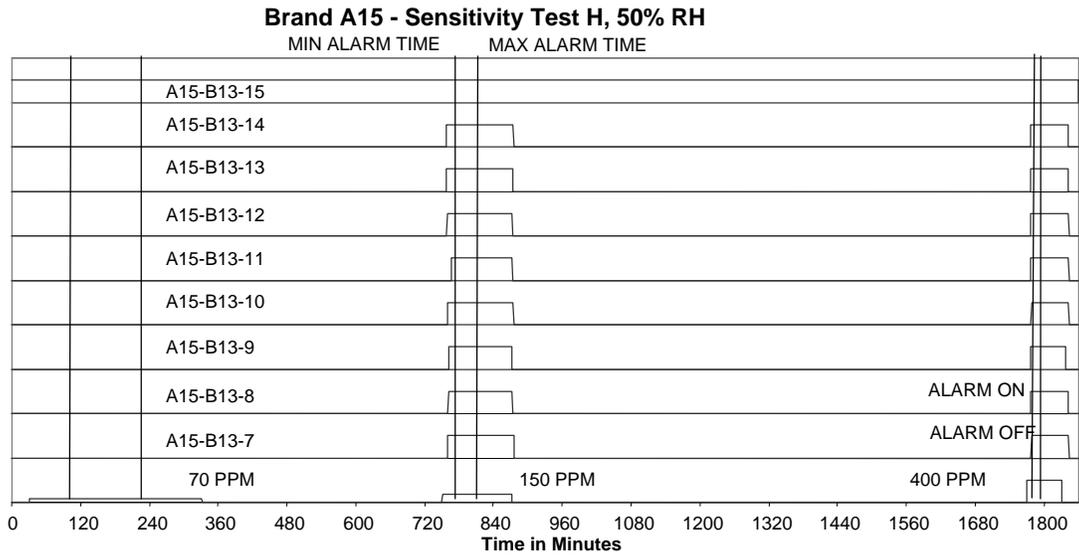


Figure 45 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. Device A15-B13-15 (the top trace) went into continuous alarm, and the others failed to activate at 70 ppm.

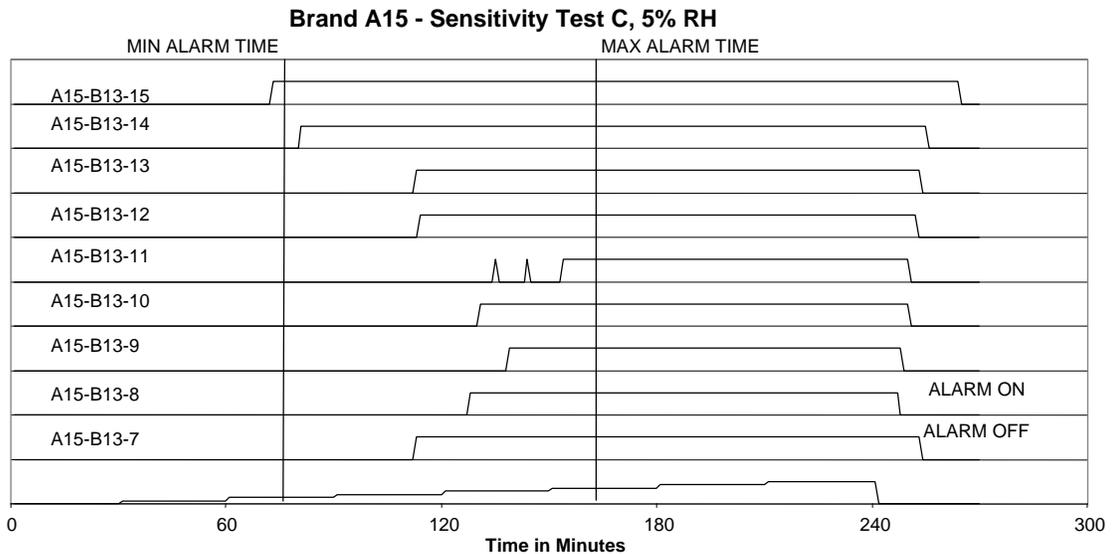


Figure 46 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. As it did in other tests, device A15-B13-11 activated intermittently before solidly alarming.

Brand A15 - Sensitivity Test C, 50% RH

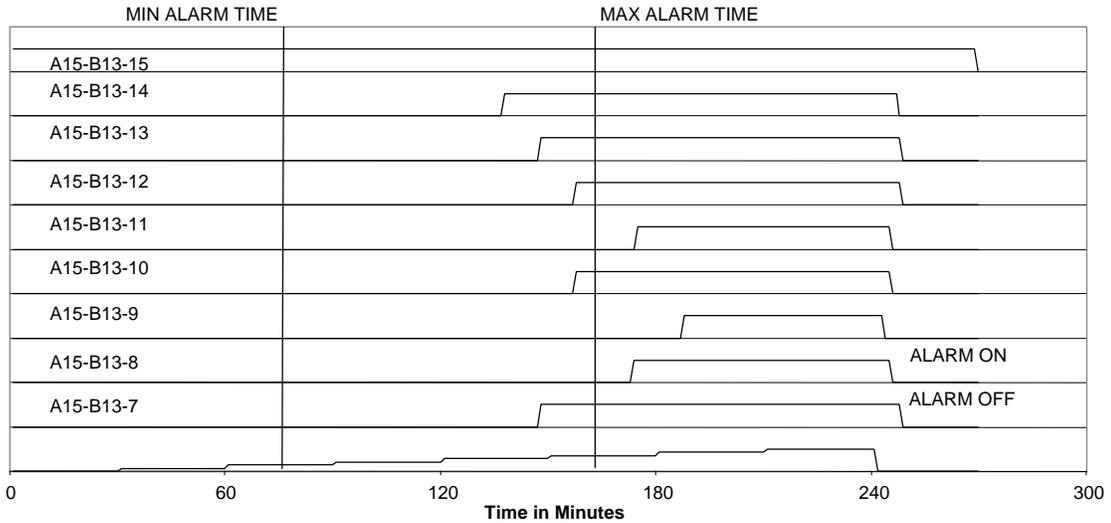


Figure 47 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb.

Brand A15 - Sensitivity Test C, 5% RH

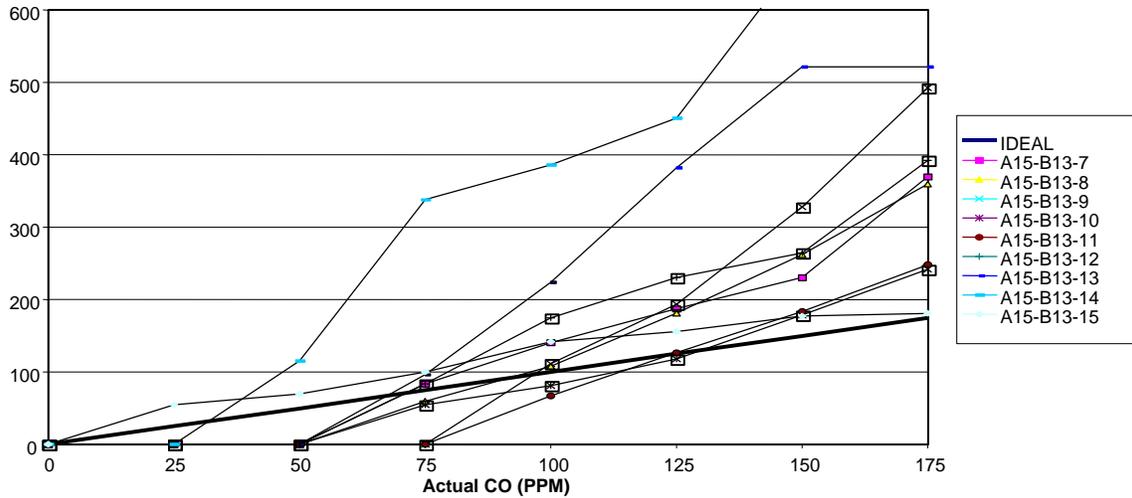


Figure 48 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. Sensitivity varies tremendously from unit to unit, with 75 ppm reported as 0 to 250 ppm.

Brand A15 - Sensitivity Test C, 50% RH

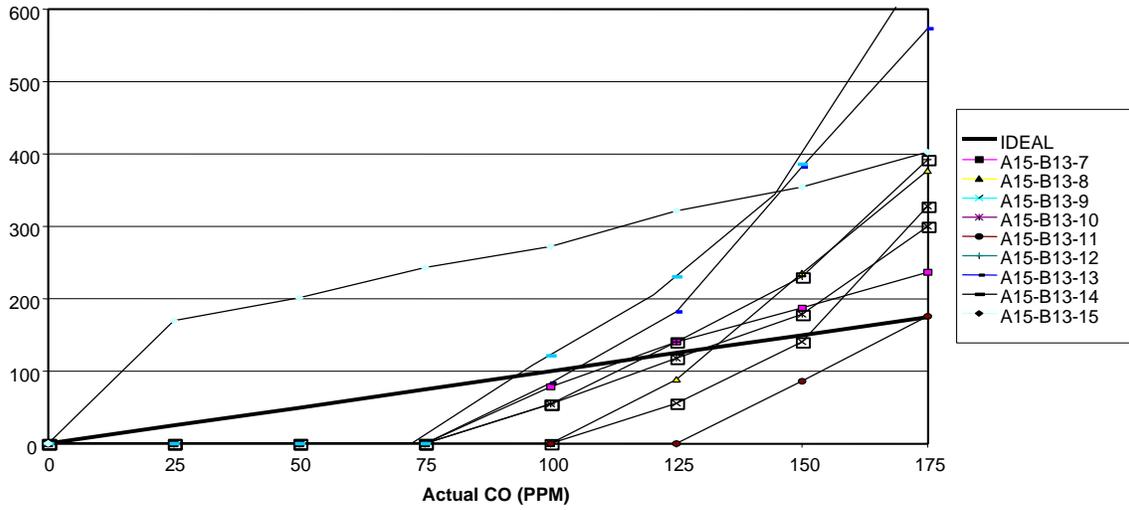


Figure 49 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. The sensitivities of these units appear to be less at 50% RH than at 5% RH, although the observed variation may result from day to day drift.

Brand A15 - Interference Test A, 50% RH

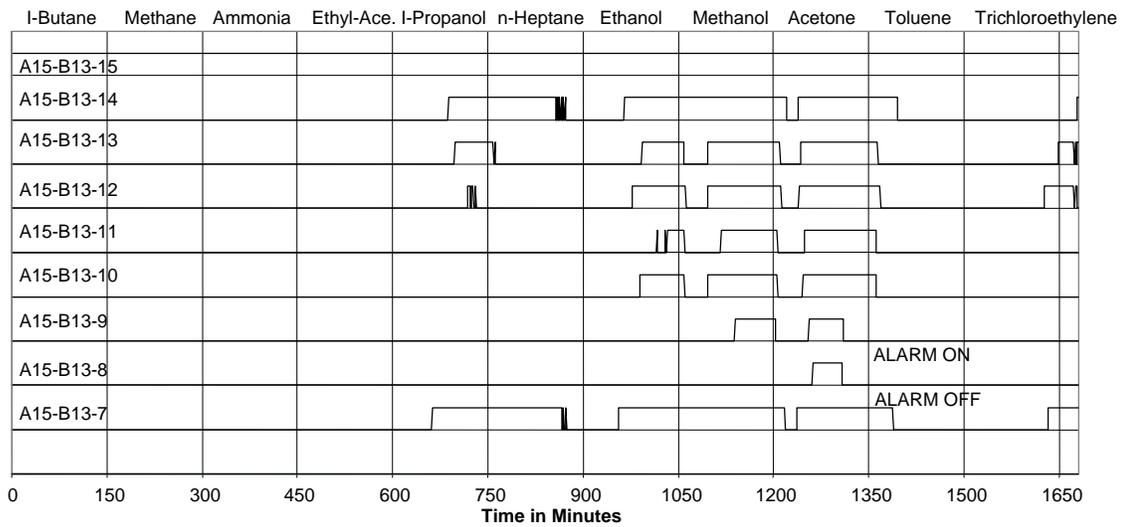


Figure 50 These alarms were variously sensitive to many of the interference gases. In particular the alcohols either broke through the filter, or it was inadequately sized.

Brand A16

Devices of this model under development had varied performance. Two of the five devices tested progressively lost sensitivity during testing, and the other alarms failed to sound in time at 400 ppm CO and 50% RH. During interference tests several alarms activated in response to i-propanol and ethyl acetate, probably as a result of filter breakthrough. Subsequent to these tests, the manufacture identified the causes of these problems.

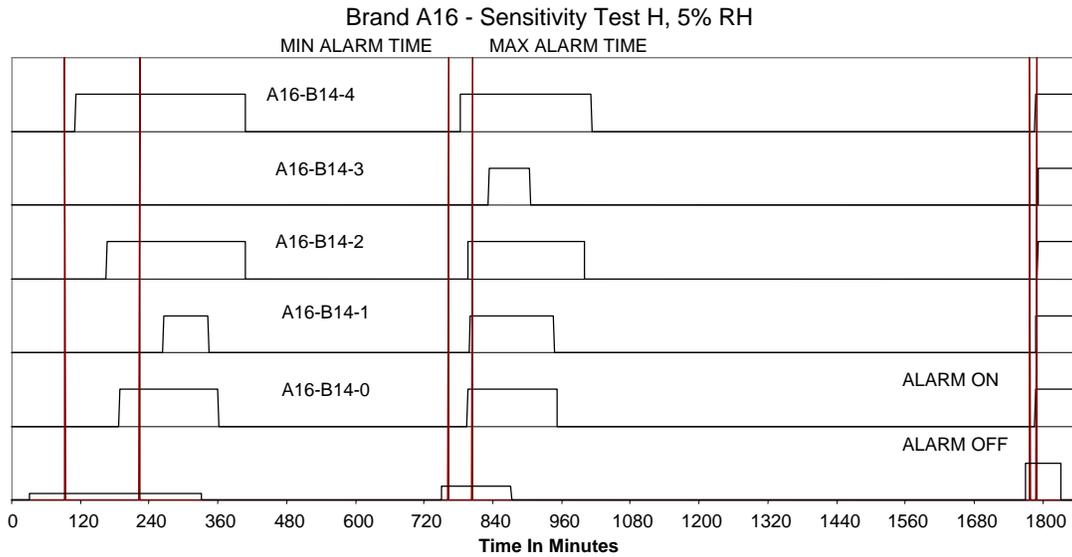


Figure 51 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained.

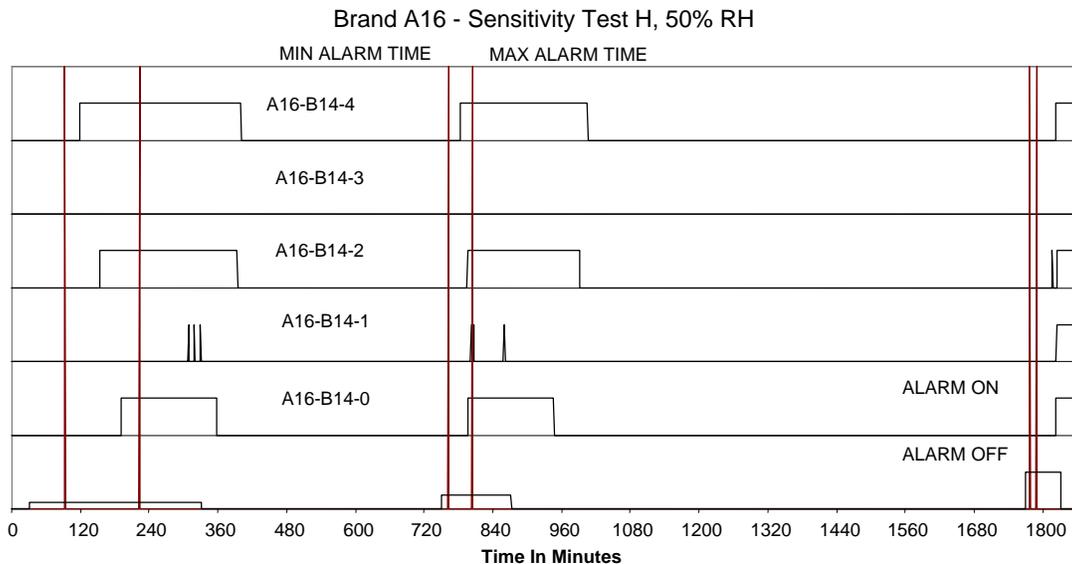


Figure 52 Two of the five alarms were inadequately sensitive at 70 and 150 ppm and all alarmed late at 400 ppm. These late alarms would have resulted in exposures to approximately 30% COHb.

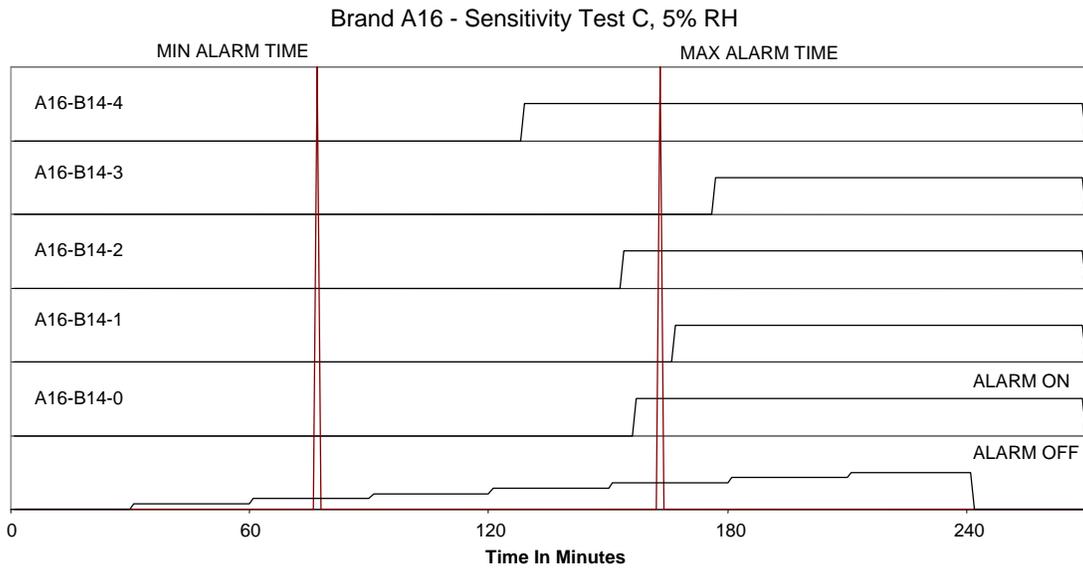


Figure 53 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. At 5% RH two of five devices alarmed late.

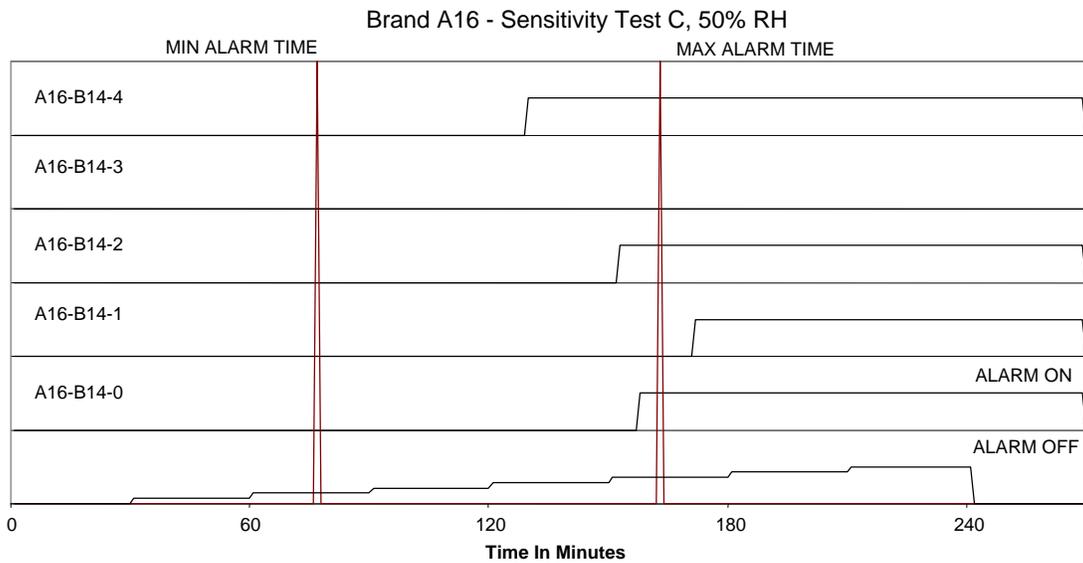


Figure 54 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. At 50% RH one device alarmed late, and one not at all.

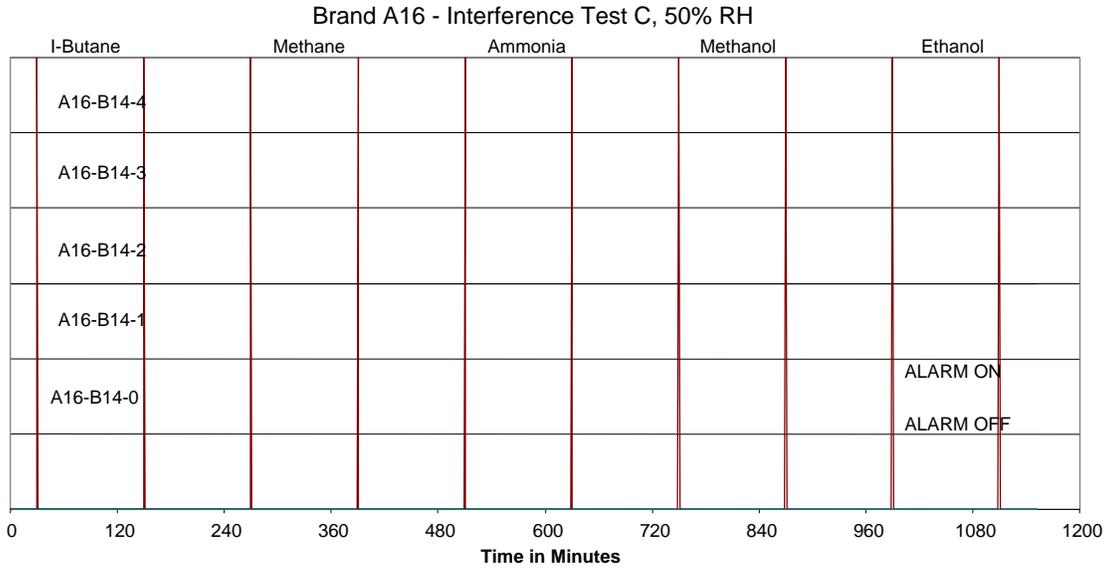


Figure 55 No brand A16 alarms were sensitive to these interference gases.

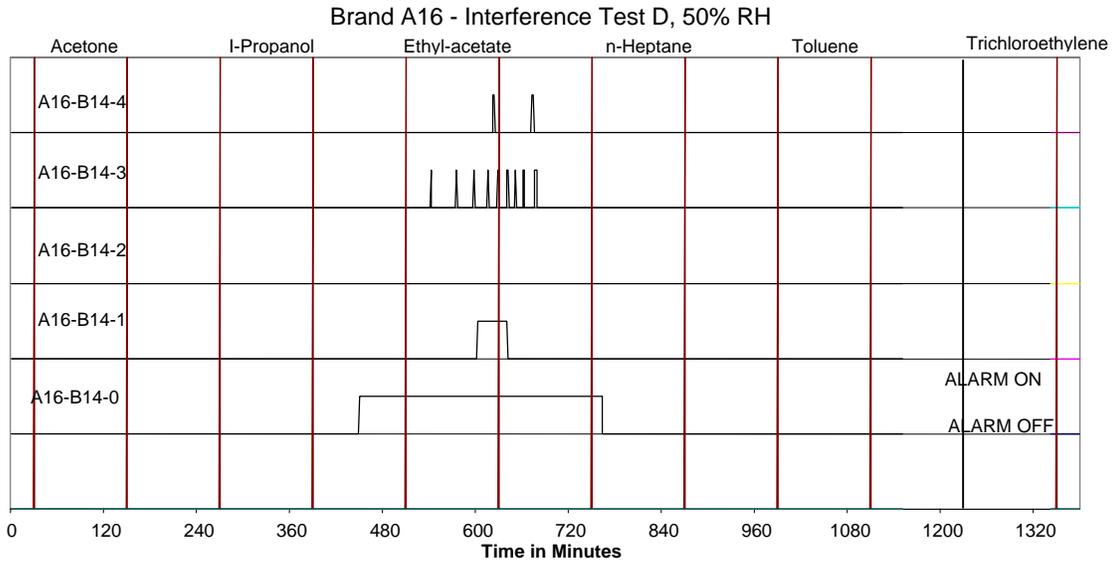


Figure 56 Several of the brand A16 alarms were sensitive to isopropanol or ethyl acetate.

Brand A17

Alarms of Brand A17 exhibited the following characteristics:

- Several spurious activations of two devices lasting several minutes occurred at the onset of 70 ppm CO. We have found such spurious activations to occur only in alarms that use counter-timer algorithms. These activations most likely result from improper variable initialization in the alarms' software that incorrectly interprets the onset of CO concentrations and does not reset internal timers properly – in short, software bugs. These spurious activations are generally seen in response to transient CO (or other gas) concentrations, and also at fixed intervals after exposure. For example, similar post-exposure spurious activations are evident in Figures 39 and 40. This problem is further discussed in “Analysis and Recommendations”.
- The alarms showed no detectable change in sensitivity to CO as humidity was varied from 5% to 50% RH.
- While there was no consistent influence of relative humidity, all alarms failed to sound in time when exposed to 400 ppm only at the greater humidity, not at the lesser humidity.
- Even though devices sounded well within the time limits when exposed to 70 and 150 ppm CO, on exposure to progressively increasing CO their sensitivity was borderline, with several alarming just under 10% COHb and several alarming just over 10% COHb. This behavior indicates an inadequately performing integration algorithm, or the use of a counter/timer algorithm.
- Devices were completely insensitive to all interference gases tested.

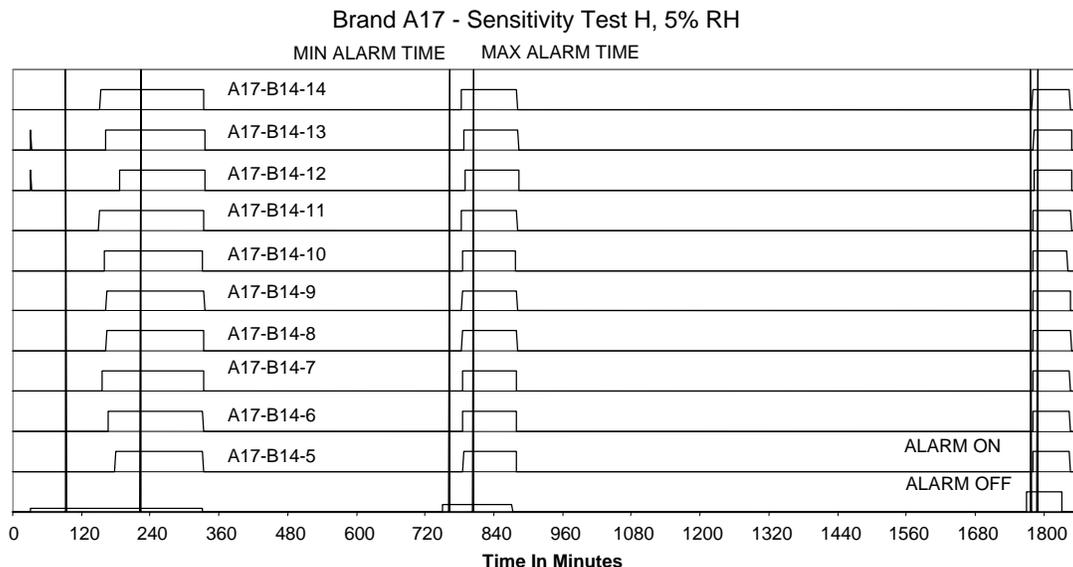


Figure 57 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms sounded within specification at 5% RH, but two sporadically activated at the outset of the experiment.

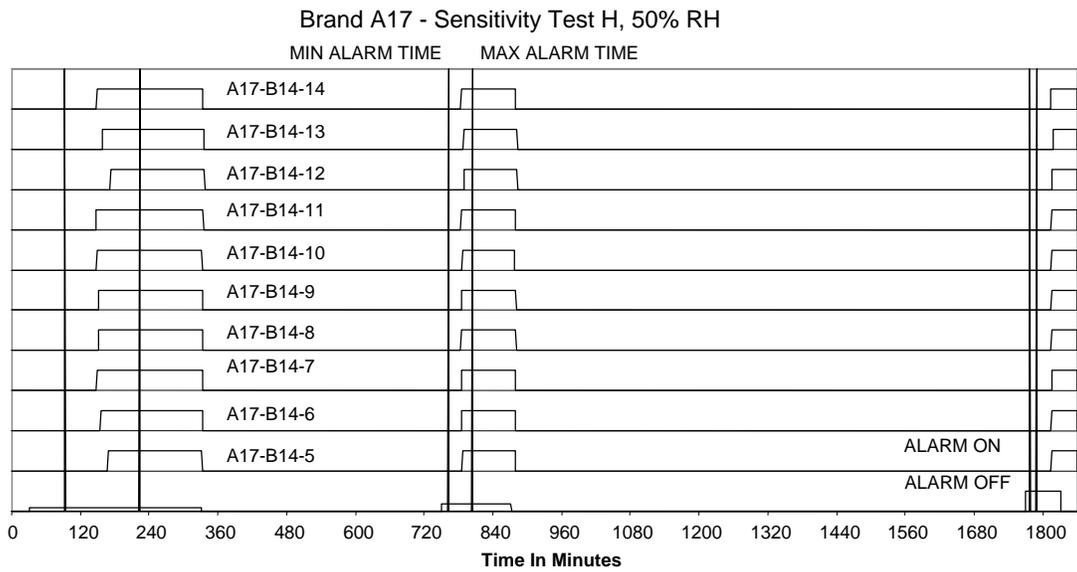


Figure 58 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms sounded within specification at 50% RH and 70 or 150 ppm CO, but they all sounded late at 400 ppm CO.

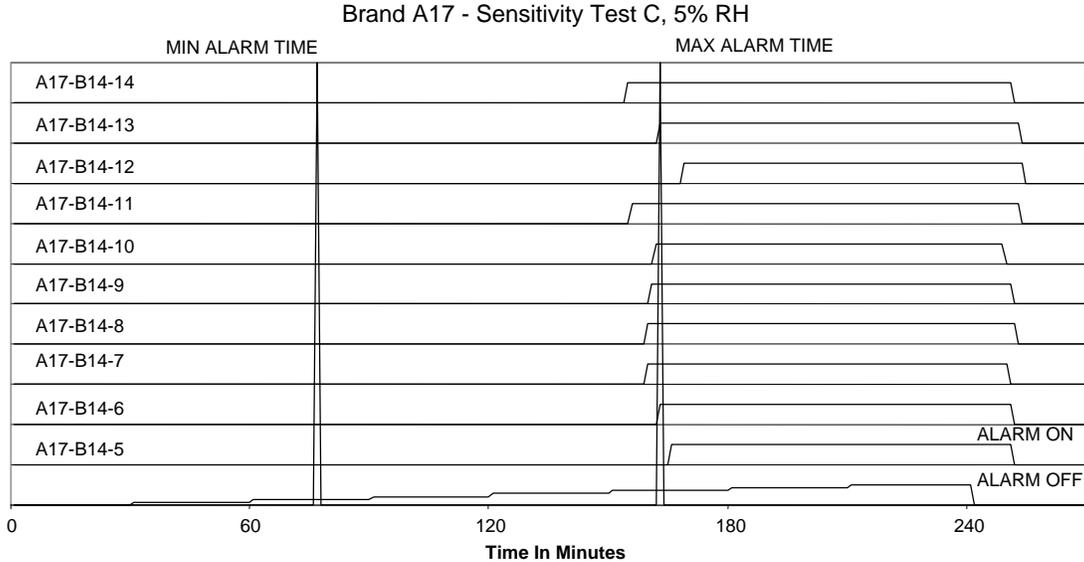


Figure 59 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. Although these devices alarmed within specification at a fixed concentration of 150 ppm, during this test several alarmed late indicating that the detection algorithm does not integrate well.

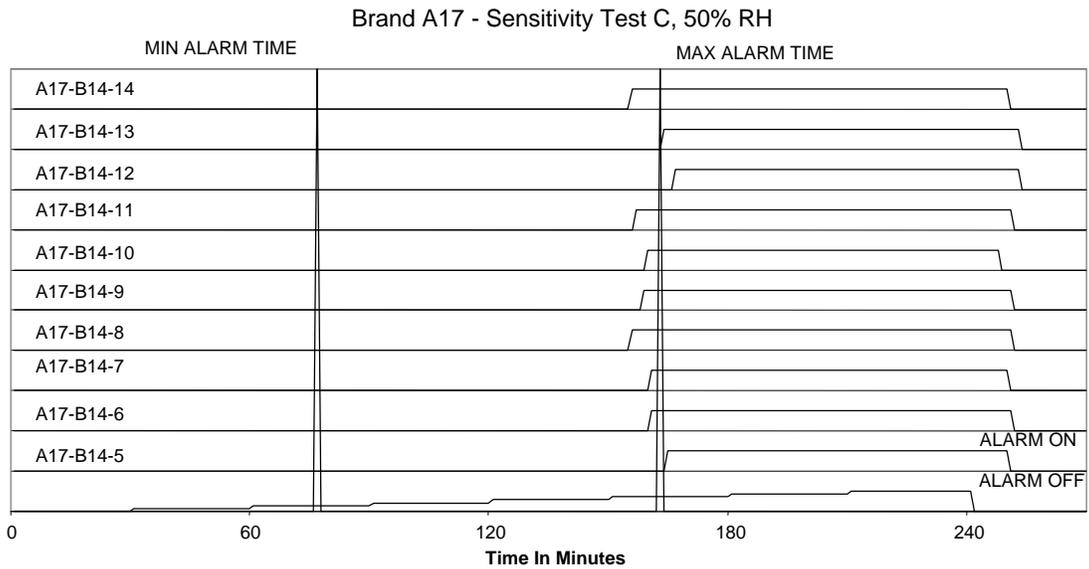


Figure 60 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. In response to gradually increasing concentration these alarms activate at nearly the same times at 50% as they do at 5%.

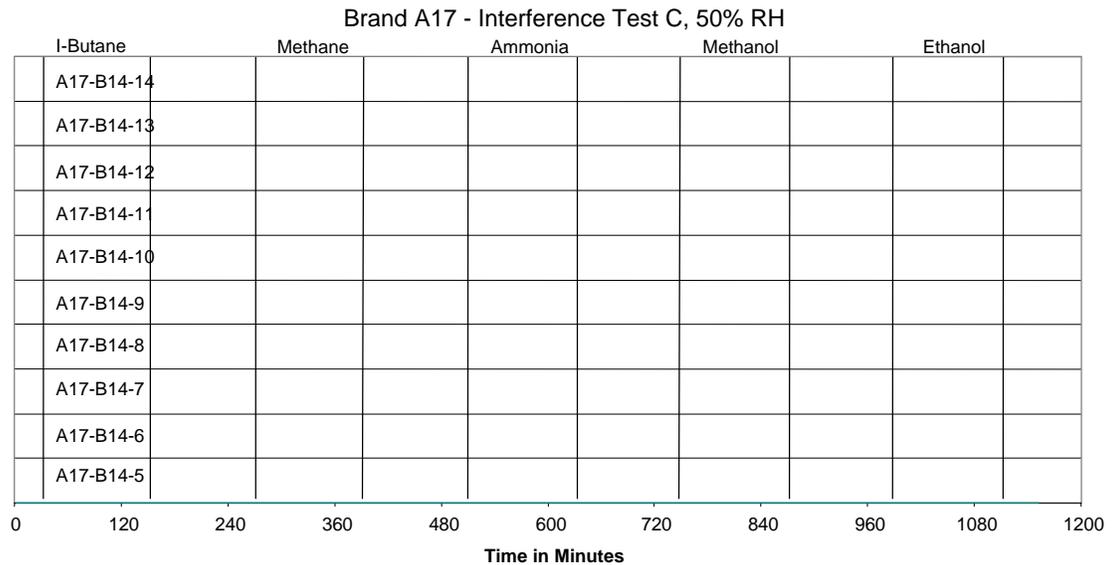


Figure 61 Alarms of brand A17 showed no sensitivity to any of the interfering gases presented.

Brand A17 - Interference Test D, 50% RH

	Acetone	I-Propanol	Ethyl-acetate	n-Heptane	Toluene	Trichloroethylene					
A17-B14-14											
A17-B14-13											
A17-B14-12											
A17-B14-11											
A17-B14-10											
A17-B14-9											
A17-B14-8											
A17-B14-7											
A17-B14-6											
A17-B14-5											
0	120	240	360	480	600	720	840	960	1080	1200	1320

Time in Minutes

Figure 62 Alarms of brand A17 showed no sensitivity to any of the interfering gases presented.

Brand A18

Devices of Brand A18 exhibited the following characteristics:

- The alarms showed no detectable change in sensitivity to CO as humidity was varied over the 5% to 50% RH range.
- Even though the alarms sounded well within the time limits when exposed to 70 and 150 ppm CO, on exposure to progressively increasing CO their sensitivity was borderline, with several just barely actuating under 10% COHb. This behavior indicates an inadequately performing integration algorithm, or the use of a counter/timer algorithm.
- Digital display accuracy was mediocre, with indications of 100 ppm ranging from 88 ppm to 138 ppm.
- Alarms were completely insensitive to all interference gases tested.

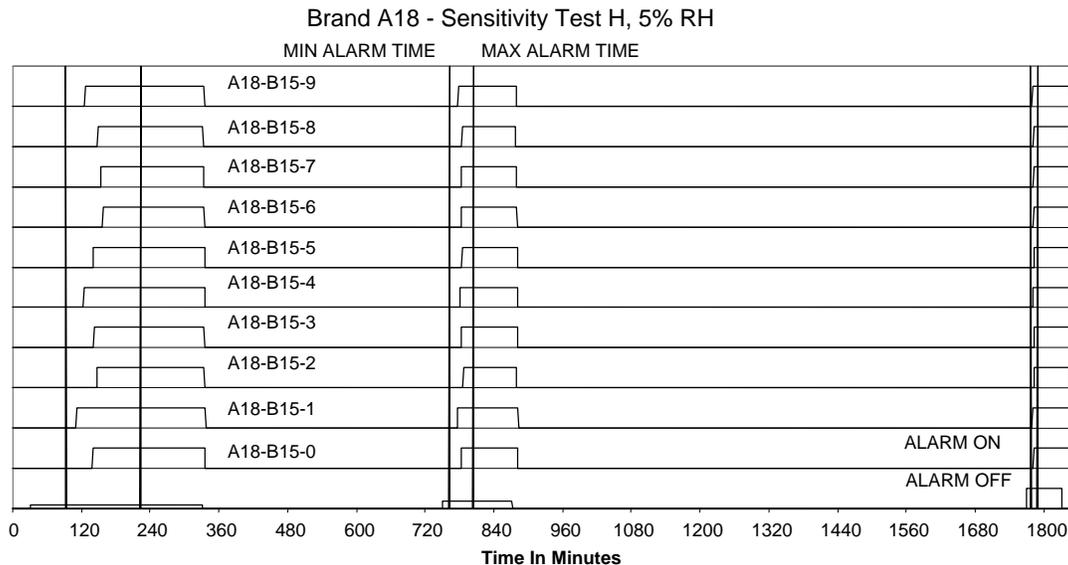


Figure 63 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms activate within specification.

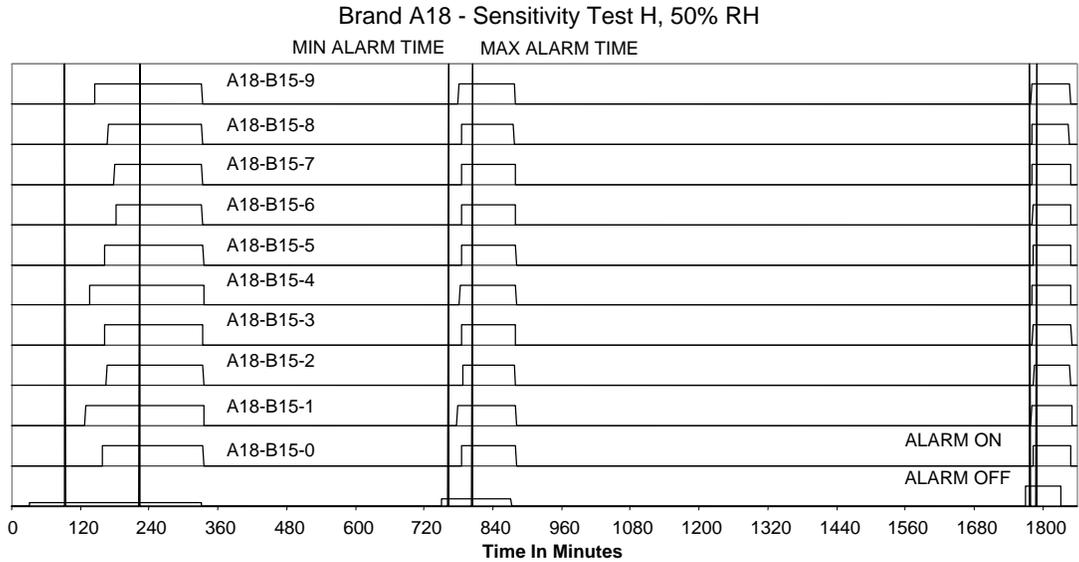


Figure 64 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms activate within specification.

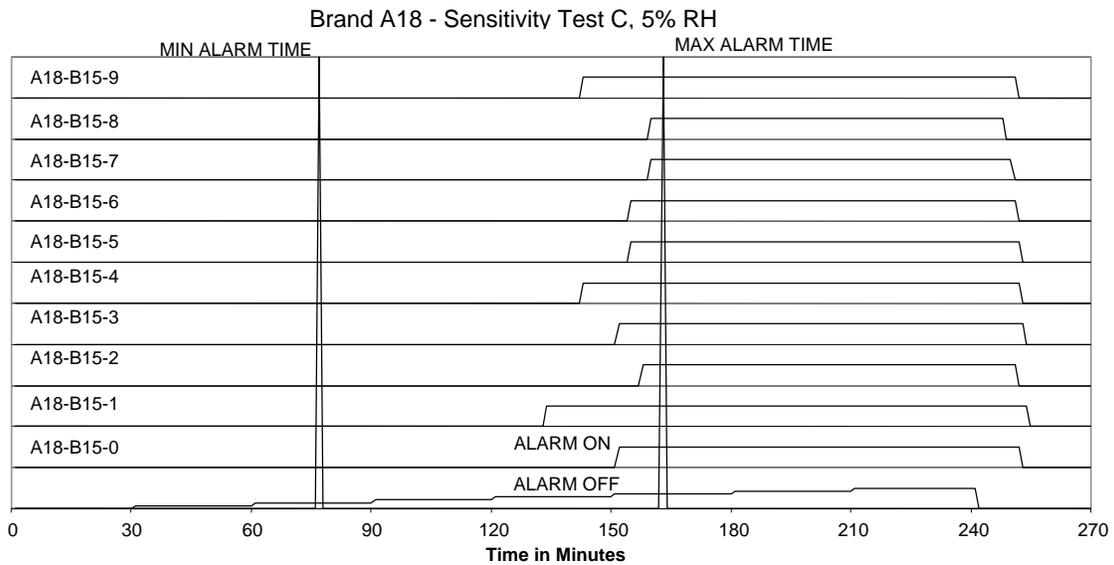


Figure 65 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms activate within specification.

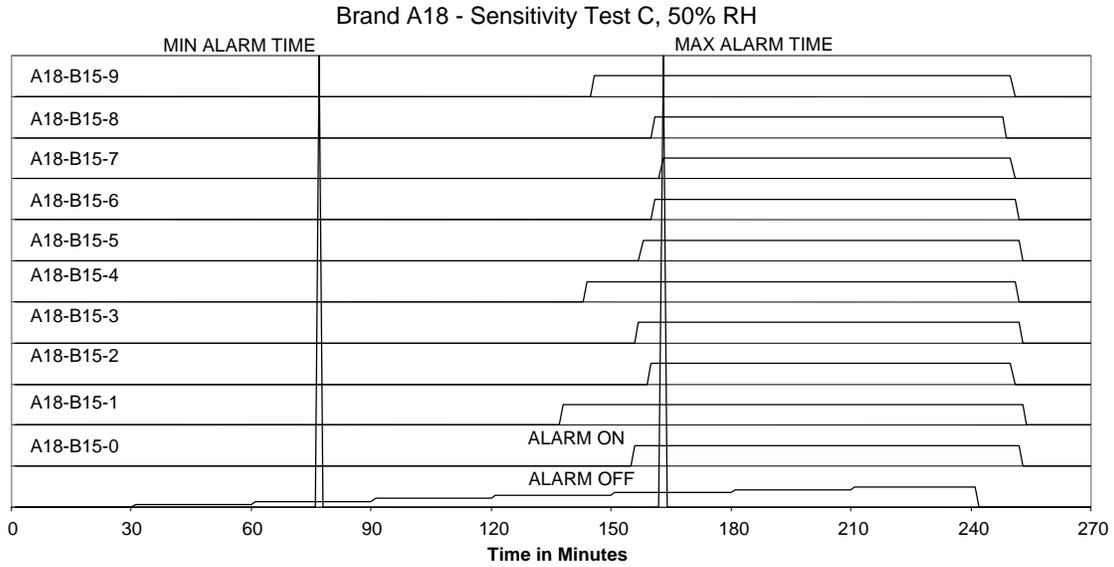


Figure 66 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms activate within specification.

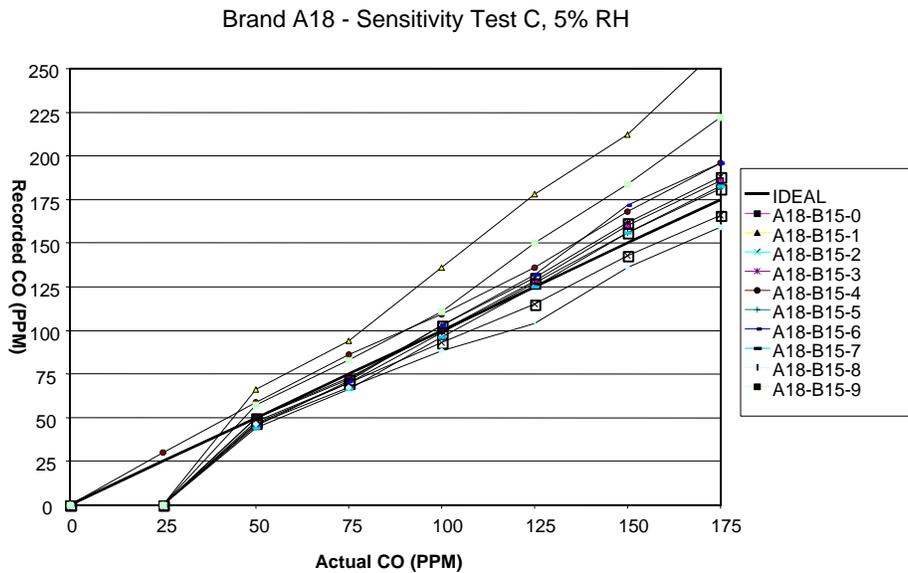


Figure 67 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. While there is considerable variation in sensitivity among alarms, most display readings are within $\pm 20\%$.

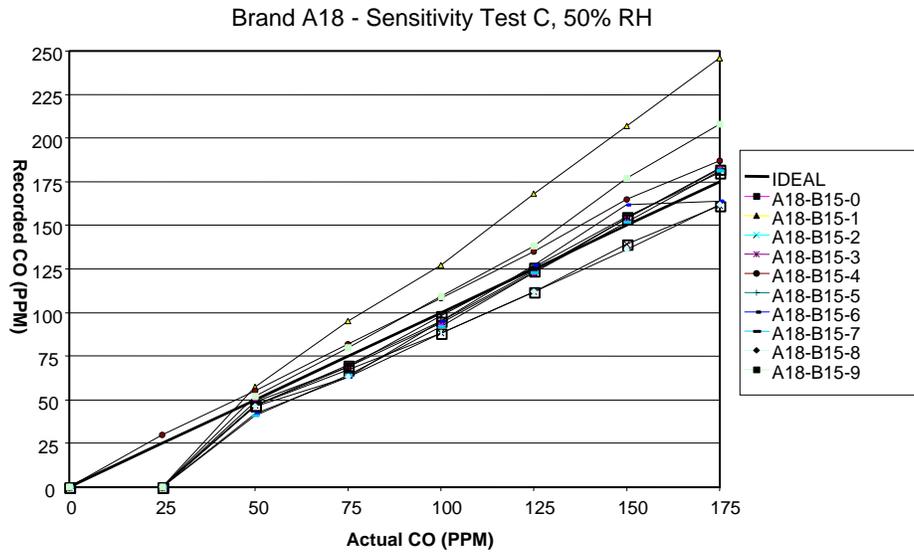


Figure 68 Digital display readings taken during Sensivity Test C are compared to the actual CO concentrations. While there is considerable variation in sensitivity among alarms, most display readings are within $\pm 20\%$.

Brand A18 - Interference Test C, 50% RH

	I-Butane	Methane	Ammonia	Methanol	Ethanol
A18-B15-9					
A18-B15-8					
A18-B15-7					
A18-B15-6					
A18-B15-5					
A18-B15-4					
A18-B15-3					
A18-B15-2					
A18-B15-1					
A18-B15-0					

Time in Minutes

Figure 69 Alarm A18 was not sensitive to any of the interference gases tested.

Brand A18 - Interference Test D, 50% RH

	Acetone	I-Propanol	Ethyl-acetate	n-Heptane	Toluene	Trichloroethylene
A18-B15-9						
A18-B15-8						
A18-B15-7						
A18-B15-6						
A18-B15-5						
A18-B15-4						
A18-B15-3						
A18-B15-2						
A18-B15-1						
A18-B15-0						

0 120 240 360 480 600 720 840 960 1080 1200 1320

Time in Minutes

Figure 70 Alarm A18 was not sensitive to any of the interference gases tested.

Brand A19

Several brands have a “supervised failure” feature in which they perform internal diagnostics and signal that the unit should be replaced by displaying “Err” if an error is found. Such a supervised failure is preferable to a false alarm or failure to alarm, in which the user does not know that an error has occurred. At the time of purchase three of thirteen Brand A19 alarms showed supervised errors by displaying “Err” on their digital displays, and one did not reset properly. The others performed extremely well throughout testing, with the following properties:

- All alarms showed no detectable change in sensitivity to CO as humidity was varied over a 5% to 50% RH range.
- Under all humidity conditions all alarms sounded well within the UL 2034 sensitivity requirements.
- On exposure to progressively increasing CO concentrations all alarms sounded well within 2.5% to 10% COHb levels.
- Under all conditions the accuracy of the digital displays rivaled that of analytical instruments, with the majority of readings within $\pm 5\%$ over the tested range of 25-175 ppm CO.
- Alarms were completely insensitive to all interference gases tested.

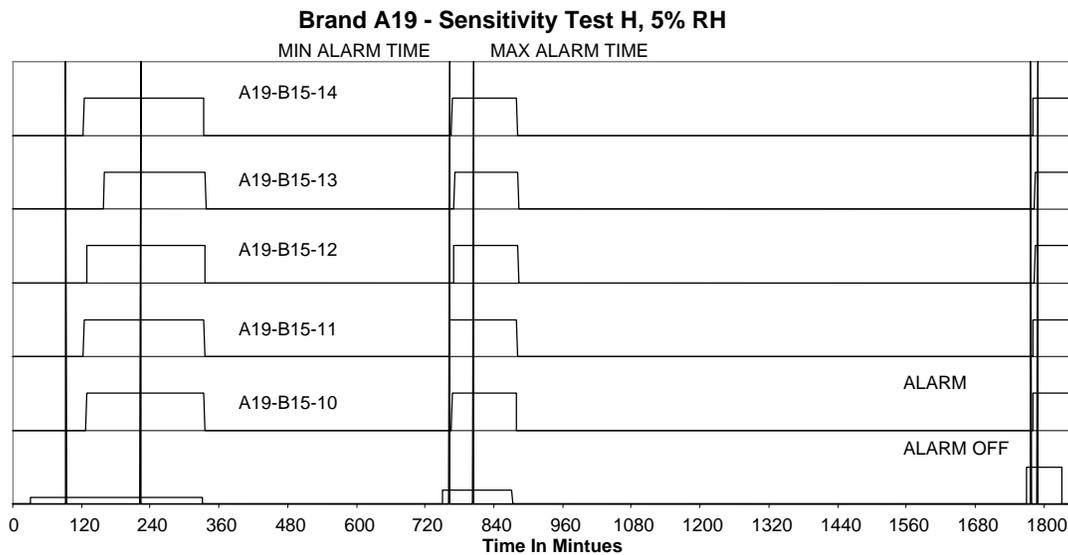


Figure 71 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms activated within specification at each test concentration.

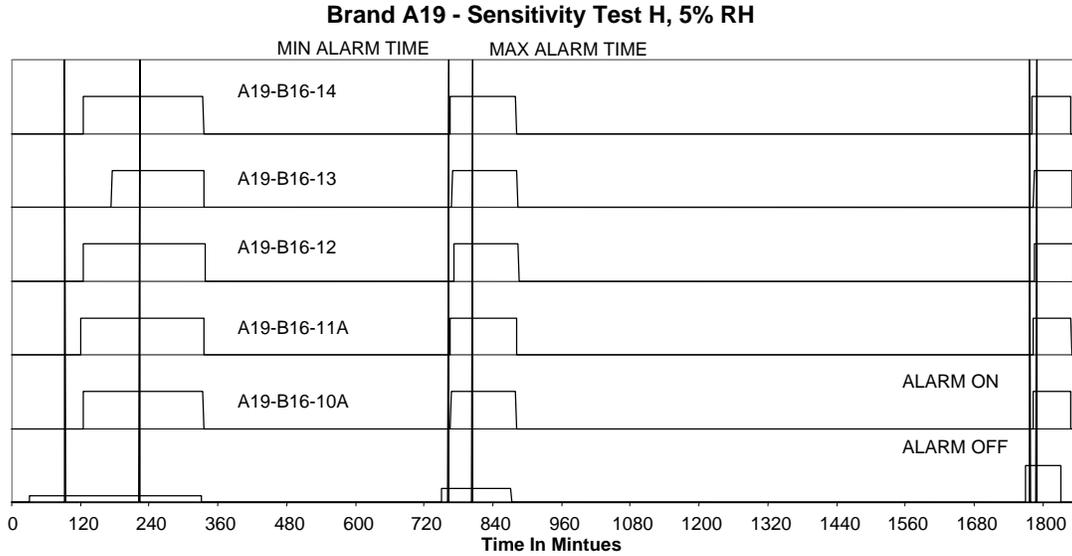


Figure 72 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms activated within specification at each test concentration.

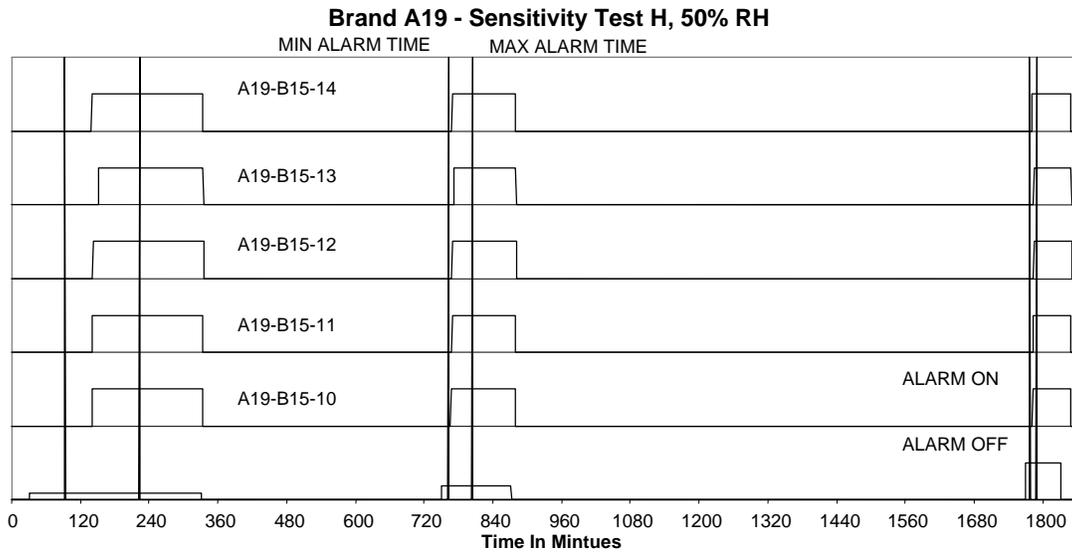


Figure 73 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms activated within specification at each test concentration.

Brand A19 - Sensitivity Test H, 50% RH

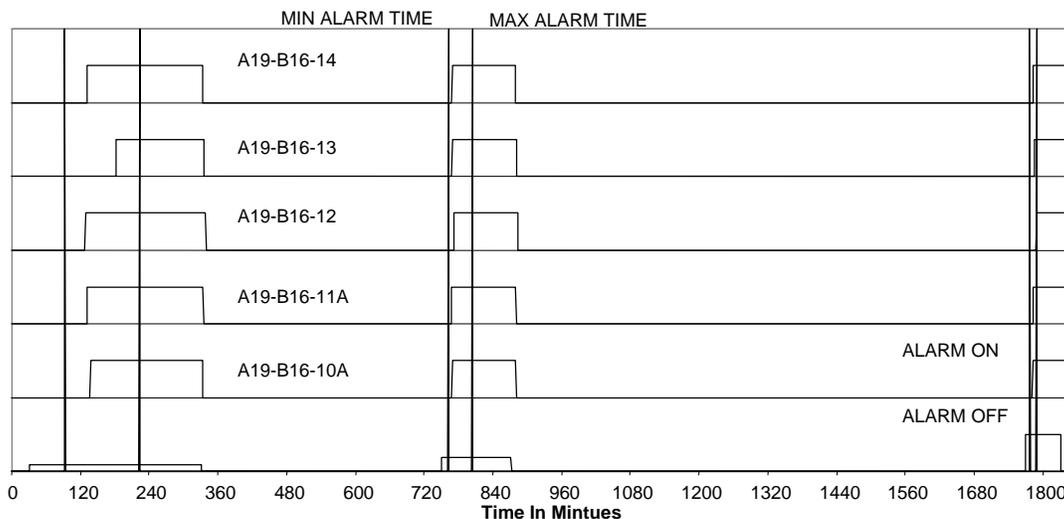


Figure 74 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms activated within specification at each test concentration.

Brand A19 - Sensitivity Test C, 5% RH

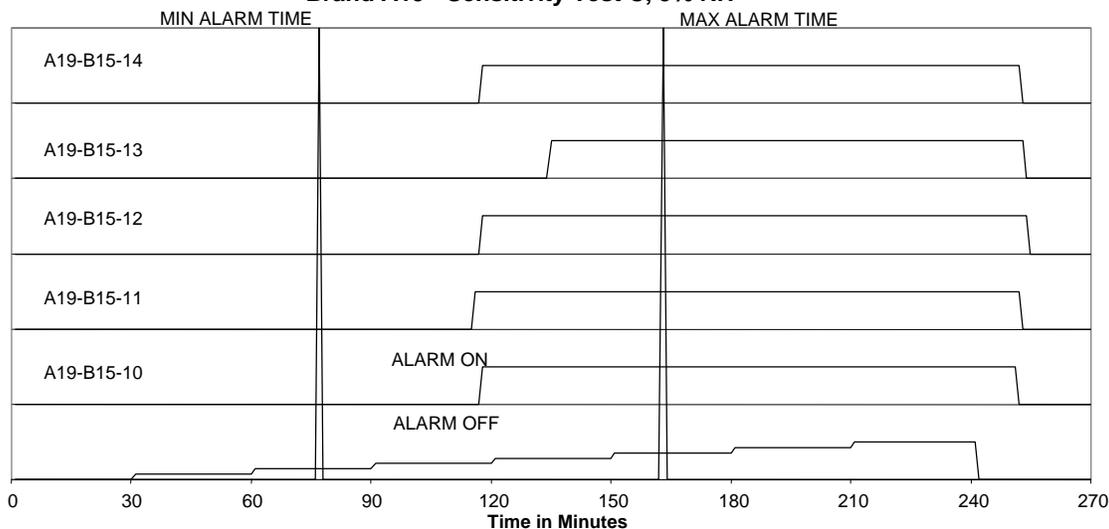


Figure 75 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms activated within specification.

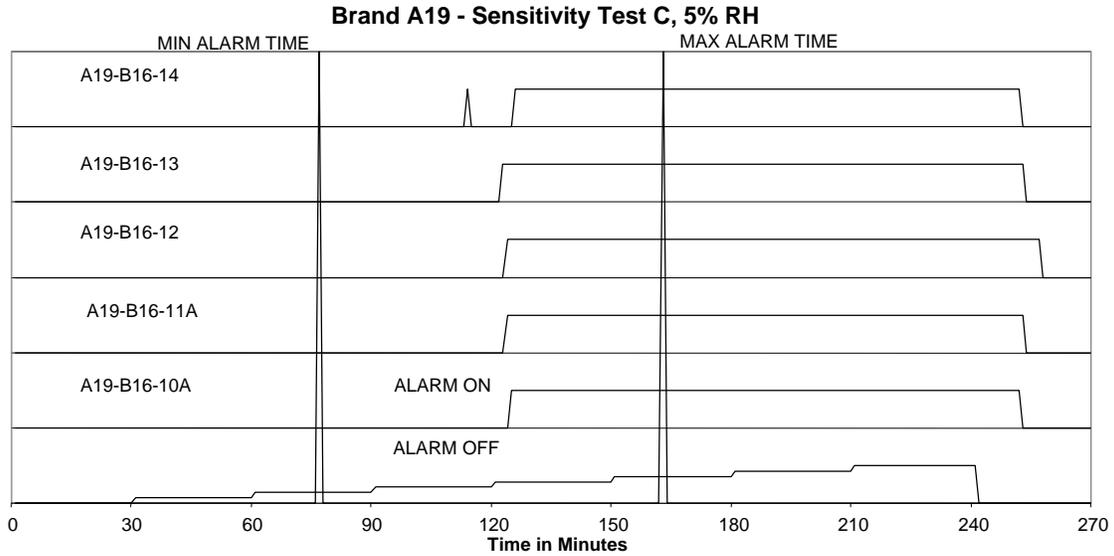


Figure 76 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms activated within specification.

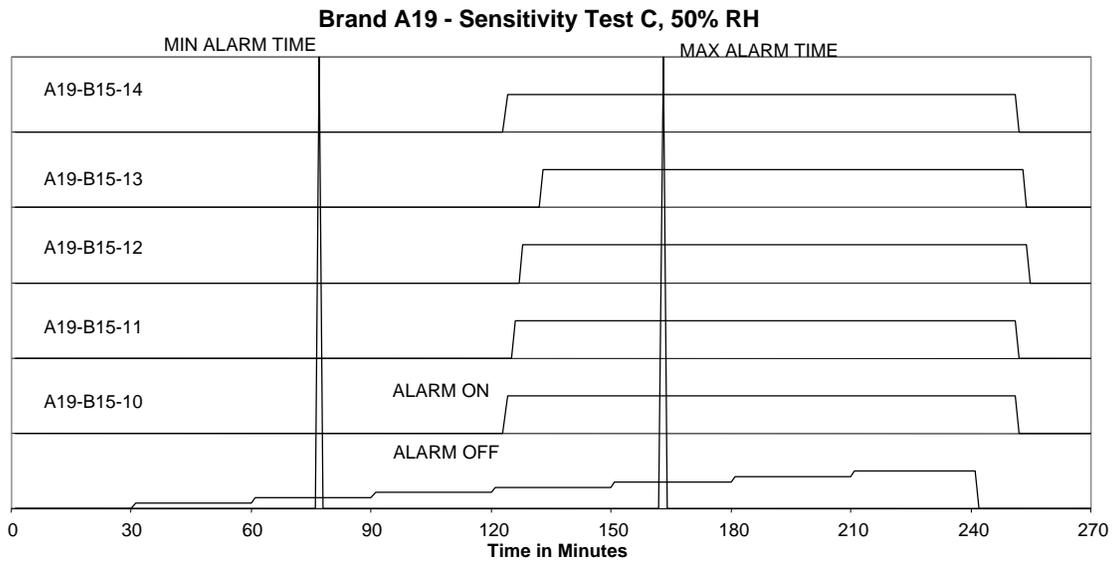


Figure 77 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms activated within specification.

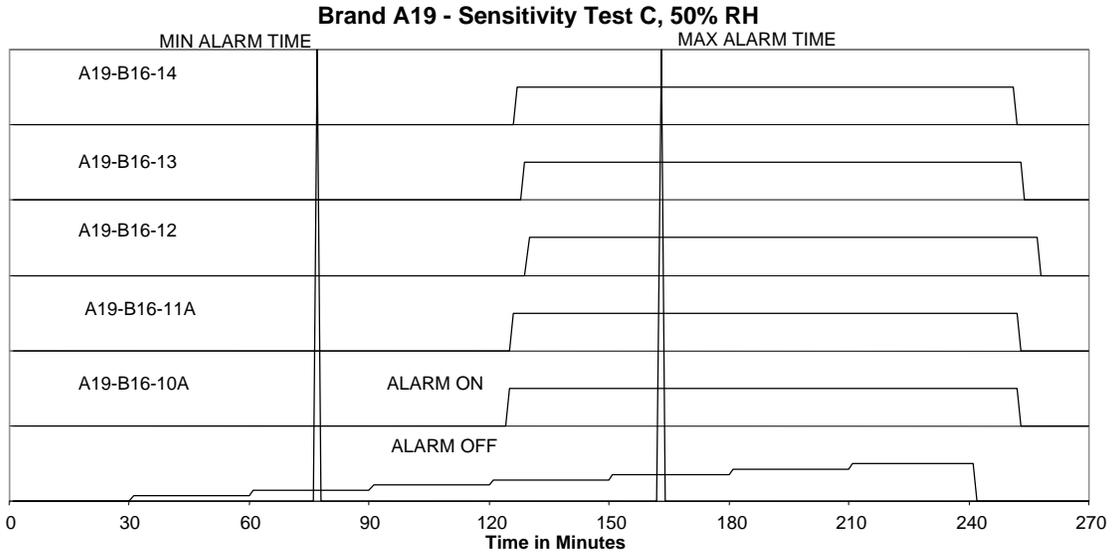


Figure 78 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms activated within specification.

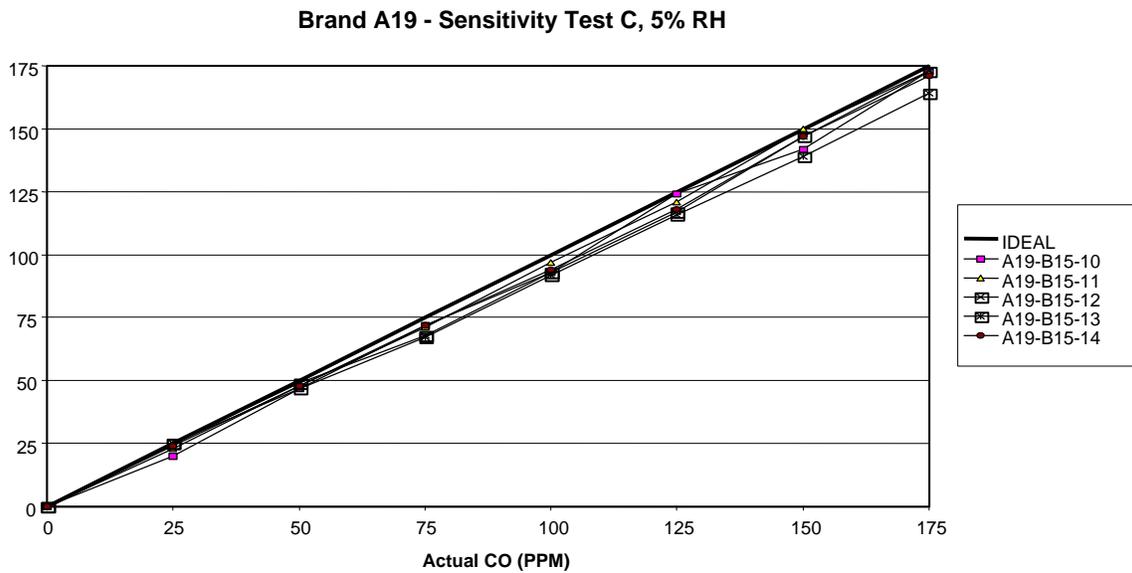


Figure 79 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. At 5% RH all display readings for these alarms are well within $\pm 10\%$ of the actual concentrations.

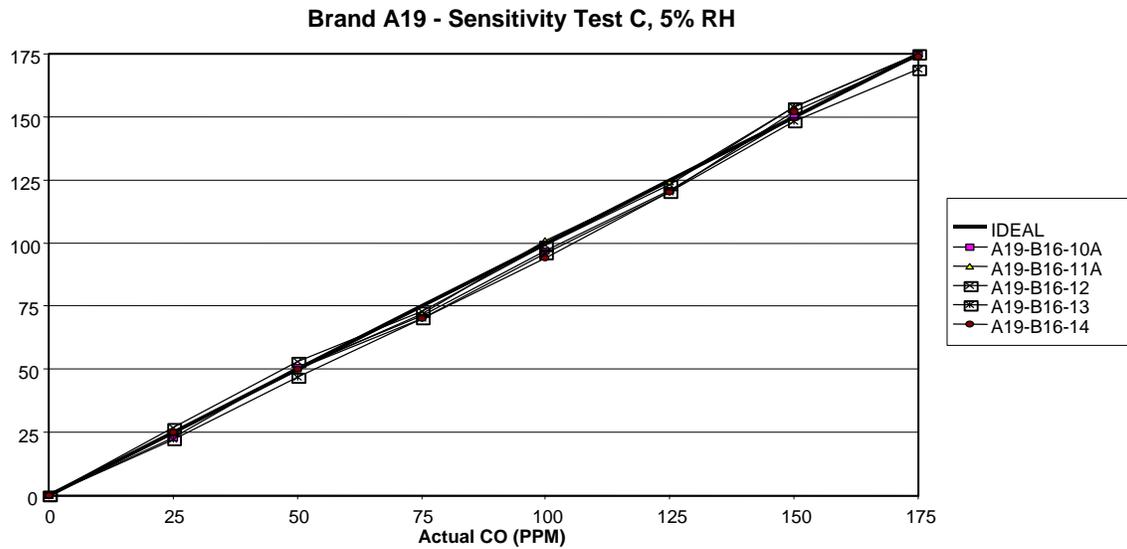


Figure 80 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. At 5% RH all display readings for these alarms are well within $\pm 10\%$ of the actual concentrations.

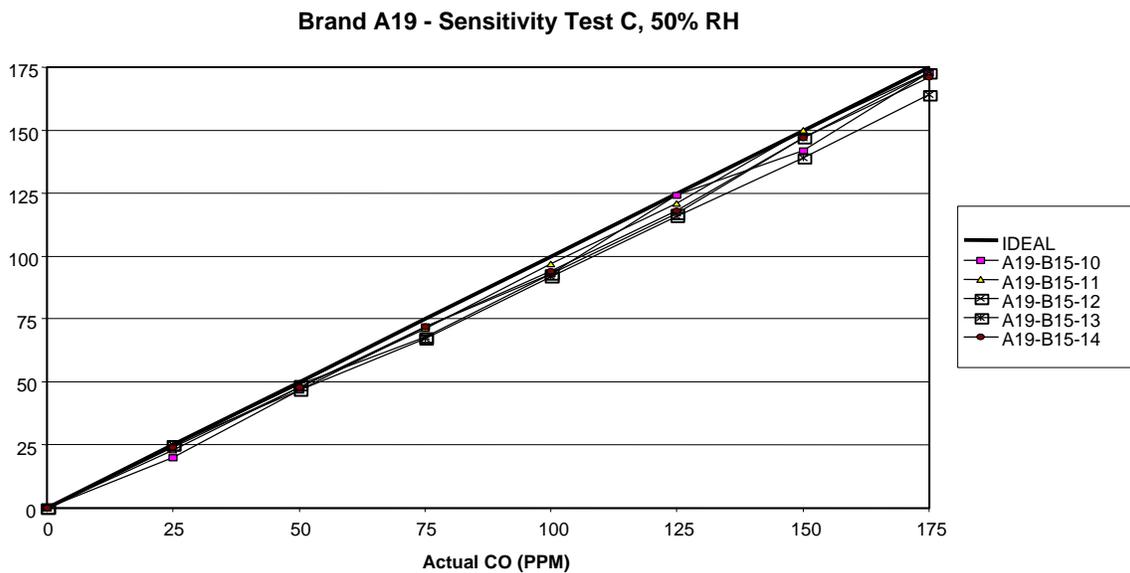


Figure 81 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. At 50% RH all display readings for these alarms are well within $\pm 10\%$ of the actual concentrations.

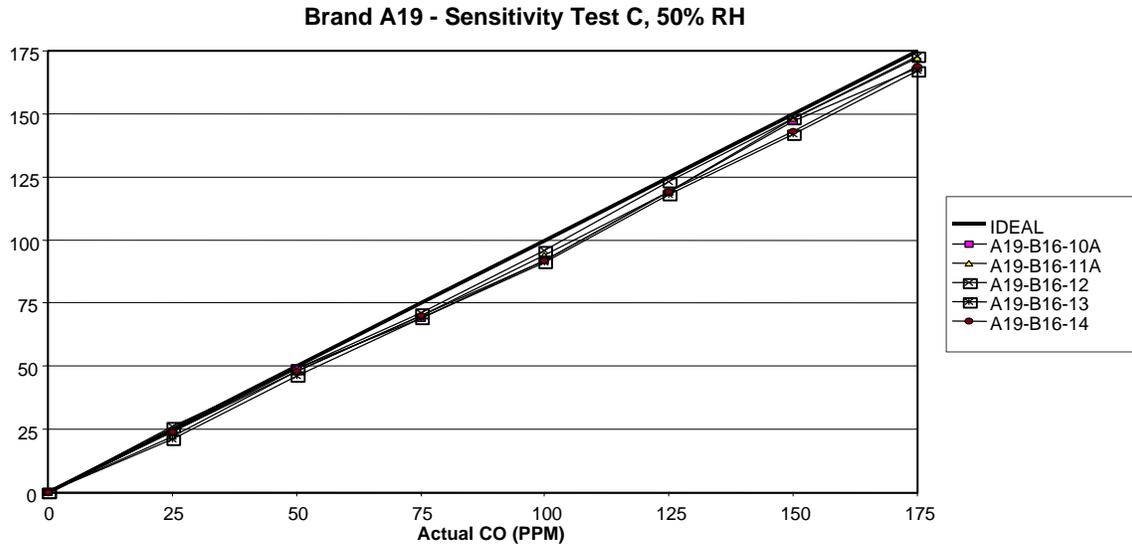


Figure 82 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. At 50% RH all display readings for these alarms are well within $\pm 10\%$ of the actual concentrations.

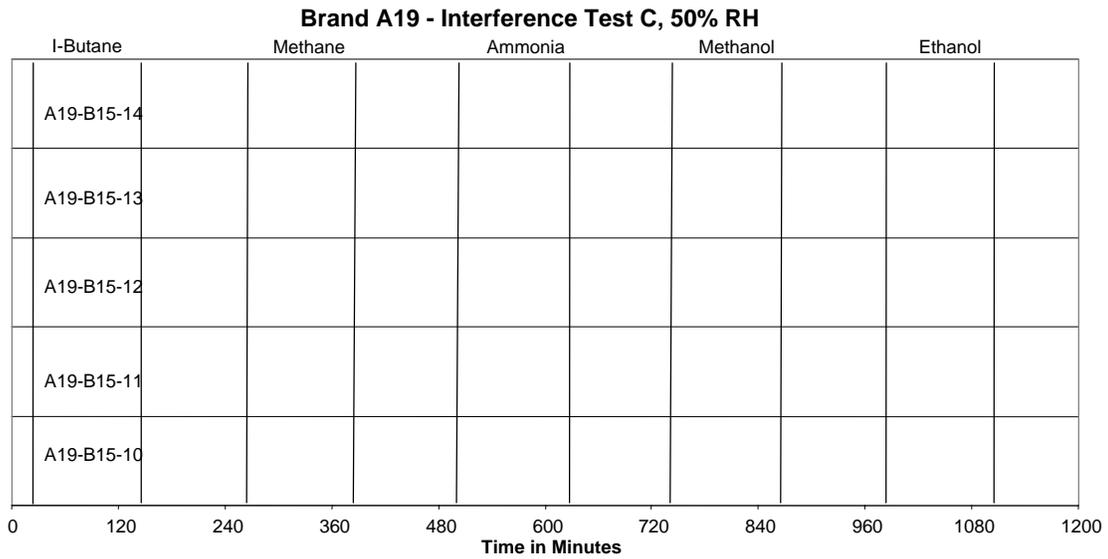


Figure 83 Alarms of Brand A19 were not sensitive to any of the interference gases tested.

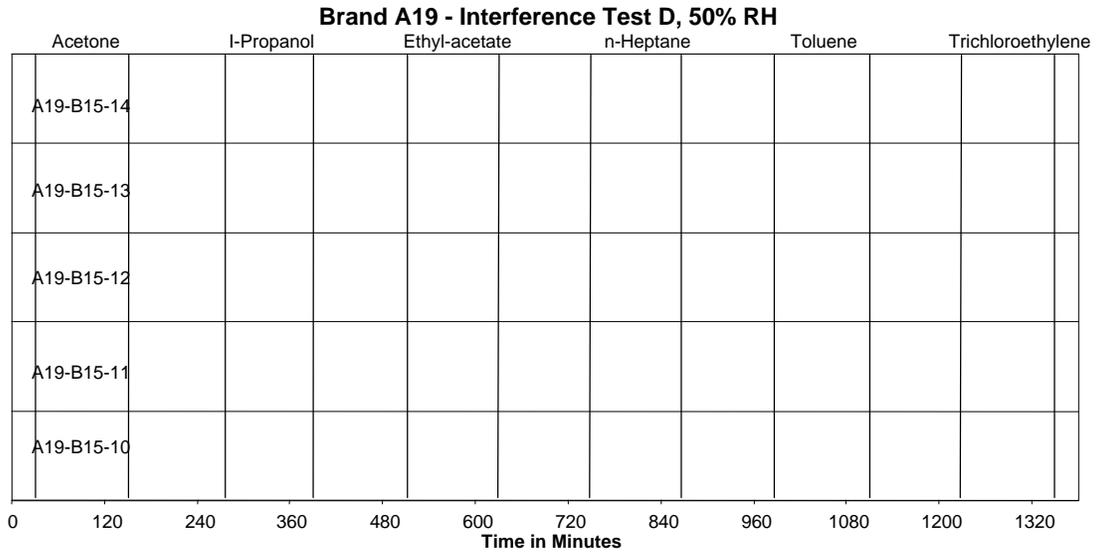


Figure 84 Alarms of Brand A19 were not sensitive to any of the interference gases tested.

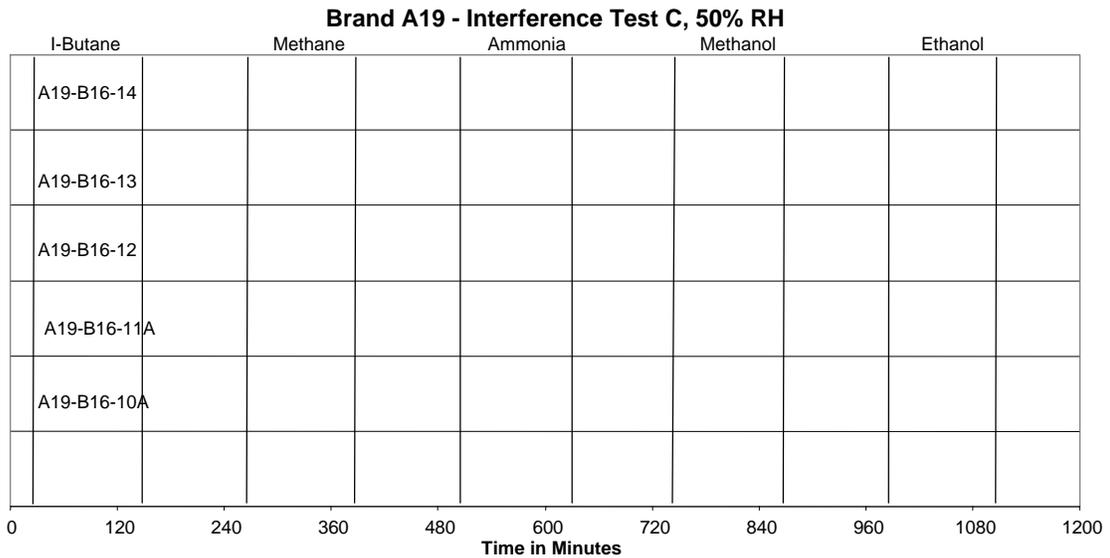


Figure 85 Alarms of Brand A19 were not sensitive to any of the interference gases tested.

Brand A19 - Interference Test D, 50% RH

	Acetone	I-Propanol	Ethyl-acetate	n-Heptane	Toluene	Trichloroethylene
A19-B16-14						
A19-B16-13						
A19-B16-12						
A19-B16-11A						
A19-B16-10A						

0 120 240 360 480 600 720 840 960 1080 1200 1320

Time in Minutes

Figure 86 Alarms of Brand A19 were not sensitive to any of the interference gases tested.

Brand A20

This brand suffered from a number of dramatic deficiencies:

- Alarms were completely insensitive to the tested concentrations of CO at 5% RH. As humidity was increased sensitivity slowly increased, but even at 50% RH alarms were sometimes insufficiently sensitive.
- While for the fixed CO concentrations of the UL sensitivity tests at 50% RH most alarms showed adequate sensitivity, for progressively increasing CO concentrations only two of ten alarms activated their alarms at less than 10% COHb.
- Digital displays were extremely inaccurate. At low humidity displays reported zero for hazardous concentrations, while at 50% RH displays underreported concentration by up to more than a factor of two.

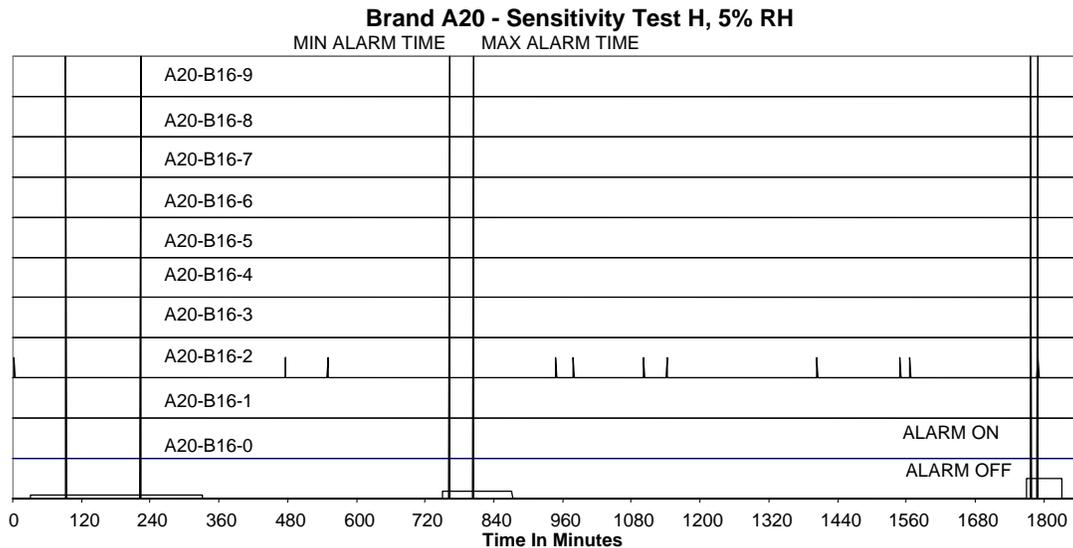


Figure 87 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. Devices of Brand A20 showed no sensitivity at 5% RH. One alarm sporadically sounded throughout the test, and the activations did not seem related to the gas presentations.

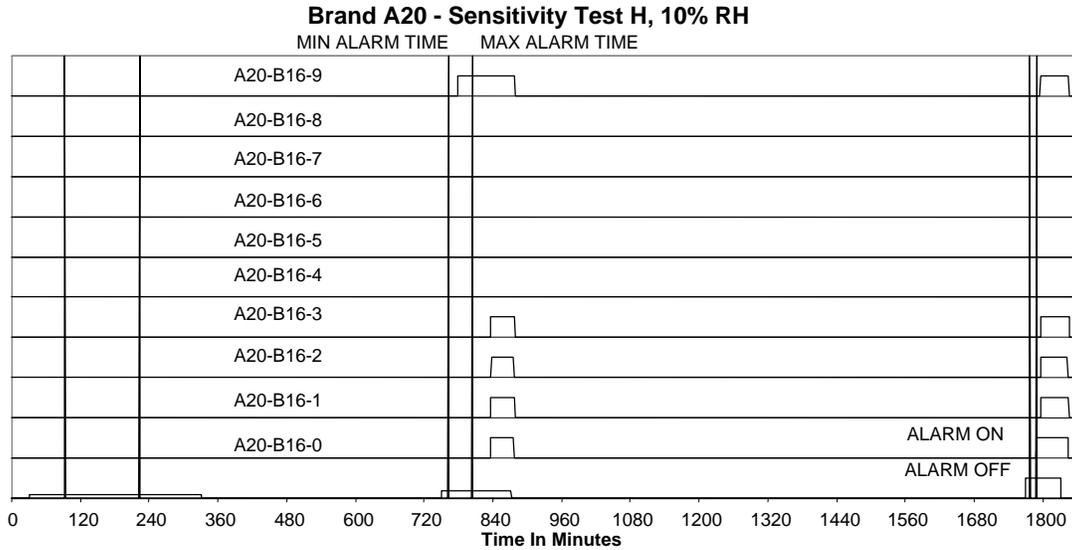


Figure 88 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. Alarms were insufficiently sensitive, with several sounding late, and many not sounding at all.

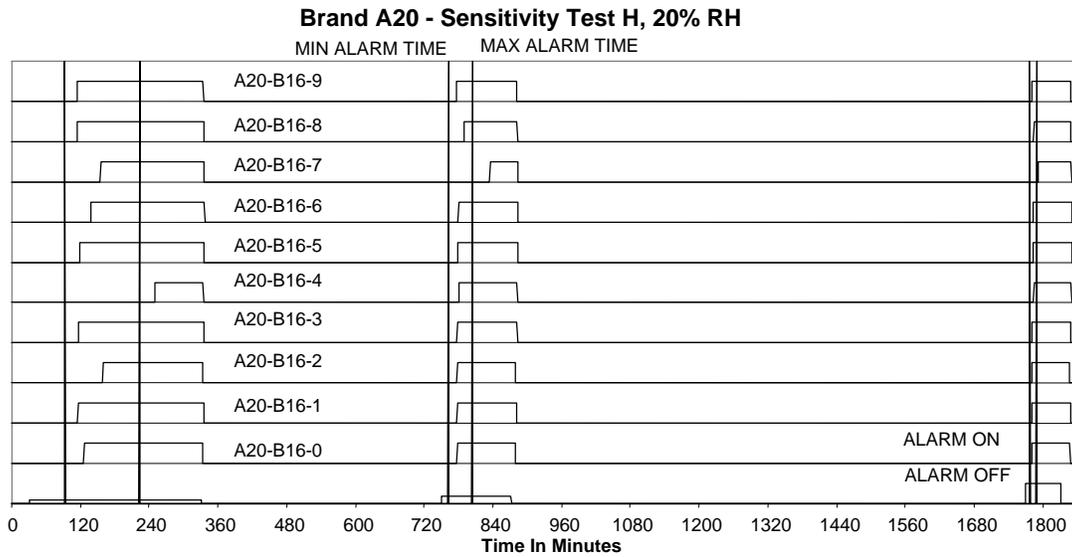


Figure 89 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. Even at 20% RH several alarms were insufficiently sensitive.

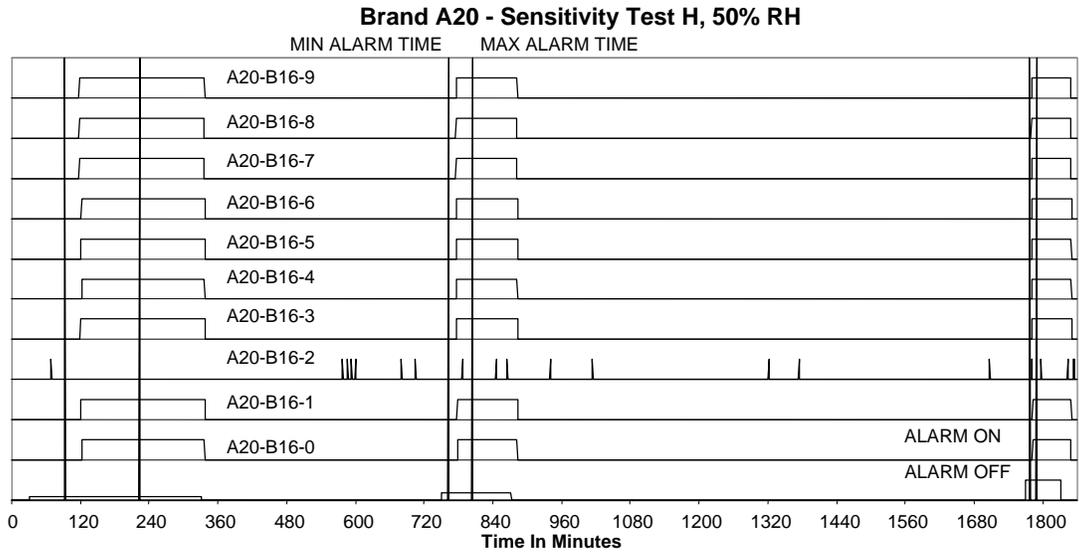


Figure 90 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms but one activated within specification. Alarm A20-B16-2 intermittently activated, just as it did at 5% RH.

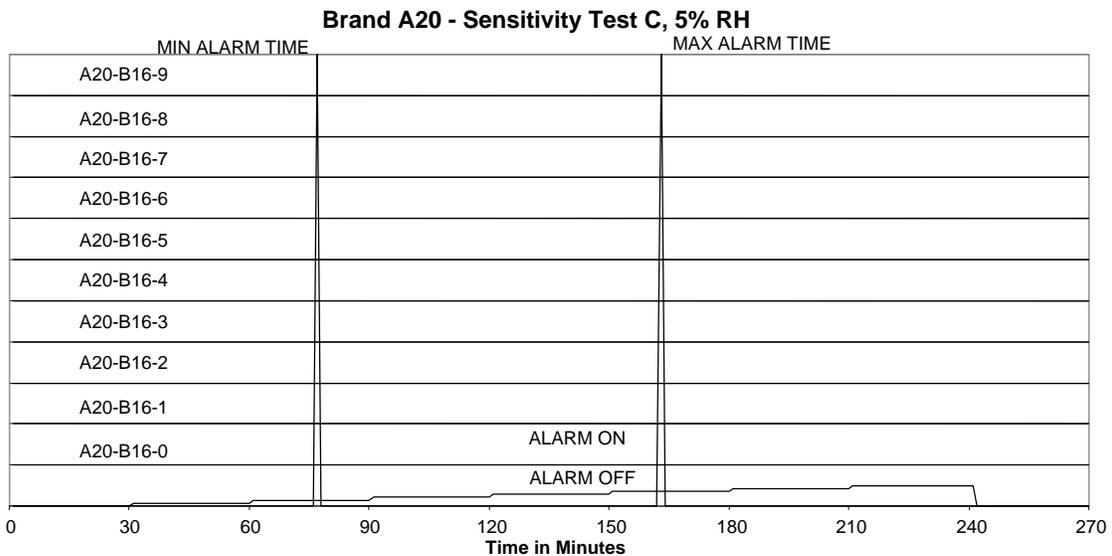


Figure 91 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. None of the alarms were sufficiently sensitive to activate during the presentation.

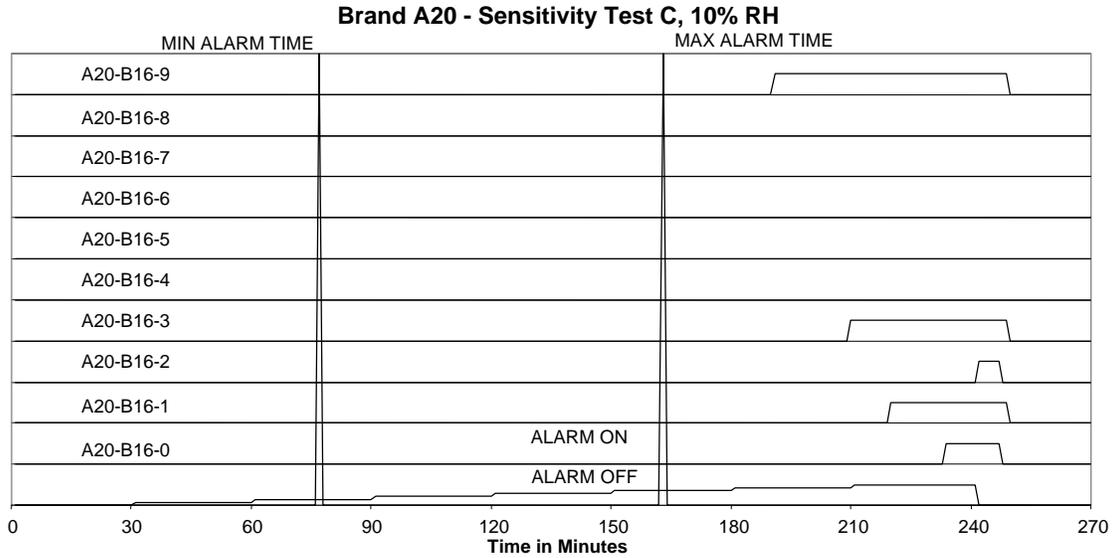


Figure 92 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. At 10% RH some of the alarms activate, albeit late, while others do not.

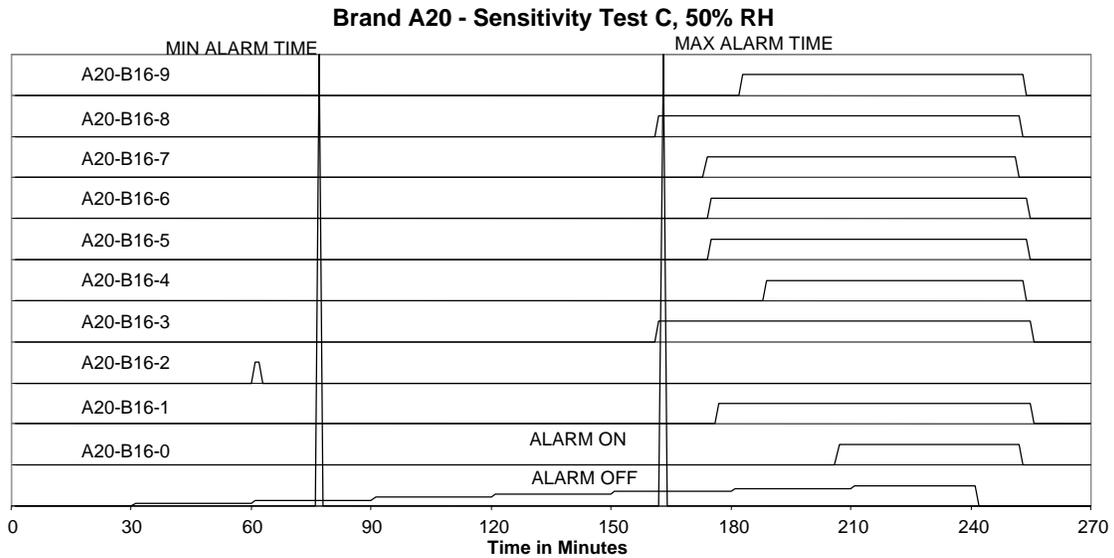


Figure 93 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. Even though they do alarm in time at the fixed UL test points, at 50% RH and on exposure to continuously increasing CO most Brand A20 alarms do not sound in time, and one fails completely.

Brand A20 - Sensitivity Test C, 5% RH

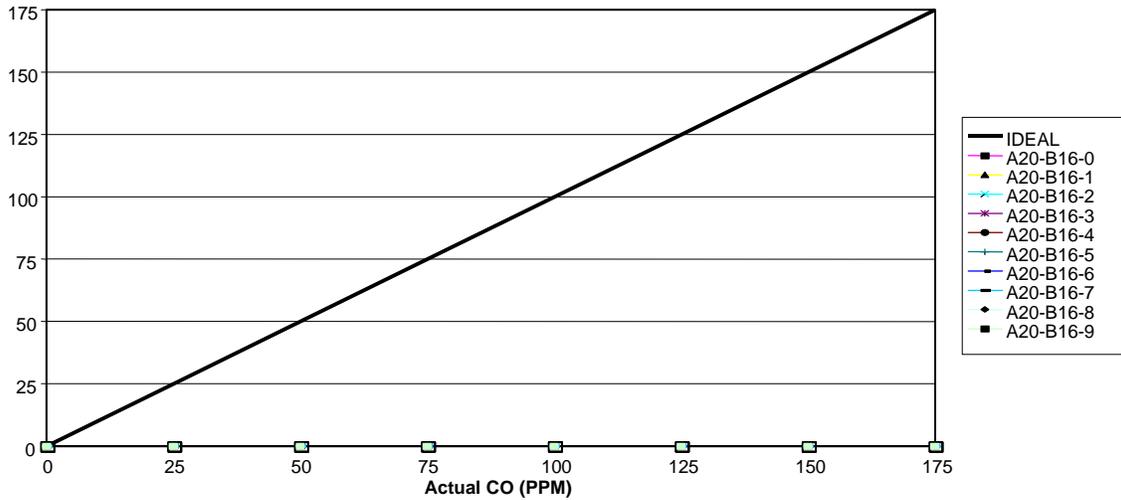


Figure 94 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. All alarms read 0 ppm throughout the test; they were completely insensitive to CO at 5% RH.

Brand A20 - Sensitivity Test C, 10% RH

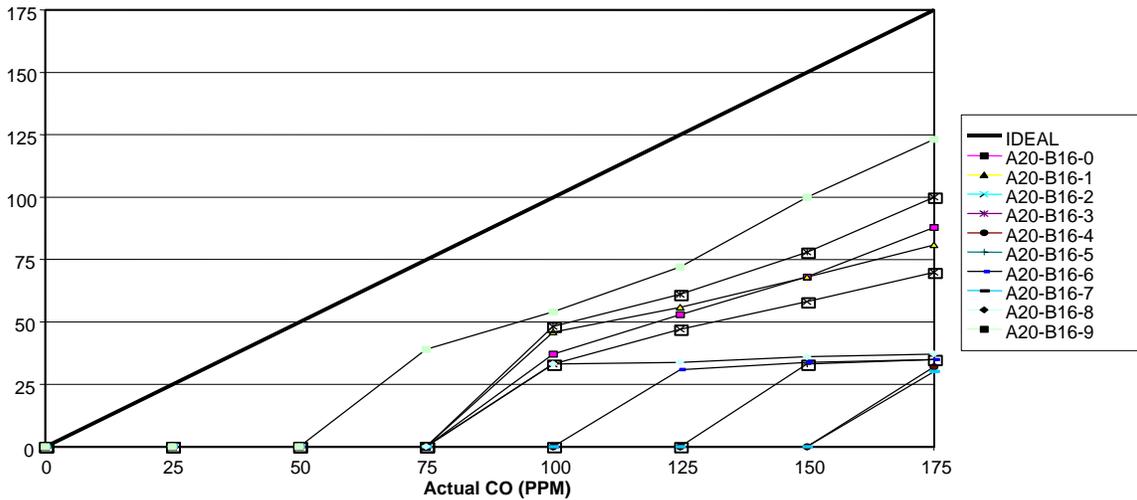


Figure 95 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. All alarms were insufficiently sensitive at 10% RH, with different alarms commencing display readings over zero at different applied concentrations. Only at concentrations of 175 ppm and greater did all alarms respond to CO. At 175 ppm five of ten alarms read 30 ppm, while the other five read between 70 and 125 ppm.

Brand A20 - Sensitivity Test C, 50% RH

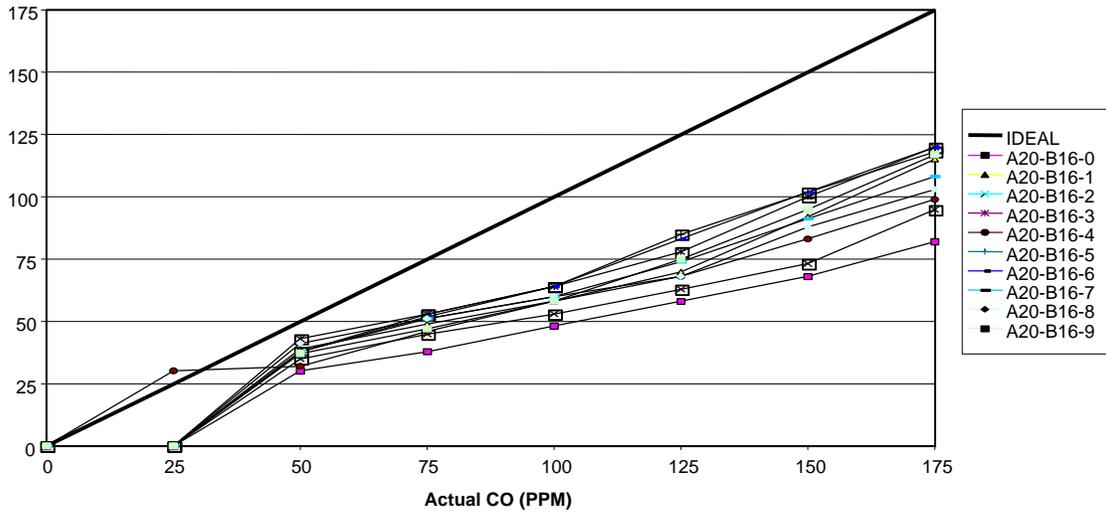


Figure 96 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations at a relative humidity of 50%. Even at 50% RH, all alarms of Brand A20 read low, by 30% to 55%.

Brand A20 - Interference Test C, 50% RH

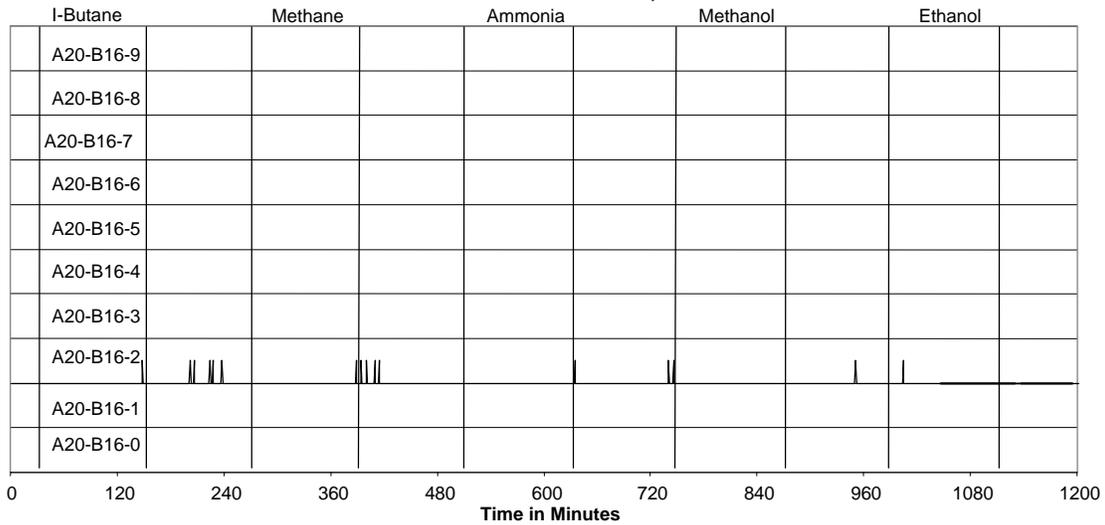


Figure 97 One alarm of ten sporadically activated throughout the interference gas presentations, while the others were not sensitive to the interference gases.

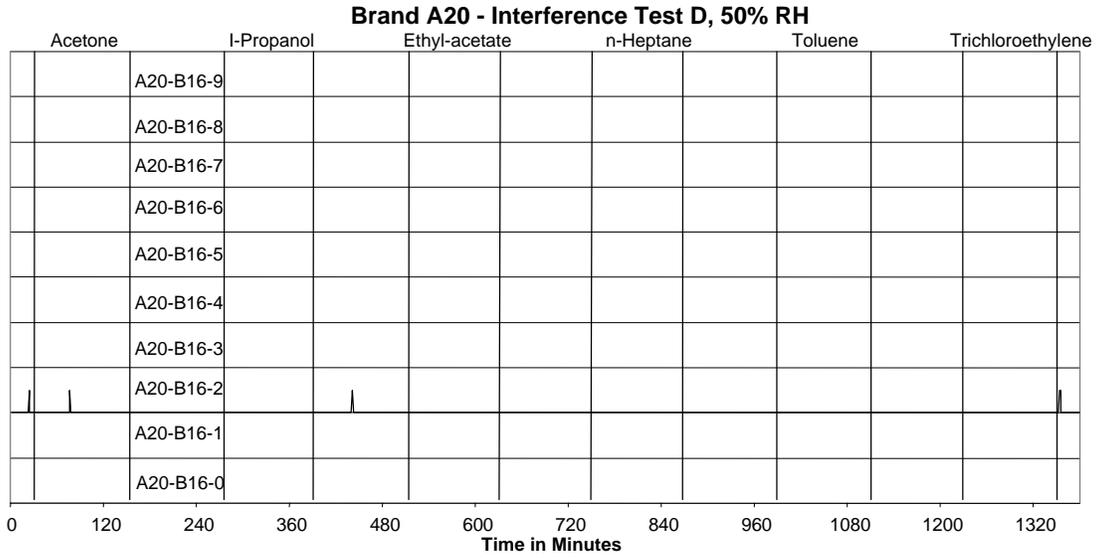


Figure 98 One alarm of ten sporadically activated throughout the interference gas presentations, while the others were not sensitive to the interference gases.

Brand A21

Devices of Brand A21 were predominantly adequately sensitive at 50% RH, but failed to alarm at 5% RH:

- Alarm activations occurred late when tested at 5% RH. At 50% RH when tested to fixed CO concentrations alarms sounded within the UL sensitivity limits.
- When tested against progressively increasing CO concentrations, activation times varied greatly from device to device. At 5% RH most devices were insufficiently sensitive, with only three of eight alarming at less than 10% COHb, while at 50% RH only one exhibited borderline sensitivity.
- The contrast between the uniformity of alarm activation times when tested with fixed CO concentrations and the variability of activation times during exposure to progressively increasing CO concentration indicates that the alarms inadequately average CO exposure.
- Alarms were not sensitive to any of the interference gases tested.

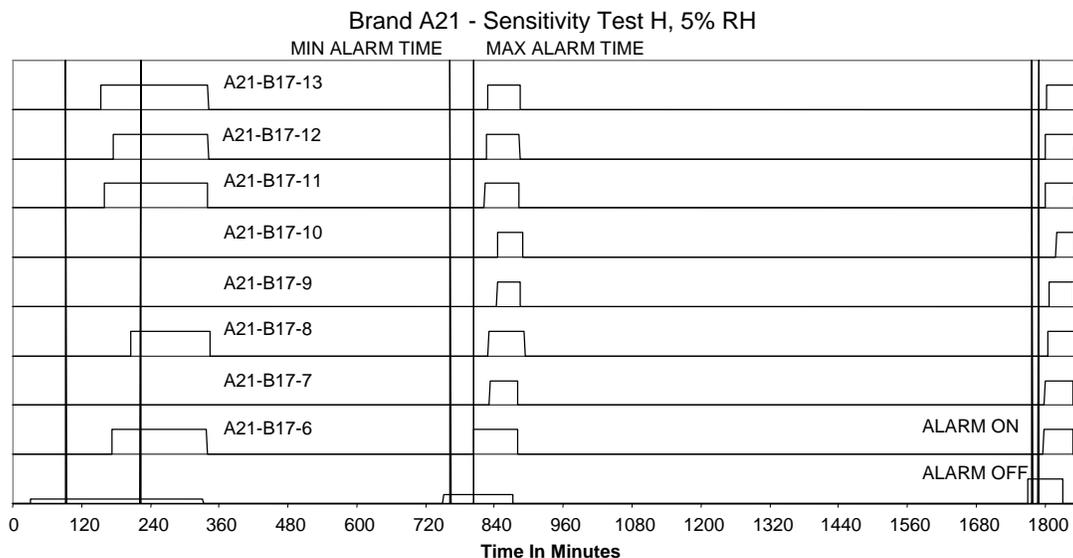


Figure 99 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. At 5% RH all devices either failed to activate or activated late in response to the UL test concentrations..

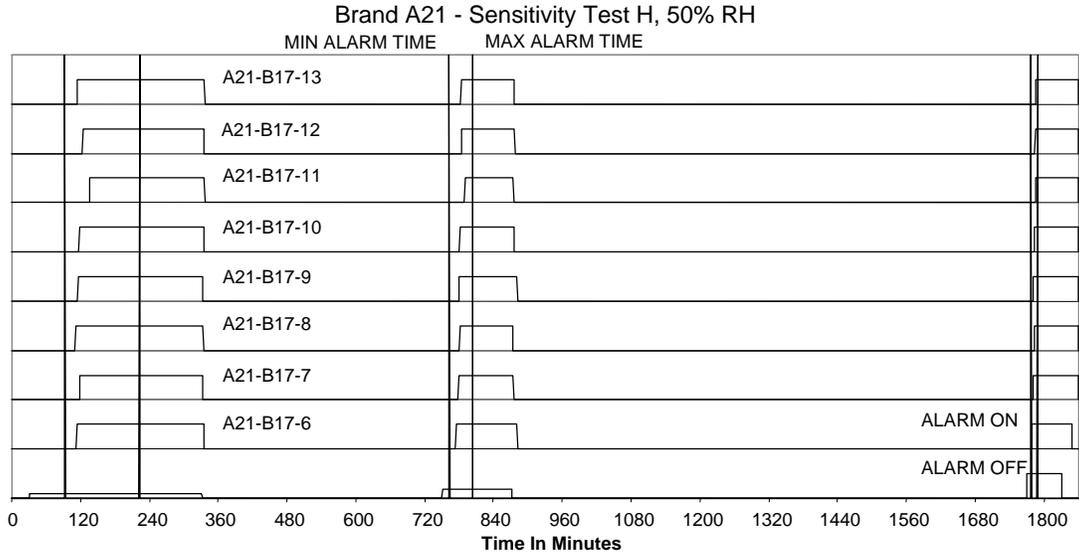


Figure 100 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. In contrast to their behavior at 5%, at 50% RH all alarms sounded within specification.

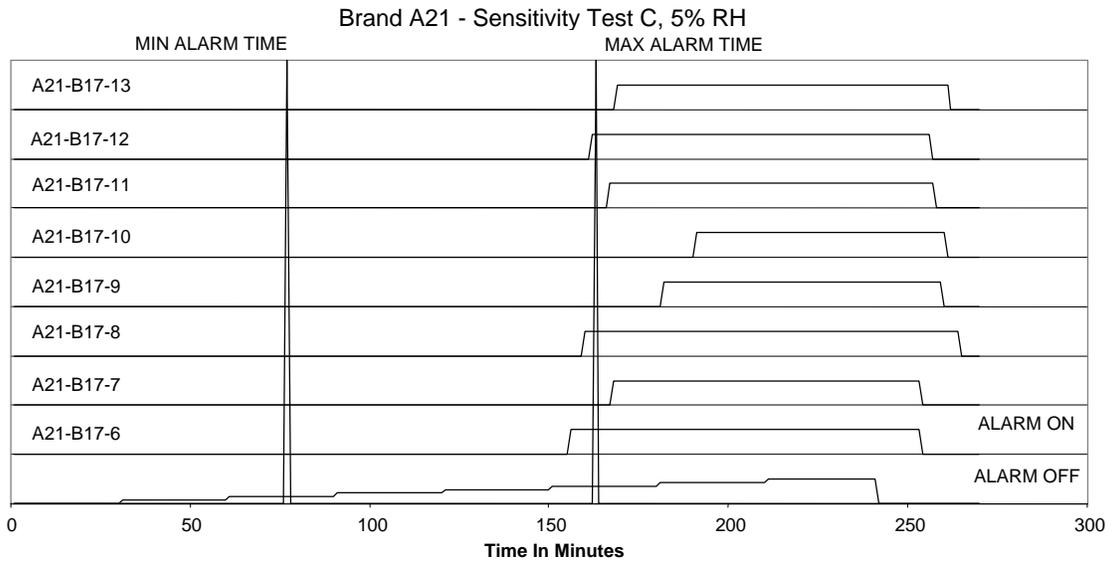


Figure 101 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. At 5% RH most alarms were insufficiently sensitive to sound within the 2.5% to 10% COHb limits.

Brand A21 - Sensitivity Test C, 50% RH

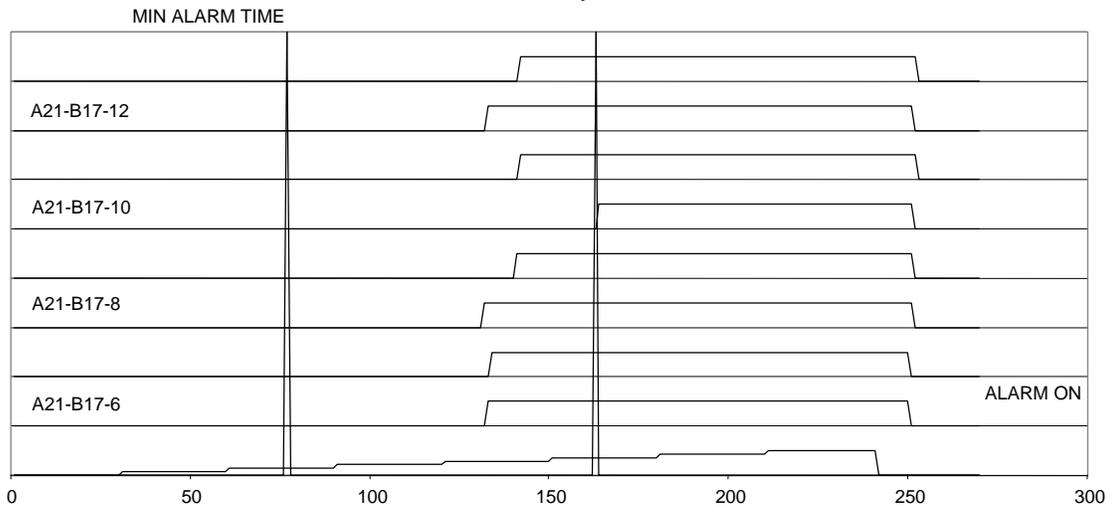


Figure 102 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. In contrast to their behavior at 5% RH, at 50% RH all alarms activated within specification.

Brand A21 - Interference Test C, 50% RH

	I-Butane	Methane	Ammonia	Methanol	Ethanol
A21-B17-13					
A21-B17-12					
A21-B17-11					
A21-B17-10					
A21-B17-9					
A21-B17-8					
A21-B17-7					
A21-B17-6					

Time in Minutes

Figure 103 No Brand A21 devices were sensitive to the interference gases presented.

	Acetone	I-Propanol	Ethyl-acetate	n-Heptane	Toluene	Trichloroethylene					
A21-B17-13											
A21-B17-12											
A21-B17-11											
A21-B17-10											
A21-B17-9											
A21-B17-8											
A21-B17-7											
A21-B17-6											
0	120	240	360	480	600	720	840	960	1080	1200	1320
Time in Minutes											

Figure 104 No Brand A21 devices were sensitive to the interference gases presented.

Brand A22

Like those of Brand A21, devices of Brand A22 were predominantly adequately sensitive at 50% RH, but failed to alarm at 5% RH:

- At 5% RH alarms were insensitive to CO for all concentrations tested and for exposures up to 30% COHb. At 50% RH when tested to fixed CO concentrations alarms sounded within the UL sensitivity limits.
- Alarms were not sensitive to any of the interference gases tested.

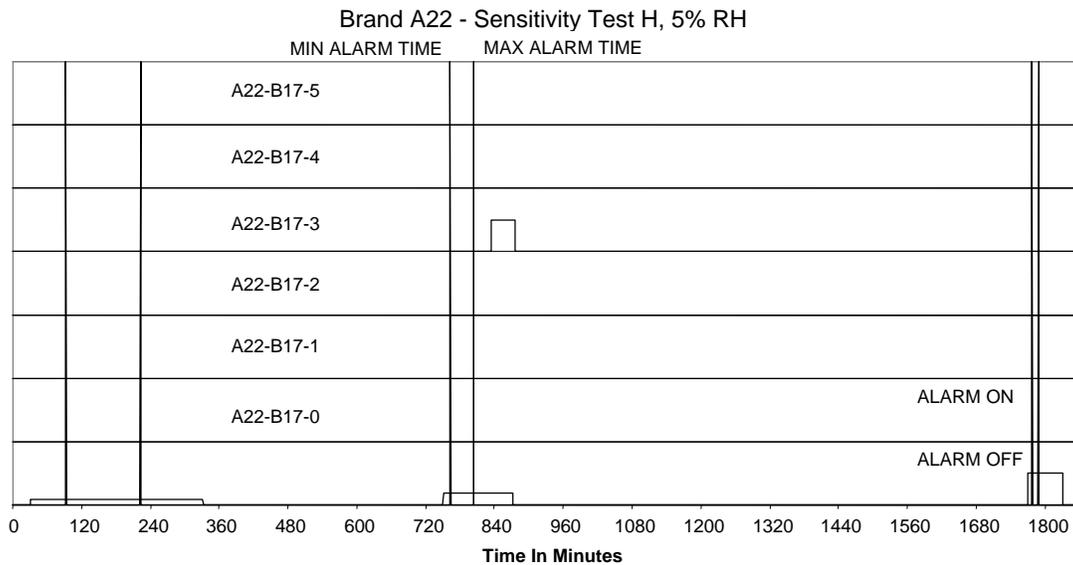


Figure 105 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. Brand A22 alarms were insensitive to CO at 5% RH, with only one activating in response to 150 ppm. No devices activated with 400 ppm at more than 30% COHb.

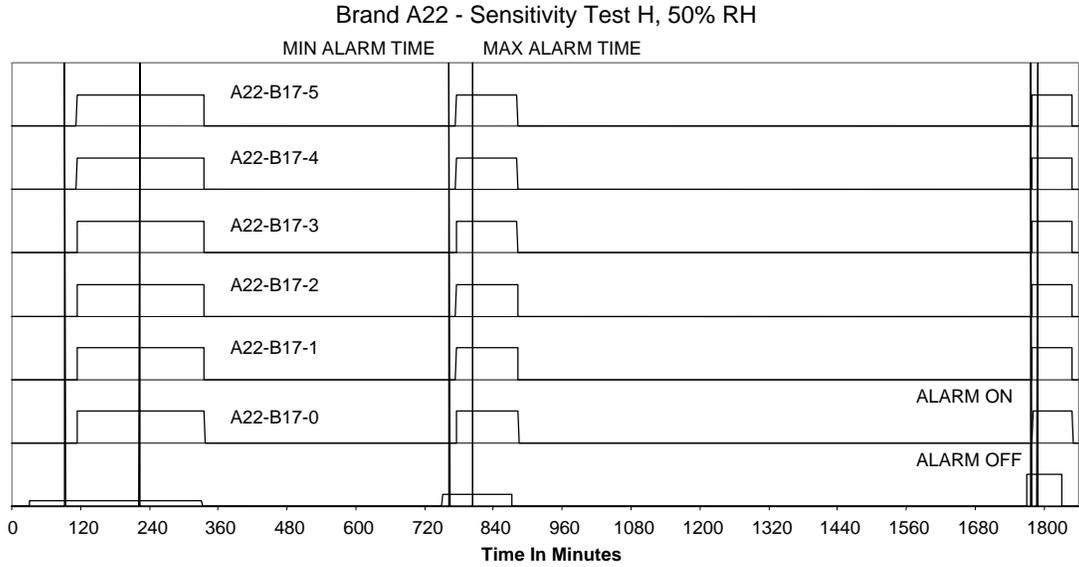


Figure 106 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. In contrast to their behavior at 5% RH, at 50% RH the alarms responded to the test concentrations within specification.

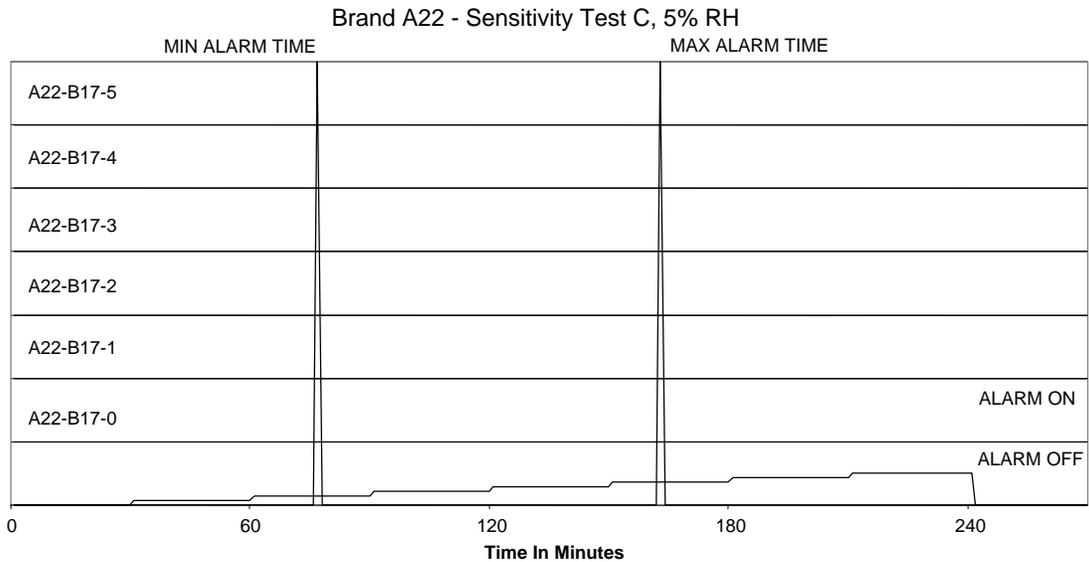


Figure 107 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. At 5% RH none of the Brand A22 alarms responded to gradually increasing CO concentration, up to a cumulative dose of 20% COHb.

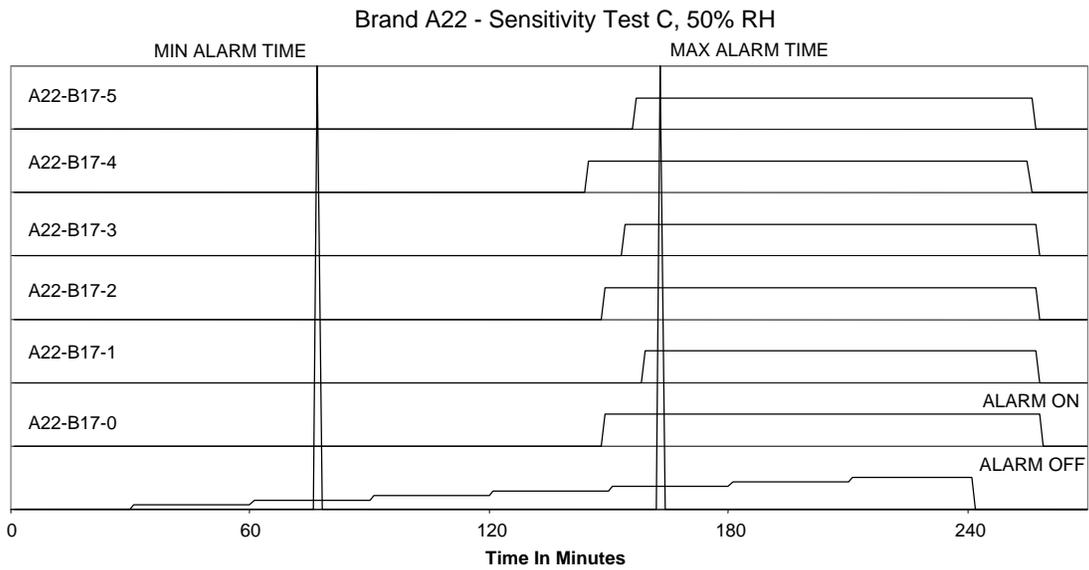


Figure 108 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. In contrast to their behavior at 5% RH, Brand A22 alarms were more sensitive at 50% RH, responding within specification to an increasing CO concentration.

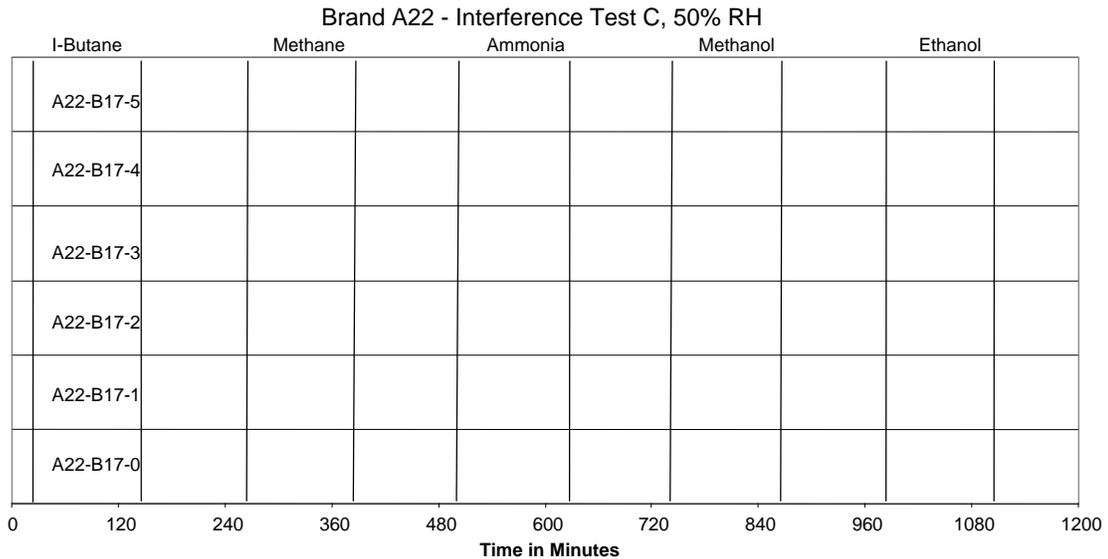


Figure 109 No Brand A22 alarms were sensitive to the interference gases presented.

Brand A22 - Interference Test D, 50% RH

	Acetone	I-Propanol	Ethyl-acetate	n-Heptane	Toluene	Trichloroethylene
A22-B17-5						
A22-B17-4						
A22-B17-3						
A22-B17-2						
A22-B17-1						
A22-B17-0						

0 120 240 360 480 600 720 840 960 1080 1200 1320

Time in Minutes

Figure 110 No Brand A22 alarms were sensitive to the interference gases presented.

Brand A23

Like Brand A14, the performance of alarms of Brand A23 was representative of electrochemical cell based alarms using counting/timing algorithms.

- At fixed CO concentrations, alarm activations were repeatable, and, except for one oddball device, occurred within the sensitivity requirements of UL 2034. However for a steadily increasing CO concentration activation times were more variable.
- One of ten alarms had dramatically less sensitivity to CO, failing to activate at less than 10% COHb for all CO presentations.
- Typical of electrochemical cell alarms, this brand showed no significant dependence of CO sensitivity on humidity.
- Alarms activated after exposure to isopropanol and to several successively presented interference gases. These activations likely resulted from filter breakthrough after becoming loaded with alcohol vapors (see Figure 116).
- After removal of the test gases during the fixed concentration tests alarms uniformly exhibited spurious activations. These activations resulted from logical errors in resetting counters or timers in the instruments' detection algorithm. Even though under the conditions of our tests these spurious activations only occurred after exposure to CO gas, their occurrence raises questions about the validity and reliability of the instrument's embedded software.

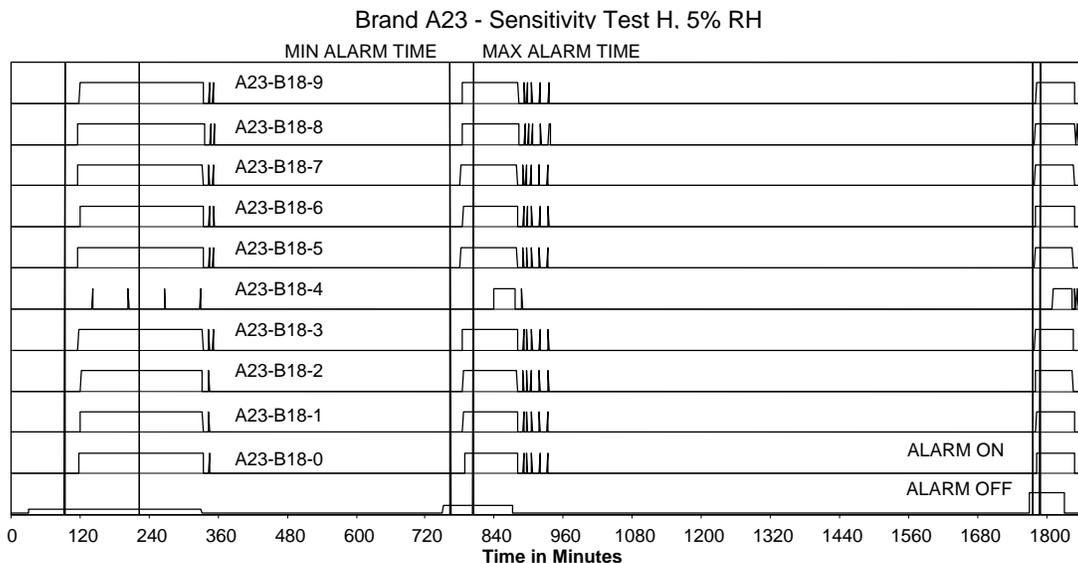


Figure 111 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms but one sounded within the 2.5% to 10% COHb limits. All alarms exhibited spurious activations after removal of the test concentration. These spurious activations are characteristic of a counter/timer algorithm that is not well implemented.

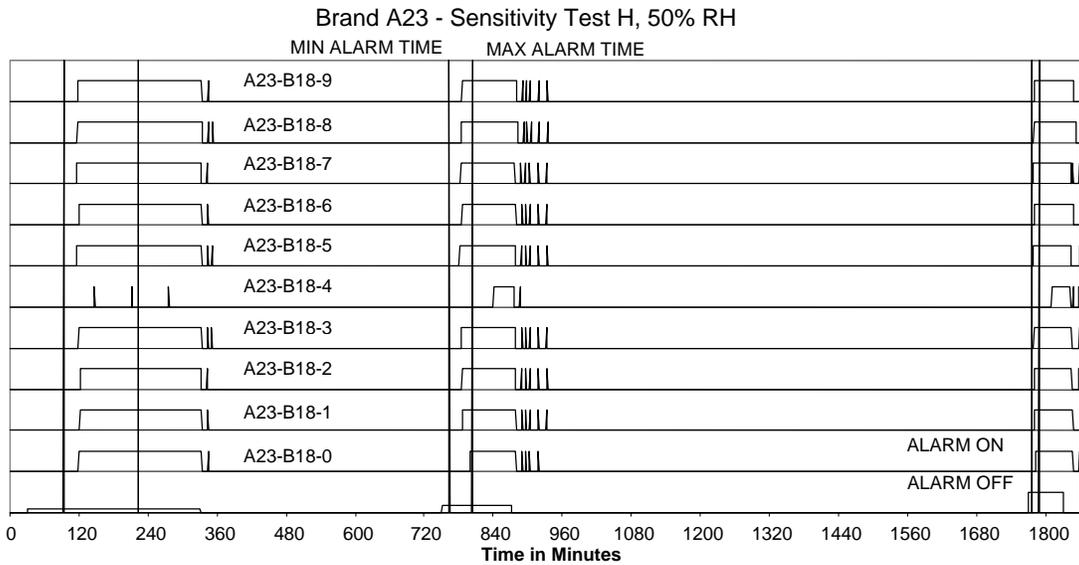


Figure 112 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. As they did at 5% RH, at 50% RH all alarms but one sounded within the 2.5% to 10% COHb limits. Again, alarms exhibited spurious activations after removal of the test concentrations.

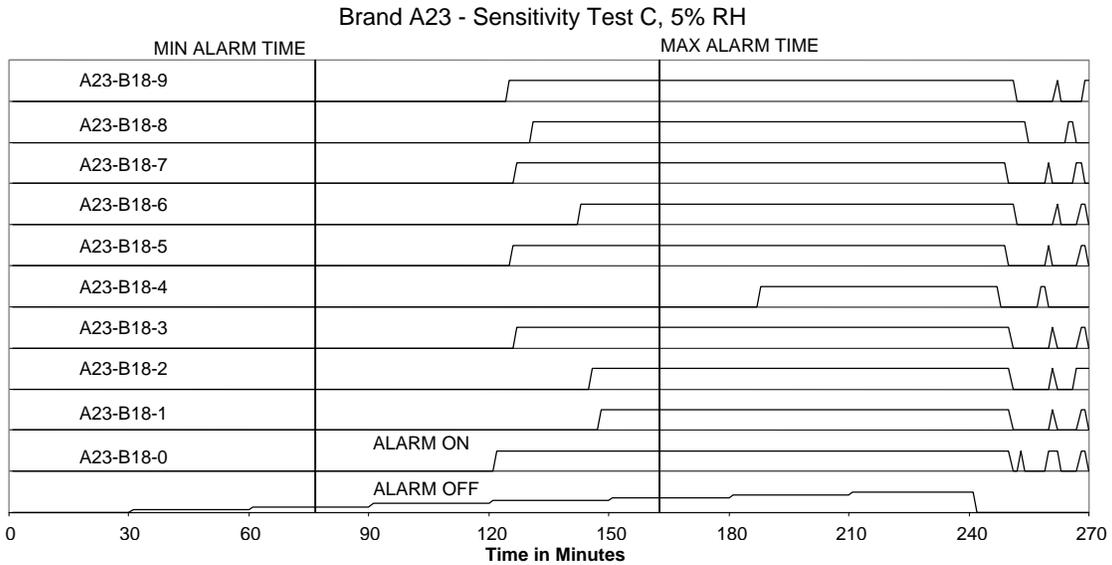


Figure 113 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms but one activated within the 2.5% to 10% COHb limits.

Brand A23 - Sensitivity Test, 50% RH

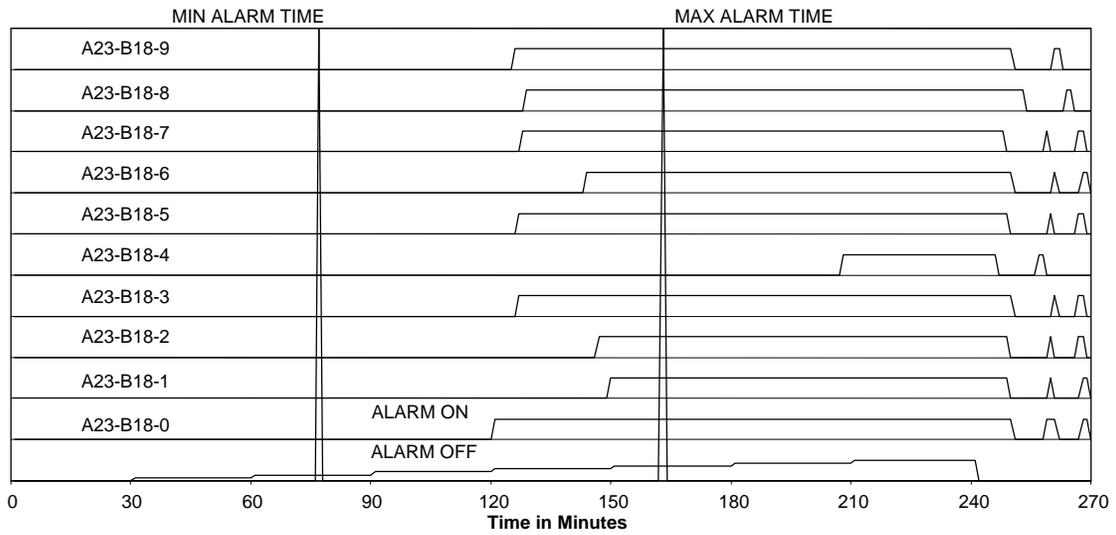


Figure 114 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. The peak COHb attained is approximately 20% COHb. All alarms but one activated within the 2.5% to 10% COHb limits.

Brand A23 - Interference Test C, 50% RH

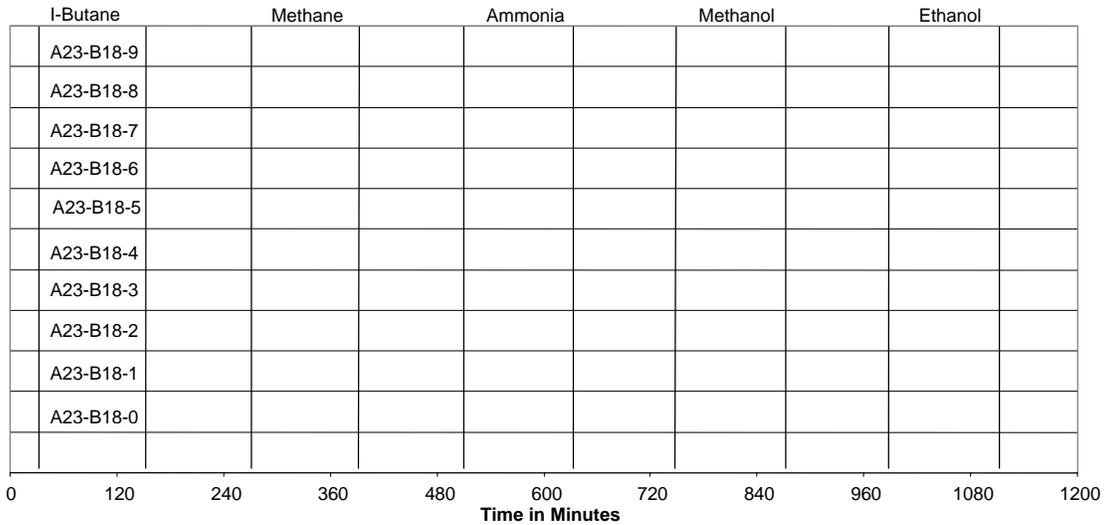


Figure 115 Brand A23 did not respond to these interference gases.

Brand A23 - Interference Test D, 50% RH

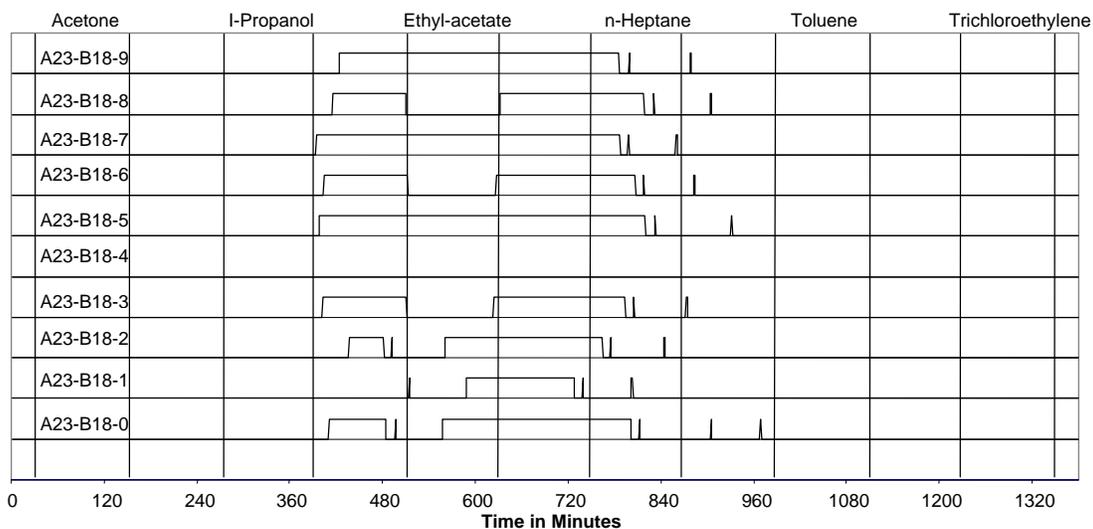


Figure 116 Most alarms of Brand A23 did respond to the interference gases isopropanol, ethyl acetate, and n-heptane.

Analysis and Recommendations

While a few of the brands tested were research devices, the majority were available commercially. In the following analysis we focus on these retail alarms.

Amid the diversity of failures exhibited in the graphs of alarm times are common behaviors shared by many of the alarm brands. In this section these failure mechanisms are categorized and tallied. Where possible, the physical mechanisms responsible for the failures are described and methods are proposed for their remediation. Thus in this section we provide an analysis of detector performance, an investigation of the causes of deficiencies, and recommendations for their remedy. We will examine these issues from the viewpoint of:

- Overall Reliability, Out-of-Box and Supervised Failures;
- Insensitivity to CO at Lower Relative Humidity;
- Inability to Integrate CO Exposure and Failure to Alarm for Varying CO Concentrations;
- Digital Display Inaccuracy; and,
- Sensitivity to Interference Gases.

Overall Reliability – Types and Numbers of Failures Observed

Table 12 summarizes the numbers of CO alarms that during the course of testing exhibited any of these failures:

- Supervised failures;
- False alarms in pure air;
- Failure to alarm at more than 10% COHb at one or more test concentrations, tallied separately for 5% RH and 50% RH; or,
- Alarm activation during exposure to interference gases.

Each entry in the table represents the number of detectors of each brand that experienced a particular type of failure. Also shown are the range of devices' digital display readings on exposure to 100 ppm CO for those brands including digital displays.

The failure categories tallied in Table 12 are not exhaustive. Only rather egregious failures are included; not listed are premature alarms, for example. These categories are also not mutually exclusive. For example, a device that failed to sound an alarm during a CO presentation may also have sounded its alarm while being exposed to clean air sometime during the several week test period. This detector would be included both in the false-alarm tally and the failure-to-alarm tally.

A summary table can not adequately capture the diversity of behavior seen in the alarms. For example, in the following table a failure-to-alarm is tallied if a device fails to alarm at any of the presented CO concentrations. While this is a serious problem, it is not as serious a problem as failure to alarm at all concentrations. Further, failure to alarm within specification (that is, at less than a cumulative exposure of 10% COHb) at 70 ppm is quite different than failing to alarm at all at 400 ppm. Because the quirks of individual brands are not easily summarized in any simple tally of failures, all the raw data for the devices

is also presented in the preceding chapter of this report. Readers are encouraged to examine the graphs of individual device's responses to enable more nuanced judgments of the performance of individual brands.

Table 12 Summary of Performance Numbers of Devices Exhibiting Failure Under Each Condition							
Brand ID ¹	Number Tested	Supervised Failures	False Alarms in Clean Air	Failures to Alarm²		Alarms with Interference Gas	Display Reading at 100 ppm CO
				5% RH	50% RH		
A4	4 ¹³	0	1	4	0	-- ³	25 – 160
A11	1	0	0	1	0	--	35 – 142
A12	6	0	0	0	0 ⁴	6	
A13 ¹¹	10	0	0	0	0	0	100 – 145 ¹⁰
A14	7	0	0	1	1	7	
A15	9	0	1	1	9 ⁵	9	0 – 400
A16	5	0	0	3	5 ⁶	4	
A17	10	0	0 ⁷	0	10 ⁸	0	
A18	10	1	0	0	0	0	85 – 138
A19 ^{11, 12}	13	4	0	0	0	0	92 – 100
A20	10	0	1	10	8	1	0 – 65
A21	8	0	0	8	1 ⁹	0	
A22	6	0	0	6	0	0	
A23	10	0	0	1	1	9	
Total Retail Alarms	81	5	3	30	28	10	
Total Alarms	109	5	3	35	35	36	
Failure Rate/Retail Alarms		6.2 %	3.7 %	37 %	35 %	12 %	
Upper Bound at 90% CL		12 %	8.3 %	46 %	43 %	19 %	

¹ Brands under development (A12, A14, A16, and A23) are shaded in the table, all others are commercially available and are unshaded.

² Failed one or more presentations of a UL 2034 Sensitivity Test concentration or during a progressive concentration test.

³ Interference gas tests were not performed for A4 and A11 devices. However, in prior tests brand A4 showed no sensitivity to the interferents.

⁴ All A12 devices did fail the CENELEC recovery test by failing to respond to 55 ppm shortly after 5000 ppm. However they did respond to all other CO presentations.

⁵ Oddly, all A15 devices (except one which continuously alarmed) failed to alarm at the 70 ppm concentration at 50% RH, but did alarm at the greater concentrations. Three failed to alarm with progressive CO concentrations at 50%.

⁶ Alarm activations were very uneven, with some alarming on time at 150 ppm but not alarming at 400 ppm until 30% COHb had been achieved.

⁷ Two devices exhibited simultaneous transient activations in clean air. On the possibility that these were caused by electrical noise they were not tallied here.

⁸ Alarms behaved consistently and well within specification at most CO concentrations, but at 400 ppm and 50% RH they alarmed consistently late, at approximately 25% COHb.

⁹ This one device just barely failed to alarm in time during the progressive concentration test.

¹⁰ Excluding one display that read 0 ppm but that was obviously nonfunctional otherwise.

¹¹ At the time of writing this report brands A13 and A19 are no longer commercially available.

¹² All retail brands except A19 are listed to UL 2034.

¹³ For the purposes of this summary retail and field units of brands A4 and A11 are summed together. It may appear that this would create selection bias: including in the tally the two field units that were withdrawn because they were defective would inflate the overall failure rate of the cohort. However, testing showed that the performance of the field units was not an aberration, but rather was identical to that of the controls of the same brand and hence characteristic of the alarm brand.

The final two rows of Table 12 show the sample failure rates and upper limits on the population failure rates at a 90% confidence level (CL) for each type of failure, for the commercially available alarms taken together. The research Brands not offered commercially were excluded from this computation. The upper limits at the 90% CL were computed using standard methods of statistical quality analysis, found in many textbooks. A detailed explanation of these methods as applied to CO alarms is provided in a prior report³¹ and in a reliability planning guide³² also included in this report as Appendix B. A similar guide is provided in the CSA 6.19-01 Standard for Residential Carbon Monoxide Alarming Devices.²²

The number of devices tested in this study would normally not be sufficient to characterize the low failure rates expected of life safety devices. For example, a population failure rate of only 1 in 100 could not be measured by testing only ten units; the most likely outcome of the test would be no noticed failures. However, the failure rates of the CO alarms tested were sufficiently great, of the order of tens of percent, that in our tests of a modest number of units there were a enough failures to accurately characterize the failure rate. It is likely however, that other failure modes were not noticed owing to their lower incidence.

The table does accurately convey the overall performance of the cohort of alarms – sample failure rates for the cohort vary from 4% to 37% depending on the failure type – but there are also many surprises, some of which are annotated in the table’s footnotes. For many alarms there is a consistent drop-off in sensitivity at low relative humidity. Even so, as a result of several countervailing effects, the overall failure rates at 50% and 5% RH are comparable. An example of one of these effects is provided by Brand A17. It worked well at 5% RH, and well for most CO concentrations at 50%, but at 400 ppm and 50% RH it consistently activated late, after a dangerous 25% COHb would have been attained in an exposed person.

It is apparent in the table that there are great differences in performance among the brands. Some brands behaved as expected: they complied with their UL specification for sensitivity to CO, they always alarmed on cumulative exposures less than 10% COHb, they were immune to the interference gases tested, and their digital displays accurately conveyed the CO concentrations. Others missed the mark on one or the other of these performance benchmarks, while some brands performed poorly.

The inter-brand diversity of performance of commercially available alarms is illustrated by the differences between Brands A19 and A20. Brand A19, while exhibiting some out-of-the-box supervised failures, met its specification in all tests, and accurately measured CO to well within 10% of value. It was the only commercially available brand that was not UL certified, yet it was among the few to fully comply with the UL Standard’s sensitivity requirements. It:

- Alarmed within the UL sensitivity specification immediately after purchase, at moderate and low relative humidity, and after exposure to interference gases;
- Always alarmed at less than 10% COHb, even when presented with steadily increasing CO concentrations (rather than with suddenly presented constant concentrations);
- Was immune to all interference gases tested;
- Produced digital display readings that were exceptionally accurate, typically displaying within $\pm 5\%$ of the true value, even after exposures to low humidity and interference gases;
- Exhibited one drawback – it had a high rate of supervised failure when purchased, with one third of alarms displaying error messages.

By way of contrast, Brand A20 was completely insensitive to CO at low humidity, and 80% of alarms failed to respond to at least one significant CO concentration at 50% humidity. Consistent with these failures, its digital display accuracy was very poor, reading between 0 and 65 ppm when exposed to 100

ppm. (The brand with the most inaccurate digital display produced readings ranging from 0 to 400 ppm when exposed to 100 ppm CO.)

Briefly, Brand A20:

- Alarmed for most, but not all, UL test concentrations at 50% RH, but became less sensitive at lower humidity, becoming unreliable at 20% RH and completely insensitive to the tested concentrations of CO at 5% RH;
- Predominantly failed to alarm at less than 10% COHb when presented with steadily increasing CO concentrations (rather than with suddenly presented constant concentrations), even at 50% RH;
- Displayed highly inaccurate concentrations failing to meet even the error specification stated in the user’s manual of +100% to -40% of reading. At 10% RH displays read low by 30% to 100%, and at 5% RH all displays failed completely.

Unfortunately, Brand, A19, was not widely available at the time this study was begun (it was available for purchase only over the internet) and is now unavailable, while A20, is well represented at numerous retail outlets.

For the cohort of commercially available alarms taken together more than a third of devices failed to alarm on exposure to one or more dangerous CO concentrations. As Table 12 shows, 37% and 35% of devices tested (or 46% and 43% at a 90% confidence level) failed to alarm at 5% and 50% relative humidity respectively. But these statistics are not equally applicable to individual brands; the brands were not uniform in their performance, but rather fell predominantly into two camps – those that worked well, and those that performed poorly.

Table 13 breaks down the failure rates of the commercial brands of Table 12 into these two categories: three brands that performed well, and six brands that performed poorly.

Table 13 Failure Rates of Well Performing vs. Poorly Performing Brands					
Brand Cohort*	Supervised Failures	False Alarms in Clean Air	Failures to Alarm		Alarms with Interference Gas
			5% RH	50% RH	
Well Performing Brands¹					
Observed Failure Rate	15 %	0 %	0 %	0 %	0 %
Lower – Upper Bound at 90% CL	7.3 – 28 %	0 – 7 %	0 – 7 %	0 – 7 %	0 – 7 %
Poorly Performing Brands²					
Observed Failure Rate	0 %	8 %	79 %	47 %	30 %
Lower – Upper Bound at 90% CL	0 – 6 %	3 – 18 %	63 – 100 %	34 – 66 %	16 – 39 %

¹ Well performing brands comprise 33 devices from three brands (Brands A13, A18, and A19).

² Poorly performing brands comprise 38 devices from six brands (Brands A4, A11, A15, A20, A21, and A22).

* One brand, A17, is excluded from the table as it does not fit clearly into either classification.

The three well performing brands always alarmed within the UL 2034 specification. Most importantly, they never failed to alarm on exposure to potentially dangerous concentrations of CO. Consequently, their failure-to-alarm rate was 0%.

Among the other six brands a significant portion of devices of each brand failed to alarm at dangerous CO concentrations. Taken together, the portion of devices of these brands that failed to alarm at one or more

dangerous concentrations was 79% at lower relative humidity and 47% at greater humidity. As a group these brands were also more sensitive to interference gases, with 30% of devices alarming, and more prone to false alarms, with 8% actuating in pure air.

Because the sample size is not large (only 33 devices in the well performing cohort) the upper bound on the failure-to-alarm rate at a 90% confidence level is rather great (7%) even though no failures were observed. Consequently, with 90% confidence we can only say that the observed failure rates of 0% are consistent with true population failure rates of 7% or less. The true rate may in fact be small, but we are only able to establish an upper bound of 7%. This is a reflection of the statistical difficulty in establishing an upper bound on low failure rates for small sample sizes.

This statistical difficulty disappears when sufficient failures are tallied, even for small sample sizes. For the poorly performing brands, failure rates are great enough that even a small sample size is sufficient to establish relatively narrow upper and lower bounds on the population failure rate. For example, the portion of devices that we observed to fail to alarm at 50% relative humidity was 47%. While a different sample of devices would have undoubtedly yielded a different failure rate, we know at a 90% confidence level that the true population failure rate lies somewhere between 34% and 66%.

One brand that does not fit neatly into the scheme of Table 13, Brand A17, performed well on most tests, but in one test at 400 ppm it actuated late, with all devices alarming after approximately 25% COHb would have been achieved in an exposed person. That all alarms of this model behaved identically prompted us to double check the experimental apparatus, but we were unable to find any experimental error that could account for the data. Our best hypothesis is that the alarms consistently actuated late as a result of some as yet undetermined circumstance of their history of exposure. If this brand had been included in the poorly performing cohort the most significant change to Table 13 would be that the cohort's failure rate for failures to alarm at 50% would have increased to 58% (46%-75% at a 90% CL) from 47% (34%-66% at 90% CL).

Out-of-Box, User-Diagnosable and Supervised Failures

Three of the ten commercially available brands showed out-of-box failures, prompt false alarms, user-diagnosable or supervised failures:

- Brand A15: One device produced an initial false alarm. When powered up it displayed 116 ppm in clean air and activated its alarm several hours later, with the display reading climbing to 160 ppm. The alarm never recovered.
- Brand A18: One device displayed an unresettable supervised error message five days after it was powered.
- Brand A19: Within two days of being turned on three devices displayed an error message and one did not display properly and could not be reset.

Our observations of prompt user-diagnosable failures in this study are consistent with the results of a prior study which found that a cohort of nine brands exhibited an overall failure rate of 10% for user-diagnosable failures, and an additional 5% for prompt-but-silent failures.³³ While that study found the defects to be distributed among five of the nine brands tested, this study found these defects concentrated among three of the ten commercially available brands tested. The overall sample size used in these studies is not sufficient to determine if the prompt defect rate has changed significantly between the two studies.

While user-diagnosable and supervised failures do not compromise the safety of the alarms, they do cause a nuisance for consumers and a warranty cost for manufacturers.

The Need for Reliability Testing

As Table 12 and Table 13 indicate that the overall failure rates for many CO alarms is in the tens of percent. As these tests were performed on newly purchased units the measured failure rates are best case estimates; we may expect that as the alarms age their failure rates are likely to increase.

The measured failure rates are consistent with those determined by other testing laboratories over the course of several years,^{7, 8, 9} including Underwriter Laboratories tests.¹⁵ These observed failure rates are surprisingly great for life safety devices.

The failures observed in these and earlier tests point out the need for distinct requirements in the standard for reliability *prediction* and reliability *testing*. Reliability *prediction* is an exercise in which a projected whole instrument reliability (in terms of MTBF) is computed by combining the known failure rates of individual components. While reliability prediction is a valuable design-time exercise, the most unreliable component of the alarm, the sensor, must be excluded because, as a newly developed chemical sensor, reliability data is not yet available for it. Consequently, the true field reliability of the alarm can not be credibly predicted. As the CPSC pointed out in their October, 1996 recommendations,

“[we] do not believe that designing to a predicted failure rate ensures detector reliability in the field.”¹⁹

For this reason reliability *testing* is also needed, both at the time-of-manufacture and throughout the projected lifetime of the device. The CPSC recommended that the UL standard,

“require each manufacturer to perform long term stability and reliability testing on each UL-listed detector model... [the criteria] including:

- *Minimum sample size;*
- *Acceptable failure rate criteria;*
- *Requirement to test detectors periodically for a total time period not less than the expected life of the detector; and,*
- *Suitable UL oversight and inspection procedures.”¹⁹*

These recommendations to UL are reiterated here. In response to these recommendations, the CSA 6.19-01 Standard for Residential CO Alarms has incorporated requirements for time-of-manufacture and in-service reliability. In Appendix B of this report we provide detailed guidance on the statistical sampling and testing methods that may be used to assure compliance with those requirements.

Insufficient Sensitivity to CO at Lower Relative Humidity

Many of the brands tested, specifically those based on semiconductor and colorimetric technologies, became less sensitive to CO when the ambient air contained low relative humidity.

Test Humidity

To determine alarms' sensitivity to humidity, CO was provided in a background of clean air containing one of several fixed *absolute* humidities. For tests at moderate humidity the carrier air contained 13,200

± 1,000 ppm water vapor (nominally 50% RH at 72°F), while for the tests at low humidity a controlled concentration of 1,320 ± 200 ppm (nominally 22%RH at 0°C, or 5% RH at 72°F) water vapor was used. In performing these tests the absolute humidity, not the relative humidity, of the carrier air was controlled. We have found that, at least for semiconductor-sensor based devices, their sensitivity to CO is influenced not by relative humidity, but by absolute humidity, that is, by the ppm concentration of water in the ambient air.

These levels, nominally 50% and 5% RH, were chosen to enable a quick assessment of the humidity dependence of the alarms. For some alarms with many failures at the lower humidity we further tested at intermediate water vapor concentrations of 2,640 and 5,280 ppm corresponding to 10% and 20% RH at 72°F.

In this study we did not test alarms at high humidity. Prior studies have shown that false alarms are exacerbated at high humidity causing field failures during morning fogs or on rainy days.⁶¹⁻⁶⁴

The Test Humidity Required by UL 2034

The lowest humidity level tested, 1,320 ppm, still provides a greater absolute humidity than that required by tests of the UL 2034 standard. UL 2034 requires that CO alarms meet their sensitivity requirement when operated at 0°C and 15±5% RH³⁴. This requirement corresponds to a lower limit of less than one half of the absolute humidity we provided (our lowest test humidity is the equivalent of 22% at 0°C). While UL 2034 requires that alarms operate at water vapor concentrations down to 600 ppm, the lowest absolute humidity we were able to reliably provide was 1320 ppm.. We were not able to duplicate the more extreme conditions of temperature required by the UL specification, but by operating solely at room temperature we were able to control absolute humidity to a level of about twice the minimum required by the UL standard.

For comparison, the absolute humidity required of the UL tests and these current tests are shown in Table 14. In short, the UL tests require operation over an absolute humidity range of nearly one hundred to one, while the humidity in these tests varied over a ten to one range, extending from a lower humidity of twice the lower test humidity of UL 2034, to an upper humidity of only one quarter of the UL 2034 upper test humidity.

Table 14 Humidity Range Required in Test				
	UL 2034 Tests*		Current Tests	
	Controlled Condition	Absolute Humidity	Controlled Condition	Absolute Humidity
“Low” Humidity Level	15 ± 5 % RH at 0 °C	0.46 mmHg (603 ppm)	1320 ± 200 ppm at 22 °C	1.00 mmHg or (1320 ppm)
“High” Humidity Level	40 ± 10 % RH at 40 °C	44. mmHg or (57,900 ppm)	13,200 ± 1,000 ppm at 22 °C	10.0 mmHg or (13,200 ppm)

* UL 2034, Section 45.1.2

It should be noted that the UL standard requires operation at even lower humidity for alarms intended for operation in recreational vehicles. UL 2034 (revised Oct. 1998) Section 69, (RV applications) requires testing for CO sensitivity at -40°C, at an unspecified humidity, after preconditioning for 72 hours. Although the humidity is unspecified, we might assume a relative humidity of 50%, which is an absolute

humidity of only 64 ppm, or 0.048 mmHg, a full order of magnitude lower than the lowest values of Table 14.

Should CO Alarms Be Able to Operate at Low Relative Humidity ?

We can only speculate as to why UL 2034 requires operation over such a wide range in humidity, and particularly to such a low humidity. Presumably, and understandably, the developers of the standard sought to assure the operation of alarms in residences subjected to low winter humidity levels. For the tests reported here, the nominal 5% level was chosen as a lower bound on levels likely to be found in northeastern, midwestern, and Canadian homes in the winter. When cool outside air containing 20-25% RH at 10° F is heated to 72° F its absolute humidity remains constant while its relative humidity drops to the 2-3% range. Air colder than 10° F is even dryer. In most residences there are multiple sources of humidity (e.g., cooking and bathing activities) that may contribute 0-2 mmHg additional water vapor pressure to the indoor environment. These sources increase the internal relative humidity to a more comfortable level (typically 15-30% RH in the winter). However, depending on the residence's air exchange rate, the level of humidity generating activity, and the time of day, the humidity fluctuates widely. Consequently 2-3% represents the lower bound on a highly variable humidity. For the purposes of this study we provide an exposure to double this lower limit as an exposure that may be representative of the low humidity that is sometimes obtained in a residence in the middle of a midwestern winter.

Consequently, the humidity range tested in this study lies within a range typical of residences in the winter. It seems reasonable that CO alarms, targeted for installation in these residences, should be able to warn of CO hazards at the expected humidity.

Some have argued that operation of alarms at low humidity is not needed because the production of CO is obligatorily accompanied by water production, guaranteeing a moderate humidity level. While this may be true of CO production from some fuels under some combustion conditions, it is not true in general. Indeed, CO is produced prodigiously, without water vapor, in many of the most common poisoning scenarios:

- Because charcoal briquets are composed primarily of carbon with little hydrogen, their burning produces insignificant water vapor but rapidly turns out deadly quantities of CO, making it, as the CPSC has pointed out, a particularly insidious cause of CO poisoning.³⁵ Fully 10% of CO poisoning fatalities from 1991 to 1995 were caused by charcoal grills alone. Coal, which also produces no water vapor, and wood heating caused additional deaths.³⁶
- In fireplace fires, particularly in their terminal stages, CO is produced with no water vapor from the remaining hot embers. It is during this stage of the fire that homeowners too frequently close the damper, fail to fully extinguish the embers, and turn in for the night, not understanding the danger from CO poisoning.
- A preponderance of non-fire CO deaths are caused by CO produced in the winter from a running automobile in an attached garage. CO in the auto exhaust diffuses into the attached residence. During this diffusion the water vapor is left behind, condensed in the cold garage and on cold exterior surfaces, subjecting the residents and their CO alarm to dangerous levels of CO unaccompanied by any increase in humidity.

The Observed Poor Sensitivity to CO at Low Humidity

Of the commercially available brands tested, five brands, A4, A11, A20, A21, and A22, significantly lost sensitivity to CO at lower relative humidity. The instruction manuals for some of these alarms tell the

user not to install them where the temperature is colder than 40°F or hotter than 100°F, implying that the devices should be functional at temperatures within that range. The 2640 ppm humidity level tested, corresponds to approximately 30% RH at 42°F, and is well within the range claimed in the unit's user's manuals. Even so, these alarms predominantly failed to sound at this humidity.

Brand A20 (see the brand-by-brand summary graphs) provides a typical example of a CO alarm whose sensitivity drops off as absolute humidity is reduced. At 50% RH nine of ten alarms activate at less than 10%COHb when tested to the UL test concentrations. But that drops off to 8 of 10, 0 of 10, and 0 of 10 at 20%, 10% and 5% RH respectively. For the progressive concentration test only two alarms activate at less than 10% COHb at 50% RH; none activate at lower humidity.

The humidity dependence of Brands A4 and A11, by a different manufacturer, are similar. For these brands we performed a test in which the humidity was continuously varied. The results of one such test is shown in Figure 117.

In this test the alarms were exposed to a fixed concentration of 200 ppm CO and their display readings recorded as humidity was first decreased from 80% RH to 5% RH, then gradually increased again. As can be seen in the Figure, the prompt response to a humidity change is dramatic. The reported CO concentration varies over a three to one range as humidity is varied, with the consequence that some alarms are greatly oversensitive at higher humidity, while others are insufficiently sensitive at lower humidity. Most of the change occurred in the 10% to 50% range.

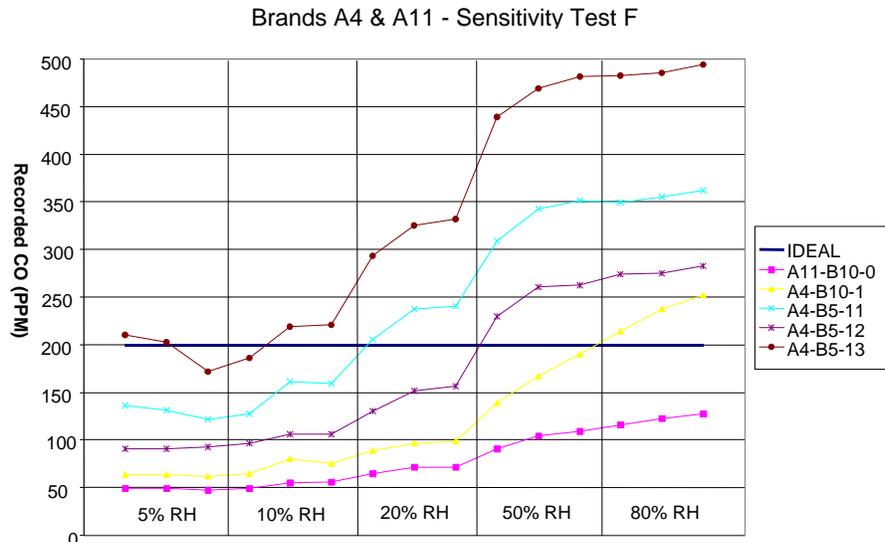


Figure 117 After conditioning at 5% RH alarm digital display readings in response to 200 ppm CO were measured at relative humidity of 5%, 10%, 20%, 50% and 80%. Each relative humidity was presented for 90 minutes and displays were read every 30 minutes. The alarm sensitivity is strongly dependent on relative humidity, with readings varying over a three to one range.

Two of the alarms subjected to this test had been withdrawn from service after a CO poisoning incident in which a car had been left running in an attached garage. Because they had failed to alarm they were suspected to be inadequately sensitive under the dry winter conditions that prevailed at the time. That suspicion is borne out by these tests that find, indeed, that these alarms are insufficiently sensitive to be protective at low humidity.

The Cause and Remediation of CO Alarms' Temperature and Humidity Dependence

It is not surprising that the sensitivity of CO sensors depends on humidity. In fact, it has been well known that both semiconductor and colorimetric reagent sensors are humidity sensitive, and that unless properly compensated this dependence would constitute a major source of unreliability for CO alarms.³⁷

While the humidity dependence of colorimetric sensors is poorly understood (at least in published scientific literature), that of semiconductor sensors has been well characterized. The general response to humidity and reducing gases of many semiconductor sensors, and particularly the tin oxide sensor most widely used in CO alarms, is of the form:

$$\frac{R_s}{R_o} \propto \left(1 + \sum_j K_j \prod_i [G_i]^{n_{ij}} \right)^{-b} \quad \text{Eqn. 2}$$

in which R_s represents the resistance response of the sensor, R_o is its response in the absence of gases, and there is a sum of products of gas concentrations G_i , each product term corresponding to a specific, independent reaction route for the oxidation of the gas on the semiconductor surface. A derivation of this response from fundamental physical principles has long been available.^{38,39} The dependence of response on water vapor is complex, but fully described by the equation. As one effect, water behaves similarly as other gases in the response, so that the sensor is independently sensitive to water vapor alone. But it may also participate in various product terms, so that the sensitivity of the sensor to a particular gas is sometimes mediated by a codependence on water. This is the case for CO detection, and the cause of the profound humidity and temperature dependence of semiconductor-based CO alarms.

As complex as the humidity dependence may be, Eqn. 2 provides a powerful basis for understanding it, and for compensating the CO response of semiconductor sensors for the effects of ambient temperature and water vapor.

The sensor's sensitivity to absolute humidity is two-fold. First, there is generally a single dominant reaction with water vapor alone, reflected in the equation as a single product term. This reaction determines the sensor's sensitivity to ambient humidity and temperature when there are no other gases present. Secondly, for the detection of CO on semiconductor surfaces, and in particular on the tin oxide sensors used in residential CO alarms, there are generally two reaction routes of different stoichiometries, but with both involving adsorbed water.^{39,40} It is well known that CO is oxidized on metal oxide surfaces in coordination with pre-adsorbed hydroxyl ions, accounting for the water dependence in the CO terms in the equation. Further, in this equation there are two terms for CO, one corresponding to a reaction involving a single adsorbed CO molecule and one involving two adsorbed molecules. For example, in some semiconductor sensors one could see the first order response at lower CO concentrations and the transition to the second order response at greater concentrations. For some popular semiconductor sensors used in commercially available alarms the first order response seems to be negligible and only the second order response is observed, for others the opposite is observed.

Consequently, considering only the response to CO and water vapor, many tin oxide sensors have a response represented by a simplified form of Eqn. 2, containing three product terms, one for the detection of water alone, and two for the two reaction routes responsible for CO sensitivity.

The sensor's dependence on ambient temperature is mediated primarily through the sensor's sensitivity to absolute humidity rather than to temperature *per se*. At a constant relative humidity, absolute humidity changes quite rapidly with ambient temperature, following an exponential curve, as,

$$[H_2O]_{ppm} = 1.40 \times 10^{11} \frac{RH}{100} e^{(-3926 / (T+231.4))} \quad \text{Eqn. 3}$$

where T is the ambient temperature in °C, RH is the relative humidity, and ambient pressure is assumed to be one atmosphere. Because the sensor is directly sensitive to absolute water vapor concentration, and because its CO sensitivity is also proportional to absolute humidity, the response is generally profoundly sensitive to ambient temperature through Eqn. 3. There are other mechanisms of temperature dependence, but they are generally masked by this much greater dependence mediated by absolute humidity.

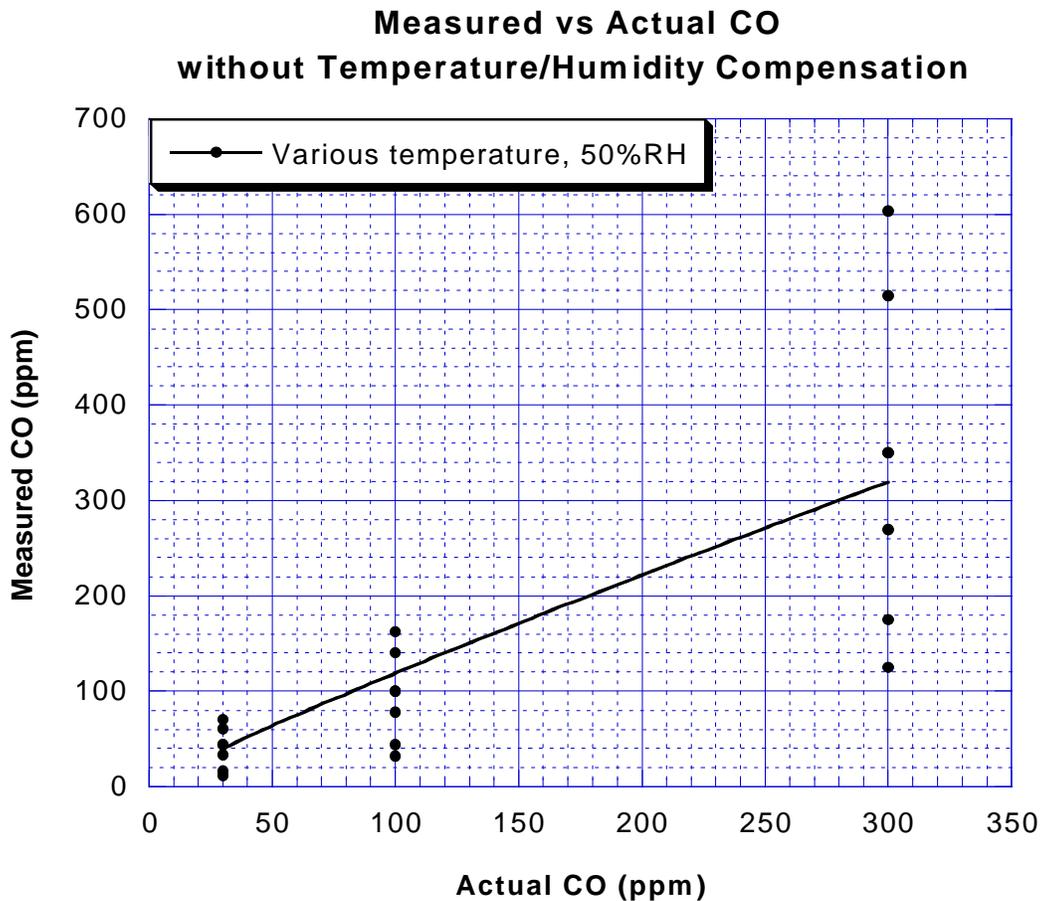


Figure 118 After a single point calibration CO concentration is computed from measured sensor response (R_s) at various temperatures (ranging from -5 to $+50$ °C) and at a nominal 50% RH. The uncompensated influence of the temperature variation is reflected in the wide range of estimated CO, spanning a factor of five to one at each presented CO concentration. The data are taken from the manufacturer's data sheet. Details of the analysis are available from Mosaic Industries.

These equations may be combined to form a temperature and humidity compensation equation for semiconductor sensors used in CO alarms. The efficacy of such a compensation scheme is illustrated by Figure 118 and Figure 119.

Figure 118 shows a typical response of a commonly used semiconductor sensor to CO at various temperatures, ranging from $-5\text{ }^{\circ}\text{C}$ to $+50\text{ }^{\circ}\text{C}$, at a nominal relative humidity of 50%. Data were taken from the sensor manufacturer's data sheet⁴¹ as sensor response as a function of ambient temperature at several CO concentrations and a fixed relative humidity of 50%.

In the uncompensated case, a single point calibration is performed using only the measured response of the sensor, R_s . The data are fit to a simplified form of Eqn. 2 that does not model the temperature and humidity dependence but does capture the nonlinear nature of the response to CO, using,

$$\frac{R_s}{R_o} \propto [CO]^{-b} \quad \text{Eqn. 4}$$

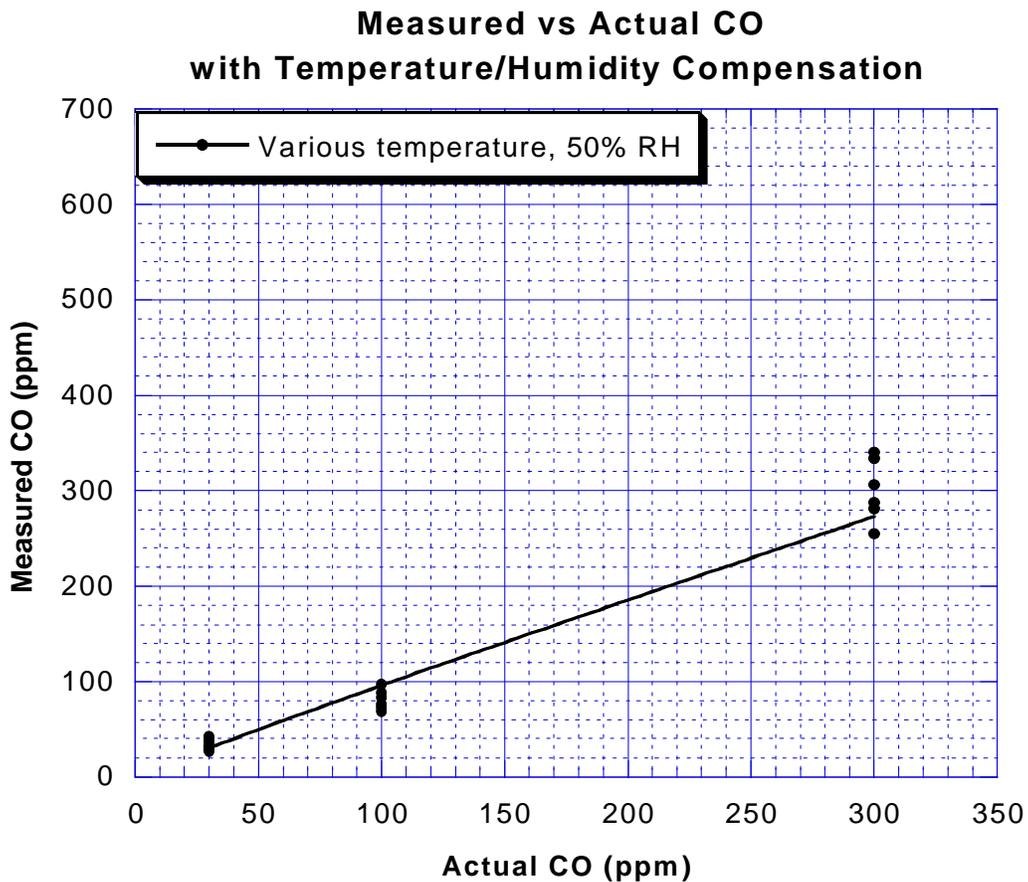


Figure 119 After a single point calibration CO concentration is computed from measured sensor response, R_s , known temperature, T , and known relative humidity at various temperatures (ranging from -5 to $+50\text{ }^{\circ}\text{C}$) and at a nominal 50% RH. This compensated measurement is far more accurate than the uncompensated measurement of Figure 118.

Carbon monoxide is measured with the sensor over the temperature range at concentrations of 30, 100, and 300 ppm and the measured concentrations plotted against the actual concentrations. As seen in Figure 118, without temperature and humidity compensation measurement accuracy is poor. At 300 ppm the measured concentrations range over a factor of five to one, roughly from 125 to 600 ppm.

Figure 119 shows the same measurement data, but with the sensor response compensated for temperature and humidity using Equations Eqn. 2 and Eqn. 3. Measurement accuracy is much better, with 300 ppm measured as between 255 and 340 ppm. Much of this residual variation likely results from an uncontrolled humidity. The data were taken from the sensor manufacturer’s data sheet, on which the humidity was listed as 50% RH. It is likely that the humidity varied point-to-point from this nominal ideal, and that that variation accounts for much of the remaining measurement error of Figure 119.

It is important to note that the apparent sensitivity of the response to ambient temperature in the Figures is accounted for solely by the variation of absolute humidity with temperature, as computed from Eqn. 3. The sensor’s direct sensitivity to ambient temperature is relatively weak, and it is masked by the much greater sensitivity to absolute humidity, which in turn depends strongly on temperature at any relative humidity.

The ranges of measured CO for each actual concentration for the two methods are given in Table 15. Uncompensated, the ratio of the greatest measurement to the least at a fixed true concentration is more than five to one (5:1); compensated, the range is reduced to 1.5 to one. These errors might be further reduced were it not for the unknown systematic and uncontrolled errors in the calibration data set.

Table 15 Temperature/Humidity Uncompensated vs Compensated CO Measurement		
Actual	Measured	
	Uncompensated	Compensated
30 ppm	12-70 ppm	28-42 ppm
100 ppm	32-163 ppm	69-97 ppm
300 ppm	125-603 ppm	255-340 ppm

Poor Integration of CO Exposure

Standards for CO alarms are predicated on the observation that CO is a cumulative poison. During exposure the human body slowly absorbs CO as it binds to hemoglobin in the bloodstream, converting it into carboxyhemoglobin (COHb). COHb slowly increases on exposure to CO, asymptotically approaching an equilibrium value proportional to the inspired CO concentration, and it slowly decreases on exposure to clean air. The clinical effects of CO poisoning closely track the concentration of COHb in the blood. Concentrations less than 10% are generally considered to be not acutely poisonous while the chance of long lasting damage or death increases steadily for greater concentrations. Consequently, a CO alarm’s activation should be keyed to cumulative exposure, that is, to a time-weighted average of CO concentration.

A consensus has developed among regulatory agencies, certification bodies, and alarm manufacturers that activation must occur before a hypothetical exposed person attains 10% COHb. Activation at this level or below should allow an exposed person to escape further danger; greater levels might incapacitate their

cognition to the degree that they would no longer be able to act to prevent further danger. UL 2034 and other CO alarm certification standards seek to assure that alarms properly integrate CO exposure and alarm at less than 10% COHb by specifying a sensitivity test. The intent of the standard seems to be that alarms should activate before the 10% level is attained, but not at so low a level that false alarms would be a problem.

To allow for some variation in the sensitivity of the alarm over time, the UL 2034 sensitivity test allows any alarm threshold level between 2.5% and 10%, for absolute concentrations greater than 150 ppm, and between 5% and 10% at lesser concentrations (to help prevent false alarms). The UL Standard assumes a particular time constant for the rate of CO uptake and sets alarm limits so that CO alarms always activate at less than 10 %COHb. Consequently, the UL 2034 sensitivity test measures times to alarm at several fixed CO concentrations, and mandates both minimum and maximum alarm times.

However, the UL sensitivity test does not simulate real exposures to CO. Because it tests to only a few, constant CO concentrations characterized by a sudden onset, it is possible for a manufacturer to build an alarm that “passes the test” without necessarily responding reliably to more realistic CO exposures. In residences, CO concentrations do not appear abruptly and remain constant as during the UL 2034 sensitivity test. Instead, CO concentrations must first increase from an insignificant concentration, and may vary and occur episodically.

A more representative but simple test of alarms’ ability to integrate CO exposure is to present a steadily increasing CO concentration, thereby simulating an actual CO poisoning scenario in which the CO concentration results from a source with a finite emission rate, rather than appearing suddenly. Indeed, this is the purpose of Sensitivity Test C used in this study.

Sensitivity Test C is more representative of realistic CO poisoning events that do not involve the abrupt appearance of a fixed concentration of CO. Instead, they are frequently initiated by a steadily increasing concentration. This concentration gradient has been predicted and observed in the many detailed theoretical and experimental studies that have been done to determine the rates of dispersion of CO in residences with a view toward developing a technical basis for alarm siting.^{67, 70, 73, 75} Predictive models have been generated for the rate of CO increase as a function of appliance type, flue blockage, and air exchange rate.⁷³ These models, as well as experimental measurements in test houses, show that in most acute poisoning scenarios CO concentrations rise steadily and nearly linearly.^{68, 69, 71, 72, 73, 74}

Detailed measurements in test houses using surrogate CO emitters, actual unvented sink water heaters, and actual radiant gas heaters with blocked flues reveal typical rates of CO increase of 50 ppm/hr with a range from 25 ppm/hr to 250 ppm/hr, over a period of approximately four hours.^{68,69,71,73,74} The progressive Sensitivity Test C of this study was designed to mimic these most likely poisoning scenarios. It uses the typical rate of CO increase of 50 ppm/hr.

Unfortunately, the results of tests using progressively increasing CO concentrations indicate that even when alarms activate in response to the unrealistically sudden application of the UL test concentrations they do not necessarily activate on exposure to other, more realistic, CO concentration profiles.

This point bears repeating: Many alarms that do conform to UL 2034 still do not alarm at less than 10% COHb when presented with increasing or varying CO concentrations. While the test points of the standard may be based on a 10% COHb level they are not sufficient to assure that alarms sound before 10%. In fact, we have found that they may not alarm until 20% or more is attained, depending on the time profile of presentation of CO. If it is a goal of the standard to assure that consumers are protected against COHb levels greater than 10%, then that goal is not being achieved.

It appears that there are two classes of alarms on the market. The first comprises several brands which do consistently integrate CO exposures and activate their alarms at a repeatable threshold for cumulative exposure, whatever the time course of CO concentration, while the second comprises brands that may activate at the fixed concentrations of the UL test, but not respond at less than 10% COHb when presented with varying CO concentration.

The brands that do reliably actuate at less than 10% COHb implement an integration algorithm that maintains a running average of CO concentrations, updated of the order of every minute. When the running average (the integral) exceeds a particular threshold the alarm activates. These algorithms actuate their alarms reliably whatever the time-profile of CO presented.

The non-integrating alarms generally use timer/counter algorithms geared to the UL tests. These algorithms start timers when particular threshold concentrations are exceeded, and actuate their alarms at threshold timer counts. The timers are reset when other threshold concentrations are subceeded. Unfortunately, it is difficult to implement a cumulative exposure alarm by setting and resetting timers. Timer/counter algorithms can be fine-tuned to behave properly for a few fixed concentrations, but they do not generalize properly to other concentrations. In particular, when confronted with varying CO concentrations these algorithms can become confused, and there are many simple, and realistic, concentration profiles for which they fail.

The Consequences of a Poor Detection Algorithm

It is sometimes easy to surmise the type of algorithm used by observing the differences in a device's sensitivity for different onsets of CO concentration. Of the fourteen brands tested, seven (Brands A14, A15, A17, A18, A20, A21, A22, and A23) show decreased sensitivity to varying CO concentrations as opposed to fixed concentrations.

For example, most devices of Brand A20 show adequate sensitivity to CO when presented with the fixed CO concentrations of the UL sensitivity test. Figure 120 shows that except for the failed device A20-B16-1 all alarms sounded uniformly at about 5% COHb when presented with 150 ppm CO at 50% RH. (The vertical lines during the 150 ppm presentation indicate the times at which 2.5% and 10% COHb would be achieved in an exposed person.) However, when presented with a gradually increasing CO concentration, as shown in Figure 121, their alarm times are highly variable, with device A20-B16-1 actuating only after 15% COHb – its sensitivity appears to have dropped by a factor of three. Of course the sensitivity of the alarm's sensor has not really changed, but the detection algorithm does not competently assess cumulative exposure if the CO concentration varies, as it undoubtedly does in realistic poisoning scenarios.

Brand A20 is not unique; many other brands exhibit this type of failure. All devices of Brand A14 alarm at about 2.5% COHb when presented with a constant 150 ppm (a UL test point), but alarm at about 10% COHb when presented with increasing concentrations. For Brand A15 eight of nine devices appear to be overly sensitive in the UL test, alarming at approximately 2% COHb at 150 ppm, but three fail to alarm at all at less than 10% COHb when subjected to the more realistic concentration profile. Likewise, Brand A17, on a %COHb basis, loses sensitivity by a factor of three between the two tests. We find that for the entire cohort of brands those that use timing/counting algorithms vary between the two tests by a factor of two to five in their apparent sensitivity.

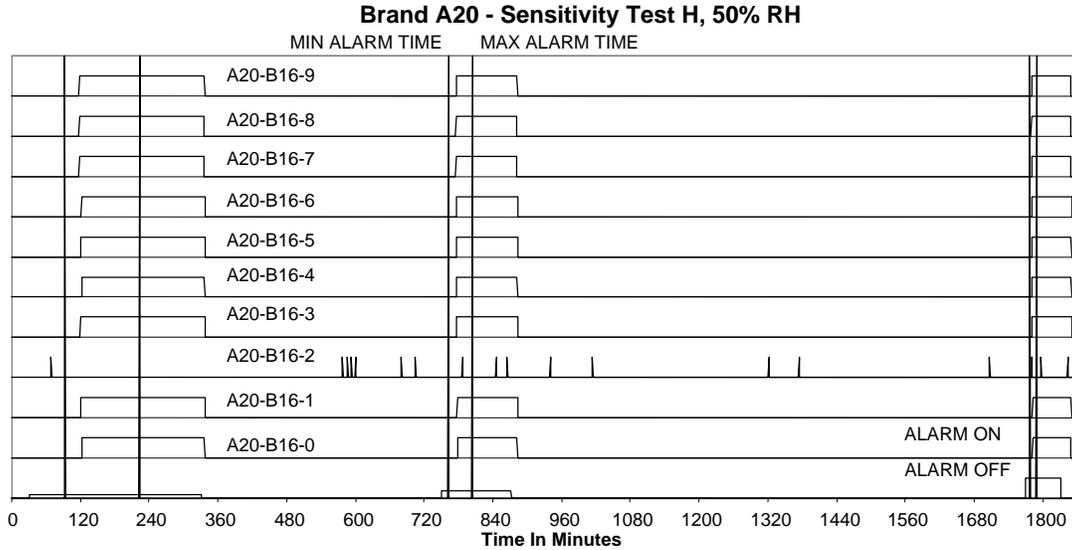


Figure 120 The bottom trace represents the times of presentation of 70, 150 and 400 ppm CO while for each presentation the two vertical lines represent the times at which 2.5% and 10% COHb were attained. All alarms but one activated within specification. Device A20-B16-2 intermittently activated.

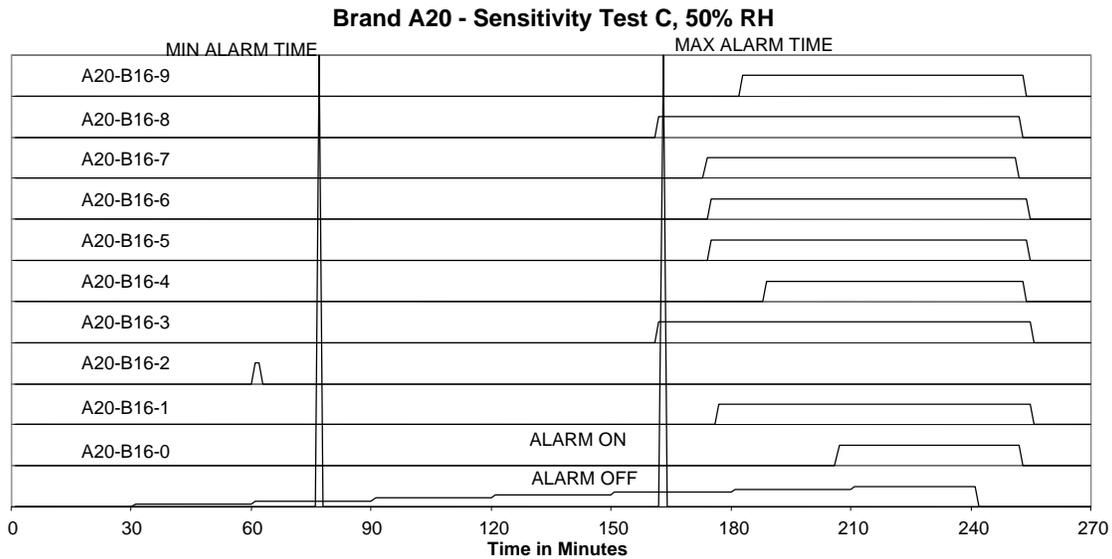


Figure 121 The lower trace shows the progressive CO concentrations from 25 ppm to 175 ppm in 25 ppm steps and the vertical bars indicate the times at which 2.5% and 10% COHb are attained. Although they do alarm in time at the fixed UL test points, at 50% RH and on exposure to continuously increasing CO, most Brand A20 alarms do not sound in time, and one fails completely.

There are other indications in the test results of timer/counter algorithms gone awry. For brands A14 and A23 for example, after the alarm ceases, there is often a period of time of intermittent alarms, always of the same pattern. These spurious activations are an artifact of the internal timing mechanism. While they do not seem to compromise the safety function of the device, as they only appear to occur after a valid

alarm activation, they do raise questions about software reliability, and suggest that further testing may uncover other problems. How would these alarms behave during more variable exposures – are their timers set and reset properly or do they become completely discombobulated?

Of course, knowing the details of the individual timing/counting algorithms used it would be possible to construct simple CO concentration profiles that show up the inadequacy of the detection algorithm. But that is beside the point. What is important is that the alarms often do not protect users from cumulative CO doses of much greater than 10% COHb for the CO concentration profiles experienced in a poisoning incident.

These tests reveal that even when alarms activate at the particular CO concentrations of the UL sensitivity test they do not necessarily activate on exposure to other life-threatening CO concentrations, and in particular, they may not activate when the CO concentration steadily increases, as it does in typical poisoning events.

Causes and Remediation

This observation, that even when alarms do comply with UL 2034's sensitivity specification, many fail to alarm at less than 10% COHb when presented with realistic CO concentration profiles, is not new. It has been discussed in CPSC Hearings⁴³, UL and CSA CO standards committee meetings, and has been observed in prior tests.⁴²

The cause has also been described and it is simple: Timing and counting algorithms do not produce an alarm threshold that corresponds to cumulative dose. In contrast, simple integration algorithms do alarm at a specific cumulative dose whatever the time course of the CO concentration. Whereas a timing/counting algorithm can be made to appear to work only for particular tests using constant concentrations, integration algorithms work for all possible presentations of CO concentration because they are biomimetic – they mimic the body's integration of carboxyhemoglobin dose.

Remediation of this problem is straightforward; it is to use integration algorithms similar to those used in the alarm brands found by this study to be the most reliable and well performing. As one manufacturer put it in testimony to the CPSC in 1996,

“... protection of consumers can be enhanced without altering the alarm thresholds of UL 2034 by the deployment of genuine time integrating detector algorithms. The overall level of protection by using such algorithms can be enhanced from a possible maximum exposure of 20% COHb to a consistent alarm at 10% COHb.

*It is recommended that UL 2034 be modified so that detectors are designed so as to interpolate between the alarm threshold points consistent with the 10% COHb exposure curve.”*⁴³

An integration algorithm maintains a running estimate of COHb levels. The fact that there is a well-established and simple mathematical relationship between CO concentrations, time, and blood COHb, invites the use of such an algorithm. Unfortunately, the use of such an algorithm is not widespread, perhaps due to a simple lack of information.

The only reliable way to assure that alarms actuate at less than 10% COHb is for actuation to occur in response to a running estimate of COHb levels. The fact that there is a well-known and simple mathematical relationship between CO concentrations, time, and blood COHb, invites the use of an algorithm in a CO alarm to continuously compute presumptive COHb in an exposed individual by time

integration, and to actuate an alarm before that level exceeds 10%. Indeed, the best performing and most reliable alarms on the market use this method; the poorest performing alarms do not.

Instead of this integration algorithm, many alarms use a single timer, or combination of timers, that initiate at particular threshold CO concentrations, and actuate the alarm after fixed time intervals.

These timing algorithms work for the fixed CO concentrations of the sensitivity test, but unfortunately they do not work in the more general and likely case of varying CO concentrations. In the field, CO concentrations do not appear abruptly and remain constant as during the UL 2034 sensitivity test. Instead, CO concentrations vary and often occur episodically.

Timing algorithms that work for the UL 2034 sensitivity test often give rise in more realistic CO exposures to consumers being exposed to COHb levels much greater than 10%. In fact, in the most simple CO exposure scenario, that of a gradually increasing CO concentration, our tests of alarms have shown that some do not alarm until 20% COHb is exceeded, even though they may comply with the UL 2034 sensitivity specification. Those alarms that truly integrate CO levels to produce a presumptive %COHb do appear to always alarm as expected, but those alarms that use other detection algorithms often do not.

A Recommended Integration Algorithm

The following discussion presents an exemplary integration algorithm designed to protect against exposure to greater than 10% COHb under all conditions of concentration exposure. Practical details of the algorithm are found in Appendix A.

The algorithm is immediately applicable to sensors that provide a prompt response to CO concentration, including semiconductor and electrochemical sensors. Some sensors on the market, specifically the colorimetric sensors, already integrate CO concentration. Unfortunately, they can integrate with only a single time constant, and do not integrate and de-integrate using two different time constants as recommended in the following algorithm. Even so, depending on the resolution of the sensor's measurement, the algorithm may be used to modify and enhance the sensor's native response by algorithmically providing the second time constant for de-integration.

The alarm algorithm converts a measurement of the CO concentration into an estimated %COHb once every operation cycle, where the operation cycle time should be less than 5 minutes. A maximum allowed cycle time ensures quick detection even at very high concentrations, and sufficiently accurate integration of CO concentrations to form good estimates of %COHb. While accuracy is improved by more frequent sampling, an advantage of long cycle times is that the alarm's processor may enter a sleep mode between samples to increase battery life. Manufacturers may make different trade-offs appropriate to the hardware platforms of their particular products.

The integration algorithm is based on the Coburn equation. Unfortunately, the form of the Coburn equation cited in the UL 2034 standard is not the most useful form so a more explicit and useful version is presented here instead. In the following discussion the difficulties of the UL-cited equation are briefly described, and a more sensible form is put forward that manufacturers can readily use to meet the sensitivity requirements of the standard. The physiological basis of the Coburn equation is detailed in Appendix A.

The form of the Coburn equation cited in UL 2034 suffers from the following drawbacks:

1. It is not clear how it can be used in a CO alarm;
2. It does not apply to varying CO concentrations;
3. It implicitly assumes an unrealistic baseline (endogenous) COHb level; and most importantly,
4. It does not apply to the general population, that is, the population the standard purports to protect.

These insufficiencies are summarized in turn:

1. The UL-cited equation is not useful as presented.

The UL 2034 Standard contains the following equation but provides no guidance regarding its significance or practical use:

$$\%COHb_t = \%COHb_o [e^{-(t/2398B)}] + 218 [1 - e^{-(t/2398B)}] [0.0003 + (\text{ppmCO}/1316)]$$

with $B=0.0404$ and t in minutes; the dimensions of other units are unspecified in the standard.

While this equation appears in conjunction with a graph of COHb vs Time (Figure 38.1 of the standard), it is not referred to in the text of the document. Providing this equation, while not referring to it, contributes to the standard's ambiguity over whether or not alarms are required to actuate before COHb levels of 10% are attained. Whatever the reasons for its inclusion, it is too inadequately specified to be useful. It contains undefined symbols with unspecified units – whether it is to be used, or how it is to be used, can not be discerned from the text of the standard.

2. It does not apply to varying CO concentrations.

The cited equation is a limited extrapolation of a steady state solution to the Coburn equation that predicts the COHb in response to a single, fixed, unchanging CO concentration. However, CO concentration is rarely constant. To be effective, a detection instrument must track the %COHb of a hypothetical exposed individual while the CO concentration continuously varies, and sound an alarm if the projected %COHb ever exceeds a threshold. For this reason it is not clear how this equation in UL 2034, that assumes only a constant CO concentration, can be used. Consequently, some brands of alarms do not use it, and instead use timing/counting algorithms designed to meet the sensitivity test of the standard at the test's specific concentrations. Unfortunately, as the tests of this study indicate, those algorithms often do not work at other CO concentrations, or for continuously varying concentration.

3. It implicitly assumes an unrealistic baseline (endogenous) COHb level.

The equation cited implicitly assumes that an exposed person's baseline, or endogenously created, COHb level is extremely low, only 0.065%. This level is far less than the normal physiological levels of 0.3 – 0.8 %COHb cited by many sources^{44, 45, 46, 47, 48, 49} for nonsmokers. Owing to the equation's inadequate description in the standard, that the endogenous level assumed is 0.065% is not apparent by inspection. Even so, it is easily derived, as shown in Appendix A. This low a level is worse than atypical, it is simply unrealistically low. Typical baseline COHb levels range from 0.2% to 0.8% with an average of 0.5% (as documented in Appendix A). Because assuming an unrealistically low baseline level biases the alarms toward lesser sensitivity, and hence lesser safety, a more realistic level should be used, as was recommended by the Consumer Product Safety Commission in October 1996.¹⁹ To date, it appears that UL has not heeded the CPSC recommendation that a more realistic baseline level be used.

4. It does not apply to the population the standard seeks to protect.

The U.S. Consumer Product Safety Commission (CPSC) found the COHb criterion (10%) seemingly implicit in the standard is “sufficiently protective”. However, the timing of alarms depends not only on the COHb criterion but also on the assumed kinetics of CO uptake. As pointed out by the CPSC in its recommendations³⁰ to UL, the kinetics of CO-uptake and elimination assumed by the UL standard is based on a single study of athletic young men engaged in strenuous exercise. Consequently it is restricted to a select, and unrepresentative, population. Aside from the technical discussion included in the CPSC recommendations, there seems to have been no effort made during the evolution of the UL standard to characterize the population to be protected with respect to their diversity of sensitivity to CO, and their varied uptake and elimination kinetics, even though the data for such characterization is readily available. As the equation cited by UL represents uptake and elimination for young men at hard work, it does not adequately model the general population, those involved in less strenuous activity, sleeping persons, or those whose activity level is already diminished by the effects of CO poisoning.

An alternative model, applicable to the general population, is put forward in the following section (and a detailed derivation is presented in Appendix A).

An Operationally Useful Coburn Equation

A more sensible, iterative form of the Coburn equation enables an alarm instrument to compute a running estimate of %COHb exposure, and to alarm if a threshold is exceeded. It is given as (see Appendix A for its derivation),

$$\%COHb_{t+\Delta t} = \%COHb_t \beta + (\%COHb_e + \alpha_o CO_{ppm}) (1 - \beta) \quad \text{Eqn. 5}$$

in which:

$\%COHb_{t+\Delta t}$ is the new estimate for %COHb;

$\%COHb_t$ is the estimated %COHb of the prior operation cycle, where $\%COHb_t$ of the first operation cycle is initialized to $\%COHb_e$;

CO_{ppm} is the CO concentration of the current operation cycle;

α_o is a numerical constant of value 0.166 %COHb/ppmCO representing the equilibrium ratio of blood COHb to inspired CO concentration;

$\%COHb_e$ is a numerical constant of value 0.5 %COHb representing endogenous CO production; and,

β is a weighting parameter that may vary from cycle to cycle and that depends on the operation cycle time, Δt , and a physiological time constant, τ_o , as $\beta = \exp(-\Delta t / \tau_o)$, in which τ_o equals either 97 minutes or 368 minutes. For example, for an update time of 3 minutes, and a τ_o of 97 minutes, the weighting parameter is $\beta = 0.9695$; for the same update time and a τ_o of 368 minutes, the weighting parameter is $\beta = 0.9919$. This weighting parameter is then switched between these two values depending on whether the %COHb is increasing or decreasing.

It is a recommendation of this report that this equation, based on sound physiological principles, be incorporated in Section 38 of the standard either to replace the equation now there, or to augment it.

This iterative equation correctly time-averages a CO dose no matter what the variation in CO concentration over time. The time-averaged dose is conveniently expressed in terms of the %COHb that would be present in a hypothetical exposed subject. Whenever this predicted %COHb exceeds a predetermined threshold (or trigger point) the alarm should then be activated. (Actually, two trigger

points with a small hysteresis is preferable to prevent alarm “flutter”.) The equation is easily implemented in low cost processors with little computational load.

This equation is made more protective of the general population by allowing the weighting factor, β , to vary depending on whether the projected COHb level is increasing (uptake) or decreasing (elimination), as discussed in Appendix A. For each iteration of Eqn. 5, if the CO concentration measurement of the current operation cycle is great enough so that $(\%COHb_e + \alpha_o CO_{ppm} > \%COHb_t)$ then a weighting parameter, β , corresponding to $\tau_o = 97$ is used, otherwise a β corresponding to $\tau_o = 368$ is used. These time constants, τ_o , correspond to the range in respiratory minute volumes (RMV) provided by the CPSC in its October, 1996 recommendations to UL (see Appendix A).

The Need for Consistent Alarm at 10% COHb

In our tests we have taken 10% COHb as the maximum alarm threshold. While UL has rejected recommended specifications that “CO alarms must actuate before 10% COHb is exceeded”, they do nevertheless assert that the test points of UL 2034 are “based” on the 10% level. CPSC’s technical analysis found this level to be “sufficiently protective”.¹⁹ As shown in Table 13, for the six poorly performing brands, a majority of devices failed to alarm in at least one test at this level at 5% RH, and about one half of them fail to alarm at this level at 50% RH. Further, in these tests as well as tests of prior models⁵⁰, we find that of the alarms that do comply with UL 2034’s sensitivity specification, many fail to alarm at 10% COHb when presented with realistic CO concentration profiles rather than the UL test’s fixed concentration. It is apparent from these tests that the devices do not meet the most essential specification of UL 2034, and yet they are certified by UL to this standard.

While prior tests by this laboratory, these tests, and a UL survey¹⁵ of retail alarms identified numerous devices that, when purchased new, failed to alarm at 10% COHb levels, UL has not withdrawn their certification. At UL 2034 Technical Advisory Panel meetings UL has consistently stated that an alarm should not be considered to “fail” the sensitivity test of the standard if it merely fails the “letter” of the standard; instead it must fail by some unspecified margin of error. While UL has refused to specify the allowed margin of error, in other communication⁵¹ UL has suggested that despite the standard’s requirement that devices alarm before 10% COHb is attained, devices should be considered not to fail if they alarm by 20% COHb.

Because UL publishes no technical basis or justification for the standard requirements it is difficult to ferret out the technical basis for the sensitivity requirements of the standard or the criteria UL uses for determining whether a device meets those requirements. In light of these questions it is advisable to present here some potential medical basis for a sensitivity requirement.

The CPSC’s recommendations to UL provided a compilation of medical references to support their statement that “the 10 percent COHb level is a reference value widely regarded by physicians as confirmation of CO poisoning, regardless of whether individuals are symptomatic, and is widely recognized as protecting healthy adults against the harmful effects of CO.”¹⁹ The CPSC then provided recommendations whose stated intent is to assure that CO alarms activate at less than 10% COHb under all conditions. The CPSC suggested additional test points to enforce the 10% COHb requirement, stating that they were “consistent with the need to protect consumers from CO exposures that pose a potential threat of death, life-threatening symptoms, residual impairment.”

There is a medical consensus that a level of 10% COHb is “sufficiently protective” (to use CPSC’s phrase) for an alarm meant to prevent acute poisoning in the general population. In the medical literature there is also abundant evidence that 20% COHb is an acutely dangerous level for the general population, and a level that is immediately life threatening, and certainly a fatal dose for some of the population.

In particular, young children, pregnant women and their unborn babies, and the elderly are at risk for CO poisoning, and CO alarm marketing is aimed squarely at these groups. These groups are not an insignificant portion of the population. About one in four American families have one or more children under the age of six⁵², and these children suffer irreparable harm at a level of 20% COHb.⁵³ Of women of childbearing age, 11 percent are pregnant at some point in any given year.⁵³ In pregnant women CO passes readily to the fetus, where, because of the greater affinity of fetal hemoglobin to CO, it accumulates to about twice the COHb levels as in the maternal blood.⁵⁴ Animal studies show that a maternal COHb of 9% is harmful, and 20% is promptly fatal to the fetus.⁵⁵ In humans, there are ample case studies demonstrating that 20% is immediately fatal to a fetus and that lower levels result in profound brain damage.^{56, 57, 58, 59}

This cursory medical review suggests that an alarm that can not respond until twice the standard's maximum alarm level for COHb does not meet the scope of the standard, which states that "CO alarms are intended to alarm at carbon monoxide levels below those that could cause a loss of ability to react to the dangers of carbon monoxide exposure"¹. The death of an exposed person would certainly cause a "loss of ability to react to the dangers of carbon monoxide exposure". It seems pointless to specify a "maximum" alarm level if certified alarms are allowed to exceed it by an amount likely to cause death.

The UL 2034 standard itself specifies a maximum alarm level – it makes sense to take the standard seriously and define a failure as noncompliance with that specification, and to refrain from certifying alarms that fail the specification. This report reiterates the CPSC recommendation that alarms must actuate at less than 10% COHb under all conditions of CO exposure.

Digital Display Inaccuracy

Complying with the UL 2034 sensitivity specification for alarm actuation requires that devices meet a minimal measurement accuracy. This accuracy is reflected not only in alarm actuation but also in the alarms' digital display reading. An alarm whose measurement accuracy is within $\pm 30\%$ would be able to alarm reliably within the 5-10% COHb limits set forth by UL 2034, while any poorer accuracy would be reflected *both* in an inability to remain within these limits and in poor digital display accuracy.

Figure 122 shows the UL test points in relation to an exposed person's carboxyhemoglobin (COHb) levels and as a function of CO concentration and time. The maximum times to alarm of the UL test points lie near the 10% COHb level, and the UL Standard implies (but in its latest revisions no longer explicitly states) that alarms must actuate at less than 10% COHb. The test points' minimum times to alarm lie near 5% COHb for the lower concentration test point, and 2.5% COHb for the greater concentration test points.

The most central dotted line in Figure 122 represents a presumptive alarm threshold level. An alarm that is equally likely to drift toward greater or lesser readings should be calibrated to alarm at the cumulative exposure represented by this central line. Calibration in the middle provides the greatest "noise margin", or immunity, to variation caused by sensitivity drift, age, varying temperature and humidity, or the influences of interfering gases. The other pairs of dotted lines in Figure 122 indicate the degradation in noise immunity as sensor errors increase from $\pm 0\%$ to $\pm 10\%$, $\pm 20\%$ and $\pm 30\%$. At $\pm 30\%$ there is no longer any noise immunity remaining; sensors with greater inaccuracy than this can not reliably alarm within the UL sensitivity specification.

Consequently, from the point of view of a device's ability to meet the UL alarm specification we can define an "adequate" digital display accuracy as roughly $\pm 30\%$. Any greater inaccuracy indicates that the device can not meet the sensitivity requirement of the standard. Of course, from the consumer's

perspective, an adequate digital display accuracy is predicated on their desired functionality for the device, their expectations for digital displays and the promises of product advertising. Considering these factors, some consumers may be disappointed with a digital display as inaccurate as $\pm 30\%$. Nevertheless, for the purposes of comparing displays among brands, we use a criterion of $\pm 30\%$.

Using this criterion, of the seven brands with digital displays one, brand A19, was extremely accurate (displaying 92-100 ppm when tested at 100 ppm), two, brands A13 and A18, were adequately accurate (displaying 85-138 and 100-145 ppm), two, brands A4 and A11, were inaccurate (displaying 25-60 and 35-142 ppm), and two, brands A15 and A20, were so inaccurate as to be virtually useless (0-65 and 0-400 ppm).

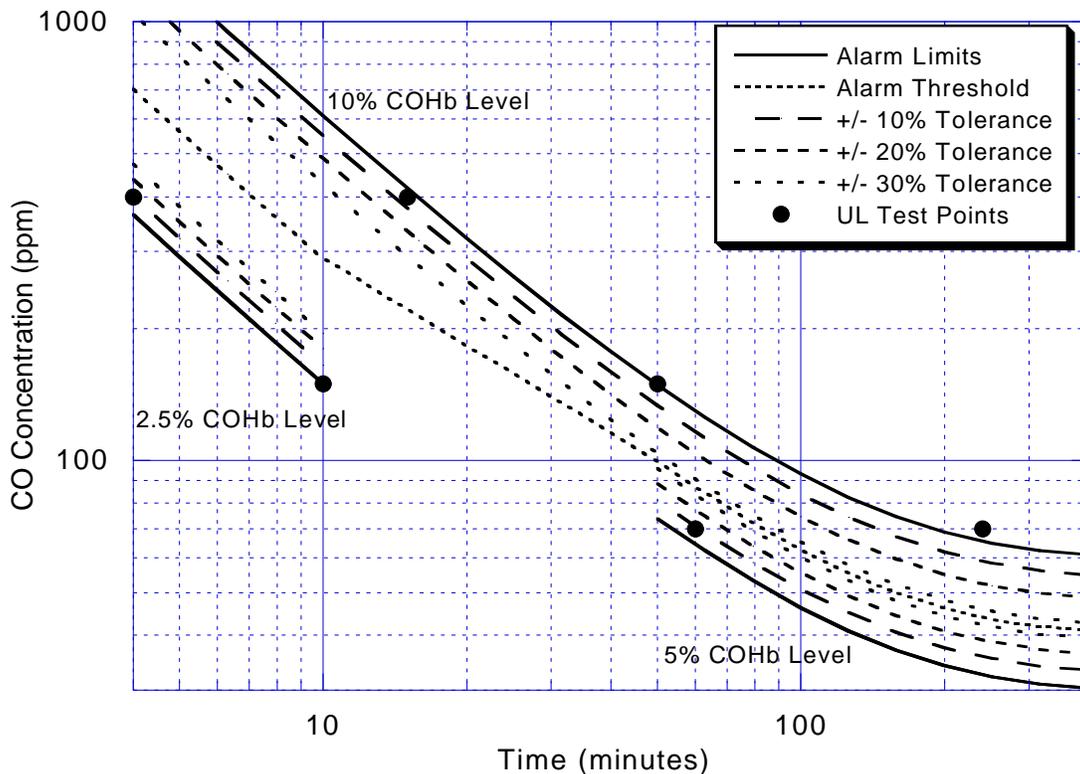


Figure 122 UL alarm limits shown in relation to carboxyhemoglobin levels (as %COHb) in an exposed person. The solid lines represent solutions of the Coburn equation for %COHb levels of 2.5, 5, and 10%, and the solid points represent the UL 2034 test points. The dotted lines show the degradation in alarm band width as a function of the inaccuracy of sensor response. For inaccuracies greater than about $\pm 30\%$ it is no longer possible to meet the UL implied specification that alarms must actuate between 5% and 10% COHb.

This range is graphically illustrated in the following figures. Figure 123 and Figure 124 show alarms' digital display readings on exposure to the progressive concentration test. During a stepwise increase in concentration digital displays were manually recorded 25 minutes into each 30 minute presentation. This allowed enough time for the CO level in the test chamber to stabilize and for the alarms' responses to attain a steady state level. Figure 123 shows an alarm (brand A15) with poorly reading digital displays. At an actual concentration of 100 ppm three of nine devices read 0 ppm while the others range up to 275

ppm. The variation is so great, and so obviously non-Gaussian in its distribution, that it is pointless to characterize the inaccuracy as a percentage.

By way of contrast, Figure 124 shows the high accuracy of the display of brand A19.

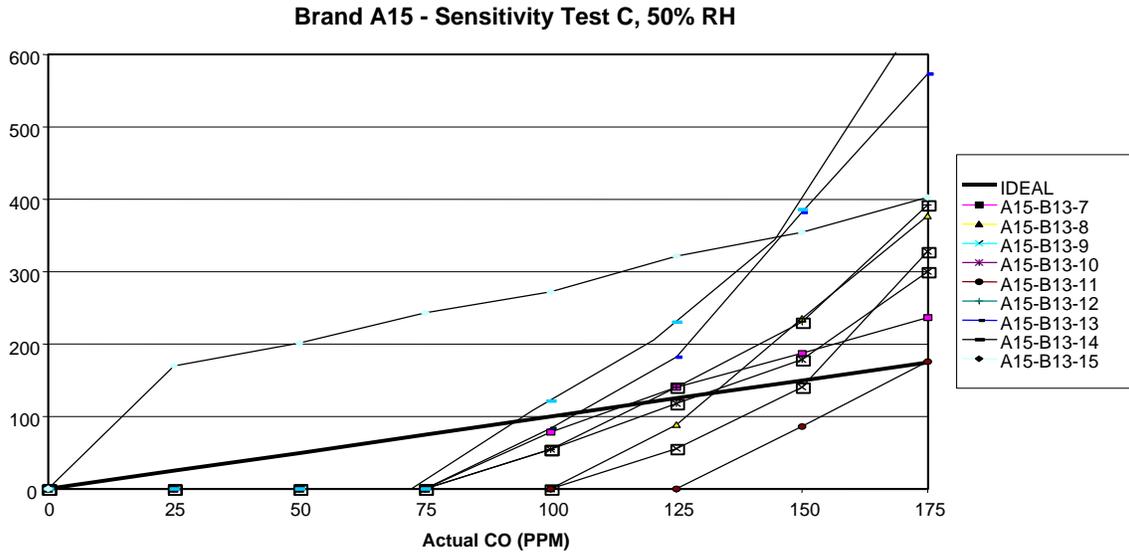


Figure 123 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. Three of nine alarms are not responsive at all up to 100 ppm.

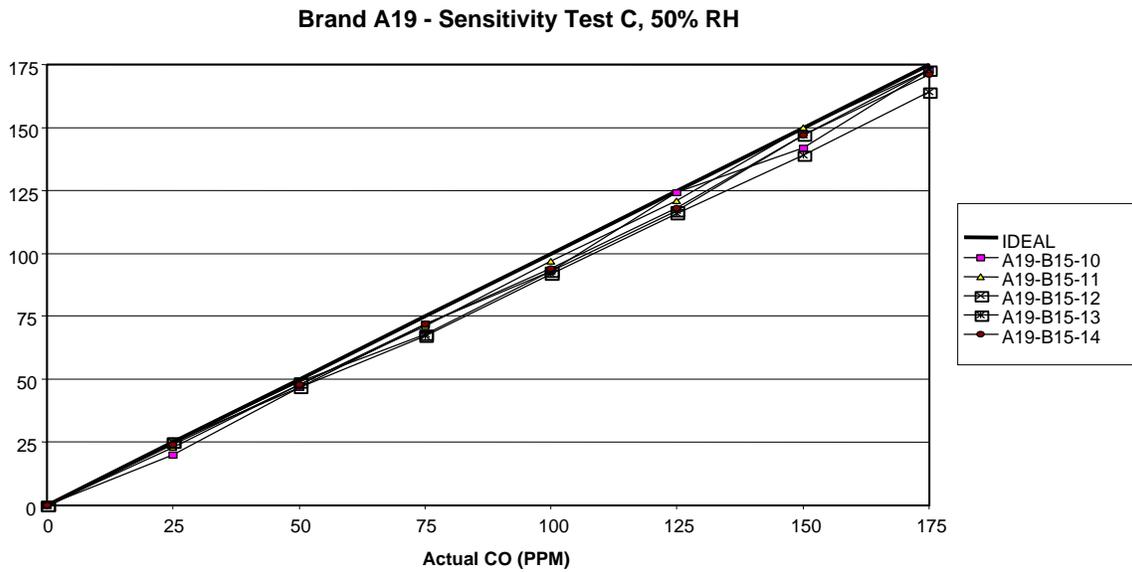


Figure 124 Digital display readings taken during Sensitivity Test C are compared to the actual CO concentrations. All display readings for these alarms are well within $\pm 10\%$ of the actual concentrations.

As these figures also illustrate, those brands with the greatest measurement variance over concentration (as slope, intercept, and point-to-point variation) also had the greatest inter-device variation. Those brands with the most linear and accurate digital displays exhibited the greatest consistency from device to device.

Recommended Display Accuracy

Consumers value digital displays, but obviously only because they have expectations for accuracy of the digital display. For this reason the Consumer Products Safety Commission recommended (recommendation 15 of Oct. 1996¹⁹) that UL specify the accuracy of digital readouts. To date UL has not heeded that recommendation.

Currently, UL allows digital displays with any level of inaccuracy. These current and past tests⁷ have demonstrated that many instruments provide displays that are off by hundreds of percent. These glaring inaccuracies only point out the inability of these alarms to meet the basic sensitivity specification of the standard. Further, consumers are faced with many displays that display essentially random numbers, with no clear dependence on the CO concentrations they purport to indicate.

The accuracy threshold for meeting the standard's sensitivity requirement is $\pm 30\%$. Displays whose accuracy is worse can not meet it. An alarm that is accurate to only $\pm 30\%$ would barely meet its sensitivity specification but have no immunity to noise, temperature and humidity variation, and sensor drift. Consequently, $\pm 30\%$ is a worst case accuracy; alarms should be designed to better this figure.

It is also clear from these tests that some alarm models are capable of much greater accuracy, with one maintaining an accuracy well within $\pm 10\%$.

At UL Technical Advisory Panel meetings from 1996 to 2001 the Gas Research Institute has recommended that the standard require a minimum accuracy for the digital displays, suggesting that $\pm 10\%$ be used. It further recommended that the reliability of display accuracy should be held to the same cumulative failure rate (failure being defined as an excursion beyond the $\pm 10\%$) as the instrument as a whole, with the same lifetime. (Assuming that recommendations for reliability and lifetime are also adopted. The current revision of UL 2034 allows any failure rate, and does not require any longevity for the alarms.) The GRI recommendation echoed a CPSC recommendation¹⁹ that digital display accuracy must be specified. These recommendations are reiterated here.

Sensitivity to Interference Gases

Of the commercially available alarms, six of eight brands tested did not alarm under exposure to any of the interference gases. Of the other two brands, one (Brand A20) had a single device alarm during interference gas exposures, but that device also false alarmed intermittently between and during other tests, so it is likely that it was not responding specifically to the interference gases.

The remaining brand, A15, had most devices alarm in response to the interference gases. The alarms did not actuate during the first three gas exposures (to i-butane, methane, or ammonia), but variously began alarming when exposed to i-propanol or ethanol, and continued to intermittently actuate throughout the remainder of the test.

Of all the brands tested, only this brand was manufactured before 1999, and it is likely that its sensor was not adequately protected by a carbon filter.

Conclusions

The majority of CO alarm brands tested in the course of this study were purchased from retailers. The following conclusions pertaining to the overall performance of commercially available CO alarms and the efficacy of their standards are based solely on this cohort of retail brands.

When CO alarms were first widely installed in 1994 it quickly became apparent that they suffered from a number of drawbacks including numerous false alarms. During the next several years tests revealed numerous failures to alarm also, particularly at lower relative humidity. After two years of CPSC sponsored workshops and hearings, in October 1996, CPSC staff made recommendations to UL for extensive modifications to the UL 2034 standard that they believed “if adopted, will provide greater protection for consumers through a more reliable and effective safety device.”^{f9} These recommendations included requirements for credible quality assurance, time-of-manufacture and lifetime testing of alarms, and an alarm threshold of 10% COHb.

The CPSC’s recommendations for reliability and quality assurance were not adopted, and subsequent tests, including those reported here, reveal a continuation of the same problems that have plagued CO alarms since their inception. Gas Research Institute sponsored field surveys and laboratory studies of CO alarms over the years have also resulted in similar recommendations for improving the standard. This study’s findings reinforce the importance of GRI’s and CPSC’s recommendations. The results of our tests show that indeed, if the CPSC’s recommendations had been adopted, they would have provided the “greater protection for consumers through a more reliable and effective safety device” intended.

The results of testing the basic functionality of CO alarms offered to consumers is disturbing. Depending on the type of failure assessed, overall failure rates for commercially available CO alarms are of the order of tens of percent – a failure rate obviously inadequate for life safety devices. Many of them perform so poorly that they will not protect consumers against CO poisoning. Clearly, many consumers are not being protected – they are not getting the level of protection they were promised when they purchased their alarm.

The inaccuracy of digital displays illustrates this point. Because consumers are willing to pay a premium for alarms with digital displays they are important marketing tools for manufacturers. This places a special onus on their functionality. We found that the accuracy of digital displays spanned a spectrum, from fully functional displays with accuracy rivaling that of analytical instruments, to displays with little relationship between display reading and CO levels. Unfortunately, consumers have no way to independently assess their functioning – they are equally likely to carry the UL mark. (This is true in general, but not in every specific. For example, the most accurate alarm overall, brand A19, is not UL certified.)

The current certification of alarms irrespective of their digital displays accuracy results in devices with accurate digital displays competing with less expensive alarms having poorly functioning displays, while consumers assume they function equally well. This inequity extends to alarms without digital displays. Those that function well, but are burdened with the added cost of quality assurance, must compete with those that function so poorly as to be incapable of protecting consumers from CO poisoning. Consumers may believe that all UL certified alarms behave similarly, and that the UL certification mark assures them of a minimum, safe level of functionality, when these tests show that it does not.

These tests found that three of ten retail brands worked well. Unfortunately, subsequent to the testing reported here, two of those “well-performing” brands (brands A13 and A19), including the best

performing brand with the most accurate digital display (brand A19), are no longer offered for sale and so are no longer available to consumers.

For six of the ten retail brands tested,

- 79% of alarms failed to alarm in response to at least one sensitivity test at 5% RH,
- 47% failed to alarm at 50% RH,
- 30% false alarmed on exposure to an interference gas, and,
- 8% false alarmed in clean air.

With the withdrawal of the best performing brands from the market, the average performance of the installed base of CO alarms is likely to be characterized by the above figures.

This study was motivated by several key questions related to CO alarm reliability, which can now be answered:

- Has alarm performance been improved by the October 1998 revision to UL 2034?

The results of this study parallel the results of earlier studies by this ourselves⁷, other US researchers^{8-9,61-66}, a British field trial¹⁶, and UL¹⁵ – the failure rates of the alarms certified to the revised standard are comparable to those of alarms certified to prior revisions. We conclude that the overall performance of alarms offered to the public has not improved with the new standard revision. Further, some UL-certified alarm brands continue to fail to meet the most basic and essential requirement of the standard, the requirement for sensitivity to CO.

- Are CO alarms susceptible to failures to alarm at low relative humidity?

The results of this study confirm and expand the observations of prior studies^{7-9,61-64} that the sensitivity of retail CO alarms depends strongly on ambient humidity. In particular, several brands of CO alarms fail to alarm at life-threatening CO concentrations when exposed to humidity typical of the indoor environment in the wintertime. While these brands are particularly susceptible to failure at low humidity, operation at moderate humidity still does not assure safe operation. Those brands most likely to fail at low humidity are also those most unreliable at moderate humidity.

- How accurate are the alarms' digital displays?

While several retail alarms show good or adequate display accuracy, many do not. The inaccuracy of several brands is so great as to indicate that the alarms do not meet their basic sensitivity specification.

Summary of Technical Recommendations

This study sheds light on the determinants for the poor performance of CO alarms, and proposes specific technical solutions for some of their most serious technical deficiencies. These technical recommendations, detailed in the prior sections, are summarized here, keyed to the pertinent technical problems:

Poor Reliability

Standard methods of quality control and assurance would prevent the high rates of failure among retail alarms seen in these tests. The recommendations of the CPSC and the Gas Research Institute for time-of-

manufacture and in-service testing are reiterated here. As suggested by the CPSC, this testing should include a statistically significant sample size, quantifiable failure rate criteria, testing over the lifetime of the alarms, a minimum required lifetime, and suitable oversight and inspection procedures. All of these recommendations have been incorporated into the reliability requirements of the CSA 6.19-01 standard. Appendix B of this report provides detailed guidance on the sample size and test frequency necessary to assure compliance with those requirements.

Insensitivity to CO at Low Humidity

The requirements of the CSA 6.19-01 and UL 2034 standards for operation at low humidity are sufficient if they are actually tested-to and enforced. Several models of alarm on the market carrying the UL certification did not meet the standard's criteria for operation at low humidity. The analysis presented on pages 94-98 proposes a method of temperature and humidity compensation that should prevent failures to alarm at low humidity for semiconductor sensor-based alarms.

Unreliable Detection of Varying CO Concentrations

As described in the Analysis section of this report, some brands of CO alarm, even when they did meet the letter of the UL 2034 standard's sensitivity test, failed to comply with its apparent intent (that alarms actuate at or before 10% COHb) under conditions of varying CO concentration. Unfortunately, CO poisoning events are characterized by varying CO concentrations, and these devices do not protect residents against dangerous CO levels. This report proposes a detection algorithm that ensures that alarms respond "biomimetically", integrating CO concentrations as an exposed person would, and always alarming at or before 10% COHb.

Failure to Alarm at 10% COHb

Under some conditions alarms fail to actuate at or below 10% COHb. While we propose technical solutions to the physical causes of these failures (e.g., poor sensitivity at low humidity, or poor integration of varying CO concentration) the failures can not be solely attributed to technical causes. In addition to these technical remedies it is also necessary that certification to the UL 2034 standard should only be provided to devices that actually meet the technical specifications of the standard. We reiterate CPSC's recommendation that all CO alarms must actuate at or below 10% COHb.

Digital Display Inaccuracy

The UL 2034 standard does not require any accuracy for alarms with digital displays, and indeed some alarms have such poor accuracy as to seemingly produce random numbers. This report echoes the standing recommendations of GRI and the CPSC that UL 2034 adopt a well specified digital display accuracy. The CSA 6.19-01 standard has incorporated a requirement that digital displays be accurate to within $\pm 30\%$ of actual concentration.

Available Standards for CO Alarms

The technical deficiencies in the performance of the CO alarms of this study have persisted since their first widespread recognition in 1994, even though technical and regulatory solutions have been available. That several brands of alarms perform well is testament to the industry's ability to produce well functioning devices at reasonable cost. That many other brands perform poorly is indicative of a failure of the standards process.

This study is only the most recent of a series of exploratory and diagnostic tests intended to ascertain the performance of CO alarms and elucidate the causes of their deficiencies. While this diagnostic testing has been successful in attaining those goals, it can not substitute for manufacturers' quality assurance testing, or due diligence on the part of certification agencies.

That some brands provide full protection while others do not indicates that the causes of poor field reliability are not merely technical, but result from the inadequacy of the UL standard to require a greater, but demonstrably achievable, level of performance, and, the failure of UL to enforce the most basic requirements of the current standard. In light of UL's refusal to incorporate the CPSC recommendations of October 1996 for alarm reliability, it is telling that the best performing alarm is not UL listed.

Many of the recommendations for improvement to the UL standard suggested here reiterate long-standing recommendations put forward as early as 1994, and cogently justified by CPSC in their 1996 recommendations. UL's delay in implementing the CPSC recommendations has resulted in the continued marketing and installation of alarms that do not protect consumers.

The UL 2034 standard is not the only North American standard for CO alarms. The CSA 6.19-01 Standard for Residential Carbon Monoxide Alarming Devices²², administered by CSA in Cleveland, Ohio, is similar to the UL standard with an important improvement: It has incorporated many of the CPSC and GRI recommendations^{19,60} for reliability. Further, unlike the UL standard the CSA 6.19 Standard was developed under an open process under which any party could make technical recommendations for consideration by a committee that is compelled by the process to consider proposals on their technical merits, that votes on the recommendations, and whose meeting minutes are published.

This open standard process, with the participation of many alarm manufacturers, resulted in a standard containing provisions that satisfy the CPSC and GRI recommendations for time-of-manufacture and in-service reliability. The CSA standard includes a number of requirements missing from the UL standard, including:

- **Time-of-Manufacture Reliability:** At the time of manufacture alarms must have a total failure rate of less than 1% based on sampling and testing at a 90% confidence level, where failure is defined as a lack of compliance with the CO sensitivity test.
- **In-Service Reliability:** Cumulative failures throughout a required minimum 3 year lifetime of alarms must be less than 23.1% for supervised failures and 14.6% for unsupervised failures.
- **Reliability Test Methods:** GRI's reliability guidelines³² have been incorporated as an Appendix to CSA 6.19-01 to assist manufacturers and certification agencies in devising appropriate test methods. These guidelines are reproduced as Appendix B of this report.
- **Reliability Prediction:** Sections of the standard that relate to component and alarm reliability have been fully rewritten to clarify the distinct requirements for reliability *prediction* and reliability *testing*, as recommended by CPSC and GRI.
- **Selectivity Test:** Existing interference gases have been uniquely identified in the standard and additional interference gas tests have been added for two-hour acute exposures to acetone, ammonia, ethanol, toluene, and trichloroethane. Recovery times for the selectivity test are now specified by the alarm's manufacturer.
- **Combustion Gas Test:** An exposure test has been added to the standard to assure that alarms remain sensitive to CO in the presence of other products of combustion, specifically CO₂, NO_x and SO₂. This test echoes a similar requirement in the European Union's CENELEC Standard and is intended in part to facilitate worldwide harmonization of the standards.

- Lifetime: A minimum 3 year lifetime is required, with lifetime defined in terms of the cumulative allowed maximum failure rate of the alarm.
- Markings and Instructions: An alarm expiration date is marked on the device so that consumers may know when it should be replaced.

The results of the tests reported here indicate that certification to UL 2034 does not assure the performance or reliability of CO alarms. Gas utilities, appliance manufacturers and other users of large numbers of alarms are urged to scrutinize the performance of alarms carefully, through a program of technical review and laboratory testing. It is recommended that whenever possible these users require certification to the CSA 6.19-01 standard.

References

- ¹ Underwriters Laboratories Inc., Standard for Safety UL 2034, “Single and Multiple Station Carbon Monoxide Alarms,” ISBN 0-7629-274-9, First Edition April 1992, Second Edition October 29, 1996, revised October 15, 1997. The revisions of October 15, 1997 were effective October 1, 1998
- ² See for example documents related to CO exposure and promoting the widespread installation of CO alarms on the United States Consumer Product Safety Commission web site: www.cpsc.gov
- ³ “Revised Standards to Improve Carbon Monoxide Alarm Performance”, Gas Research Institute Digest (GRID), Vol. 21 No. 3, Fall 1998.
- ⁴ “Carbon Monoxide Response Survey Analyses: Utility Data -- Final Report”, GRI-96/0409, February 1997, prepared by Resource Strategies, Inc., Madison, Wisconsin
- ⁵ “Carbon Monoxide Response Survey Analyses: Utility Data -- Supplemental Report”, GRI-97/0408, April 1998, prepared by Resource Strategies, Inc., Madison, Wisconsin.
- ⁶ “Residential Carbon Monoxide Alarm Population: Six Cities Study”, GRI-98/0273 and GRI-00/0014, Resource Strategies, Inc., Madison, Wisconsin.
- ⁷ “Performance Testing of Residential CO Alarms”, by Mosaic Industries, Inc., Newark CA, GRI-98-0284, December 1998
- ⁸ “Chamber Tests of Residential CO Alarms”, GRI-98/0140, May 1998, prepared by Roger L. Hedrick, GARD Analytics, Inc., Park Ridge, IL.
- ⁹ “Carbon Monoxide Alarm Performance Testing, Phase 1: Basic Performance, Task 1: Steady-State Concentration Tests”, GRI-95/0220. 1995, prepared by M.W. Paquette, D. Bilodeau, C. Stone, G. King, D. James, and A. Steiner, ETL Testing Laboratories, Inc.
- ¹⁰ “Carbon Monoxide: Communicating Risk and Responsibility”, prepared for the Gas Research Institute by Teri L. Vierma, Resource Strategies, Inc., Madison, WI, March, 1998, GRI-98/0316.
- ¹¹ “Carbon Monoxide Response Survey Analyses: Utility Data – Final Report”, GRI-96/0409, February 1997, prepared by Resource Strategies, Inc., Madison, Wisconsin.
- ¹² “Carbon Monoxide Response Survey Analyses: Utility Data – Supplemental Report”, GRI-97/0408, April 1998, prepared by Resource Strategies, Inc., Madison, Wisconsin.
- ¹³ Smith, Charles L., “Smoke Detector Operability Survey, Report on Findings”, Directorate for Economic Analysis, Consumer Product Safety Commission, Bethesda, MD, October 1994.
- ¹⁴ “Residential Carbon Monoxide Alarm Population: Six Cities Study”, GRI-98/0273 and GRI-00/0014, Resource Strategies, Inc., Madison, Wisconsin.
- ¹⁵ Tests at Underwriters Laboratories have shown significant noncompliance of UL-listed CO alarms with the UL Standard. While these tests were discussed at an August 1999 Technical Advisory Panel Meeting for UL 2034, they do not seem to have been otherwise published.
- ¹⁶ “Joint Industry Project on Carbon Monoxide Issues: Long-term Reliability of Domestic CO Alarms”, Prepared by Advantica Technologies Ltd., for the British HSE (Health and Safety Executive), Report 360/2001, ISBN 0 7176 2085 9
- ¹⁷ *Workshop on Carbon Monoxide Detectors, Arlington Virginia, June 2, 1994*, Final Report, August 1994, Richard C. Bratcher and Hannelore Chromy, Science and Technology Management, Inc., 2511 N. 124th St., Suite 205, Brookfield, WI 53005, 414-785-5940.

- ¹⁸ For example, “Consumer Product Safety Commission: Carbon Monoxide Detectors; Public Hearing”, Federal Register, Vol. 60, No. 205, Tuesday, October 24, 1995.
- ¹⁹ “Recommendations of the Consumer Product Safety Commission (CPSC) Staff Concerning Underwriters Laboratories (UL) Inc. Standard 2034 for Single and Multiple Station Carbon Monoxide Detectors (First Edition – 1992; New and Revised Requirements – August 1995)”, CPSC, October 25, 1996, Elizabeth W. Leland, Project Manager, CO Detection.
- ²⁰ “Test Protocols for Residential Carbon Monoxide Alarms, Phase 1”, GRI-96/0055, September 1996, prepared by Paul K. Clifford and Michael G. Dorman, Mosaic Industries, Inc., Newark CA.
- ²¹ “IAS U.S. Requirements for Carbon Monoxide Alarms for Residential Use, No. 6-96”, June 1, 1998, International Approval Services – U.S., a Division of CSA, Cleveland OH
- ²² “CSA 6.19-01, Residential Carbon Monoxide Alarming Devices”, Canadian Standards Association, March 2001.
- ²³ Draft “Standard Guide for Testing Systems for Measuring Dynamic Responses of Carbon Monoxide Detectors to Gases and Vapors”, ASTM D-22 Committee, c/o George Luciw, 610-832-9710, gluciw@astm.org
- ²⁴ J. E. Peterson and R.D. Stewart, “Predicting the carboxyhemoglobin levels resulting from carbon monoxide exposures”, *J. Appl. Physiol.*, Vol. 39 (1975), pp. 633-638.
- ²⁵ J.E. Peterson, Industrial Health, Prentice-Hall Publ., Englewood Cliffs, N.J., (1977)
- ²⁶ R.D. Stewart et al, “Experimental Human Exposure to Carbon Monoxide”, *Arch. Environ. Health*, Vol. 21, (1970), pp. 154-164.
- ²⁷ Richard A. Wadden and Peter A. Scheff, Indoor Air Pollution: Characterization, Prediction, and Control, John Wiley & Sons Publ., (1983), ISBN 0-471-87673-9, Chapter 2.
- ²⁸ “Criteria for a recommended standard...Occupational Exposure to Carbon Monoxide”, U.S. Department of Health, Education and Welfare, NIOSH, 1972, HSM 73-11000.
- ²⁹ R.F. Coburn, R.E. Forster, and P.G. Kane, “Considerations for the physiological variables that determine the blood carboxyhemoglobin concentration in man”, *Journal of Clinical Investigation*, Vol. 44 (1965), pp 1899-1910.
- ³⁰ “Recommendations of the Consumer Product Safety Commission (CPSC) Staff Concerning Underwriters Laboratories (UL) Inc. Standard 2034 for Single and Multiple Station Carbon Monoxide Detectors”, U.S. Consumer Product Safety Commission, ed. Elizabeth W. Leland, Project Manager, CO Detection, October 25, 1996.
- ³¹ GRI-96/0055, “Test Protocols for Residential Carbon Monoxide Alarms, Phase 1”, Mosaic Industries, Inc., September 1996, Chapter 6, pgs 6-6 – 6-16.
- ³² “GRI Guidelines for Assuring the Time-of-Manufacture and In-Service Reliability of CO Alarms”, Gas Research Institute, included as Appendix B of this report.
- ³³ GRI-98/0284, “Performance Testing of Residential CO Alarms”, Mosaic Industries Inc., December 1998.
- ³⁴ Underwriters Laboratories Inc., Standard for Safety UL 2034, “Single and Multiple Station Carbon Monoxide Alarms,” ISBN 0-7629-274-9, First Edition April 1992, Second Edition October 29, 1996, revised October 15, 1997, Section 45.1.2

- ³⁵ See for example CPSC Press Release # 85-057, "Burning Charcoal Causes 83 Deaths From Carbon Monoxide", and CPSC's "Requirements for Labeling of Retail Containers of Charcoal; Proposed Amendments", Federal Register: August 10, 1995 (Volume 60, Number 154), Proposed Rules, Page 40785-40792.
- ³⁶ "Non-fire Carbon Monoxide Deaths and Injuries Associated with the Use of Consumer Products: Annual Estimates", September 1998, CPSC
- ³⁷ GRI-96/0055, "Test Protocols for Residential Carbon Monoxide Alarms", prepared by Mosaic Industries, Inc. for the Gas Research Institute, September 1996, pg. 4-9 and 4-25.
- ³⁸ "Homogeneous Semiconducting Gas Sensors: A Comprehensive Model", P. K. Clifford, Proceedings of the International Meeting on Chemical Sensors, 1983.
- ³⁹ "Mechanisms of Gas Detection by Metal Oxide Surfaces", Paul K. Clifford, PhD dissertation, Carnegie-Mellon University, 1981.
- ⁴⁰ "Characteristics of Semiconductor Gas Sensors I. Steady State Gas Response", P. K. Clifford and D. T. Tuma, Sensors and Actuators, 3 (1982/83) 233-254.
- ⁴¹ "TGS 2440 – for the detection of Carbon Monoxide", available from Figaro, Inc., Osaka, Japan and Wilmette, Illinois.
- ⁴² "Performance Testing of Residential CO Alarms", GRI-98/0284.
- ⁴³ "Contribution to US Consumer Product Safety Commission Hearings on Residential Carbon Monoxide Detectors, February 1996: Comparison of US and British Standards for Residential Carbon Monoxide Detectors and Realistic Technology Improvements Reducing Alarm COHb Levels", P. McGeehin of Capteur Sensors & Analysers Ltd., Abingdon, UK
- ⁴⁴ J. E. Peterson and R.D. Stewart, "Predicting the carboxyhemoglobin levels resulting from carbon monoxide exposures", *J. Appl. Physiol.*, Vol. 39 (1975), pp. 633-638.
- ⁴⁵ J.E. Peterson, Industrial Health, Prentice-Hall Publ., Englewood Cliffs, N.J., (1977)
- ⁴⁶ R.D. Stewart et al, "Experimental Human Exposure to Carbon Monoxide", *Arch. Environ. Health*, Vol. 21, (1970), pp. 154-164.
- ⁴⁷ Richard A. Wadden and Peter A. Scheff, Indoor Air Pollution: Characterization, Prediction, and Control, John Wiley & Sons Publ., (1983), ISBN 0-471-87673-9, Chapter 2.
- ⁴⁸ "Criteria for a recommended standard...Occupational Exposure to Carbon Monoxide", U.S. Department of Health, Education and Welfare, NIOSH, 1972, HSM 73-11000.
- ⁴⁹ R.F. Coburn, R.E. Forster, and P.G. Kane, "Considerations for the physiological variables that determine the blood carboxyhemoglobin concentration in man", *Journal of Clinical Investigation*, Vol. 44 (1965), pp 1899-1910.
- ⁵⁰ "Performance Testing of Residential CO Alarms", GRI-98/0284.
- ⁵¹ UL ballot to the ANSI Canvass of the CSA 6.19 Standard for Residential CO Alarms.
- ⁵² U.S. Bureau of the Census, "Statistical Abstract of the United States: 1994", 114th edition, Washington DC, 1994.
- ⁵³ Binder JW and Roberts RJ, "Carbon Monoxide Intoxication in Children", *Clinical Toxicology*, 1980, v 16, no. 3, p. 287-295.

- ⁵⁴ see for example, Longo LD, “The Biological Effects of Carbon Monoxide on the Pregnant Woman, Fetus and Newborn Infant”, *Am. J. Obstet. Gynecol.*, 1977 v. 129 no. 1 p. 69-103; or Longo LD, “Carbon Monoxide in the Pregnant Mother and Fetus and its Exchange Across the Placenta”, *Annals New York Academy of Sciences*, v 174, p. 313, 1976.
- ⁵⁵ Longo, LD and Hill, EP, “Carbon Monoxide Uptake and Elimination in Fetal and Maternal Sheep”, *American Journal of Physiology*, v 232, p. H324, 1977.
- ⁵⁶ Phillips P, “Carbon Monoxide Poisoning During Pregnancy”, *British Medical Journal*, v1 p 14, 1924
- ⁵⁷ Muller GL and Graham S, “Intrauterine Death of the Fetus Due to Accidental Carbon Monoxide Poisoning”, *New England Journal of Medicine*, v 252, no. 25, p. 1075, 1955
- ⁵⁸ Goldstein DP, “Carbon Monoxide Poisoning in Pregnancy”, *American Journal of Obstetrics and Gynecology*, V92, No. 4, p. 526, 1965
- ⁵⁹ Ginsberg MD and Myers RE, “Fetal Brain Injury After Maternal Carbon Monoxide Intoxication”, *Neurology*, V26, p. 15, 1976
- ⁶⁰ “GRI Recommendations to the (Interim CSA) 6.19 Technical Committee for Carbon Monoxide Detectors and the Technical Advisory Panel for UL 2034, October 1999”, available from the Gas Research Institute.
- ⁶¹ Wilson, A.L., “Correlation of Indoor Residential Carbon Monoxide Detector Alarm Frequencies and Ambient CO Levels”, *Proceedings of the Specialty Conference on Engineering Solutions to Indoor Air Quality Problems*, Research Triangle Park, NC, July21-23,1997, AWMA, pp137-149.
- ⁶² Wilson, A.L., “Summary Report: Evaluation of CO Detectors Exposed To Outdoor Air: Phase I (Feb.-Dec. 1997)”, Prepared by Wilson Environmental Associates for the Southern California Gas Company, December, 1998.
- ⁶³ Wilson, A.L., “Evaluation of CO Detectors Exposed To Outdoor Air: Phase II-Final Report (Dec. 1997 to June 1999)”, Prepared by Wilson Environmental Associates for the Southern California Gas Company, July, 1999.
- ⁶⁴ “Re: Results of Chamber Tests for Kidde Lifesaver Models 9CO-1 and 9C)-1C”, Letter-style report from A.L. Wilson of Wilson Environmental Associates to Daryl Hosler of the Southern California Gas Company, January 25, 1999. The tested CO alarms “go into warning mode... at a few hours exposure to high humidity and very low CO but will not alarm when there is enough CO exposure to be deadly.”
- ⁶⁵ “Re: Kidde Premium Plus Nighthawk CO Alarm (Model KN-COPP-3)”, Letter-style report from A.L. Wilson of Wilson Environmental Associates to Daryl Hosler of the Southern California Gas Company, March 21, 2000, (Seven units of the tested brand of residential CO alarms responded accurately to CO levels from 0 to 200 ppm but gave false values when exposed to portable telephone, cellular telephone and ham radio signals.)
- ⁶⁶ “Re: Kidde Premium Plus Nighthawk CO Alarm (Model KN-COPP-3)”, Letter-style report from A.L. Wilson of Wilson Environmental Associates to Daryl Hosler of the Southern California Gas Company, July 7, 2000. (Five additional units of the brand tested earlier (see Ref 65) responded accurately to CO levels from 0 to 200 ppm but gave false values when exposed to portable telephone, cellular telephone and ham radio signals.)
- ⁶⁷ “Evaluation of Carbon Monoxide Detectors in Domestic Premises, Literature Review”, David Ross, Martin Smith, Mike Spearpoint, Debbie Smith, and Richard Colwell, Building Research Establishment Ltd., HSE Research Report 246/1999, ISBN 0 7176 1704 1.

- ⁶⁸ “Evaluation of Carbon Monoxide Detectors in Domestic Premises, Part 1: Testing Carbon Monoxide Detectors in situ”, Prepared by the Building Research Establishment Ltd., HSE Research Report 1999, ISBN 0 7176 2496 X.
- ⁶⁹ “Evaluation of Carbon Monoxide Detectors in Domestic Premises, Part 2: Recommendations for the Siting of Carbon Monoxide Detectors”, Prepared by the Building Research Establishment Ltd., HSE Research Report 1999, ISBN 0 7176 2496 X.
- ⁷⁰ “Evaluation of Carbon Monoxide Detectors in Domestic Premises, Summary Report”, Prepared by Dr. D.I. Ross, the Building Research Establishment Ltd., HSE Research Report 236/1999, ISBN 0 7176 2482 X.
- ⁷¹ “Evaluation of Carbon Monoxide Detectors in Domestic Premises, Investigation of the Movement of Carbon Monoxide with an Home”, Prepared by the Building Research Establishment Ltd., HSE Research Report 238/1999, ISBN 0 7176 2486 2.
- ⁷² “Evaluation of Carbon Monoxide Detectors in Domestic Premises, Experimental and Computational Investigation of the Movement of Carbon Monoxide within a Single Room”, Prepared by the Building Research Establishment Ltd., HSE Research Report 237/1999, ISBN 0 7176 2485 4.
- ⁷³ “A model to predict build-up of carbon monoxide from a faulty appliance in a single room”, Steve Bullman, Advantica Technologies Limited, Ashby Rd., Loughborough, Leicestershire, LE11 3GR, UK, British HSE (Health and Safety Executive) Research Report 341/2001, ISBN 0 7176 2017 4.
- ⁷⁴ “An experimental study of the build-up and dispersal of combustion products containing carbon monoxide in a house”, prepared by British Gas plc for the British Health and Safety Executive, British HSE (Health and Safety Executive) Research Report 141/1977, ISBN 0 7176 1399 2.
- ⁷⁵ “Carbon Monoxide Dispersion in Residential Buildings: Literature Review and Technical Analysis”, Prepared by Andrew K. Persily, Building and Fire Research Laboratory, National Institute of Standards and Technology, U.S. Dept. of Commerce, for the National Fire Protection Research Foundation, October 1996, Report No. NISTIR 5906
- ⁷⁶ “News from CPSC: CPSC, Kidde Safety Announce Recall of Carbon Monoxide Alarms”, Release # 99-082, Originally issued March 19, 1999; Revised November 7, 2001, Consumer Product Safety Commission

Appendix A

The Use of the Coburn Equation in the Design of CO Alarms

This Appendix provides a simple equation for computing carboxyhemoglobin levels (%COHb) from carbon monoxide (CO) concentrations, for any CO concentration profile in time. The equation is derived from fundamental, and well tested physiological principles, as embodied in the Coburn equation. This Appendix also clarifies the practical use of the Coburn equation in CO Alarms for integrating CO concentrations and triggering alarm activations.

The Coburn equation is a physiologically based model of the relationship between inspired CO and the carboxyhemoglobin (%COHb) it produces in an exposed person's bloodstream. This equation can be used in a CO alarm to integrate measured CO concentrations over time and predict the %COHb that would be present in a hypothetical exposed subject. Whenever the predicted %COHb exceeds a predetermined threshold (or trigger point) an alarm can then be sounded.

Presently, manufacturers of CO alarms use a variety of *ad hoc* integration formulas and threshold-activated timers to determine %COHb or alarm activation times. Some of these formulas are explicitly based on a version of the Coburn equation cited in UL Standard 2034 and some simplify that equation for easy implementation. Still others, particularly threshold-activated timers, are designed to satisfy UL 2034 at the test concentrations but may deviate significantly from the behavior suggested by the Coburn equation at other CO concentrations, or they may cause significantly different recovery times than the UL equation.

The following sections of this Appendix:

- Introduce the Coburn equation and derive a simple, readily used form for it;
- Discuss the three parameters of the Coburn equation (the baseline %COHb, the equilibrium constant, and the averaging time constant) and their practical importance;
- Show that with a single set of values for its parameters, the Coburn equation models COHb uptake and elimination in response to all CO profiles well;
- Provide a simple modification to the Coburn equation to model a more diverse population, or a single individual at varying levels of metabolic activity; and,
- Reduce the Coburn equation to a simple practical algorithm for complying with the specification of UL 2034. This algorithm may be used to compute %COHb from any time-varying CO concentration, and it is easily implemented in CO Alarms.

Those seeking practical guidance for incorporating an appropriate integration formula in a CO alarm need not wade through this entire document but can proceed directly to the fourth section.

The Coburn Equation

Coburn, Foster and Kane established the physiological basis for CO adsorption and elimination by the blood and described it using an absorption model in which hemoglobin is assumed to be the predominant absorber¹. (References for Appendix A are listed separately at the end of the Appendix.) Their model

produces a single differential equation often called the Coburn-Foster-Kane equation, or more simply, the Coburn equation. This equation describes both the uptake and the elimination of CO as,

$$\begin{aligned} d[\text{CO}]/dt = & V_{\text{CO}}/V_{\text{B}} \\ & + p\text{CO}/(V_{\text{B}}(1/D_{\text{L}} + P_{\text{L}}/V_{\text{A}})) \\ & - [\text{CO}] p\text{O}_2 / (M V_{\text{B}} [\text{O}_2] (1/D_{\text{L}} + P_{\text{L}}/V_{\text{A}})) \end{aligned} \quad \text{Eqn. A-1}$$

in which $d[\text{CO}]/dt$ is the rate of change of CO in the blood (as ml gaseous CO/ml liquid blood)/min), $p\text{CO}$ is the inspired partial pressure of CO (in mmHg), and $[\text{CO}]$ is the concentration of CO in the bloodstream (as ml gaseous CO/ ml liquid blood). The meaning of the other symbols are given in Table 1, along with typical quantitative values for the physiological parameters and references to the studies that determined those values.

The Coburn equation is a single compartment, single time constant model for the uptake of CO by the human body. This equation is itself an approximation: Because CO is fat-soluble, body fat acts as a second compartment for CO absorption; hence a two-compartment, two-time-constant model better describes uptake and release of CO by the body over long periods of time. Further, muscle myoglobin acts as a third storage compartment. Even so, the single compartment model provides a good approximation for CO uptake during exposures of under a day's duration. Any inaccuracies introduced by using a single-component model instead of a multi-compartment model are dwarfed by the inaccuracies introduced by applying a single set of model parameters to an entire population. A single component model is more than accurate enough for use in CO alarms.

In order to convert Eqn. A-1 into a more readily understandable form we use these useful identities^{2,3}:

The partial pressure of CO is related to its concentration as, $p\text{CO} = 7.6 \cdot 10^{-4} \text{ mmHg/ppm} * \text{CO}_{\text{ppm}}$	Eqn. A-2
The half life of CO in the blood is a fixed fraction of its exponential decay time, $\tau_{1/2} = \tau \text{Ln}(2) = 0.693 \tau$	Eqn. A-3
%COHb is proportional to blood CO as, $\% \text{COHb} = 100 * [\text{CO}] / [\text{CO}]_{\text{sat}}$	Eqn. A-4
Hemoglobin-bound oxygen in the blood decreases with bound CO as, $[\text{O}_2] = [\text{O}_2]_{\text{sat}} - [\text{CO}]$	Eqn. A-5
Oxyhemoglobin is proportional to blood oxygen concentration as, $\% \text{O}_2\text{Hb} = 100 * [\text{O}_2] / [\text{O}_2]_{\text{sat}}$	Eqn. A-6
Arterial oxygen saturation, %O ₂ Hb or SaO ₂ , depends on blood hemoglobin concentration, HB, and mean capillary oxygen partial pressure, pO ₂ , as, $[\text{O}_2]_{\text{sat}} = (1.34 \text{ ml O}_2 / \text{g Hb}) \text{HB} + (3 \times 10^{-5} \text{ ml O}_2 / \text{ml blood/mmHg}) p\text{O}_2$	Eqn. A-7
or more simply as (see ref. [A-4]), $[\text{O}_2]_{\text{sat}} = 0.20 \text{ ml O}_2 / \text{ml blood}$	Eqn. A-8

Table A-1. Parameters of the Coburn Equation				
Symbol	Meaning	Units	Range and Typical Averages	Ref.
α_o	Ratio of equilibrium %COHb to inspired COppm	%COHb/COppm	0.166	2, 3
[CO]	CO concentration in the arterial blood as a ratio of volumes of gaseous CO to blood	ml/ml blood	$0 < [CO] < 0.20$	2
[CO] _{sat}	CO concentration in the arterial blood at 100 %COHb	ml/ml blood	0.20	2
COppm	Inspired CO concentration in ppm by volume.	ppmv	$0 - 10^6$	
%COHb	Portion of blood hemoglobin bound to CO, expressed as % of saturation. %COHb ₀ indicates %COHb at time=0 and %COHb _t indicates the value at a particular time.	percentage	$0 < \%COHb < 100$	3
%COHb _e	Endogenous COHb, i.e., steady state portion of blood hemoglobin bound to CO in the absence of inspired CO	percentage	0.3 - 0.7 0.33 0.4 0.76	15, 16 8 9 2
D _L	Diffusion capacity of the lungs for CO. Increases from 30 for healthy sedentary persons to 60 for heavy exercise.	ml/(min*mmHg)	17- 60 30 (sedentary) 60 (moderate/heavy exercise) 60 (used in UL 2034)	17 2 17 18
HB	Blood hemoglobin concentration	gHb/ml blood	.125 - .175 0.145	3 3
HC	Hematocrit, or ratio of red cell volume to total blood volume	percentage	40-54% (for men) 47% (for men)	3 3
M	Haldane constant, ratio of binding rate of CO with Hb to that of O ₂ with Hb	dimensionless ratio	220 @ pH=7.4 218	 2
[O ₂] _{sat}	O ₂ concentration in the arterial blood at 100 %O ₂ Hb	ml/ml blood	0.20	2, 3
%O ₂ Hb, SaO ₂	Portion of arterial blood hemoglobin bound to O ₂ , expressed as % of saturation. Appx 97% in the absence of nonendogenous CO.	percentage	$0 < \%COHb < 100$ 97	3
[O ₂]	O ₂ concentration in the arterial blood as a ratio of volumes of gaseous O ₂ to blood	ml/ml blood	$0 < [O_2] < 0.20$ 0.20	2, 3
pCO	Inspired CO partial pressure	mmHg	0 – 760 0.076 (at 100 ppm)	
P _L	Pulmonary air pressure, atmospheric pressure (760 mmHg) minus the partial pressure of water vapor at body temperature (47 mmHg)	mmHg	713	2
pO ₂	Mean pulmonary capillary oxygen partial pressure for sedentary to moderate exercise and at a normal ventilation/perfusion ratio at an inspired O ₂ partial pressure of 150 mmHg.	mmHg	100	2, 3
τ	Exponential characteristic time for the variation of COHb in the blood. The half life, $\tau_{1/2}$ is related to τ as, $\tau_{1/2} = \tau \ln 2 = 0.693 \tau$	minutes	366-462 (sedentary) 400 (sedentary) 97 (used in UL 2034) 83 (heavy exercise)	2 5 5 5
V _A	Alveolar ventillation rate	ml/min	4,000 (sleeping) 6,000 (sedentary) 30,000 (used in UL 2034) 43,000 (heavy exercise) 60,000 (very heavy exercise)	2 2, 3 18 19 3
V _B	Total blood volume	ml	5500.	2
V _{CO}	Whole body endogenous rate of CO production	ml/min	0.007-0.025 0.007	2

The form of the Coburn equation of Eqn. A-1 describes the rate of CO adsorption and desorption in terms of ml CO/ml blood and the inspired CO partial pressure in mmHg. Using the identities of Eqns. A-2-8 these units can be converted to the more familiar units of percent carboxyhemoglobin (%COHb) and

inspired CO concentration in ppm (CO_{ppm}), and the equation rearranged to make it more understandable, as,

$$d(\%COHb)/dt = (\%COHb_e + \alpha CO_{ppm} - \%COHb) / \tau \quad \text{Eqn. A-9}$$

in which an endogenous carboxyhemoglobin level, $\%COHb_e$, an equilibrium constant, α , and a time constant, τ are defined as,

$$\%COHb_e = 100 \tau V_{CO} / (V_B [CO]_{sat}) \quad \text{Eqn. A-10}$$

$$\alpha = \alpha_o (1 - \%COHb / 100) \quad \text{Eqn. A-11}$$

where,

$$\alpha_o = (7.6 \times 10^{-2} \text{ mmHg/ppm}) M / pO_2 \quad \text{Eqn. A-12}$$

and,

$$\tau = \tau_o (1 - \%COHb / 100) \quad \text{Eqn. A-13}$$

where,

$$\tau_o = M V_B (1/D_L + P_L/V_A) [O_2]_{sat} / pO_2 \quad \text{Eqn. A-14}$$

Eqn. A-9 has been found experimentally to model the uptake and elimination of CO well. It describes the rate of increase of $\%COHb$ on exposure, the establishment of an equilibrium between blood and ambient CO levels, the elimination of CO after exposure, and even the integration of discontinuous exposures or varying CO concentration to produce varying $\%COHb$ ^{2,5}.

This Coburn equation would be a first order linear differential equation if the portion of hemoglobin occupied by oxygen (represented by $[O_2]$ in Eqn. A-1) were constant. As it is, $[O_2]$ decreases with CO uptake, particularly at high $\%COHb$, causing both α and τ to vary with $\%COHb$ as shown in Eqns. A-11 and A-13,

$$\alpha = \alpha_o (1 - \%COHb / 100) \quad \text{and} \quad \tau = \tau_o (1 - \%COHb / 100)$$

Because the $\%COHb$ appears implicitly in these parameters the Coburn equation is said to be *nonlinear*, and it is not possible to transform the differential equation into a closed form analytical solution. Even so, a computer simulation of the equation, solving it as a difference equation, is easy, and such a simulation can be implemented in a small processor, even in a CO alarm⁶.

Even so, we will follow Coburn's widely accepted derivation of an approximate analytical solution. It is this approximate analytical solution that is used in the UL 2034 Standard for CO alarms. To solve this equation for $\%COHb$ as a function of time, Coburn *et al* assume that $[O_2]$ remains approximately constant (i.e., that $\%COHb$ remains small enough) so that α is approximated with α_o and τ with τ_o . This approximation results in a first order *linear* equation that remains quite accurate in describing the dynamics of CO uptake and elimination for $\%COHb$ less than 30%. Using this approximation, Eqn. A-9 is further simplified to,

$$d(\%COHb)/dt = (\%COHb_e + \alpha_o CO_{ppm} - \%COHb) / \tau_o \quad \text{Eqn. A-15}$$

Here we've replaced α with α_o and τ with τ_o . This approximation makes possible an analytical solution for $\%COHb$ as a function of time. If we assume a constant CO concentration the analytic solution is,

$$\%COHb_t = \%COHb_o \exp(-t/\tau_o) + (\%COHb_e + \alpha_o CO_{ppm}) (1 - \exp(-t/\tau_o)) \quad \text{Eqn. A-16}$$

in which $\%COHb_t$ is the percentage COHb in the blood at time t and $\%COHb_o$ is the initial percentage COHb at time zero. This is the equation cited in UL 2034⁷, which uses the following parameter values:

$$\begin{aligned}\alpha &= 0.166 \\ \%COHb_e &= 0.065\% \\ \tau_o &= 97 \text{ min.}\end{aligned}$$

But CO concentration is rarely constant. A detection instrument must track the %COHb of a hypothetical exposed individual while the CO concentration continuously varies and sound an alarm if the projected %COHb exceeds a threshold. For this reason the differential form, Eqn. A-15, is more useful than the analytical solution that assumes a constant CO concentration of Eqn. A-16.

The Parameters of the Coburn Equation

Equation 15 makes it clear that the Coburn equation has three terms, for the rates of endogenous CO production, the uptake of CO, and the elimination of CO respectively. The equation is fully specified by three parameters, %COHb_e, α_o, and τ_o. We'll discuss the importance of these parameters in turn:

%COHb_e – Endogenous CO Production Contributes to Baseline %COHb

The first term in the equation represents the baseline COHb resulting from endogenous CO production. Even in the absence of inhaled CO, the body generates a small amount of CO through its own metabolic processes, more specifically, by the catabolism of heme protein. In healthy persons the natural turnover of heme accounts for a baseline %COHb_e of about 0.2-0.8%. Some measurements of *nonsmokers* reveal approximately 2 ppm of CO in each exhaled breath, corresponding to a %COHb_e of 0.33%⁸. Other measurements reveal a mean of 0.4% for nonsmokers⁹. Stewart *et al*², from which many of the values of Table 1 are taken, use a V_{CO} that implies a baseline level of 0.23%, but measure an actual average level of 0.76% for 73 measurements of 18 healthy, sedentary subjects.

There are various disease processes that can greatly increase endogenous COHb production. For example, patients with hemolytic anemia may have a %COHb_e as great as 4-6%¹⁰, and the metabolism of inspired or ingested chlorinated hydrocarbons can generate significant endogenous CO. This is the mechanism by which the inspiration of methylene chloride is so dangerous¹¹. Traumatic anemia, as caused in military recruits by marching long distances or in long distance runners, may also elevate baseline COHb.

The endogenous production of CO in healthy persons is not an important contribution to any ill effects *per se*, but it does provide a baseline COHb level that should be considered in the determination of total CO exposure¹². Consequently, in order to apply the Coburn equation to the general population we must use a realistic value for %COHb_e. While the baseline level is expected to vary widely among individuals, the level used in the version of the Coburn equation cited by UL 2034, 0.065%, is not realistic, being an order of magnitude smaller than any measurements of sedentary subjects.

Small, person-to-person variation in %COHb_e does not appreciably affect predictions of the Coburn equation for a general population, as the small variations are compared to an alarm level that is much greater in magnitude. Consequently, using a single value of %COHb_e to model the response of an entire population to CO is justified.

A more realistic value for %COHb_e should be used than that provided in the UL 2034. We recommend a change in UL 2034 that incorporates a value typical for sedentary persons:

$$\%COHb_e = 0.5 \qquad \qquad \qquad \text{Eqn. A-17}$$

α_o – The Equilibrium Constant for %COHb

The second term, the equilibrium constant, determines the rate of COHb production as CO is inspired as well as the equilibrium level eventually achieved. The competition between CO and oxygen for hemoglobin is reflected in this term as the ratio of the partial pressure of inspired CO to the partial pressure of O₂, with the influence of CO being weighted heavily by the Haldane constant, the ratio of CO binding rate to O₂ binding rate (see Eqn. A-12). The Haldane constant is large, approximately 218, indicating that CO competes effectively for Hb.

The rate of COHb accumulation depends primarily on the amount of CO inhaled, i.e. the concentration breathed times the breathing rate, and the equilibrium constant, α_o . The breathing rate enters the equation through the time constant, τ_o . While this characteristic decay time may vary considerably from subject to subject and with level of activity, the equilibrium constant, α_o , varies much less. However, it may be expected to depend on barometric pressure, increasing with altitude. It is also likely to depend on the type of hemoglobin; different hemoglobins (i.e., adult hemoglobin, fetal hemoglobin, and recessive and dominant sickle cell hemoglobins) may be expected to have different binding constants for CO. In particular, it has been shown that fetal hemoglobin binds CO more tightly than adult hemoglobin, so the unborn fetus, newborns and pregnant women are at greater risk from CO poisoning¹³. Even so, α_o can be treated as nearly constant for healthy persons (from Eqn. A-12):

$$\alpha_o = 0.166 \text{ \%COHb} / \text{ppm} \quad \text{Eqn. A-18}$$

Along with the small endogenous contribution, this equilibrium constant determines the equilibrium %COHb produced while inspiring a constant CO concentration. For example, for a inspired CO concentration of 100 ppm, the equilibrium %COHb is given by,

$$\text{\%COHb} = \text{\%COHb}_e + \alpha_o * 100 \text{ ppm} = 17.1\% \quad \text{Eqn. A-19}$$

τ_o – An Exponential Time Constant Governs the Rates of CO Uptake and Elimination

The *rate* at which an equilibrium %COHb is achieved, or decays, is governed by a single time constant (for a constant level of metabolic activity in a single person). While the uptake rate is proportional to the external CO concentration and the inhalation rate divided by the time constant, the elimination rate is proportional to the internal concentration (the current %COHb) and the exhalation rate divided by the time constant. The greater the breathing rate, the more CO enters the lungs to react with hemoglobin, and the more oxygen enters the lungs to displace the hemoglobin-bound CO. This breathing rate enters the equation through τ_o , and influences the rates of CO uptake and elimination equally. It is the most important determinant of the half life of COHb, or $\tau_{1/2}$, where $\tau_{1/2} = 0.693 \tau_o$.

The time constant for uptake and elimination, τ_o , or equivalently the half-life, $\tau_{1/2}$, is the one parameter of the Coburn model that varies the most from person to person. Because it is inversely proportional to the breathing rate it depends on the exposed person's level of exertion, excitement, anemia, and general physical condition. The variation with exertion is quite large; pulmonary ventilation rate can increase by a factor of ten from rest to heavy exercise.

Consequently, the absolute rates of CO uptake and elimination may vary over a range of six to one depending primarily on the level of physical activity. Our own simulations of the nonlinear Coburn equation show that the half life can vary from 277 minutes for a resting individual, to as low as 48 minutes for a vigorously exercising person¹⁴. Even among sedentary persons the time constant may vary considerably depending on how "sedentary" or "nervous" the subject is. For example, in one study measurements of $\tau_{1/2}$ for 18 sedentary nonsmoking subjects in 39 trials ranged from 128 to 409 minutes, with an average of 320 minutes².

How Well Does the Coburn Equation Model COHb Uptake and Elimination ?

The linear Coburn equation predicts a person's CO uptake and elimination closely. It can even be used to approximate CO uptake and elimination in a population, as long as all members of the population are engaged in similar levels of physical activity. Figures A-1 and A-2 show the solution of Eqn. A-15 compared to data for numerous measurements of COHb in 18 sedentary subjects.

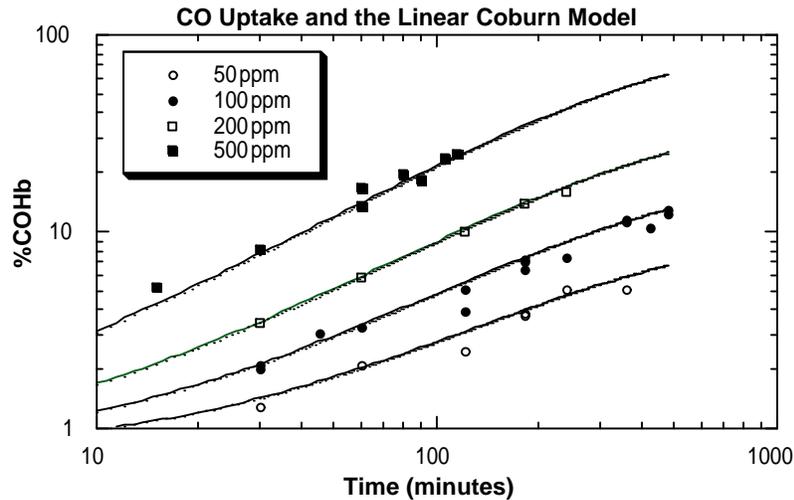


Figure A-1. The increase in %COHb with exposure to CO. The points represent averages for several subjects while the lines represent computer simulations by Mosaic Industries of the linear Coburn equation, Eqn. A-19. Data are taken from reference A-2.

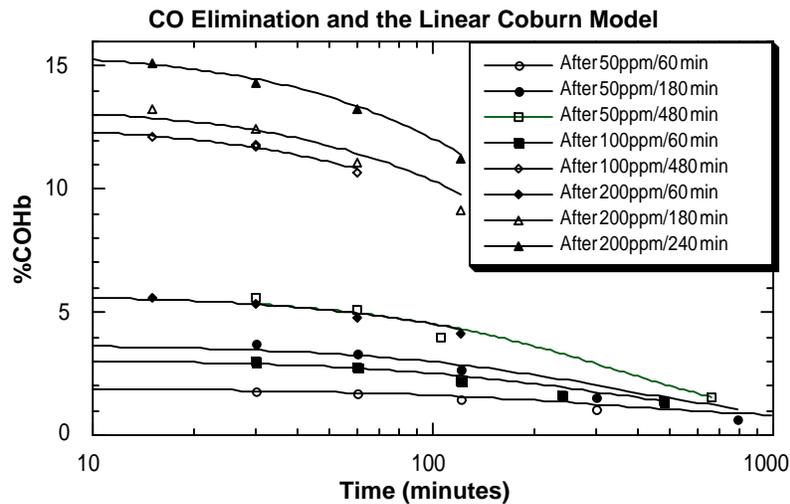


Figure A-2. The decay of %COHb after various cumulative exposures to CO while breathing clean air. The points represent averages for several subjects while the lines represent computer simulations by Mosaic Industries of the linear Coburn equation, Eqn. A-19. Data are taken from reference A-2.

The data is taken from Peterson and Stewart². Each point represents the average %COHb for several subjects exposed to constant CO for a period of time, while the lines represent a computer simulation of Eqn. A-15 (actually, the difference form of Eqn. A-20 is used with updates at one minute intervals) using $\%COHb_e = 0.76$, $\alpha_o = 0.166$, and $\tau_o = 359$ minutes. The data of Fig. A-1 represent %COHb after exposure to constant concentrations of approximately 50, 100, 200, and 500 ppm CO. The data of Fig. A-2 represents the decay of %COHb during exposure to clean air after the subjects had been exposed to the listed cumulative doses.

As shown in the figures, the single Coburn equation, with a single set of the three parameters, accurately describes both CO uptake and elimination. While the data shown in Figures A-1 and A-2 result from exposures to *constant* CO concentrations, the Coburn equation accurately models %COHb as a function of time in response to *any* profile of CO exposure. For example, Peterson and Stewart² show the conformance of the equation to measurements during exposures to steadily increasing CO concentration. The Coburn equation provides a method of time-averaging a CO dose no matter what the variation in CO concentration over time. The time-averaged dose is conveniently expressed in terms of %COHb.

Thus, for a single individual at a single metabolic level, a single set of parameters in the Coburn accurately model CO uptake and elimination.

Using the Coburn Equation to Model a More Diverse Population

As emphasized earlier, the absolute rates of CO uptake and elimination vary from person to person, and may vary over a range of six to one depending primarily on the level of physical activity. Owing to this variation, a single time constant is insufficient to model CO uptake and elimination in an individual with varying metabolic rate, or across a population. Instead, a variation in the time constant is required.

Certain demographic groups are at increased risk of harm from CO exposure. Particularly at risk are children, pregnant women and their unborn babies, the elderly and infirm, people with cardiovascular disease or respiratory problems, and others with chronic medical conditions. For example, low doses of CO produce significant effects on cardiac function during exercise in subjects with coronary artery disease²⁰, and even ambient CO levels are positively associated with hospital admissions for congestive heart failure.²¹ At particular risk are the unborn; when a pregnant woman is exposed to carbon monoxide, her fetus will eventually come to have a higher COHb level than her own maternal circulation. Initially the rise in COHb in the fetus lags behind that of maternal COHb, since the CO has to pass through the maternal circulation before reaching the fetal circulation. Eventually, however, the higher affinity of the fetus' hemoglobin (a mixture of fetal-Hb and adult-Hb) for CO means the fetal COHb level will exceed the maternal COHb level, by a factor of about two to one. For this reason and others, the damaging effects of CO on the fetus are likely to be greater.

Furthermore, even among the healthy, the rate of CO uptake varies over a factor of five or six. The primary physiological determinant of the rate of CO uptake and release is an individual's level of physical activity, usually measured in mets, and indicated by the respiratory minute volume (RMV) or alveolar ventilation rate. The RMV ranges used by the CPSC in its Recommendations to UL 2034 (September 1996)²² are:

- 6-10 L/min, for sedentary activity
- 10-20 L/min, for light activity
- 20-30 L/min, for moderate activity
- 30-40 L/min, for heavy activity

Unfortunately, a single RMV can not adequately characterize CO absorption in a person throughout the day (they may be moderately active during the day, but asleep at night), nor can it be applied to subpopulations in vastly different states of health (an elderly person with cardiac insufficiency is likely to be relatively inactive, while an active child will have a moderately high specific RMV). Further, as a person experiences the symptoms of CO poisoning their metabolic level drops, changing the dynamics of further CO uptake and elimination.

How can an integration formula that relies on a single time constant to characterize the rate of CO uptake/elimination be applied to individuals within the population to be protected with such different responses to CO and with time-varying activity levels? Well, it simply can not. Currently, the UL standard for residential CO alarms (UL 2034)²³ seeks to resolve this dilemma by choosing a single time constant, characteristic of a heavy work level (30 L/min).

The advantage of this choice is that it results in a worst case (or conservatively protective) scenario for the rate of CO uptake; the model assumes that CO is absorbed rapidly and %COHb climbs quickly. Unfortunately, that choice also results in the best case (or most unprotective) scenario for the rate of CO elimination, with %COHb dropping quickly after exposure. Unfortunately, the level of COHb attained during varying CO concentrations is underestimated, and the length of time the COHb is present is also severely underestimated. The effects of CO poisoning depend not only on the COHb level, but also how long that level is maintained. Ironically, using the shorter time constant results in modeling the entire population, including those who are at greatest risk because of cardiac insufficiency, with those who are most athletically healthy.

It is better to use an integration model characterized by a conservatively protective time constant for CO *absorption*, and a different, but also conservatively protective time constant for CO *desorption*.

The practical algorithm recommended here assumes that %COHb rises as quickly as it would in a person engaged in heavy activity, but falls as slowly as it would in a sedentary person. But what should the two time constants be? They are taken from CPSC's recommendations to UL 2034 of September 1996.

For CO uptake, the model uses a time constant determined by a respiratory minute volume of 30 L/min, taken from the low end of the heavy activity level (the same level implicitly used by UL 2034 in its citation of the Coburn equation) combined with a pulmonary diffusion capacity for CO typical of high ventilation rates (30 ml/mmHg). For CO elimination, the model uses a time constant determined by a respiratory minute volume of 6 L/min, taken from the low end of the sedentary level (but still much greater than the sleeping level of 4 L/min) combined with a pulmonary diffusion capacity for CO typical of low ventilation rates (60 ml/mmHg). The way in which these values determine the time constants is described in Appendix A.

The uptake time constant, $\tau_o = 97$ min, is equivalent to the time constant implicit in the form of the Coburn equation now cited in the UL standard, while the elimination time constant, $\tau_o = 368$ min, is typical of that of a sedentary person, or a person who has been exposed to CO.

A Practical Algorithm for Implementing the Coburn Equation in CO Alarms

As we have seen, the Coburn equation provides a means of time-averaging a CO dose no matter what the variation in CO concentration over time. The time-averaged dose is conveniently expressed in terms of the %COHb that would be present in a hypothetical exposed subject. Whenever this predicted %COHb exceeds a predetermined threshold (or trigger point) an alarm can then be sounded.

A more directly applicable form of the equation is found by converting either Eqn. A-15 or Eqn. A-16 to a difference form as,

$$\%COHb_{t+\Delta t} = \%COHb_t \beta + (\%COHb_e + \alpha_o CO_{ppm}) (1 - \beta) \quad \text{Eqn. A-20}$$

in which:

$\%COHb_{t+\Delta t}$ is the new estimate for $\%COHb$;

$\%COHb_t$ is the estimated $\%COHb$ of the prior operation cycle, where $\%COHb_t$ of the first operation cycle is initialized to $\%COHb_e$;

CO_{ppm} is the CO concentration of the current operation cycle;

α_o is a numerical constant of value 0.166 $\%COHb/ppmCO$ representing the equilibrium ratio of blood COHb to inspired CO concentration;

$\%COHb_e$ is a numerical constant of value 0.5 $\%COHb$ representing endogenous CO production; and,

β is a weighting parameter that may vary from cycle to cycle and that depends on the operation cycle time, Δt , and the physiological time constant, τ_o , as $\beta = \exp(-\Delta t/\tau_o)$, in which τ_o equals either 97 minutes or 368 minutes. For example, for an update time of 3 minutes, and a τ_o of 97 minutes, the weighting parameter is $\beta = 0.9695$; for the same update time and a τ_o of 368 minutes, the weighting parameter is $\beta = 0.9919$.

The time interval between updates, Δt , must be chosen to be much less than the physiological time constant, τ_o , CO_{ppm} is the average CO concentration in the interval t to $t + \Delta t$, and the $\%COHb$ at time zero, $\%COHb_o$, assuming no prior CO exposure, is initialized to the endogenous baseline as,

$$\%COHb_o = \%COHb_e \quad \text{Eqn. A-21}$$

Iterations of this single equation accurately compute carboxyhemoglobin levels whatever the temporal profile of the inspired CO concentration.

This equation is made most protective of the general population by allowing the weighting factor, β , to vary depending on whether the projected COHb level is increasing (uptake) or decreasing (elimination). For each iteration of Eqn. A-20, if the CO concentration measurement of the current operation cycle is great enough so that $(\%COHb_e + \alpha_o CO_{ppm} > \%COHb_t)$ then a weighting parameter, β , corresponding to $\tau_o = 97$ is used, otherwise a β corresponding to $\tau_o = 368$ is used. These time constants, τ_o , correspond to the range in respiratory minute volumes (RMV) provided by the CPSC in its October, 1996 recommendations to UL.

This equation can be readily implemented in CO alarms as it requires only a few arithmetic operations for each iteration. For example, the projected $\%COHb$ in Figures A-1 and A-2 used an update time of one minute, and parameter values of $\%COHb_e = 0.76$, $\alpha_o = 0.166$, and a single $\tau_o = 359$ minutes (corresponding to a single person at a single metabolic level), and implemented the update scheme,

$$\text{New_}\%COHb = 0.00279 (0.76 + 0.166 CO_{ppm}) + 0.99721 \text{ Old_}\%COHb \quad \text{Eqn. A-22}$$

On 8- and 16-bit processors the variables can be scaled so that the arithmetic can be accurately implemented using 16-bit integers.

Consider an example in which we assume that the update time is one minute. The running average CO concentration as $\%COHb$ is kept in variables initialized to $\%COHb_e$ as,

$$\text{New_}\% \text{COHb} = \text{Old_}\% \text{COHb} = \% \text{COHb}_e = 0.5\% \quad \text{Eqn. A-23}$$

where 0.5% is taken as the typical endogenous carboxyhemoglobin level (Eqn. A-17). The update algorithm then requires that the following sequence of equations (Eqns. A-24 to A-27) be iterated once per minute, using current values of the CO concentration, expressed in ppm. First, we test whether

$$0.5 + 0.166 \text{ CO}_{\text{ppm}} > \text{Old_}\% \text{COHb} \quad \text{Eqn. A-24}$$

Then, either equation A-25 or A-26 is used to update an estimate of the %COHb. If the test of Eqn. A-24 is true, we implement,

$$\text{New_}\% \text{COHb} = (0.5 + 0.166 \text{ CO}_{\text{ppm}}) / 97 + \text{Old_}\% \text{COHb} (1 - 1/97) \quad \text{Eqn. A-25}$$

and if not, we instead implement,

$$\text{New_}\% \text{COHb} = (0.5 + 0.166 \text{ CO}_{\text{ppm}}) / 368 + \text{Old_}\% \text{COHb} (1 - 1/368) \quad \text{Eqn. A-26}$$

Then we update the iteration as,

$$\text{Old_}\% \text{COHb} = \text{New_}\% \text{COHb} \quad \text{Eqn. A-27}$$

Each minute, a test compares New_%COHb to the alarm threshold and activates the alarm if New_%COHb is greater than the threshold.

When implementing the update of Eqn. A-25 or A-26 on small processors scaled integer arithmetic must generally be used. Scale factors will depend on the magnitude of the sensor signal, the A/D method employed, and the update time interval. Scale factors and the update time interval can also be chosen to optimize one or another of the arithmetic operations.

While the implementation of this equation will depend on processor used and the scale factors of the A/D, we offer an example meant to give a flavor for the integer algorithms possible. We will choose scale factors that reduce multiplications and divisions to factors of two so that they can be implemented on simple processors using shift operations. Consider the following scale factors:

- The update time is chosen to be 45 seconds rather than one minute. This time is generally under the complete control of the instrument designer and can be chosen with wide latitude.
- CO_{ppm} is unity scaled, that is 100 ppm CO is represented by the integer 100. This scale is readily attainable with a 10-bit A/D convertor and a sensor whose full scale response is 1000 ppm.
- %COHb is scaled so that the integers 0 – 65536 represent 0 – 42.5 %COHb. This scale factor is chosen to reduce multiplications and divisions to shift operations.
- The time constant for elimination, 368 seconds, is more conservatively set to a nearby value, 388 seconds (an even factor of 4 greater than the lesser, more critical time constant of 97 seconds).

Under these conditions the test of Eqn. A-24 becomes,

$$771 + 256 * \text{CO}_{\text{ppm}} > \text{Old_}\% \text{COHb} \quad \text{Eqn. A-28}$$

Because the multiply by 256 is a multiply by 2⁸ it is implemented by simple byte selection. The update formulas of Eqns. A-25 and A-26 become,

$$\text{New_}\% \text{COHb} = 6 + \text{Old_}\% \text{COHb} - \text{Old_}\% \text{COHb} / 128 + 2 * \text{CO}_{\text{ppm}} \quad \text{Eqn. A-29}$$

and,

$$\text{New_}\% \text{COHb} = 1.5 + \text{Old_}\% \text{COHb} - \text{Old_}\% \text{COHb} / 512 + \text{CO}_{\text{ppm}} / 2 \quad \text{Eqn. A-30}$$

In these computations the multiplications and divisions are by powers of two and are handled by simple shift operations. The addition of 1.5 in Eqn. A-30 is implemented by adding either 1 or 2 on alternate iterations. A comparison of New_%COHb to 10000 (representing 6.5% COHb) determines whether the alarm condition is activated. Using slightly different up and down thresholds (for example 6.25% and 6.75%) is recommended to implement a small hysteresis in alarm activation. Alternatively, a time delay that prevents too rapid alarm extinction can be implemented.

In the likely event that the scale factor of the A/D operation, in which the CO_{ppm} value is acquired, is controlled by hardware considerations, other scale factors should be changed. The ability to arbitrarily scale the %COHb values within a 16-bit integer provides the degree of freedom necessary to guarantee that multiplications and divisions can always be reduced to simple shift operations.

Appendix A References

- A-1. R.F. Coburn, R.E. Foster, and P.G. Kane, "Considerations for the physiological variables that determine the blood carboxyhemoglobin concentration in man", *Journal of Clinical Investigation*, Vol. 44, (1965), pp 1899-1910.
- A-2 J.E. Peterson and R.D. Stewart, "Absorption and Elimination of Carbon Monoxide by Inactive Young Men", *Arch. Environ. Health*, Vol. 21, (1970), pp 165-171.
- A-3. See any basic reference on clinical general or respiratory physiology. A good introductory book available on-line is "Pulmonary Physiology in Clinical Practice, the Essentials for Patient Care", by Lawrence Martin, M.D., Chief of the Pulmonary Division of Mt. Sinai Medical Center, Cleveland Ohio, © February 1999, available on the internet at www.mtsinai.org/pulmonary/books/physiology.
- A-4. Peterson and Stewart [A-2] use 1.38 rather than the 1.34 of [A-3] in Eqn. A-5 but neglect the second, partial pressure term for the solubility of oxygen in the blood plasma, resulting in the same value of [O₂]_{sat}.)
- A-5. "Test Protocols for Residential Carbon Monoxide Alarms, Phase 1", Gas Research Institute Report GRI-96/0055, by Paul K. Clifford and Michael G. Dorman, Mosaic Industries, Inc., Sept. 1996. Figures 2.1 and 2.2 show the conformance of computer simulations of the nonlinear Coburn equation to clinical measurements.
- A-6. For results of such a simulation using the data of Peterson and Stewart [A-2] see [A-5].
- A-7. Underwriters Laboratories Inc., Standard for Safety UL 2034, "Single and Multiple Station Carbon Monoxide Alarms," ISBN 0-7629-274-9, First Edition April 1992, Second Edition October 29, 1996, revised October 15, 1997. The revisions of October 15, 1997 were effective October 1, 1998. Note that the UL Standard uses different notation. They cite the following equation: %COHb_t = %COHb_o [e^{-(t/2398B)}] + 218 [1-e^{-(t/2398B)}] [0.0003 + (ppmCO/1316)] with B=0.0404. This expression is equivalent to Equation A-16 with α = 0.166, %COHb_e = 0.065%, and τ_o = 97 min. (The %COHb_e of 0.065% is unrealistic; an appropriate value would be 0.5%.)

The UL equation represents the particular solution of the more general Coburn model of Eqn. A-15 for constant CO concentration. For use with arbitrary concentration profiles, the more general form, Eqn. A-15, or its difference equation form, Eqn. A-19, should be used.

- A-8. Peter S. Lee, Richard F. Majkowski and Thomas A. Perry, "Tunable diode laser spectroscopy for isotope analysis -- Detection of isotopic carbon monoxide in exhaled breath", *IEEE Transactions on Biomedical Engineering*, Vol. 28, No. 10, October 1991, pp. 966-973.
- A-9. This data is widely quoted, for example in Richard A. Wadden and Peter A. Scheff, *Indoor Air Pollution: Characterization, Prediction, and Control*, John Wiley & Sons Publ., 1983, ISBN 0-471-87673-9, Chapter 2, and in "Criteria for a recommended standard...Occupational Exposure to Carbon Monoxide", U.S. Department of Health, Education and Welfare, NIOSH, 1972, HSM 73-11000.
- A-10. R.D. Stewart, "The effect of carbon monoxide on humans", *Journal of Occupational Medicine*, Vol. 18, (1976), pp. 304-309.
- A-11. Marc B. Schenker, "Indoor Air Pollution, Eight Questions Physicians Often Ask", *Consultant*, August 1991, pg 59-64.
- A-12. "Criteria for a recommended standard...Occupational Exposure to Carbon Monoxide", U.S. Department of Health, Education and Welfare, NIOSH, 1972, HSM 73-11000
- A-13. "From the Centers for Disease Control and Prevention: Leads from the Morbidity and Mortality Weekly Report: Carbon Monoxide Levels During Indoor Sporting Events -- Cincinnati, 1992-1993", *Journal of the American Medical Association*, Vol. 261 No. 6, February 9, 1994, pp. 419.
- A-14. See figure 2.2 and reference 33 of [A-5].
- A-15. "Review of the national ambient air quality standards for carbon monoxide: 1992 Reassessment of scientific and technical information, OAQPS staff paper", U.S. Environmental Protection Agency, Report No. EPA-452/R-92-004, (August 1992), pp 16 and 20.
- A-16. "Indoor Air Pollution, An Introduction for Health Professionals", a joint publication by the American Lung Association, American Medical Association, U.S. Consumer Product Safety Commission, and U.S. Environmental Protection Agency Indoor Air Division.
- A-17. "Handbook of Air Pollution" by James P. Sheehy *et al*, U.S. Dept. of Health, Education, and Welfare, Public Health Service Publication No. 999-AP-44.

The half life of COHb while breathing fresh air varies from 277 minutes for a resting individual, to 91 minutes for the level of exercise assumed by UL 2034, to 63 minutes for the "heavy work" level of the Mosaic model. Further details of the model are available from Mosaic Industries, Inc.
- A-18. The work level, ventilation rate, and CO diffusivity assumed by UL 2034 is taken from S. Steinberg and G.D. Nielsen, "A proposal for evaluating human exposure to carbon monoxide contamination in military vehicles", U.S. Army Human Engineering Laboratory, AMCMS Code 672716.H700011, March 1977.
- A-19. The model predictions were calculated by Mosaic Industries using computer simulations of the nonlinear Coburn-Forster-Kane single-compartment differential equation as quoted by [A-9] and in [A-12].

The nonlinear model differs from the linear model in that it accounts for the decrease in oxyhemoglobin as COHb increases instead of approximating it as a constant. For the resting nonsmoking subjects the model assumed an average alveolar ventilation rate of 6 l/min and a diffusion capacity of the lungs for CO of 30 ml/min-mm Hg; other parameters of the model were adjusted to fit the data. Model predictions for individuals engaged in heavy work are based on the same model parameters except that alveolar ventilation is assumed to increase to 40 l/min and the diffusion capacity to 60 ml/min-mm Hg. (For comparison, the UL COHb/exposure curves are also based on a diffusion capacity of 60 ml/min-mm Hg but an alveolar ventilation of 30 l/min.) The values used in the Mosaic simulations result from assuming that while progressing from a sedentary metabolism to a work rate of 700 kg-m/min the breathing rate increases from 10 to 21.2 breaths/min, tidal volume increases from 0.75 l to 2.03 l, dead volume remains at 0.15 l, and minute ventilation increases from 7.5 l/min to 43 l/min.

This heavy work exercise level of 700 kg-m/min and its consequent breathing rate are still less than one half of the maximum work load and ventilation rate of which healthy young men are capable. The relationship we use between work levels, tidal volumes and ventilation rates is taken from [A-17]. The half life of COHb while breathing fresh air varies from 277 minutes for a resting individual, to 91 minutes for the level of exercise assumed by UL 2034, to 63 minutes for the “heavy work” level of the Mosaic model. Further details of the model are available from Mosaic Industries, Inc.

- A-20. Allred, E.N., Bleecker, E.R., Chaitman, B.R., Daham, T.E., Gottlieb, S.O. (1991) *Effects of carbon monoxide on myocardial ischemia*, Environ. Health Perspectives, 91, 89-132
- A-21. Morris, R.D., Naumova, E.N., Munasinghe, R.L. (1995) *Ambient air pollution and hospitalization for congestive heart failure among elderly people in seven large US cities*, Am. J. Publ. Health, 85, 1361-1365.
- A-22. “Recommendations of the Consumer Product Safety Commission (CPSC) Staff Concerning Underwriters Laboratories (UL) Inc. Standard 2034 for Single and Multiple Station Carbon Monoxide Detectors (First Edition – 1992; New and Revised Requirements – August 1995)”, CPSC, October 25, 1996, Elizabeth W. Leland, Project Manager, CO Detection.
- A-23. Underwriters Laboratories Inc., Standard for Safety UL 2034, “Single and Multiple Station Carbon Monoxide Alarms,” ISBN 0-7629-274-9, First Edition April 1992, Second Edition October 29, 1996, revised October 15, 1997. The revisions of October 15, 1997 were effective October 1, 1998.

Appendix B:

Guidelines for Assuring the Time-of-Manufacture and In-Service Reliability of CO Alarms

This Appendix is intended to provide practical guidance for testing CO alarms to demonstrate compliance with standards for time-of-manufacture and in-service reliability. In particular, it is meant to assist alarm manufacturers in identifying sample sizes and frequencies for testing. Throughout this Appendix the CSA 6.19-01 Standard for Residential Carbon Monoxide Alarms is cited as an example for its reliability requirements. Even so, the methods described are not limited to those particular requirements, but are equally useful when designing tests to other requirements, and should be applicable with little modification to future UL Standard requirements.

This Appendix contains the following sections:

Definitions

Specific terms, nomenclature and symbols used throughout this document are defined. Consult these definitions if you have any questions about the precise meaning of particular variables in equations.

Introduction

The measurement of failures and reliability are generally discussed.

Assuring Time-of-Manufacture Reliability

This section presents a statistical basis for choosing the number of manufactured alarms to be tested, and it provides a step-by-step procedure for assessing time-of-manufacture reliability.

Assuring In-Service Reliability

This section presents a statistical basis for choosing the number of alarms and the times they must be operated before testing, and it provides a step-by-step procedure for assessing in-service or lifetime reliability.

Statistical Derivation

While the sampling methods described in this document are commonly used in statistical literature on reliability testing, they are rarely explained in full. Because good explanations are hard to find, a full derivation is provided in this section. Understanding this section is not needed to apply the foregoing methods of quality assurance – it is only provided for those interested in understanding the philosophical and mathematical basis for the statistics used

Definitions

CSA Standard

CSA 6.19-01 Residential Carbon Monoxide Detectors, its Amendments and Revisions, available from the Canadian Standards Association, 178 Rexdale Boulevard, Etobicoke, Ontario M9W 1R3.

Device

“Device”, “unit”, “detector”, and “alarm” are used interchangeably in this document to indicate CO alarms.

Device Time (T_{DUT})

The number of devices undergoing in-service reliability tests multiplied by the time that they are powered up and operated under simulated residential conditions. Device time is measured in either device-quarters or device-hours. For calculating device time, a year is defined as 365.25 days or 8766 hours, a quarter as 2191.5 hours, and a month as 730.5 hours. As a practical matter, quarters can be taken as any contiguous three-month periods, and the small errors introduced by the months’ differing numbers of days can be ignored.

DUT

An abbreviation for Devices Under Test

F_{LS}

The observed lifetime failure rate of a sample of devices, expressed as the cumulative fraction of devices failed at the end of a presumptive three year lifetime.

F_{L90}

An upper bound, evaluated at a 90% confidence level, on the lifetime failure rate.

F_{TOM}

The observed time-of-manufacture failure rate expressed as the ratio of failed devices to the number of devices actually tested. $F_{TOM} = N_{fS} / N_{DUT}$

F_{TOM90}

An upper bound, evaluated at a 90% confidence level, on the time-of-manufacture failure rate for an entire population of alarms. It is expressed as the ratio of an estimated upper bound on the number of failures in a tested sample to the total number of devices in the sample as,

$$F_{TOM90} = N_{f90} / N_{DUT}$$

where N_{f90} is an upper bound on N_{fS} found from a statistical table (Table B-1).

Failure

As defined in the Reliability Appendix to the CSA Standard, a failure is classified as either *Supervised* or *Unsupervised*.

A *Supervised Failure* is the sounding of a trouble signal, resulting from any condition other than low battery power, immediately prior to which the device was fully functional and

operating within the sensitivity specification of the CSA Standard Section 7.5, for tests at CO concentrations of 70, 150, and 400 ppm, but excluding the test at 30 ppm.

An *Unsupervised Failure* is any failure that is not a supervised failure; that is, whenever a device does not satisfy the sensitivity specification of the CSA Standard Section 7.5, for tests at CO concentrations of 70, 150, and 400 ppm, but excluding the test at 30 ppm, without sounding a trouble signal.

Lifetime

The operating life of an alarm is defined as the time during which the cumulative supervised failures remain fewer than 23.1% of alarms, and the cumulative unsupervised failures remain fewer than 14.6%. These cumulative failure rates, if they occur over a three year period, indicate mean-times-between-failure for supervised and unsupervised failures of 100,000 hours and 166,667 hours respectively. Alarms must have a minimum operating life of three (3) years in residential service. Measurements of in-service reliability are made with a view to estimating alarms' average reliability over their lifetime and assuring that their lifetime is greater than the minimum required.

MTBF

An average Mean-Time-Between-Failure of CO alarms over their lifetime. We can not know the actual MTBF unless all devices produced are tested over their entire lifetime, an impossible task if product is released. Reliability assessment instead *estimates* MTBF by testing a limited number of devices (a sample) for a limited time and computing an upper bound on the number of failures at a 90% confidence level. The upper bound on the MTBF, the $MTBF_{90}$, is computed from the observed MTBF of the sample, $MTBF_S$.

$MTBF_S$

The mean-time-between-failure observed of a sample of devices.

$MTBF_{90}$

An upper bound on the actual, but unknown, mean time between failure at a 90% confidence level for all alarms of a particular model. $MTBF_{90}$ is estimated quarterly based on the number of failures observed in a sample of devices held for progressively longer duration.

N_{DUT}

The number of devices under test.

N_f

The number of failures expected in a population, reflective of the true population failure rate.

N_{fS}

The number of devices in a sample observed to fail a test.

N_{f90}

An upper bound on the number of failures at a 90% confidence level. It is computed from the number of observed failures, N_{fS} .

No

The number of unfailed devices in a sample.

Quarter

A quarter is defined as one quarter of a year, or 2191.5 hours. As a practical matter, quarters can be taken to be any integral three month periods, and the small errors introduced by the months' differing numbers of days can be ignored. Assessments of in-service reliability are made on a quarterly basis.

Sample

A collection of alarms drawn from final production that are subsequently subjected to tests.

Δt

Elapsed time, in hours or quarters.

T_{DUT}

see Device Time

Test

A test (tests, testing) is defined as the sensitivity test of the CSA Standard Section 7.5, for tests at CO concentrations of 70, 150, and 400 ppm, but excluding the test at 30 ppm.

Introduction

Reliability is the measure of how well CO alarms meet their operational specifications, particularly their sensitivity to CO gas. An alarm's lifetime is characterized by three phases: *infant mortality*, *useful life*, and *wear-out*, each with its own reliability figure. When a detector is first manufactured a variety of failure mechanisms show up under relatively moderate stress levels and therefore appear early. This period of early failures, termed infant mortality, can be reduced through proper manufacturing controls and sensor screening techniques. Immediately after production detectors should be "burned-in", that is, operated under normal conditions, for a period of time expected to be greater than their "infancy". This allows the effective detection of many of the early failures and the removal of failed units. After burn-in, the failure rate of the remaining units is termed the *time-of-manufacture* failure rate and it is measured by statistical sampling and testing. This failure rate is the fraction of units that after burn-in, and after having eliminated infant mortalities, still do not meet their sensitivity specification. It is expressed as a percentage.

After the period of infant mortality is a period of useful life during which only occasional random failures appear. During this lifetime the failure rate is usually low and approximately constant. The sources of failure during this time are varied, including environmental stimulus, inherent physical defects, irreversible chemical change, sensor drift, and thermal degradation. Taken together, all these failure mechanisms can be characterized by a single number, the lifetime failure rate, or its inverse, the mean-time-between-failure (MTBF).

The final period is that in which the detectors literally "wear-out". The causes of wear-out are numerous, including sensor drift, exposure-related degradation, and the depletion of chemical reagents. If the sensing technology is sufficiently well matched to the application and the detector well designed, this period occurs beyond the lifetime required by the consumer.

The reliability of alarms is assured by assessing their failure rates at the time-of-manufacture and throughout their lifetime. Because it is generally not feasible or too expensive to test the entire population of manufactured units, statistical methods for sampling and testing fewer units are used. The results of testing this *sample* of devices are subject to statistical fluctuations. If we were to repeatedly draw the same number of devices for a sample, after testing them we would find different numbers of failures, not because of any inability to count the failures properly, but because samples distributed randomly contain numbers of failures that fluctuate from sample to sample. If we were to repeat our tests numerous times, with different devices each time, and the measured failure rates were below a particular value a certain percentage of the time, say 90%, then we would treat that value as an upper bound on the failure rate at a 90% confidence level (CL). Statistical methods allow us to perform only one set of tests and still determine this upper bound. Because the entire population of devices is not tested we can not really *determine* the actual population failure rate, but we can, based on the sample failure rate, *estimate* an upper bound for it, and know with a certain confidence that it is less than the upper bound.

This document provides guidelines for determining the numbers of units that must be sampled and tested to find an upper bound for the population failure rate at a 90% confidence level. Data in Table B-5 is also provided for those wishing to test at confidence levels other than 90%.

Assuring Time-of-Manufacture Reliability

Overview

Time-of-manufacture reliability testing requires that a sufficient number of newly manufactured CO detectors be tested in order to provide a statistically valid assessment of their conformance to their sensitivity specification.

The CSA Standard requires that alarms ready for sale must adhere to an upper bound on their failure rate. Briefly, at the time of manufacture and after burn-in if applicable, devices must have a total failure rate, for Supervised and Unsupervised Failures *combined*, of less than 1% when estimated at a 90% confidence level. This failure rate must be calculated quarterly or for each major change in design, production or components, whichever comes first. Please consult the reliability appendix of the revised CSA Standard for Carbon Monoxide Alarms (CSA-6.19-01) for the precise requirements.

Customarily, after manufacture 100% of all units produced are allowed to “burn-in” for a period of time, after which obviously nonfunctioning units are culled from the population for removal or repair. Failure during burn-in may be defined in terms of spontaneous alarm, failure to activate when the test button is depressed, or other obvious malfunctions indicative of a defective unit. Burn-in of 100% of production is consistent with the requirements of CSA 6.19-01 and with the specifications of many component manufacturers including sensor manufacturers. The remaining units after burn-in, those that would normally be sold to the public, should then be randomly sampled for reliability testing. If no burn-in is implemented, or if not all units are burned-in, the population to be sampled should in any case consist of the pool of detectors to be sold to the public.

The statistical methods underlying reliability testing presume a constant time-of-manufacture failure rate for the CO detector. Anything that may alter this failure rate (either by increasing or decreasing it), including changes in production method, components or design, handling of components or finished goods, or systematic changes in product quality or consistency over time owing to learning curve

efficiencies, would define a new population for testing purposes and require a reassessment of reliability by restarting sampling and testing. Additionally, time-of-manufacture reliability must be assessed at least as frequently as on a quarterly basis.

Time-of-manufacture testing suggests the use of classical statistical methods for binomial experiments, in which the experiment comprises a number of repeated trials, each trial results in a “pass” or a “fail”, and the probability of failure is constant trial to trial.

Time-of-manufacture reliability assurance requires that a sample of devices be tested and the number of failures tallied. In a manufacturing environment the size of the production population can not be specified up-front and testing must proceed over a growing population. In this case, statistical methods must be able to use a tally of observed failures (incidents of noncompliance with a given set of performance criteria) to extrapolate a maximum failure rate for the entire population at a chosen confidence level. This approach, commonly used in statistical literature on reliability testing but rarely explained in full, is independent of population size. Because good explanations of the method are rare, a full derivation is provided in the final section of this Appendix.

To estimate the time-of-manufacture failure rate a sample of devices is randomly chosen from each quarter’s production. They may be chosen on a weekly basis if the total quarter’s production is known before hand. Manufacturers must exercise care in establishing a procedure for random sampling. The procedure used may be highly specific to the manufacturer. For example, the practical difficulty in randomly sampling from a production run of 100,000 units is greater than for a production run of 1,000 units. Nevertheless, random sampling helps avoid systematic bias in product quality across the sample.

The sampled devices are tested and the number not passing all the required tests is tallied. The observed sample failure rate is then,

$$F_{TOMS} = N_{fs} / N_{DUT} \quad \text{Eqn. B-1}$$

in which N_{fs} is the number of devices of the sample failing to comply with at least one sensitivity test and N_{DUT} is the total number of devices tested.

In choosing the sample size, N_{DUT} , there is a fundamental trade-off between the cost of testing and the accuracy of the measured failure rate. The greater the sample size, the more failures are tallied and the more accurately the actual failure rate can be estimated, but at greater cost. Naturally, we would like to test the minimum number of devices. We therefore use statistics to determine the minimum number of units needed to assure that the failure rate does not exceed the maximum allowed at a 90% confidence level (CL). The size of this *statistical* sample depends on the target reliability and the number of failures expected.

Owing to statistical fluctuations, for any observed failure rate the actual population failure rate may be less or greater, and this discrepancy is most pronounced when only small numbers of failures are observed. An analysis of the appropriate statistics (see the final section of this Appendix for details) shows that for any number of observed failures there is an upper limit on the expected number of failures at a 90% confidence level, *whatever the number of devices tested*. These limits are shown in Table B-1. For numbers of failures less than 100 but not shown in the table, values for the upper bound should be interpolated from the tabulated values; for failures greater than 100 the final row of the Table provides an approximate equation.

We see from Table B-1 that even if no failures are observed in our sample we can not simply assume that there is a zero population failure rate. Instead, we can be 90% sure that the failure rate is less than that

which would result from observing 2.3 failures. So rather than using the observed failure rate of Eqn. B-1 we can compute an upper bound on the failure rate at the 90% confidence level as,

$$F_{\text{TOM } 90} = N_{f90} / N_{\text{DUT}} \quad \text{Eqn. B-2}$$

in which N_{f90} is the upper bound on the number of failures corresponding to the number of failures actually observed, from Table B-1.

Table B-1. Upper Bound on the Expected Number of Failures at a 90% Confidence Level, Based on the Number of Failures Observed	
Failures Observed N_{fS}	Upper Bound N_{f90}
0	2.30
1	3.89
2	5.32
3	6.68
4	8.00
5	9.28
6	10.54
7	11.77
8	13.00
9	14.21
10	15.41
11	16.60
12	17.78
13	18.96
14	20.13
15	21.30
20	27.05
30	38.33
40	49.39
50	60.35
100	114.06
$N_{fS} > 100$	$N_{fS} + 0.82 + 1.28 \sqrt{(N_{fS} + 0.41)}$

The CSA reliability standard requires that this failure rate be 1% or less (for supervised and unsupervised failures combined),

$$F_{\text{TOM } 90} = N_{f90} / N_{\text{DUT}} \leq 0.01 \quad \text{Eqn. B-3}$$

The minimum sample size needed is computed from Eqn. B-3 by assuming that no failures will be observed, and using the minimum N_{f90} as,

$$N_{\text{DUT}} \geq N_{f90} / 0.01 = 2.30 / 0.01 = 230. \quad \text{Eqn. B-4}$$

Even if no failures are observed the minimum number of alarms that must be tested each quarter is 230. This is a minimum; it may be wise to test more. The additional number of units tested should be based on a trade-off between the marginal costs of their testing and the number of failures expected to be observed.

Devices may be tested on a batch basis, where the marginal cost of testing a complete batch over an incomplete batch is low. In this case it may be wise to round up the number of units to fill a complete batch.

If in testing these 230 (or more) devices no failures are observed then the time-of-manufacture failure rate will have been established to be 1% (or less) at a 90% confidence level.

But what if some failures *are* observed? In that case, we will not have assured the reliability at the required level, but it is still possible that we may be able to do so by testing more units. Owing to statistical fluctuations, even if the true failure rate is sufficiently small, more failures than are allowed at the 90% confidence level might be observed when the sample size is small. This is possible because for small numbers of observed failures statistical fluctuations (also called counting error) are a large portion of the observed failure rate. Consequently, the ratio of the upper bound on the number of failures to the observed number of failures is large. As the number of observed failures increases, usually as a result of increasing the sample size, the counting error as a portion of the observed failure rate decreases, and a better estimate of the actual failure rate is made. For larger samples the ratio of the upper bound on the number of failures to the observed number of failures diminishes, approaching unity.

For this reason, if the number of failures exceeds the allowed number when testing a small number of samples, it may not continue to exceed it when a greater sample size is tested. Increasing the sample size may yet demonstrate compliance. However, if the population failure rate really is too large, then increasing the sample size will also increase the number of observed failures, and it will become progressively more apparent that compliance will not be demonstrated.

Consequently, if compliance is not demonstrated at the minimum sample size (of 230 units) the sample size may be increased. For example, suppose one failure had been observed in testing 230 devices. We see in Table B-1 that if one failure is observed the upper bound on the number of failures at the 90% confidence level is 3.89 failures. Using this value in Eqn. B-4 we find that if we test a total of 389 devices and still observe only this one failure we will have demonstrated compliance to the 1% failure rate at a 90% CL. Therefore we must test an additional 159 devices ($230+159 = 389$) in the hope that no more failures are observed. If however more failures are observed then we will still not have demonstrated compliance and the sample size may yet again be increased. The required sample sizes for different allowed numbers of failures are given in “Time-of-Manufacture Sampling Plan” of Table B-2. These numbers are derived from Table B-1 using Eqn. B-3.

If compliance is still not demonstrated with the larger sample the number of devices tested may be increased again. As the sample size increases the costs of testing increase while the estimate of the failure rate becomes more accurate. But if as the sample is increased the number of failures continues to remain too great, just when should we abandon the strategy of increasing the sample size?

When this strategy is abandoned is a judgement call; there is always a chance, however small, that increasing the sample size will eventually result in a demonstration of compliance. But there is an increasingly greater chance that it will not. We suggest that the sample size not be further increased once it becomes *more likely than not* (i.e., at a 50% confidence or more) that the population failure rate is greater than allowed. This is determined by computing the observed sample failure rate, $F_{TOM S}$, from Eqn. B-1. If it is greater than the population failure rate allowed (1% in this case) then the number of samples should no longer be increased, and instead we accept the writing-on-the-wall, that the units are *not-in-compliance*. In this case it will be more efficient to evaluate the reasons for failure rather than to expand testing. Of course, relaxing either the target failure rate or the confidence level would achieve favorable results in tests where failures are observed, but these criteria are set by the CSA requirements and can not be altered.

Table B-2. Time-of-Manufacture Sampling Plan				
Allowed Sample Failure Rate, F_{TOMS}	Number of Failures Allowed¹	Number of Failures Allowed for a Sample Size Increase²	Average Weekly Sample³	Total Sample Size
0.00 %	0	2	19	230
0.26 %	1	3	32	389
0.38 %	2	5	44	532
0.45 %	3	6	56	668
0.50 %	4	8	67	800
0.54 %	5	9	77	928
0.57 %	6	10	88	1054
0.59 %	7	11	98	1177
0.62 %	8	13	108	1300
0.63 %	9	14	118	1421
0.65 %	10	15	128	1541
0.66 %	11	16	138	1660
0.67 %	12	17	148	1778
0.69 %	13	18	158	1896
0.70 %	14	20	168	2013
0.70 %	15	21	178	2130
0.74 %	20	27	225	2705
0.78 %	30	38	319	3833
0.81 %	40	49	412	4939
0.83 %	50	60	503	6035
0.88 %	100	114	951	11406

- 1 The number of failures allowed at the total sample size that demonstrates compliance with a maximum 1% failure rate at a 90% confidence level.
- 2 The number of failures beyond which it is unlikely that compliance will be demonstrated even if the sample size is increased. For an observed number of failures greater than the “Number of Failures Allowed” and less than or equal to the “Number of Failures Allowed for a Sample Size Increase”, it is likely that compliance will eventually be demonstrated if the sample size is increased sufficiently.
- 3 The average weekly sample assumes that there is an average of 12 weeks of production in each quarter.

This number of failures beyond which it is not worth while to increase the sample size is also shown in Table B-2. For example, suppose that 230 newly manufactured units are tested and one failure is observed. This is more than the zero failures allowed to demonstrate compliance with the time-of-manufacture reliability requirement. Is it worth testing more units in the hope that the one failure observed was a “statistical fluke”? Even though the observed number of failures is greater than the number allowed at the 90% CL, the observed failure rate is still less than the maximum allowed. From Eqn. B-1, $F_{TOMS} = N_{fS} / N_{DUT} = 1/230 < 1\%$. Consequently, it is worth while increasing the sample size in the hope that at a larger sample size the number of observed failures falls below the number allowed. If there had been three or more failures then these would have been greater than the “Number of Failures Allowed for a Sample Increase”, and it is likely (with more than a 50% chance) that increasing the sample size would just result in a more precise measurement of a failure rate that is too great. But with only one observed failure, even though compliance was not demonstrated, it is still possible that compliance might

be shown at a greater sample size. Suppose we attempt this by increasing the number of devices under test to 389. If no more failures are observed when we test the additional 159 units then we will have demonstrated compliance because one failure is allowed in tests of 389 or more units.

Continuing this example, if more failures are observed at the 389 unit level we will not have shown compliance. Suppose two more failures were observed so that the total number of failures were three. Would it be worth attempting a larger scale test? As long as the number of failures is less than 4 there is a better than 50% chance that increasing the sample size will yet demonstrate compliance. So we test an additional 279 units to bring the total to 668, a level at which three failures are allowed. If no more failures are observed we will have won our bet and finally demonstrated compliance. If more failures are observed we will not have demonstrated compliance, and if as many as four more failures are observed (for a total of seven) then it is not worth while to increase yet again the number of units under test.

Procedure

All data pertaining to the sampling procedure and test results must be recorded on a controlled test sheet and maintained in a file for future reference.

Each quarter a sample of devices is selected for testing. Devices should be sampled uniformly throughout the quarter in average weekly samples as given in Table B-2. After testing, the units can be resubmitted to production stock. If there is good reason to believe that the population failure rate is well below 1% then at the beginning of this reliability program units may be sampled on a schedule that results in a minimum of 230 units sampled by the end of the first quarter. If there is reason to believe that the population failure rate is an appreciable portion of 1% then a sample size should be chosen from Table B-2 using the estimated failure rate as the “Allowed Sample Failure Rate”. For successive quarters, units should be sampled on a schedule that results in a total sample size chosen from Table B-2 using the observed sample failure rate of the prior quarter as calculated from Eqn. B-1 as the “Allowed Sample Failure Rate”.

Samples of manufactured units may be held and batch tested anytime during the quarter. However, if compliance is not demonstrated at the target sample size the sample size may need to be increased, and it is best not to wait until the end of the quarter to do this. As soon as a failure is observed the sampling rate may be increased so that the next appropriate sample size is attained before the end of the quarter.

If the number of observed failures is less than or equal to the “Number of Failures Allowed” of Table B-2, the devices are deemed to be *in-compliance* for the quarter.

If the number of failures observed is greater than the “Number of Failures Allowed” of Table B-2, but not greater than the “Number of Failures Allowed for a Sample Size Increase”, then either the devices may be deemed to be *not-in-compliance*, or, the total sample size may be increased in an effort to demonstrate compliance at a greater sample size, at the manufacturer’s discretion. If for the larger sample the number of observed failures remains in this interval, the sample size may be increased again.

If, at any sample size, the number of observed failures is greater than the “Number of Failures Allowed for a Sample Size Increase” of Table B-2 (or, equivalently, the observed sample failure rate is greater than 1% as computed by Eqn. B-1), it is recommended that the sample size not be increased further, and that the devices be deemed to be *not-in-compliance*.

If the devices are *not-in-compliance* corrective action must be taken to determine the cause of the failure rate. The certifying agency must be notified and approve an action plan to reduce the failure rate. The

action plan should include the diagnostic tests, disposition, redesign, or rework done to the failed model. Depending on the action plan, the reliability testing may need to be restarted.

As long as the devices remain *in-compliance* and the model is produced, quarterly testing should be continued.

Assuring In-Service Reliability

Overview

Throughout their lifetime some CO alarms may fail abruptly or drift slowly, eventually failing by no longer meeting critical specifications. While we will assume that the failure rate is constant throughout the alarms' lifetime, we know that this is not strictly true. Near the beginning of their lifetime there is often an increased failure rate, an "infant mortality", as marginal components fail, and near the end of their lifetime the failure rate may be enhanced by aging effects. Because the failure rate is not constant, reliability tests over a short period of time usually do not lead to good estimates of the true lifetime failure rate. *Consequently, the goal of a program of in-service reliability is to determine the best estimate possible of the alarm's lifetime failure rate, as soon as possible in the product life cycle, and then to progressively refine that estimate using greater test durations, until measurements are made over the entire lifetime.* Although the assumption of constant failure rate is a poor one, by progressively refining the estimated failure rate over time, the lifetime failure rate is eventually determined.

Failure rates are measured as a number of failures per unit time, and are often expressed in percent failures per thousand hours, percent failures per month, or its inverse, mean-time-between-failures (MTBF) which is expressed in hours. Throughout this guide we use MTBF.

The CSA Standard requires that alarms must adhere to an upper bound on their cumulative failure rate over a presumptive three year lifetime. Briefly, it requires that the average MTBF over the lifetime of the alarm be not less than 100,000 hours for Supervised Failures or 166,667 hours for Unsupervised Failures when estimated at a 90% confidence level. Compliance must be verified through quarterly testing that continues for a minimum period of three years. This reliability procedure must be reinitiated whenever there is a significant change in the design or manufacturing process. Please consult the reliability appendix of the revised CSA Standard for Carbon Monoxide Alarms (CAN/CGA-6.19-01) for the precise requirements.

Table B-3 places into perspective the consequences of various failure rates. The first column is the mean-time-between-failure (MTBF). The rows in bold faced type indicate the $MTBF_{90}$ allowed by the CSA Standard for supervised and unsupervised failures. The second column is the monthly failure rate as a percent per month. The third column gives the portion of installed alarms that are in failure for each failure rate, making the optimistic assumption that consumers promptly replace alarms at the end of a three year lifetime so that the average age of the installed detectors is only 18 months and none are older than three years. The final column gives the cumulative portion of alarms failing over their lifetime.

At the end of a presumptive three year lifetime the CSA Standard allows 14.6% of alarms to have failed without indicating their failure, and an additional 23.1% of alarms to have failed with a warning indication, for a total of 37.7% failed alarms.

Mean-time-between-failure is estimated by tallying the failures in a sample over time. To develop the mathematical basis for our calculation of MTBF we first express our assumption of a constant failure rate as an equation. For a fixed number of devices under test, N_{DUT} , the rate at which the number of failures,

N_f , increases over time is given by the quotient of the number of unfailed devices at any given time and the MTBF as,

$$d/dt N_f(t) = (N_{DUT} - N_f(t)) / MTBF \quad \text{Eqn. B-5}$$

Table B-3. Field Consequences of Various Mean-Times-Between-Failure			
MTBF (hours)	Failures per Month (%/month)	Portion of Field Units in Failure ^c	Cumulative Lifetime Failures ^d
500,000.	0.15 %	2.6 %	5.1 %
250,000.	0.29 %	5.1 %	10.0 %
166,667. ^a	0.44 %	7.5 %	14.6 %
100,000. ^b	0.73 %	12.1 %	23.1 %
50,000.	1.45 %	22.2 %	40.9 %
20,000.	3.58 %	44.4 %	73.1 %

- a) The CSA reliability limit for *unsupervised* failures.
- b) The CSA reliability limit for supervised failures.
- c) The portion of field units in failure at any time, making the ideal assumptions of a uniform distribution in the age of alarms, their continuous replacement at the end of a 3 year lifetime, and an average age of installed units of only 18 months. The failed portion is given by the average of the integral of Eqn. B-7 over the interval of $\Delta t = \{0, 3 \text{ yrs}\}$. It is approximately one half of the cumulative portion failed over the alarms' entire lifetime, given directly by Eqn. B-7.
- d) Assuming a lifetime of 3 years.

If instead of using a fixed number of devices, failed devices are replaced so that the number of unfailed devices under test, N_O , remains constant during the test, then the rate of failure is constant and given by,

$$d/dt N_f(t) = N_O / MTBF \quad \text{Eqn. B-6}$$

We can solve these differential equations to find the number of accumulated failures after time Δt as,

$$N_f / N_{DUT} = 1 - \exp[-\Delta t / MTBF] \quad \text{Eqn. B-7}$$

and,

$$N_f / N_O = \Delta t / MTBF \quad \text{Eqn. B-8}$$

Because it is impractical to promptly replace failed devices during the test time, the more generally useful expression is Eqn. B-7. This equation is used to estimate the number of expected failures from a known MTBF. For example, using a MTBF of 166,667 hours, the portion of detectors failing over the course of a presumptive three year lifetime ($\Delta t = 26298$ hours) would be 14.6%. Consequently, if we had 100 devices under test we would likely observe around 15 failures.

In reliability testing we want to estimate the MTBF from the number of failures observed in a test. We can solve Eqn. B-7 for the MTBF for our sample as,

$$MTBF_S = -\Delta t / \ln[1 - N_{fS} / N_{DUT}] \quad \text{Eqn. B-9}$$

in which N_{fS} is the actual number of failures we observe in time Δt out of the number of devices in the sample N_{DUT} , and \ln is the natural logarithm.

Equation B-9 is the fundamental equation to use when we know a cumulative test time and number of failures and we want to estimate the MTBF. It is easily applied when there is a single cohort of devices all held for the same amount of time. However, it is difficult to apply if different numbers of devices are held for different lengths of time. In that case Eqn. B-9 may be approximated with a more easily applied form as,

$$MTBF_S = \alpha \Delta t (N_{DUT} / N_{fS}) \quad \text{Eqn. B-10}$$

in which α is a parameter adjusted to make Eqn. B-10 agree with Eqn. B-9 at a particular cumulative failure rate. For the lifetime cumulative failure rates for supervised and unsupervised failures allowed by the CSA Standard the best values for α are 0.878 and 0.925 respectively,

$$\text{Supervised Failures:} \quad MTBF_S = 0.878 * \Delta t * (N_{DUT} / N_{fS}) \quad \text{Eqn. B-11}$$

$$\text{Unsupervised Failures:} \quad MTBF_S = 0.925 * \Delta t * (N_{DUT} / N_{fS}) \quad \text{Eqn. B-12}$$

These approximations work well over a wide range of accumulated failures, but for fewer accumulated failures than the maximum allowed they slightly underestimate the MTBF, while for greater accumulated failures they slightly overestimate it.

We use these approximations because they can be easily extended to account for devices held for different amounts of time. (When we use this approximation we are implicitly assuming that the intrinsic failure rate over the lifetime of the device is approximately constant.) For different numbers of units held for different times a total device time is computed as,

$$T_{DUT} = \sum_i \Delta t_i N_{DUT_i} \quad \text{Eqn. B-13}$$

and a total number of failures as,

$$N_{fS} = \sum_i N_{fS_i} \quad \text{Eqn. B-14}$$

in which the index, i , pertains to different samples of devices held for different amounts of time. Using these definitions, Equations B-11 and B-12 become,

$$\text{Supervised Failures:} \quad MTBF_S = 0.878 \frac{\sum_i \Delta t_i N_{DUT_i}}{\sum_i N_{fS_i}} \quad \text{Eqn. B-15}$$

and

$$\text{Unsupervised Failures:} \quad MTBF_S = 0.925 \frac{\sum_i \Delta t_i N_{DUT_i}}{\sum_i N_{fS_i}} \quad \text{Eqn. B-16}$$

For example, if 100 alarms are held for a first quarter, and in addition to these 50 more alarms are held for a second quarter, the total device time is given by Eqn. B-13 as,

$$T_{DUT} = 100 * 2 + 50 * 1 = 250 \text{ device-quarters, or,}$$

$$T_{DUT} = 100 * 2 * 2191.5 + 50 * 2191.5 = 547,875 \text{ device-hours}$$

If at the end of the first quarter the 100 alarms are tested and two unsupervised failures detected, and at the end of the second quarter 148 alarms are tested (all 150 minus the two that had already failed) and an additional unsupervised failure detected, then the total number of unsupervised failures would be $N_{fS} = 3$.

The observed $MTBF_S$ for the test is computed using Eqn. B-16 as

$$MTBF_S = 0.925 * 547,875 / 3 = 168,928 \text{ hours.}$$

As was the case with time-of-manufacture reliability, the number of failures observed will fluctuate from sample to sample about an ideal average with a standard deviation (called the counting error) roughly equal to the square root of the ideal average. Consequently, if we were to have drawn a different sample of devices and had tested them we would have observed a somewhat different number of failures, N_{fS} . We would like to predict the mean time between failure at a 90% confidence level. This $MTBF_{90}$ is not based solely on the observed number of failures, but on the upper limit on the number of failures at a 90% confidence level. Consequently, in Eqns. 15 and 16 we replace N_{fS_i} with N_{f90_i} , from Table B-1, as,

For Supervised Failures:	$MTBF_{90} = 0.878 \frac{\sum_i \Delta t_i N_{DUT_i}}{\sum_i N_{f90_i}}$	Eqn. B-17
--------------------------	--	-----------

and

For Unsupervised Failures:	$MTBF_{90} = 0.925 \frac{\sum_i \Delta t_i N_{DUT_i}}{\sum_i N_{f90_i}}$	Eqn. B-18
----------------------------	--	-----------

In these equations N_{f90} is the upper limit on N_{fS} at the 90% confidence level, as determined from Table B-1, and the index i pertains to different cohorts of devices held in operating condition for different lengths of time.

Continuing the above example, if of the 150 alarms held for a total device time of 547875 device-hours we had observed three unsupervised failures, the upper limit on the number of failures is found from Table B-1 to be 6.68 failures. The $MTBF_{90}$ is computed as,

$$MTBF_{90} = 0.925 * 547875 / 6.68 = 75866 \text{ hours.}$$

Given a sample of operating devices, the length of time they have been operating, and the number of failures observed, Equations B-17 and B-18 allow us to compute their $MTBF_{90}$. We need however to determine a minimum sample size to be tested. Our goal is to estimate the alarms lifetime failure rate in the minimum elapsed time possible (one quarter) and then to successively refine that estimation as the alarms are held in operating condition for longer periods of time, up to their minimum lifetime of three

years. Consequently, we need to compute the minimum device time required to estimate the failure rate at a 90% confidence level in one quarter.

Assuming no failures are observed in the sample, the minimum device time required is calculated from Equations B-17 and B-18 using the target $MTBF_{90}$ and the minimum upper limit on the number of failures, N_{f90} , of 2.30, as,

$$\text{For Supervised Failures: } T_{DUT} = N_{f90} MTBF_{90} / 0.878 \quad \text{Eqn. B-19}$$

$$\text{or, } T_{DUT} = 2.3 * 100,000. / 0.878 = 261,959. \\ \text{from which } N_{DUT} = T_{DUT} / \Delta t = 119.5 \quad \text{Eqn. B-20}$$

and,

$$\text{For Unsupervised Failures: } T_{DUT} = N_{f90} MTBF_{90} / 0.925 \quad \text{Eqn. B-21}$$

$$T_{DUT} = 2.3 * 166,667. / 0.925 = 414,415. \\ \text{from which } N_{DUT} = T_{DUT} / \Delta t = 189.1 \quad \text{Eqn. B-22}$$

We see that in order to verify that the mean-time-between-failure is at least 166,667 hours at a 90% confidence level we need to operate 190 (rounded up from 189.1) devices continuously for one quarter, at the end of which they are tested and found to have no failures. Because the device-time required for unsupervised failures is greater than that for supervised failures we use the greater number of units. If we were testing only for supervised failures we would require 120 units.

The results of calculations of Eqns. 19-22 for observed failures from 1 to 100 are presented in the “In-Service Reliability Sampling Plan” of Table B-4.

If no failures are observed we conclude that the devices are *in-compliance* and we need not continue to hold the entire sample for a second quarter. We only require that for the next quarter’s test we continue to maintain a cumulative device time of 190 device-quarters while maximizing the age of the sample of retained devices. Consequently, for the second quarter we need only retain,

$$N_{DUT} = T_{DUT} / \Delta t = 190 \text{ device-quarters} / 2 \text{ quarters} = 95 \text{ devices}$$

In fact, as long as no failures are observed we may continue to reduce the number of retained devices on a schedule given by,

$$N_{DUT} = T_{DUT} / \Delta t \quad \text{Eqn. B-23}$$

in which T_{DUT} is given in device-quarters from Table B-4 and Δt is given in quarters. Quarter by quarter we may continue to reduce the number of devices retained. As we do so, we maintain a constant cumulative device-time, but the age of the retained devices steadily increases toward their specified lifetime, providing us with an increasingly more realistic measure of lifetime reliability. As long as we continue to see no failures, by the end of three years we will need only to retain 16 devices.

It is not necessary to reduce the size of the retained sample. In fact, more accurate measures of device reliability will be made by retaining as many devices as possible. In addition, having more retained devices than the minimum required simplifies further testing if more failures are observed.

What if a failure is observed when testing at the end of the first quarter? Just as in the procedure for time-of-manufacture reliability, if compliance is not demonstrated at the minimum device time the device time may be increased.

For example, suppose one unsupervised failure had been observed in testing the 190 devices. We see from Table B-4 that if one failure is observed the required device time increases to 320 device-quarters. From Eqn. B-23 we see that we must retain at least $320/2 = 160$ devices. Now, by the end of the second quarter we will have 160 devices all having been held for the entire two quarters. If when all retained devices are tested at the end of the second quarter no other failures are observed, then we will have demonstrated compliance with the required $MTBF_{90}$. We may then continue the test with a constant device time of 320 device-quarters, reducing the number of devices retained to $320/3 = 107$ for the third quarter and further reducing the number quarter by quarter if we desire. If, however, more failures had been observed the device time would again be increased for the third quarter.

Table B-4. In-Service Reliability Sampling Plan

Number of Failures Allowed ¹	Number Allowed for an Increase in Device Time ²	Minimum Device Time for Required $MTBF_{90}$			
		Supervised Failures		Unsupervised Failures	
		Device-Hours	Device-Quarters	Device-Hours	Device-Quarters
0	2	262,000	120	414,000	190
1	3	443,000	203	701,000	320
2	5	606,000	277	959,000	438
3	6	761,000	348	1,204,000	550
4	8	911,000	416	1,441,000	658
5	9	1,057,000	483	1,672,000	763
6	10	1,200,000	548	1,899,000	867
7	11	1,341,000	612	2,121,000	968
8	13	1,481,000	676	2,342,000	1069
9	14	1,618,000	739	2,560,000	1169
10	15	1,755,000	801	2,777,000	1267
11	16	1,891,000	863	2,991,000	1365
12	17	2,025,000	925	3,204,000	1462
13	18	2,159,000	986	3,416,000	1559
14	20	2,293,000	1047	3,627,000	1655
15	21	2,426,000	1107	3,838,000	1752
20	27	3,081,000	1406	4,874,000	2224
30	38	4,366,000	1993	6,906,000	3152
40	49	5,625,000	2567	8,899,000	4061
50	60	6,874,000	3137	10,874,000	4962
100	114	12,991,000	5928	20,551,000	9378

1 The number of failures allowed at the cumulative device time that demonstrates compliance with a minimum $MTBF$ of 100,000 hours for supervised failures or 166,667 hours for unsupervised failures at a 90% confidence level.

2 The number of failures beyond which it is unlikely that compliance will be demonstrated even if the device time of the test is increased. For an observed number of failures greater than the "Number of Failures Allowed" and less than or equal to the "Number Allowed for an Increase in Device Time", it is likely that compliance will eventually be demonstrated if the device time is increased sufficiently.

To take another example, in a test of 190 devices for one quarter no unsupervised failures occur and one supervised failure occurs. From Table B-4 we find that the minimum number of device-quarters required while allowing zero unsupervised failures is 190, and the minimum number required while allowing one supervised failure is 203. We therefore require that the device time be increased to the greater of these, to 203 device-quarters. We must retain at least $203/2 = 102$ of the devices for the second quarter. Suppose that during the second quarter there are no additional failures. In this case we will have demonstrated compliance with the required $MTBF_{90}$ for both supervised and unsupervised failures.

In some cases it will not be possible to increase the device time simply by retaining devices; additional devices may need to be added to the sample. For example, suppose that only the minimum number of devices, 190, are tested in the first quarter, and that two unsupervised failures occur. From Table B-4 we see that in the next quarter 438 device-quarters are needed. When held for an additional quarter the original 190 devices can account for only 380 of those device-quarters. Unless a larger sample was initially drawn and operated in a simulated residential environment, 58 newly manufactured additional devices must be added. The sample now contains devices of mixed age, and its cumulative device time is calculated using Eqn. B-13.

If at any time the tested sample is augmented with new devices, reliability tests should be continued for longer than the three year lifetime, until the average age of the sample of devices achieves the three year lifetime.

Of course, if we are unable to show compliance in a particular quarter owing to too many observed failures, increasing the device time of the test is no guarantee that compliance can be demonstrated in the next quarter. If the population failure rate is too great, more failures are likely to become apparent in tests at the greater device times. Just as in the case of time-of-manufacture reliability assurance, there is a level of observed failures for each device time that signals the unlikelihood of demonstrating compliance even if the device time is increased. This number of failures is the number that results in a calculation (using Eqns. 15 & 16) of a $MTBF_S$ of less than the required minimum $MTBF_{90}$. These numbers of failures allowed for an increase in test time if compliance is not achieved are also tabulated in Table B-4.

For example, suppose in the first quarter 190 devices are tested and four unsupervised failures are found. Because this exceeds the “Number Allowed for an Increase in Device Time” from Table B-4 it is pointless to continue into the second quarter by increasing the test time. The devices should be deemed *not-in-compliance*. The observed sample mean-time-between-failure, $MTBF_S$, is given by Eqn. B-16 as

$$MTBF_S = 0.925 * 414,000 / 4 = 95,700 \text{ hours}$$

This time is less than the target $MTBF_{90}$ indicating that it is unlikely that an increase in the device time will demonstrate compliance.

Procedure

All data pertaining to the sampling procedure and test results must be recorded on a controlled test sheet and maintained in a file for future reference.

A minimum of 190 packaged, market-ready devices is initially selected for testing. These are serialized, placed in a simulated residential environment, and powered up. At the manufacturer’s discretion devices may be installed in actual residences. The devices must be continuously monitored so that false alarms or supervised failures are recorded. The history of each device, including time of power-up and times of supervised failures, must be logged.

At the conclusion of each quarterly period the devices are tested to the Sensitivity Test, and the number of supervised and unsupervised failures, and the total device time tallied. The sample mean-time-between-failure, $MTBF_S$, and the lower limit on the mean-time-between-failure at a 90% confidence level, $MTBF_{90}$, must be computed using Eqns. 15-18. The numbers of observed failures are compared to the values in Table B-4 to determine whether the devices are in compliance, and whether the total device time should be increased for the next quarter.

For successive quarters, devices should be retained or augmented on a schedule that results in a total device time chosen from Table B-4 using the observed failures of the prior quarter.

If the number of observed failures is less than or equal to the “Number of Failures Allowed” of Table B-4, the devices are deemed to be *in-compliance* for the quarter, otherwise they are *not-in-compliance*.

If the number of failures observed is greater than the “Number of Failures Allowed” of Table B-4, but not greater than the “Number Allowed for an Increase in Device Time”, then either the total sample size may be increased in an effort to demonstrate compliance at a greater device time in the next quarter, or the manufacturer may forfeit compliance. If for the greater device time the number of observed failures remains in this interval, the device time may be increased again for the next quarter.

If for any quarter, the number of observed failures is greater than the “Number Allowed for an Increase in Device Time” of Table B-4 (or, equivalently, the observed sample $MTBF_S$ is less than the target $MTBF_{90}$ for either supervised or unsupervised failures), it is recommended that the device time not be increased further, and that the devices be deemed to be *not-in-compliance*.

If the devices are *not-in-compliance* corrective action must be taken to determine the cause of the failure rate. The certifying agency must be notified and approve an action plan to reduce the failure rate. The action plan should include the diagnostic tests, disposition, redesign or rework done to the failed model. Depending on the action plan the reliability testing may need to be restarted.

As long as the devices remain *in-compliance* and the model is produced, quarterly testing is continued.

Statistical Derivation of Table B-1

Assuring Time-of-Manufacture or In-Service Reliability makes use of a statistical table that provides an upper limit on the number of estimated failures at a 90% confidence level given an observed number of failures (Table B-1). The section provides a derivation for the values of that Table. Understanding this section is not needed to apply the foregoing methods of quality assurance – it is only provided for those interested in understanding the philosophical and mathematical basis for the statistics used.

Because it is generally not feasible to test the entire population of manufactured units, statistical methods for sampling and testing fewer units are used. Each of these tests can be considered a binomial experiment in which the outcome is classified as either a “pass” or a “failure”. Tests of each sampled unit are assumed to be independent of the outcome of tests of other units, and the probability of failure of each unit is assumed to be the same. Using a tally of observed failures from the sample, an upper bound on the entire population’s failure rate is found at a chosen confidence level.

Statistically assessing the failure rate places a lower bound on the number of units required for testing. The number of units needed for testing is estimated as follows: In a large population of devices there will be some fraction, F , that will fail a test of any particular characteristic. We will call F the “population failure rate”. The goal of any particular test is to establish an upper bound on F at a given confidence

level (e.g. 60%, 80%, or 90%) by testing a smaller subpopulation, or sample, of only N_{DUT} units. However, the sample size, N_{DUT} , must be sufficiently large to confidently place an upper bound on the failure rate. The number of units *expected* to fail a test, N_f , is given by the size of the subpopulation or sample tested times the population failure rate as,

$$N_f = N_{DUT} F \quad \text{Eqn. B-24}$$

Although N_f is the *expected* number of failures, the *actual* number of failures, N_{fS} , observed in any particular sample of N_{DUT} units may be more or less. There is approximately a 50% chance that the observed failures would be more numerous than the expected number and a 50% chance that they would be less numerous. In fact, any number of failures from 0 to N_{DUT} might be observed but with different probabilities.

How are we to know, given the number of failures observed, what is the actual population failure rate? The answer is that we can not know. Even if we hypothesize a particular population failure rate, it is not really meaningful to ask the question, “What is the probability that this population failure rate is correct?” The reason is that there is no statistical universe of values from which the failure rate is drawn. There is only a single population failure rate, the correct one, we don’t know it, and we don’t really measure it by testing only a sample. Instead, there is a single population, characterized by a single, unknown failure rate, and a statistical universe of samples, each with their own failure rate, that are drawn from it.

Even so, we can turn the question around and ask, “Assuming the correctness of a particular population failure rate, what is the probability that the observed number of failures could have occurred?” If the probability of observing the actual number of failures, N_{fS} , is too small, we can conclude that the population failure rate under consideration is “unlikely” to be the correct one. Looking at it another way, our intuition tells us that the actual number of failures should not be too improbable for the correct assumption of population failure rate. Moreover, we can establish a likely upper bound on the correct population failure rate by asking the question “Assuming that if the population failure rate were to equal a certain upper bound, what is the probability of observing more than the actual number of failures?” If that probability is great, then we can have confidence that the population failure rate is truly less than the hypothetical upper bound.

In other words, we identify the probability of observing more than the actual number of failures in a sample given an upper bound on the population failure rate as the *likelihood* of the true population failure rate’s being less than the upper bound given the actual number of failures. This is as close to a notion of a “probability that a population failure rate is correct” as we can come. Although this definition of likelihood is entirely based on intuition, and has no formal mathematical basis in and of itself (and consequently, whatever the philosophical foundation of statistics, it is not a branch of mathematics!), it turns out to be the basis of many extremely useful methods of statistics.

So, assuming a particular population failure rate, F , we identify the cumulative probability of observing *more* than an actual number of failures as the likelihood of, or confidence level for being correct when using that assumed value as an upper bound on the true population failure rate. This identification is the core assumption for our statistical assessment of failure rate. Consequently, if the expected number of failures given an assumed failure rate of F were actually to be observed, that is, if N_{fS} were to equal N_f , then the test would have established that the failure rate is truly less than or equal to F with only about 50% confidence. If the observed number of failures were found to be much less than the number expected from a particular assumed failure rate then we would have a high confidence that the actual failure rate is less than or equal to the assumed failure rate. Our assumed failure rate could then be used as an upper bound on the actual failure rate with great confidence.

To establish an upper bound on the failure rate, we will consider the probability distribution for observing differing numbers of failures given a particular population failure rate. This distribution is given by the Binomial probability distribution. The probability of observing a particular number, N_{fS} , or fewer failures, $P(\leq N_{fS})$, is given by one minus the cumulative binomial probability for observing more than N_{fS} failures with a probability of F each in a subpopulation of N_{DUT} units, and is related to the Incomplete Beta function as,

$$P(\leq N_{fS}) = 1 - \sum_{N_{fS}+1}^{N_{DUT}} [N_{DUT}! / j! (N_{DUT} - j)!] F^j (1-F)^{N_{DUT}-j} = 1 - I_F(N_{fS} + 1, N_{DUT} - N_{fS}) \quad \text{Eqn. B-25}$$

where $I_F(N_{fS} + 1, N_{DUT} - N_{fS})$ is the Incomplete Beta function.

Unfortunately, tabulations of the Incomplete Beta function are not as commonly available as those of the Gamma or Chi-Squared functions so it is preferable to approximate the above in terms of those distributions. This may be done by assuming that the number of observed failures is much less than the total number of units tested, $N_{fS} \ll N_{DUT}$, generally a good assumption, so that the Poisson distribution may be used as an approximation for the Binomial distribution. The probability that the number of observed failures, N_{fS} , is equal or fewer than the expected number, N_f , is given by the cumulative Poisson distribution function, $P(N_{fS} \leq N_f)$, which in turn is given by the complement of the incomplete gamma function, often called GAMMAQ, or the chi-squared functions as,

$$P(N_{fS} \leq N_f) = \text{GAMMAQ}(N_{fS}+1, N_f) = P_{\mathbf{C}^2}(2 N_f, 2 N_{fS}+2) \quad \text{Eqn. B-26}$$

This probability equals one minus the confidence we have that the failure rate is actually less than or equal to the value responsible for the expected number of failures, N_f . In the above equation $P_{\mathbf{C}^2}(2 N_f, 2 N_{fS}+2)$ represents the probability that an observed chi-squared, \mathbf{C}^2 , will exceed a value of $2N_f$ by chance for $2 N_{fS}+2$ degrees of freedom. Note that the standard interpretation of this distribution in terms of chi-squared and degrees of freedom is not particularly meaningful in this application. Rather, the chi-squared distribution is used for its availability; it is the most widely tabulated form of the Gamma function. Even so, the chi-squared distribution is usually tabulated not as probability as a function of squared error, $P_{\mathbf{C}^2}$, but rather as *reduced*-chi-squared as a function of probability, \mathbf{C}_n^2 , (where reduced-chi-squared, \mathbf{C}_n^2 , is chi-squared divided by the number of degrees of freedom, n). Recast in these terms, Equation B-26 becomes

$$N_{DUT} F = N_f \leq (N_{fS}+1) \mathbf{C}_n^2(1-C, 2 N_{fS}+2) \quad \text{Eqn. B-27}$$

Thus the failure rate is bounded for a given confidence level as

$$F \leq (N_{fS}+1) \mathbf{C}_n^2(1-C, 2 N_{fS}+2) / N_{DUT} \quad \text{Eqn. B-28}$$

in which,

- N_{DUT} is the number of units undergoing test,
- F is the hypothesized population failure rate,
- N_{fS} is the actual number of failures observed in a test of N_{DUT} units,
- C is our degree of confidence in the bound on the failure rate, and,
- $\mathbf{C}_n^2(1-C, 2 N_{fS}+2)$ is the reduced chi-squared distribution at a significance of $1-C$ and a number of degrees of freedom $n = 2 N_{fS} + 2$.

In a manner similar to that of the derivation above, we can also establish a lower bound on the population failure rate given the observed number of failures. This done by assuming a particular population failure rate, F , and identifying the probability of observing *fewer* than the actual number of failures as the likelihood of, or confidence level for being correct when using the assumed value as a lower bound on the true population failure rate. In this case a lower bound is found for the population failure rate as,

$$F \geq N_{fS} C_n^2 (C, 2 N_{fS}) / N_{DUT} \quad \text{Eqn. B-29}$$

Knowing this lower bound is generally less useful than knowing the upper bound, but it is useful for estimating minimum failure rates from field data.

Using these equations, Table B-5 lists upper and lower bounds on the expected number of failures, $N_f = N_{DUT} F$, at various confidence levels and numbers of failures observed after testing any sample size. As expected, at the 50% confidence level these bounds are approximately the actual numbers of failures observed. Also, for any number of failures observed the upper bound on the expected number increases with increasing confidence level.

Table B-5 Upper and Lower Bounds on the Expected Number of Failures at Various Confidence Levels Given an Observed Number of Failures						
Number of Failures Observed	Confidence Level (CL)					
	99%	95%	90%	80%	60%	50%
	UpperBound / Lower Bound					
0	4.6 / 0.0	3.0 / 0.0	2.3 / 0.0	1.6 / 0.0	0.92 / 0.0	0.69 / 0.0
1	6.6 / 0.01	4.7 / 0.05	3.9 / 0.11	3.0 / 0.22	2.0 / 0.51	1.7 / 0.69
2	8.4 / 0.15	6.3 / 0.36	5.3 / 0.53	4.3 / 0.82	3.1 / 1.4	2.7 / 1.7
3	10 / 0.44	7.8 / 0.82	6.7 / 1.1	5.5 / 1.5	4.2 / 2.3	3.7 / 2.7
4	12 / 0.82	9.2 / 1.4	8.0 / 1.7	6.7 / 2.3	5.2 / 3.2	4.7 / 3.7
5	13 / 1.3	11 / 2.0	9.3 / 2.4	7.9 / 3.1	6.3 / 4.2	5.7 / 4.7
6	15 / 1.8	12 / 2.6	11 / 3.2	9.1 / 3.9	7.3 / 5.1	6.7 / 5.7
8	17 / 2.9	14 / 4.0	13 / 4.7	11 / 5.6	9.4 / 7.0	8.7 / 7.7
10	20 / 4.1	17 / 5.4	15 / 6.2	14 / 7.3	12 / 8.9	10.7 / 9.7
15	27 / 7.5	23 / 9.2	21 / 10	19 / 12	17 / 14	15.7 / 14.7
20	33 / 11	29 / 13	27 / 15	25 / 16	22 / 19	20.7 / 19.7
50	69 / 35	63 / 39	60 / 41	57 / 44	52 / 48	50.7 / 49.7
100	126 / 78	118 / 84	114 / 87	109 / 92	103 / 97	101 / 99.7
$N_{fS} > 100$	$N_{fS} +/- 2.33\sqrt{N_{fS}}$	$N_{fS} +/- 1.65\sqrt{N_{fS}}$	$N_{fS} +/- 1.28\sqrt{N_{fS}}$	$N_{fS} +/- 0.84\sqrt{N_{fS}}$	$N_{fS} +/- 0.25\sqrt{N_{fS}}$	N_{fS} / N_{fS}

The final row of the table, for $N_{fS} > 100$, is found by approximating the Poisson distribution with a Gaussian distribution. In this case we obtain a upper and lower bounds of

$$F_{upper} = N_{fS} + \beta^2/2 + \beta \sqrt{(\beta^2/4 + N_{fS})} \quad F_{lower} = N_{fS} + \beta^2/2 - \beta \sqrt{(\beta^2/4 + N_{fS})} \quad \text{Eqn. B-30}$$

where β is the number of standard deviations from the mean of a Gaussian distribution, the integral to which is equal to the confidence level (i.e., $\beta=0.0, 0.25, 0.84, 1.28, 1.65$, and 2.33 for $CL = 50, 60, 80, 90, 95$ and 99%). If N_{fS} is great enough that the standard deviations of the distributions at the lower and upper bounds are comparable, then these bounds are simplified to $N_{fS} +/- \beta \sqrt{N_{fS}}$ as shown in the table.

The values of Table B-1 are taken from the upper bounds at the 90% confidence level of Table B-5.

