CLIN. CHEM. 33/2, 261–268 (1987)

Accuracy and Precision of Breath-Alcohol Measurements for a Random Subject in the Postabsorptive State

G. Simpson

The accuracy of estimates of blood-alcohol concentration based on measurements of breath-alcohol concentration in a randomly selected subject by a random quantitative evidential breath-alcohol analyzer is evaluated with respect to the breath analyzer itself, its calibration, and the biological variables of the subject being tested. There are no suitable experimental data for rigorous determination of the overall accuracy, so I estimate it from the CV of the available data. I find that the uncertainty in these breath-analyzer readings for a random subject in the postabsorptive state is at least ±15%, ±19%, or ±27%, depending on whether ±2 CV, the experimental range, or ±3 CV, respectively, is used to express the overall uncertainty. Over 90% of this uncertainty is due to biological variables of the subject, and at least 23% of subjects will have their actual blood-alcohol concentration overestimated. Manufacturers' specifications for the accuracy and precision of these instruments are inconsistent with the experimental values reported in the literature and I recommend that an appropriate amount of uncertainty be reflected in the results from these breath analyzers, especially when they are used for law-enforcement purposes.

Additional Keyphrases: *forensic medicine · analytical error · intra- and inter-individual variation · instrumental error · blood alcohol concentration*

Breath testing is being increasingly applied in the workplace and in research, but it is most commonly used in connection with law enforcement, and the accuracy of the results is quite important to the subject being tested. How accurately is the blood-alcohol concentration (BAC) of a random subject estimated from measurement of breath-alcohol concentration (BrAC) by a typical quantitative evidential breath-alcohol analyzer? Studies done in the past to determine this accuracy have involved homogeneous samples of subjects under laboratory conditions—which are not representative of the actual population or conditions (*1–6*)—or field trials—which are more representative but are otherwise flawed (*5–9*). While performance standards and guidelines for these analyzers have been established by state legislatures and federal agencies (*10, 11*), they deal with the accuracy and precision of measured breath-alcohol concentrations in terms of standardized alcohol-vapor samples, rather than with how accurately the subject's BAC is estimated. Analytical measurements usually have well-defined accuracy and precision, which are useful for indicating the reliability of the measurements. However, BAC estimates based on measurement of BrAC are less well defined because they are made by using a conversion factor, which is known to have a significant amount of uncertainty (*1–3, 5, 6*), and this uncertainty often is not taken into account in forensic applications.

There are suitable data in the literature to provide only an estimate of the accuracy of BAC results from these analyzers, and that is the purpose here. The approach involves selection of representative data from the literature, which are subjected to standard methods of evaluation of analytical results.

Methods

To determine the accuracy of BAC results estimated from BrAC measurements for a random subject by a random instrument, one must select samples of subjects, instruments, and instrument operators that are representative of operational or field use. If the results from a sample are to be applied to a population, the sample must be selected so that it is representative of that population. This is accomplished by the use of random sampling (*12*), but even more importantly, "random sampling leads to probability models for distributions. Since the conclusions to be drawn about populations by means of samples are to be based upon probabilities, samples must be selected in such a manner that the rules of probability can be applied to them (*12*)."

Because the criteria for random sampling have rarely been met in previous studies, the results of these studies are not applicable to a random subject. The minimum error involving a random subject must therefore be estimated from the precision of the data in the literature. Precision is used to estimate indeterminate error or uncertainty, and such an estimate may or may not reflect the presence of determinate or systematic error. While estimates based solely on precision must always be considered guardedly (*13*), this is one of the few approaches that the existing data allow.

The total uncertainty in BAC estimated from BrAC measurements can be broken down into three independent parts: the instrument itself, the calibration of the instrument, and the biological variables of the subject. The standard deviation (SD) or coefficient of variation (CV) can then be determined for each of these parts, and the total SD or CV is calculated from the variance, $(SD)^2$ or $(CV)^2$. Because the variance is additive, $SD_T = [(SD_I)^2 + (SD_C)^2 + (SD_S)^2]^{1/2}$, where SD_T is the total SD, SD_I is the SD of the instrument itself, SD_C is the SD of calibration, and SD_S is the SD of the biological factors involving the subject. Similarly, $CV_T = [(CV_I)^2 + (CV_C)^2 + (CV_S)^2]^{1/2} \times 100$. If the functional form of T is a sum or a difference, then SD^2 is additive; if it is a product or a quotient, then CV^2 is additive (*13*).

Results

Instrumental Uncertainty

The closest approximation to a study of a representative sample of quantitative evidential breath-alcohol analyzers is the work by Caplan et al., who measured the in vitro accuracy and precision of 90 new Model 1000 Breathalyzer instruments (*14*). The average SD was reported to be 0.0037 g/210 L, and this includes contributions from both the instrument and its calibration with a breath-alcohol simulator (SD_{IC}). For purposes of illustration, the contribution

587 F North Ventu Park Rd., Suite 531, Newbury Park, CA 91320.

Received April 28, 1986; accepted November 21, 1986.

from each of these is selected to be approximately one-half of 0.0037 g/210 L. SD_{IC} can be expressed as CV_{IC} = 3.7%, because the mean alcohol concentration was 0.100 g/210 L. The value selected for the contribution from the instrument itself is CV_I = 2.5%.

Calibration Uncertainty

Breath analyzers are routinely calibrated with a commercially available device known as a breath-alcohol simulator (*1, 5, 6, 14, 15*). There are a number of these devices on the market, five of which appear on the list of conforming products published by the U.S. Department of Transportation (*10*). A sample of simulators representative of those used in the field should be evaluated, but because such experiments have not been done, the accuracy of the alcohol-vapor samples produced by simulators will be estimated from the precision reported in the literature. To calibrate with a simulator, one places a known concentration of alcohol in water in the reservoir and warms it to 34.0 ± 0.2 °C (*6, 14, 15*). The concentration of alcohol in the vapor above the solution is then calculated from k_{aw}, the partition ratio for alcohol between air and water as measured by Harger et al. (*16*), and the vapor is admitted to the sample compartment of the instrument being calibrated. The sources of error in such a calibration procedure are: determination of the concentration of the aqueous alcohol solution, the experimental accuracy and precision in the value of k_{aw}, and the fluctuation in the concentration of alcohol vapor produced by the thermostatically controlled heating element (*15*). The concentration of the aqueous alcohol solution used in the simulator can be determined by any of several methods, including gas chromatography (GC) and titration with potassium dichromate. The CV to be used here for this contribution is 1.5%. The CV for k_{aw} reported by Harger et al. (*16*) is 2.1% at 35 °C. These data were selected because of their wide use in the calculation of alcohol-vapor concentrations produced by simulators, and because they are used in the Dept. of Transportation publication involving simulator performance (*10*). The variation in alcohol vapor concentration produced by the thermostated heater, assuming that the manufacturer's specification of 34.0 ± 0.2 °C is met, is determined by the change in vapor pressure of alcohol with temperature. The value reported by Dubowski is 6.8%/°C (*15*), and multiplying (0.2)(0.068) × 100 = 1.4%. This relative deviation can be converted to a CV from the relationship, CV = K_M(range) (*13*). If K_M, as defined in ref. *13*, for 10 determinations is used, then CV = (0.33)(0.028) × 100 = 0.9%. The total CV for the concentration of alcohol vapor produced by a simulator can be calculated from

$$CV_C = [(0.015)^2 + (0.021)^2 + (0.009)^2]^{1/2} \times 100 = 2.7\%$$

Analytical Uncertainty

Inclusion of the contributions to variability from both the instrument and calibration yields

$$CV_{IC} = [(0.025)^2 + (0.027)^2]^{1/2} \times 100 = 3.7\%,$$

the value reported for the Model 1000 Breathalyzer (*14*). This is defined as "analytical uncertainty," and it is primarily derived from instrumental error (*13*). The values constituting this result were selected to illustrate the possible sources of error, and this is but one of a number of ways by which reasonable numbers can be assigned, but the resulting value for CV_{IC}, 3.7%, will be the same. It is a useful number, not only because it should be indicative of the reliability of the Model 1000 Breathalyzer results under field conditions but also because it apparently represents a maximum CV among quantitative evidential breath analyzers and it will therefore serve to indicate the maximum contribution of instrumental uncertainty to the overall uncertainty.

Uncertainty from Biological Factors

Many variables are related to biological aspects of the subject (*1, 3, 5, 16*). Collectively, these can be expressed by what is known as the "blood–breath ratio," which is commonly programmed into breath analyzers as a conversion factor of 2100:1. The values used here are those of Dubowski and O'Neill, which are reported as a mean, ±SD, CV, and range of 2280, ±241.5, 10.6%, and 1706–3063 for a sample of healthy adult males in the fully postabsorptive state under laboratory conditions (*3*). From these data, CV_S = 10.6% and this is the CV associated with the biological factors of the subject being tested. It is related primarily to method error, which arises from the use of 2100:1 for all subjects, regardless of what their actual blood–breath ratio happens to be.

Total Uncertainty

The total CV is $CV_T = [(0.025)^2 + (0.027)^2 + (0.106)^2]^{1/2} \times 100$ = 11.2%. This would be the expected variation in the readings from a random Model 1000 Breathalyzer for 68 out of 100 healthy men tested under laboratory conditions in the fully postabsorptive state, assuming a normal distribution (*3*), if the breath analyzer were calibrated at 2280:1. If 95 out of 100 of these subjects are to be included, then the range of variation is ±2 CV_T = ±22%. However, breath analyzers in the U.S. are calibrated at 2100:1, which means that ±2 CV_T corresponds to a variation of about ±16%. And because CV_{IC} = 3.7% is apparently a maximum CV among breath analyzers, a smaller value of CV_{IC} = 2% will be used here to give a more conservative estimate of the total uncertainty. This results in $CV_T = [(0.02)^2 + (0.106)^2]^{1/2} \times 100$ = 10.8%, so that the uncertainty in a reading from a randomly selected breath analyzer for 95 out of 100 random subjects will be at least ±15%. If contributions from CV_{IC} are neglected, then CV_T = 10.6%, instead of 10.8%, so that the uncertainty from biological variables accounts for more than 90% of the total, even when the maximum contribution from CV_{IC} is used.

Discussion

Instrument Uncertainty

For determination of the accuracy and precision of a randomly selected breath analyzer, alcohol-vapor standards of known concentration would be measured by a sample of breath analyzers that is representative of those used in the field. Ideally, the alcohol-vapor standards themselves would be traceable to the National Bureau of Standards. Only by using a well-characterized sample can the accuracy and precision of the instrument itself be separated from the accuracy and precision of the calibration method. However, previous studies on the in vitro accuracy and precision of these instruments have made use of breath-alcohol simulators for the production of standard alcohol-vapor samples (*1, 6, 14, 15, 17*) so that the concentration of these vapor samples is somewhat uncertain and CV_I is not readily separable from CV_C.

The Model 1000 Breathalyzer is no longer marketed in the U.S., but it remains on the list of conforming products

published by the Dept. of Transportation (*10*), and it is difficult to say how many of the thousands of these analyzers that have been purchased over the years are still being used in law enforcement. The 90 instruments tested were new and certified as ready for use by the manufacturer (*14*), and they were calibrated and tested with alcohol-vapor samples generated by simulators. Of the 90 instruments tested, it was found that 15 could not meet the Dept. of Transportation guidelines for accuracy and precision. After repairs were made, the instruments were tested again and as a group of 86 they did meet the Dept. of Transportation guidelines. However, the authors concluded: "Experience in the State of Maryland indicates that every instrument to be introduced into routine use should be tested. Although demonstration and other field units may perform satisfactorily, there is no assurance that any manufacturer can consistently provide a production instrument whose results can be expected to warrant the trust of the judiciary and the public." This illustrates the importance of selecting a representative sample of breath analyzers. Earlier tests of individual Model 1000 Breathalyzers (see, e.g., *6*), have indicated reliable performance, but the study of a more nearly representative sample yields quite different results. The performance of other breath analyzers, based on tests of individual instruments (*17*), appears to be considerably better than that of the Model 1000 Breathalyzer, but tests conducted on individual instruments under laboratory conditions by trained scientific personnel are not a reliable indicator of actual field performance. It simply is not known what the actual performance of these instruments would be if 90 of each type were tested under field conditions.

Calibration Uncertainty

The performance of breath-alcohol simulators has been studied by Dubowski (*15*), who concluded: "Presently available commercial simulators used singly are not precision calibrating devices and should not be so employed. The temperature regulation of some commercial simulators is occasionally and unpredictably inadequate. The effect of temperature fluctuations on effluent alcohol concentrations is marked and often unappreciated. Use of simulators in tandem offers a simple means of improving their performance. Redesign of simulators for certain applications is indicated, and some commercial simulators require modification for any use." The results of Emerson et al. (*6*) also do not inspire confidence. They report: "In summary, six of the twelve simulators needed repair during the seven month trial which, considering the simple nature of the device, we feel is an unacceptably high failure rate. The simulator is the fundamental reference for the breath analysis, and any simulator which is to be used in future should be much more reliable." The most frequent malfunction reported in these Smith and Wesson Mark IIA simulators was overheating (*6*).

Despite such reports, remarkable performance from simulators has been and still is being claimed. Jones (*18*) referred to the vapor samples produced by a Smith and Wesson simulator as "primary standards," and he reported an in vitro CV of 0.67% for measurement of concentrations of simulator samples by a GC Intoximeter. However, k_{aw} from the results of Harger et al. (*16*) was used to calculate the concentration of these vapor samples (*18*), and the reported precision of 2.1% at 35 °C was apparently ignored. While an "independent" headspace sample was generated at 25 °C, the concentration of this and the simulator sample were calculated by using k_{aw} from Harger et al., and all vapor samples were measured with the GC Intoximeter being calibrated; an independent method was not used to measure and confirm the calculated vapor concentrations. In a 1984 article by Martin et al. (*19*), simulator samples were also used as standards without independent confirmation of the actual vapor concentrations, yielding a SD of 0.017 g/L; this, too, is smaller than the CV = 2.1% for k_{aw} reported by Harger et al. (*16*). In recent work, Caplan et al. (*14*) and Goldberger and Caplan (*17*) made use of a GC to determine the concentrations of simulator solutions by means of headspace samples. But the k_{aw} from Harger et al. was used to calculate the alcohol vapor concentrations produced by the simulator (Dr. Y. Caplan, private communication), so that it is not appropriate that most of the reported values for precision (*14, 17*) are less than the 2.1% reported by Harger et al. If the data of Harger et al. are used to calculate vapor concentrations, then the precision of k_{aw} must also be included, and all reported values for the CV of these breath analyzers should have been at least 2.1%.

Other values for k_{aw} have been reported. The data of Jones (*20*) indicate a relative deviation in the experimental value of k_{aw} to be +2.4%, −1.3% at 25 °C, and additional uncertainty will result upon extrapolation to 34 °C. The relative deviation reported by Harger et al. (*16*) is 3.6% at 35 °C. Dubowski (*15*) reported a value for k_{aw} derived from the data of Harger et al. and six other studies, four of which were published between 1911 and 1925, but the precision in this derived value is not reported.

Recently, Dubowski has stated (*5*) that ". . . several of the shortcomings of the 1979 devices have been remedied in the current generation of simulators." However, no references are cited to support this, nor are any of the improvements in design or performance described. In fact, it is stated in the sentence just preceding (*5*): "The simulator device for breath-alcohol analyzer calibration and control analysis has not changed substantially in design or commercial execution . . ." A 1983 paper (*21*) describes some of the effects of poor simulator design, while a 1985 article by Jones (*22*) reports constant-temperature regulation of 0.05 °C for a Smith and Wesson Mark IIA simulator—a substantial improvement as compared with earlier Mark IIA simulators. This improvement should nearly eliminate the concentration variations due to the thermostat, so that $CV_C = [(0.015)^2 + (0.021)^2]^{1/2} \times 100 = 2.6\%$, instead of the 2.7% found earlier. And if alcohol standards from the National Bureau of Standards are used to prepare simulator solutions, as described by Dubowski (*5*), the contribution from concentration measurement might also be eliminated, leaving $CV_C = 2.1\%$, assuming the data of Harger et al. for k_{aw} are used. Use of an alcohol-vapor sample from the National Bureau of Standards could eliminate the need to use a simulator, and virtually all calibration uncertainty. But until all of these improvements are implemented, the actual performance of simulators under field conditions will remain poorly characterized and appropriate studies on a sample of simulators representative of those in the field need to be done. It appears quite possible that the results of such studies will show that CV_C is greater than CV_I.

Analytical Uncertainty

Even though there are numerous possible sources of error, the value for CV_{IC}, 3.7%, compares favorably with the CV associated with direct blood analysis (*8, 23*), indicating that BrAC or standard alcohol-vapor samples can be analyzed

about as reliably as direct blood samples. Reliable measurement of BrAC does not necessarily indicate a reliable estimate of BAC unless the subject's blood–breath ratio is known and unless a proper "deep-lung" sample of breath is obtained. The breath analyzer must be designed to collect a proper deep-lung sample, and some analyzers have sampling systems that do a better job of this than others. However, even a well-designed sampling system can be circumvented voluntarily by a subject who uses certain breathing techniques just before the breath test (*24*). This "sampling error" should be minimal for the data I have used here, because trained, cooperative subjects were used, but measurements done in the field can be expected to contain larger contributions from this source of error.

Uncertainty from Biological Factors

To estimate a subject's BAC from measurement of his BrAC, a conversion factor must be used. While there are some differences in the units used to report BAC and BrAC, effectively a subject's measured BrAC is multiplied by 2100 to yield the BAC estimate (*3*). This is the value used for the blood–breath partition ratio, and it has been adopted on the basis of many in vitro and in vivo studies. The model used to describe this ratio is based on Henry's Law, and it has been thoroughly discussed elsewhere (*1, 5, 16, 25, 26*). All of the biological variables of a subject are contained in his or her partition ratio at a given time. To determine how accurately the BAC is estimated from a BrAC measurement, the subject's actual BAC must be known at that time. If the actual value of the subject's BAC is not known, then the accuracy of the BAC estimated from a BrAC measurement cannot be determined; it can only be estimated from the CV for population data, and this requires the use of statistical methods.

To determine the probability that a randomly selected subject has a certain value for the partition ratio, a random sample from the population of people who drink and drive must be selected. Simultaneous measurements of BAC and BrAC must then be made under conditions that are representative of those found in the field. Moreover, because drinking and eating habits vary greatly from one country to another, and even within the same country, it is necessary to select a sample representative of the particular population of interest. While there have been many studies in which BrAC and BAC have been simultaneously measured (*3, 5, 26*), the samples used were not representative of the population that drinks and drives in the U.S., nor were the conditions representative of those found in the field (*1–8, 27*). Dubowski has recently pointed out that (*5*) "hundreds, if not thousands, of laboratory and field studies have been carried out on the performance of quantitative evidential breath-alcohol analyzers; but very few meet appropriate criteria for adequacy of study design and execution, or employed adequate and proper statistical data treatment. Hence few such studies have been published . . ." He also recently compiled a table showing the range of values found for the blood–breath ratio (*5*). Even though there are contradictory results presented in this table,[1] it is clear that the variability in the blood–breath ratio is greater under field conditions than under laboratory conditions. And while the results from field studies are more representative of a generalized population, such studies have also been flawed in a variety of ways. These include the use of time-corrected BACs (*6–8*), which are known to be of questionable validity (*3, 28*); no use of random sampling methods (*6, 7, 9*); significant amounts of elapsed time between either the arrest and measurement of BAC and BrAC, or between measurement of evidential BAC or BrAC samples and measurement of test BAC or BrAC samples (*6–9*); and no adequate determination of whether the subjects were in fact in the postabsorptive state (*6–9*). There are of course many practical difficulties involved in the design and execution of a proper study, some of which may be insurmountable, and this means that there are certain limitations that must be placed on the applicability of these results to a random subject.

Because several important variables are not controlled during field studies, the most reliable data for the blood–breath ratio are from the results of laboratory studies (*1–3, 27*). Jones (*1*) reports a mean, ±SD, CV, and range of 2180, ±189, 8.7%, and 1837–2863 for 21 healthy men in the postabsorptive state. Calculation of the SD from the data of Alobaidi et al. (*2*) results in 2231, ±279, 12.5%, and 1414–3133 for five adult subjects in the postabsorptive state (a sixth subject showed no correlation between BAC and BrAC and was excluded by these authors as an unexplained anomaly). The results of Dubowski and O'Neill were reported (*3*) to be 2280, ±241.5, 10.6%, and 1706–3063 for "397 paired specimens and 142 corresponding blank values" from healthy men in the fully postabsorptive state. Because of the number of subjects studied, the data of Dubowski and O'Neill (*3, 27*) should be the most reliable, but there are difficulties with the application of these results to a random subject or to any generalized population.

A sample of healthy adult males is not representative of the population that drinks and drives, nor are the technical training and skills of the laboratory personnel who did these experiments representative of the training and skills of the personnel who normally operate and calibrate breath analyzers in the field. Consequently, the statement (*29*) made on the basis of these results that "the 2100:1 conversion tends to underestimate the actual BAC in about 86% of the population by a mean of about 8%, because the functional alcohol partition ratio between blood and breath in healthy adult males was found to be 2.28×10^3," is at best misleading. It is misleading not only because the sample and the conditions are not representative, but because the percentage of the population claimed to be underestimated is not correct. The percentage can be calculated from the area under this normal error curve and the report that 86% would have underestimated BACs implies that 14% would have their BACs overestimated. This is consistent with the approximation that 2100 is about one SD less than the mean of 2280, which means that the BAC of 86% of subjects will be underestimated if the breath analyzer is calibrated at 2038. Because breath analyzers in the U.S. are calibrated at 2100, the correct percentage of healthy men having underestimated BACs would be about 77%, and about 23% would have their BACs overestimated, not 14%. The data of Schmidt et al. (*30*) indicate that a similar percentage of their subjects had their BACs overestimated 120 min after drinking stopped. A recent field study by Pribor et al. (*9*) indicates that about 19% of the subjects had their BAC overestimated.

[1] For example, the entries for Isaacs et al. and Emerson et al. involve the same data from a 1977–78 field study, yet the experimental ranges of 1800–3350 and 1155–3045 reported by Dubowski are very different. The reported range of 1155–3045 is the result of an incorrect calculation; the correct range calculated from the results of Emerson et al. is about 1448–3818. The entry for Gatt involves data from just one of three different types of breath analyzers used in a 1981–82 study done by Isaacs et al. (*7*) and consequently does not represent the complete findings of this study.

It is not clear why the results from this field study show a smaller percentage of subjects having overestimated BACs than that found in the laboratory studies of Dubowski and O'Neill (*27, 29*) and Schmidt et al. (*30*). The experimental design of the study probably accounts for a major part of this unexpected result.

An important variable not controlled in a random subject under field conditions is the phase of alcohol metabolism or distribution that he or she happens to be in at the time of the breath test (*1, 3*). It has been established that a subject's blood–breath ratio tends to be less than 2100:1 in the absorptive state and greater than 2100:1 in the postabsorptive state (*1–3, 25, 26*). This means that BAC estimates from BrAC measurements will tend to be falsely high in the absorptive state and falsely low in the postabsorptive state.[2] This is demonstrated by the data of Jones (*1*), in which the range for the blood–breath ratio is 1837–2863 for the postabsorptive state and 990–2863 when both the absorptive and postabsorptive states are included. While it may be useful to distinguish between these two states of alcohol distribution for academic reasons, there is little evidence to justify the conclusion that a random subject will be postabsorptive when tested under field conditions. Nevertheless, Jones (*1*) stated that "in practice, breath alcohol analysis for medicolegal purposes may be expected to be carried out after the peak concentration has been reached," but no references were cited to support this, and the very subjects he studied were not considered to be postabsorptive until two hours after drinking had stopped. Martin et al. (*19*) reported that it took between 60 and almost 120 min to reach the peak venous BAC, and Schmidt et al. (*30*) found that a large proportion of their subjects had their actual BACs overestimated until 120 min after drinking stopped. Dubowski has concluded that (*3*) "... alcohol absorption is not always complete within 60 to 90 minutes, as often claimed," and that "... it is not possible to establish whether an individual is in the absorption or elimination phase ... from the results of two consecutive blood or breath alcohol measurements, however timed." Indeed, in a field study done in Canada on 243 arrested drivers, it was reported (*23*) that: "... very few are in the elimination phase. Most of them were found in the plateau and absorption phase."

It has been known for some time that the blood–breath ratio varies, not only from one subject to another but also within the same subject as a function of time (*1–3*). The SD reported by Jones (*1*) consists of nearly equal contributions from both *inter*- and *intra*-subject variability, and he defined this as "biological variation," pointing out that it would be even greater under field conditions. If other authors have not included intra-subject variation in the SD of the blood–breath ratios, then the SDs they report are too small by a significant amount.

So the results from laboratory studies are not applicable to a random subject under field conditions because important variables have been carefully controlled in the former, and results from field studies are not reliable because of inadequate control of important variables and because of other flaws. The best solution to this problem is the adoption of the most nearly reliable data, which are from controlled laboratory studies, even though the results represent the minimum expected variability[3] in the blood–breath ratio for a random subject under field conditions. The variability is greatest when subjects are in the absorptive state, and this is the subject of a forthcoming publication.

Total Uncertainty

The value of $CV_T = 10.8\%$ indicates that the uncertainty in a breath-analysis result will be about $\pm 22\%$ ($\pm 2\ CV_T$) for 95 out of 100 subjects if the analyzer is calibrated at 2280:1. For an analyzer calibrated at 2100:1, the uncertainty will be about $\pm 15\%$ for 95 out of 100 subjects, and this implies that about 23% of these subjects will have their actual BAC overestimated. The relative deviation from the data of Dubowski and O'Neill (*3, 27*) can be found from the experimental range of 1706–3063, and its value is about $\pm 19\%$. Because these data are based on results from a homogeneous sample under laboratory conditions, the values may be regarded as minimum values. How does the uncertainty expressed by these values compare with the specifications for accuracy and precision published by the manufacturers of quantitative evidential breath analyzers? Manufacturers commonly claim an accuracy and repeatability of "better than" 0.01% BAC, or 0.10 g/210 L, which is equivalent (*6, 31*). Such a specification is ambiguous, however, because it does not state where in the measurement range this amount of error occurs, nor does it state what the magnitude of the error is, it just states "better than." If the specified absolute deviation of 0.01% BAC is converted to a relative deviation, then at an analyzer reading of 0.08% BAC the accuracy and repeatability are 12.5%, while at a reading of 0.40% BAC they are 2.5%, yet the data of Dubowski and O'Neill indicate a relative deviation of 19%. It appears that only at low BACs do the manufacturers' claims for accuracy and precision even approach the 19% derived from the data of Dubowski and O'Neill and, moreover, the 0.01% BAC claimed by the manufacturers is a maximum deviation, while the experimental value of 19% is a minimum because it is derived from a homogeneous sample under laboratory conditions. At higher BACs, it follows from the manufacturers' specifications that the accuracy and precision increase until they equal and then surpass that of direct blood analysis. In general, manufacturers' specifications for the accuracy and precision of scientific instruments are conservative, for obvious reasons, but it is especially important that they be conservative when the measurements are used as evidence in a legal proceding. There now are abundant experimental results that show that the manufacturers' specifications for accuracy and precision of their breath analyzers are far too optimistic, and appropriate steps need to be taken to correct this.

The experimental range reported by Dubowski and O'Neill (*3, 27*) indicates a relative deviation of about $\pm 19\%$ in the results from breath analyzers, while the value expressed by $\pm 2\ CV_T$ is about $\pm 15\%$. Are these reasonable minimum values for the expected uncertainty in breath analyzer readings? Jones found that even under carefully controlled laboratory conditions, there was a variation of $\pm 26\%$ of the mean BAC in the estimates of BAC from a Model 900 Breathalyzer for 55 healthy men when the absorptive and postabsorptive states were included, using 95% confidence limits (*32*).

[2] Arterial BAC is estimated more accurately than venous BAC by BrAC measurements during the absorptive state. But venous blood, or in some cases capillary blood, is most often the sample used for direct measurement of BAC in arrested drivers, and this means that the accuracy of BAC estimates must be based on direct measurements of venous blood, or in some cases capillary blood.

[3] It is certainly possible that some other homogeneous groups—e.g., healthy women—would show less variability in the blood–breath ratio. Since this has not yet been shown to be the case, though, it is assumed here that healthy men represent the minimum variability when compared with the general population.

Implications

Even though there is significant variability in the partition ratio in the general population, the use of 2100:1 as a conversion factor for everyone has been reaffirmed as recently as 1984 (National Highway Traffic Safety Administration). Its continued use implies that every subject has a partition ratio of 2100:1, that this value is a constant, with no uncertainty in it.[4] None of these implications is valid and the variability as expressed by CV_T is simply ignored. The consequences of assuming that such implications are valid are most serious for subjects having BAC estimates that are at or near the presumptive or per se statutory limits for BAC of 0.08 g/deciliter (0.08%) or 0.10 g/deciliter (0.10%) BAC, or the equivalent in units of g/210 L, because they will be made liable for unjust prosecution. While the decision as to just what percentage of the population should be made liable for unjust prosecution is problematic, a 23% minimum, which results from the use of 2100:1 for all subjects, seems excessive. Adoption of the range 1797–2763 (±2 SD from the data of Dubowski and O'Neill) would result in possible unjust prosecution of only about 2.5% of healthy men in the fully postabsorptive state who have BAC estimates from breath analyzers at or near the statutory limit.

BrAC is now commonly reported in grams per 210 L of breath, which corresponds to a BAC reading in grams per 100 mL of blood for a breath analyzer with which the 2100:1 conversion factor is used; i.e., a BrAC reading of 0.10 g/210 L corresponds to a BAC of 0.10 g/100 mL. An increasing number of jurisdictions no longer convert BrAC to BAC by means of the breath analyzer (*3, 5*), in conformity with the Uniform Vehicle Code. Instead, the instrument is calibrated to read out direct BrAC, and this has gained support from some prominent researchers (*1, 3, 5, 25, 33*) as well as from the National Safety Council. At first glance, it appears that such a calibration solves the problem of variability in the blood–breath ratio, because the instrument no longer uses the 2100:1 ratio to convert BrAC measurements to BAC results. It is indeed possible in principle to eliminate the 2100:1 conversion factor, but the only acceptable way to do this is by doing the appropriate studies that might establish the direct BrAC at which driving abilities are impaired. Because this direct BrAC has not yet been established, the 2100:1 ratio has simply been used to convert existing statutory limits of BAC to new limits of direct BrAC. Using this approach does not solve the problem of variability in the blood–breath ratio because the conversion is effectively being done by the legislature when they define the statutory limit of BrAC to be 0.08 g/210 L or 0.10 g/210 L. Whether the conversion is done by the legislature or by the breath analyzer, it is still being done, and the uncertainty in this factor should be reflected in both cases. Ignoring this uncertainty is not only contrary to accepted methods for the evaluation of analytical results, but the use of such an approach also causes difficulties from a legal standpoint, especially regarding per se statutes (*34*).

Isaacs et al. (*33*) have also advocated the use of direct BrAC, and to support this position they cited the example of urinary alcohol testing. They concluded that: "In Great Britain, the analysis of a urine sample for alcohol has not needed to be converted to an equivalent BAC since 1967, and yet it has been generally accepted that there is a variation in the blood:urine ratio within the population as a whole. This is an entirely parallel situation [to direct BrAC reporting]." However, in the U.S. a different approach has been used, as shown by the statement (*3*) that "there is massive documentation that the blood alcohol concentration cannot be established sufficiently reliably for forensic purposes from the alcohol concentration of a pooled bladder urine specimen because of the extensive variability of the blood:urine ratio of alcohol . . . These problems have been recognized in policy statements of the National Safety Council's Committee on Alcohol and Drugs (1979) and of the National Highway Traffic Safety Administration (1975), which list blood and breath as specimens of choice for alcohol analysis and discourage use of urine for alcohol determination in law enforcement." It is very curious that in the U.S. the problem of variability in the blood–urine ratio is resolved by simply abandoning urine alcohol testing, while the problem of variability in the blood–breath ratio is resolved statutorily by using the 2100:1 conversion to establish direct BrAC measurements in the Uniform Vehicle Code (*5*). If certain fundamental principles of science and law are to be maintained, the uncertainty in both of these conversion factors must be acknowledged. This can be done by amending existing statutes to require that an appropriate amount of uncertainty be reflected in the results of alcohol analyses.

Further Observations

It is becoming clear that there are large *inter-* and *intra-*subject variations in the human pharmacokinetics of alcohol (*1–3, 5, 19, 28, 35*). Significant short-term fluctuations in BAC and BrAC have been documented (*3*), even though in some cases these reported variations have been attributed to instrumental error (*26*). Jones (*1*) was similarly critical of the work by Alobaidi et al. (*2*) because they measured BrAC ". . . by means of a helium–neon infrared analyzer which apparently had not been compared with standard and well-established breath testing equipment." It is not clear from the data in reference *2* that the variability is due to instrumental error, yet Jones gives no rationale for his conclusion that it is. It was apparently not recognized that measurements made by completely different instrumentation and calibration methods can serve as a valuable tool for the detection of determinate error. The CV calculated from the data in (*2*) for the postabsorptive state is similar to that reported by Dubowski and O'Neill (*3, 27*), and the substantial variations reported in reference *2* and in other work (*6, 19, 30*) are not likely to be the result of instrumental error but the result of variations in the blood–breath ratio, the predominant source of error in breath analyzer readings.

How should the variability of the 2100:1 ratio be taken into account? Dubowski has stated (*3*): "It is evident from considerations of quantitative human biology that a single ratio or conversion factor will not apply to all persons." Others have pointed out that the use of population-average pharmacokinetic parameters largely ignores individual variations and is consequently not a reliable means of predicting a value for a random individual (*36, 37*). It is fairly clear, for example, that it is not possible to make a reliable estimate of an individual's weight just by assuming it is the same as the population average weight. Experimen-

[4] It has been argued that there is approximately a 9% allowance for the uncertainty in breath analyzers in which the 2100:1 ratio is used, because the average blood–breath ratio for most people is said to be about 2300:1. However, the 2300:1 value involves either the knowledge or the assumption that most people are postabsorptive when tested, and this simply has not been shown to be the case under field conditions. Furthermore, 9% is not a large enough correction to account for even the minimum variations in the blood–breath ratio found in a homogeneous sample, let alone the general population.

tal results also illustrate the inadequacy of applying the population-average blood–breath ratio to all subjects. The postabsorptive mean of 2231 from the data of Alobaidi et al. (2) agrees closely with the mean found in many other studies, but it was concluded that breath analysis is not reliable for the quantitative determination of BAC because of large inter- and intra-subject variability in the blood–breath ratio. Jones (*1*) reported a mean, ±SD, CV, and range of 2180, ±189, 8.7%, and 1837–2863, and these results reinforce the view that there is not excessive variation in the blood–breath ratio and that most subjects will have their actual BAC underestimated by breath testing. A closer look at these postabsorptive data reveals that 76% of the subjects had at least one measurement of their blood–breath ratio that was less than 2100:1 and that 19% had at least one result less than 1900:1. The data of Schmidt et al. (*30*) show that the average difference between BAC as estimated from BrAC and directly measured BAC is about 0.80 g/deciliter (0.008%) for BAC at 60 min after drinking stopped, which also supports the view that these blood–breath differences are not excessive. However, inspection of the graphical data reveals that between 65% to 80% of the subjects had their actual BAC overestimated and that 45% to 50% of the subjects had their BACs overestimated by 0.01% BAC or more.

Because a single value for the blood–breath ratio based on the population average for the postabsorptive state does not adequately apply to a random subject, a conversion factor based on some measure of the variability of the blood–breath ratio should be used instead. The minimum range has been determined experimentally and can be used to express the variability in this ratio. The CV may also be used, but since a representative probability distribution is not yet available for the general population, no probabilities should be used in connection with the CV. Because the experimental range from laboratory studies is indicative of values that are possible in a homogeneous sample, such values are even more likely to occur in a heterogeneous population. Consequently, the experimental range, or even $\pm 3\ CV_T$, should be used to establish the conversion factor. The relative uncertainty in breath-analyzer readings corresponding to the experimental range is about ±19% and that from $\pm 3\ CV_T$ is ±27%. Use of these values is consistent with giving the defendant the benefit of the doubt about the minimum uncertainty in breath-analyzer results.

CV_T = 10.8% consists almost entirely of contributions from CV_S, and it is large in the context of analytical measurements. Is there a value of CV_S beyond which an analytical method should be considered too inaccurate or too unreliable for a particular use? BAC is frequently estimated from calculations based on body weight, fluid contents, and alcohol dose (*3, 28, 38, 39*). Nomograms have been developed from these calculations, but Dubowski and O'Neill have argued that they should not be used because of the large variability, CV = 16.5%, in actual results for peak BAC after a given dose of alcohol (*3, 39*). Estimates of BAC from BrAC measurements also have a large CV, but in this case Dubowski has effectively recommended that the legislature use the 2100:1 conversion to implement direct BrAC measurements (*3, 5, 25*), and this recommendation has been adopted in the Uniform Vehicle Code (*5*). It would follow from this reasoning that there is some value for CV_S between 10.6% and 16.5% at which methods for the estimation of BAC become unreliable.

But the most troubling question is: "Why would BAC be estimated from a BrAC measurement when a direct blood measurement is both possible and readily available?" It is a huge leap, both quantitatively and qualitatively, to assume that there is any equivalence between a directly measured BAC and a BAC estimated from a BrAC measurement. When BAC is measured directly, it is done by well-proven analytical methods with accuracy and precision that can be rigorously determined. When BAC is estimated from a BrAC measurement, the result is an estimate based on a faulty probability distribution. Given the choice, it would seem that if a conclusion is to be made about the BAC of a random subject, especially when the conclusion can have serious consequences, it would be far preferable to make it on the basis of a direct measurement rather than on the basis of a probability derived from an inappropriate distribution. Because such estimates of BAC will probably continue, though, 95% confidence limits for results from laboratory studies represent a minimal safeguard in the selection of the amount of uncertainty to be used for the partition ratio and the breath-analysis results.

Recommendations

The accuracy of BAC estimates from BrAC measurements can be improved if recommendations made in the past are implemented. These include the measurement of and correction for breath-temperature variations (*25, 35*), calibration with tandem simulators for each subject (*15, 29*), and an objective determination of whether the subject is actually in the postabsorptive state (*40*). Manufacturers of quantitative evidential breath alcohol analyzers need to use explicit, unambiguous specifications for accuracy and precision that are consistent with results of research published in the scientific literature. Subjects should be informed that breath-analyzer results can be less accurate than direct BAC measurements, and a direct blood test should be advised when the breath result is at or near statutory limits. The practice of reporting an appropriate amount of uncertainty along with breath test results should also be adopted.

Conclusions

For many years, claims have been commonly made about the accuracy and precision of quantitative evidential breath alcohol analyzers—by scientists, expert witnesses, and by manufacturers—that are not supported by experimental results in the literature. The actual minimum uncertainty in a random breath-analyzer result for a random subject can be reasonably expressed by using either $\pm 2\ CV_T$, the experimental range, or even $\pm 3\ CV_T$. These result in relative uncertainties of ±15%, ±19%, or ±27%, respectively. Over 90% of this uncertainty is ascribable to variables involving the subject, even when the least-reliable conforming analyzer is used. Breath-analysis results are analytical measurements, and the traditional rules for reporting analytical results should be followed. Results of two consecutive breath tests in the postabsorptive state are not sufficient information to determine reliably the BAC of a random subject for evidentiary purposes unless the results exceed the statutory limit and the appropriate uncertainty.

I am greatly indebted to Gordon Smyth, Assistant Professor, Department of Mathematics, Statistics and Applied Probability Program, and John H. Kennedy, Professor of Chemistry, Department of Chemistry, both of the University of California, Santa Barbara, for helpful discussions and for review of the manuscript.

References

1. Jones AW. Variability of the blood:breath ratio in vivo. J Stud Alc 1978;39:1931–9.

2. Alobaidi TAA, Hill DW, Payne JP. Significance of variations in blood: breath partition coefficients of alcohol. Br Med J 1976;2:1479–81.

3. Dubowski K. Absorption, distribution and elimination of alcohol: highway safety aspects [Review]. J Stud Alc, Suppl 10, 1985:98–108.

4. Tsukamoto S, Karasawa S, Sudo T, et al. Experimental study on ethanol concentration ratios of breath to body fluid. Nihon Univ J Med 1983;25:269–80.

5. Dubowski K. Recent developments in alcohol analysis. Alcohol Drugs Driving 1986;2:13–46.

6. Emerson VJ, Hollyhead R, Isaacs MDJ. The measurement of breath alcohol. J Forens Sci Soc 1980;20:3–70.

7. Isaacs MDJ, Jacobs JM, Emerson VJ, Broster CG, Hunt DJ. Field trial of three substantive breath alcohol testing instruments. London: Her Majesty's Stationery Office, HMSO publication 1982. ISBN 0 11 340783 1.

8. Bonnichsen R, Goldberg L. Large-scale breath-blood comparisons under field conditions: methods, evaluation techniques and results. In: Goldberg L, ed. Alcohol drugs and traffic safety. Vol II. Stockholm: Almqvist and Wiksell International, 1981:796–810.

9. Pribor HC, Campbell CR, Hebb JH. Ethanol and vehicular safety; the Nashville experience: Part II. Breath-/blood-alcohol correlation. Lab Management 1985;23:21–7.

10. Standard devices to measure breath alcohol. Department of Transportation, National Highway Traffic Safety Administration Highway Safety Programs. Fed Register 1984;49:48854–72.

11. State of California, Administrative Code, Title 17. 1975;75:88.4.14.

12. Hoel PG. Elementary statistics. 4th ed. New York: John Wiley and Sons, 1976:133–6.

13. Skoog DA, West DM. Fundamentals of analytical chemistry. New York: Holt, Rinehart & Winston, 1963:33–66.

14. Caplan YH, Yohman DT, Schaefer JA. An in vitro study of the accuracy and precision of Breathalyzer Models 900, 900A and 1000. J Forens Sci 1985;30:1058–63.

15. Dubowski K. Breath alcohol simulators: scientific basis and actual performance. J Anal Toxicol 1979;3:177–82.

16. Harger RN, Raney BB, Bridwell EG, Kitchel MF. The partition ratio of alcohol between air and water, urine and blood; estimation and identification of alcohol in these liquids from analysis of air equilibrated with them. J Biol Chem 1950;183:197–213.

17. Goldberger BA, Caplan YH. Infrared quantitative evidential breath-alcohol analyzers: in vitro accuracy and precision studies. J Forens Sci 1986;31:16–19.

18. Jones AW. The precision and accuracy of a gas chromatograph intoximeter breath alcohol device. Part I—in vitro experiments. J Forens Sci Soc 1978;18:75–80.

19. Martin E, Moll W, Schmid P, Dettli L. The pharmacokinetics of alcohol in human breath, venous and arterial blood after oral ingestion. Eur J Clin Pharmacol 1984;26:619–26.

20. Jones AW. Determination of liquid/air partition coefficients for dilute solutions of ethanol in water, whole blood and plasma. J Anal Tox 1983;7:193–7.

21. Russell JC, Jones RL. Breath alcohol concentration and analysis in the presence of chronic obstructive pulmonary disease. Clin Biochem 1983;16:182–7.

22. Jones AW. Electrochemical measurement of breath-alcohol concentration: precision and accuracy in relation to blood levels. Clin Chim Acta 1985;146:175–83.

23. Adrian IW. The measurement of breath alcohol concentration and the inherent physiology of the lung. In: Goldberg L, ed. Alcohol, drugs and traffic safety. Vol II. Stockholm: Almqvist and Wiksell International, 1981:622–36.

24. Jones AW. How breathing technique can influence the results of breath alcohol analysis. Med Sci Law 1982;22:275–80.

25. Mason MF, Dubowski K. Breath alcohol analysis: uses, methods and some forensic problems: review and opinion. J Forens Sci 1976;21:9–41.

26. Jones AW. A historical and experimental study of the breath/blood alcohol ratio. In: Israelstam S, Lambert S, eds. Alcohol, drugs and traffic safety. Toronto: Addiction Research Foundation of Ontario, 1975:509–26.

27. Dubowski K, O'Neill B. The blood/breath ratio of ethanol [Abstract]. Clin Chem 1979;25:1144.

28. Hume DN, Fitzgerald EF. Chemical tests for intoxication: what do the numbers really mean? Anal Chem 1985;57:876A–86A.

29. Dubowski K. Alcohol analysis: clinical laboratory aspects. Part I. Lab Management 1982;20:43–54.

30. Schmidt G, Schneider V, Wustefeld H. Vergleichende Untersuchungen zur Atemalkoholbestimmung. Blutalkohol 1986;23:133–7.

31. Operator's Manual, Intoxilyzer 4011A. CMI Inc., Minturn, CO.

32. Jones AW. Precision, accuracy and relevance of breath alcohol measurements. Mod Probl Pharmacopsych 1976;11:68–78.

33. Isaacs MDJ, Emerson VJ, Fuller NA, Hollyhead R. Blood: breath ratios obtained in trial of three breath alcohol testing instruments. In: Goldberg L, ed. Alcohol, drugs and traffic safety. Vol II. Stockholm: Almqvist and Wiksell International, 1981:442–55.

34. Thompson SG. The constitutionality of chemical test presumptions of intoxication in motor vehicle statutes. San Diego Law Rev 1983;20:301–38.

35. Gatt JA. The effect of temperature and blood:breath ratio on the interpretation of breath alcohol results. New Law J 1984;134:249–52.

36. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis–Menten model: routine clinical pharmacokinetic data. J Pharmacokin Biopharmacokin 1980;8:553–71.

37. Martin E, Moll W, Schmid P, Dettli L. Problems and pitfalls in estimating average pharmacokinetic parameters. Eur J Clin Pharmacol 1984;26:595–602.

38. Dubowski K. Human pharmacokinetics of ethanol. I. Peak blood concentrations and elimination in male and female subjects. Alc Tech Rep 1976;5:55–63.

39. O'Neill B, Williams AF, Dubowski KM. Variability in blood alcohol concentrations. J Stud Alc 1983;44:222–30.

40. Fitzgerald EF, Hume DN. The single chemical test for intoxication: a challenge to admissibility. Mass Law Rev 1981;66:23–27.