MAC Attack?

IN this issue of the Journal, Lu et al. explore the arcane subject of modeling binary data using population analysis, a technique that determines the response of the typical individual, as well as inter- and intraindividual variability. They demonstrate that when there are small numbers of observations per individual, the population approach to data analysis results in a hugely biased estimate of the Hill coefficient in logistic regression. The article mentions minimum alveolar concentration (MAC) only in passing, but their findings raise the question, Is MAC fundamentally flawed?

MAC is among the most useful concepts in anesthetic pharmacology. MAC establishes a common measure of potency for inhaled anesthetic drugs: the partial pressure at steady state associated with 50% probability of movement to noxious stimulation (e.g., incision). We use the concept of MAC to provide uniformity to our dosage of inhaled anesthetic drugs, establish the relative amounts of drug for different endpoints (e.g., MACawake, MACBAR, MACknife), characterize drug interactions (e.g., MAC-reduction), and guide our search for mechanisms of anesthetic action (the concentration responsible for biologic effects must be similar to MAC).

One of the great mysteries of anesthetic action is that MAC is so consistent. The inhaled anesthetic drugs are unique in pharmacology in their incredibly small amount of pharmacodynamic variability. Within a population, MAC varies by not more than 10–15% among individuals. MAC varies from species to species by approximately the same amount as it does from individual to individual. Someday, when we understand the mechanism of inhaled anesthetic action, we will look back on this low variability in MAC and think “it was so obvious that the mechanism had to be X, because only that could have accounted for the low variability.”

Lu et al. demonstrate that the type of study used to determine MAC in humans might produce highly biased underestimates of variability. By definition, MAC in humans is the concentration associated with 50% probability of response to initial incision. It is limited to initial incision to provide a uniform experimental design. However, because there is only one initial incision in a patient, you only get one lousy bit of information per patient: response or no response. There is no room for partial responses—either the patient responded or didn’t. It takes eight patients to make a single byte of data.

A consequence of the minimal data in each observation is that estimates of MAC and its variability are vulnerable to bias. Paul and Fisher observed that the classic “up–down” experimental design to determine MAC could be expected to produce errors in MAC of 10%, and that variability in MAC was systematically underestimated. In a previous manuscript, Lu and Bailey demonstrated that when patient-to-patient differences are ignored, and the data are treated as arising from one giant rat (called the naive pooled data approach), the steepness of the concentration versus response curve is grossly underestimated. Figure 1 shows the probability versus response relationship in many individuals (thin lines) and the apparent curve that would result from treating the data as though arising from one individual (thick line).

In the current article, the authors ask the question, Could population analysis describe representative individuals (fig. 1, thin lines) and thus correct the “error” of the thick line in figure 1? Their results are quite disconcerting. They demonstrate that it takes at least 10 observations per subject to get an unbiased estimate of the Hill coefficient with the population approach. To understand the reason for this, consider a study with only two observations per patient. With two observations, there are four possibilities for the concentration versus response relationship as shown in figure 2. The thin curve has a very large Hill coefficient. This curve perfectly predicts the observations in 2A–C but provides a perfectly terrible fit of the observations in 2D. The thick curve provides OK fits of all the data points (similar to the thick curve in fig. 1). However, if A, B, C, and D were all individuals in the same study, a population approach would average the nearly infinite Hill coefficients of A, B, and C (thin lines), with something more modest to fit D. The average of three near-infinities and something less than infinity still yields an enormous value for the Hill coefficient. Because the Hill coefficient is directly related to the SD of MAC, could the low variability in MAC be an artifact of the data analysis?

Fortunately, the early MAC studies preceded modern population analysis techniques and simply used the giant rat analysis technique. More recent studies continue to use the giant rat analyses technique. As a result, virtually all MAC studies estimate the response shown by the thick line in figure 1 and do not attempt to estimate the response in individuals (fig. 1, thin lines). This is a good thing. First, clinicians want to set their doses at
concentrations at which the majority of individuals are anesthetized, which is the dose determined using the giant rat analysis technique. Second, the probability of response versus concentration curve is by definition steeper in each individual than in the population as a whole. Because the population as a whole shows variability of just 10% or less, in each individual the curve must be almost vertical, with individuals moving from 100% chance of responding to zero chance of responding, with very small increments in concentration. This agrees with clinical practice.

Although the observations of Lu et al. thus do not invalidate the conclusions of MAC studies to date, they convincingly demonstrate that studies with only a single observation per subject will never establish the concentration versus response curve in individuals, at least not by using population analysis techniques. More important, the article by Lu et al. reinforces the previous message of Paul and Fisher: In human MAC studies, each individual literally contributes one bit of data. As a result, modest differences in MAC values between two groups in a study, or when compared with historical controls, may be an artifact unless very careful statistical measures are used to compare the groups.

Omeed Sani, M.D.*, Steven L. Shafer, M.D.†
*University of Szeged, Szeged, Hungary. †Palo Alto Veterans Administration Health Care System, Palo Alto, California. steven.shafer@stanford.edu

References


Fig. 1. Individual concentration versus probability of no response curves (thin lines) which (in theory) can be estimated using population modeling, and an overall concentration versus probability of no response curve (thick line), which results when interindividual differences are ignored (the naive pooled data approach).

Fig. 2. Four possible alignments of two data points. In graphs A, B, and C, the perfect fit for individuals would have an infinite Hill coefficient. In graph D, the infinite Hill coefficient results in a very poor fit. The naïve pooled data approach to all of the data results in a fairly shallow curve (thick line) and thus a small Hill coefficient. Were these data from four separate individuals, the average of the infinite slope from individuals A, B, and C, with the shallow slope for individual D would still yield a very high value for the Hill coefficient.
Principles of Successful Sample Surveys

Every few weeks, we receive an article based on some form of survey. Unfortunately, most of those submissions are fatally flawed, usually because the survey was performed in a fashion that calls any conclusions into question. The survey in this month’s issue, from the American Society of Anesthesiologists Committee on Transfusion Medicine, is likely to be of great interest to many anesthesiologists, but there are problems with its design and conduct. I asked Dr. Burmeister, who has long experience in the design and conduct of surveys, to provide a basic introduction to survey design, from the perspective of a professional survey expert. I would strongly urge anyone considering carrying out a survey in the future to read his comments very closely. Surveys can be a very valuable way of collecting important information; however, like all good experiments, they are rarely as easy or straightforward to perform as they might seem.

Michael M. Todd, M.D., Editor-in-Chief, anesthesiology@uiowa.edu

THE results of sample surveys are important in our personal and professional lives. Gallup polls and results from various news agency polls tell us what is important politically. Nielsen studies determine what television shows are presented on national networks. Similarly, results of samples drawn from professional organizations help formulate and evaluate recommendations important to the practice of medical specialties. An example of such a study is included in this issue of the Journal. It is essential that such studies be properly conducted. If they are not, serious consequences could result.

Perhaps the most important necessity for a valid study is the existence of a complete sampling frame, which is a listing of all individuals constituting the population of interest. The identification of the sampling frame may be difficult, even when the population of interest is a professional organization. The list of members is often dynamic, including those who may no longer be active and excluding the very recent additions to membership. If a simple random sample is to be selected from a membership list, as done in the Nuttall et al. study, all reasonable efforts should be made to ensure that the list is current.

If not, the results of the survey are potentially biased, assuming former members on the list and new subjects not included are different in characteristics and opinions from those included in the population of interest.

Another aspect of sample surveys that rivals sampling frames in importance is the survey instrument. Questions must be of established reliability and validity. Occasionally, well-established questionnaires can be used. It is more likely, however, that at least some questions may be newly developed for the proposed study. In such cases, it is advisable to seek consultation from experienced survey personnel and to complete pilot studies to evaluate the survey instrument.

Adequate sample size is another requirement for a successful sample survey. There is no easy determination of adequate sample size; for example, sample size is not determined by selecting a specific percentage of the population. Instead, it depends on the desired precision of characteristics to be estimated and the confidence level assigned to achieving the specified precision. Most sample size determinations, as done by Nuttall et al., are based on precision by specifying the half-width of a confidence interval, or margin of error. If hypotheses are to be tested, the necessary sample size is increased because statistical power is an additional consideration. It should be noted that Nuttall et al. did not actually include power in their sample size computation. If sample designs less efficient than simple random sampling are used, the sample size is also increased.

Simple random sampling may not be the most efficient of the possible sampling plans. For example, if the characteristics being estimated vary by age, experience, gender, subspecialty, and so forth, it may be advantageous to stratify the population of interest and select simple random samples within each stratum. Doing so could improve the precision of the estimates, or reduce the necessary sample size for a specified level of precision. On the other hand, because of the lack of an up-to-date sample frame, it may be necessary to cluster the population and to select samples of the clusters, rather than a simple random sample of individuals. Clusters are often geographic in nature, such as state or local professional organizations. Individuals comprising a given cluster are often similar, which decreases precision and increases the necessary sample size.

The use of alternative sampling plans not only affects the sample size, but it also affects the estimation of precision and alters the width of confidence intervals. Effective stratification will increase the precision of estimates; however, the use of cluster sampling almost always decreases the precision. The overall effect on the precision of complex sample designs, using both stratification and cluster sampling and, perhaps, other types of sampling, is difficult to predict. Therefore, it is essen-
tial to analyze the collected data in a manner consistent with the sample design. This often requires the use of challenging computer programs such as SUDAAN (Survey Data Analysis). However, the assumption of sample random sampling when, in fact, a more complex sample design was used can lead to misleading results.

Misleading results are also a consequence of nonresponse. Nonresponse bias is equal to the proportion of nonresponse multiplied by the difference of the responders and nonresponders; consequently, there is no absolutely acceptable level of response. Increasing the initial sample size to accommodate a relatively low expected response rate, as done by Nuttall et al., does not eliminate nonresponse bias. A better use of resources would be to decrease the initial sample size and increase efforts to contact initial nonresponders.

Because it is nearly impossible to eliminate nonresponse bias when studying human populations, it is essential to describe the potential nonresponse bias and minimize the fraction of nonresponders. Description of the potential nonresponse bias is not always possible; however, demographic characteristics of the entire population may be available from membership files, Bureau of the Census data, or other sources. If these data are available, then characteristics of the sample can be compared to those of the population. Of course, even nearly identical demographic characteristics would not rule out potential nonresponse bias, as it is likely that the demographic characteristics would only be moderately correlated, at best, with the characteristics of interest.

Therefore, strategies to reduce the proportion of nonresponders are highly recommended. At minimum, a sampling of the nonresponders should be attempted. It would be naïve to hope that doing so would eliminate nonresponse bias; after all, nonresponders have their reasons for nonparticipation. A simple second contact will result in only some increase in participation; however, those who respond initially and after second contact can be compared to gain some insight relative to potential nonresponse bias.

As noted above, it is preferable to use strategies to reduce the initial level of nonresponse. Such strategies might include endorsements from community leaders or leadership boards, use of short questionnaires, incentives, and so forth. However, it must be concluded that such strategies will not eliminate nonresponse bias. The best we can hope for is a reasonable reduction in such bias.

In summary, the survey is a very important research tool that is challenging to do well. It certainly has inherent limitations, but it also has the potential to make important contributions. The purpose of this summary of challenges is to serve as a reminder to readers of the Journal and potential authors of the many considerations necessary to complete good survey research.

Leon F. Burmeister, Ph.D. College of Public Health, The University of Iowa, Iowa City, Iowa. leon-burmeister@uiowa.edu

References

Cardiac Arrest following Regional Anesthesia with Ropivacaine

Here We Go Again!

The current issue of the Journal contains the first reports of cardiac arrest associated with ropivacaine administered for surgical regional anesthesia.1,2 In the first case report, by Chazalon et al., progressive bradycardia and asystole occurred in a 66-yr-old woman who received ropivacaine, 6.67 mg/kg, for lower extremity blocks.1 In the second case report, Huet et al. describe sudden cardiac asystole in a 66-yr-old, 100-kg man after the administration of ropivacaine 1.88 mg/kg for a lumbar plexus block.2 Fortunately, resuscitation was successful in both patients and there were no sequelae.

Ropivacaine was introduced into clinical practice in the United States in the early 1990s as a possible safer alternative to bupivacaine. The decision to bring another long-acting amide local anesthetic to market is inextricably linked to the history of bupivacaine use in the United States. Twenty-four years ago, George Albright, M.D., then Assistant Professor, Department of Anesthesia, Stanford University School of Medicine, Stanford, California, published an editorial in the Journal alerting practitioners to six anecdotal cases of almost simultaneous seizures following intravascular injection of what were then the newer long-acting amide local anesthetics to market.3

Ropivacaine differs from other local anesthetics in that it has a narrower margin between the dose or plasma concentration required to produce seizures as compared to those resulting in cardiovascular collapse.4–6 This accrues from the fact that supraconvulsant doses of bupivacaine, but not lidocaine or mepivacaine, may induce lethal ventricular arrhythmias out of proportion to the drug’s anesthetic potency.4–6 Two theories have been proposed to explain this phenomenon. First, both bupivacaine and lidocaine block cardiac sodium channels rapidly during systole; however, during diastole, bupivacaine dissociates off these channels at a much slower rate than lidocaine.7 As a result, at normal heart rates, diastolic time is sufficiently long for dissociation of lidocaine, but a bupivacaine block intensifies and depresses electrical conduction, causing reentrant type ventricular arrhythmias. Second, high blood concentrations of bupivacaine may cause a ventricular arrhythmia through a direct brainstem effect.8

Ropivacaine is structurally similar to mepivacaine and bupivacaine. Unlike formulations of other local anesthetics in clinical use, ropivacaine is prepared as the single levorotatory isomer rather than as a racemic mixture of the levo and dextro forms of the drug. This is important because the levorotatory isomer has less potential for systemic toxicity than the dextrorotatory isomer.9 In vitro and animal studies have demonstrated that ropivacaine is intermediate between lidocaine and bupivacaine in its depressant effects on cardiac excitation and conduction as well as in its potential to induce reentrant type ventricular arrhythmias and death.10,11

The relationship between the anesthetic potency of ropivacaine and its margin of safety has been a source of controversy. The in vitro potency of ropivacaine is approximately 25% less than that of bupivacaine. This is not surprising considering that ropivacaine has a shorter aliphatic chain, a propyl group, attached to the pipe-chol ring as compared to a butyl group for bupivacaine. Lower potency would be important only if greater doses of ropivacaine as compared to bupivacaine were required to achieve comparable anesthetics in women receiving epidural analgesia during labor is almost twice that of bupivacaine.12,13 Thus, it is important to consider that although at equal doses ropivacaine seems to have a wider margin of safety than bupivacaine, the potential for systemic toxicity will also be affected by the relative total dose required for an individual block. It is noteworthy that 0.75% rather than 0.5% ropivacaine was used for regional anesthesia in both of the reported cases.1,2

Patients affected by severe systemic toxicity with ropivacaine may respond more readily to conventional resuscitation than those intoxicated with bupivacaine. In one study comparing both drugs, cardiac resuscitation was less difficult and fewer animals died after supraconvol-

Accepted for publication August 20, 2003. Dr. Santos received research funding from ASTRA (Wilmington, Delaware) for studies of ropivacaine, and from CHIROSCIENCE (in 1999 Chiroscience merged with Celltech, Slough, Berkshire, United Kingdom) for studies involving levobupivacaine. Dr. Santos is also a member of the Levo/ropivacaine Advisory Board (Norwalk, Connecticut, United States).
sant doses of ropivacaine. It is reassuring that both patients in the reported cases responded quickly to conventional resuscitation efforts. 

Whereas refractory ventricular tachycardia/fibrillation would be the expected arrhythmia associated with bupivacaine cardiotoxicity, in the reported cases both patients intoxicated with ropivacaine developed progressive bradycardia, hypotension, and asystole. Thus, it appears that cardiac rate and rhythm disturbances in humans are different with ropivacaine than bupivacaine. However, it is interesting to note that both patients received hydroxyzine for premedication. Hydroxyzine is a first-generation antihistamine drug that elicits dose-dependent slowing of cardiac repolarization and prolongs the Q wave to T wave interval. Whether hydroxyzine somehow modifies the arrhythmogenic effects of long-acting amide local anesthetics, to our knowledge, has not been studied.

We believe that in contemporary practice, the greater risk of life-threatening systemic toxicity from long-acting amide local anesthetics probably now resides with peripheral rather than epidural anesthesia. Epidural anesthesia, even with the use of bupivacaine, had become very safe long before the introduction of ropivacaine. In pregnant women, among whom the problem of bupivacaine cardiotoxicity was most prevalent, the case fatality rate decreased from 8.6 per million regional anesthetics for cesarean delivery between 1979 and 1984, to 1.9 per million regional anesthetics for cesarean delivery from 1985 to 1990. This was accomplished through education and modifications in epidural technique, such as adherence to maximum recommended dosage guidelines, use of the lowest possible concentration and volume consistent with effective anesthesia, use of an appropriate test dose to reduce the risk of unintended intravascular injection, heightened vigilance and monitoring while performing a block, and, perhaps most important, slow fractional dosing of local anesthetic. Although these principles have been uniformly embraced for epidural anesthesia, the same may not be possible for peripheral nerve blocks for the following reasons. First, peripheral nerve blocks typically require the administration of large volumes of local anesthetic (30–40 ml) to achieve satisfactory anesthesia. Second, most peripheral nerve blocks are performed as a single injection through a needle located precisely in the proximity of the nerve(s) to be blocked, and the temptation is ever-present to deliver the required dose of local anesthetic rapidly before the patient moves and the injection of local anesthetic is misplaced because of needle movement. Third, for many peripheral nerve blocks, injection is made in the vicinity of large arteries and veins. Local anesthetics used for peripheral nerve blocks must be judiciously selected for individual patients having specific procedures. It seems that we also need to develop more reliable injection and monitoring techniques to reduce the risk of systemic toxicity associated with individual peripheral nerve blocks. If we do not, we are doomed to repeat history, even with the use of newer, relatively less toxic amide local anesthetics.

Linda S. Polley, M.D.,* Alan C. Santos, M.D., M.P.H.† "Women’s Hospital, Ann Arbor, Michigan. Division of Obstetric Anesthesiology, University of Michigan Health System, Ann Arbor, Michigan. †Department of Anesthesiology at St. Luke’s-Roosevelt Hospital Center, College of Physicians and Surgeons of Columbia University, New York, New York. lpolley@umich.edu

References

3. Albright GA: Cardiac arrest following regional anesthesia with etidocaine and bupivacaine. ANESTHESIOLOGY 1979; 51:285–7
10. Molter R, Covino BG: Cardiac electrophysiological properties of bupivacaine and lidocaine compared with those of ropivacaine, a new amide local anesthetic. ANESTHESIOLOGY 1990; 72:322–9