

## Are Estimates of MAC Reliable?

Matthias Paul, M.D.,\* Dennis M. Fisher, M.D.†

**Background:** Potency of inhaled anesthetics (minimum alveolar concentration [MAC]) is typically studied in humans using an “up-down” approach in which the (quantal) response to skin incision is assessed only once for each individual, so that each individual’s MAC is never determined. The authors examined the influence of interindividual variability and study design issues (e.g., the number of patients enrolled in a study) on the accuracy of MAC estimates.

**Methods:** The typical sequence of a MAC study was simulated. The authors varied and tested the impact of several factors: anesthetic concentration used to start a study; number of “crossovers” (successive patients having different responses to skin incision) to terminate a study; concentration increment between consecutive patients; interindividual variability; and “measurement error.” For each factor, simulations were replicated 500 times, and the resulting estimates were summarized.

**Results:** Starting an experiment below or above the “true” value led to slightly biased MAC estimates; in contrast, variability was underestimated with starting concentrations close to the true value. More than six crossovers improved MAC estimates minimally but increased variability estimates toward true values. A larger increment size affected MAC minimally and increased variability estimates toward true values. A larger interindividual variability led to more “outlier” estimates for MAC. Under many conditions, several of 500 replicates yielded MAC estimates that deviated more than 10% or even more than 25% from the “true” value.

**Conclusion:** Individual experiments may yield inaccurate MAC estimates. This inaccuracy is minimized as the number of crossovers increases; however, improvement diminishes as the number of crossovers exceeds six.

MOST studies that report potency of inhaled anesthetics (minimum alveolar concentration<sup>1</sup> [MAC]) in humans use an “up-down” method<sup>2</sup> in which the presence or absence of movement in response to a surgical incision is assessed, and the target anesthetic concentration for each patient is based on the movement response of the previous patient.<sup>3,4</sup> Because each individual is assessed only once, the actual transition point between presence

and absence of movement for that individual is never determined. Instead, when two successive individuals given different anesthetic concentrations differ in their response, the midpoint of these anesthetic concentrations is recorded. For example, if a patient moves in the presence of an end-tidal concentration of 0.8%, and the next patient, given 0.9%, does not move, this change in response is termed a “crossover” with a midpoint concentration of 0.85%. When four (or more) crossovers have occurred, the experiment is typically terminated, and the midpoint values are averaged to determine MAC. In some instances, the SD of these midpoint values is reported as an estimate of interindividual variability in MAC.<sup>3</sup> This up-down study design for quantal (dichotomous) data differs from the design typically used for drugs that have a quantitative (*i.e.*, range of values) response, for which a fixed design, typically a certain number of individuals studied at each of several predetermined doses, is usual. The up-down approach is typically adopted because it permits a small number of individuals to be studied.<sup>2</sup> We question whether the up-down approach results in accurate estimates of MAC and its interpatient variability. Limitations in the data available from humans prevent accurate assessment of these issues in the absence of a large prospective study. However, using simulation, we can assess the impact of interindividual variability, experimental error, and other issues on the accuracy of MAC experiments conducted using the up-down approach.

### Methods

We simulated (and replicated) MAC experiments using the following approach. First, we assumed that each patient has an individual value for MAC that is a function of the “typical value” for the population (“population MAC,” assumed in most of our simulations to be 1% inhaled anesthetic) and an SD (permitted to vary from one tenth to three tenths of the typical value). Second, we assumed that several factors (*e.g.*, machine measurement error and differences between end-tidal and effect site concentration) influenced the “accuracy” of measured end-tidal values as a representation of effect site concentrations (we use this term, rather than brain concentrations, because the effect of inhaled anesthetics on spinal  $\alpha$  motor neurons may contribute to observed immobility as a response to skin incision<sup>5</sup>). Third, we assumed that in a small number of patients (those in whom individual MAC differed minimally from the actual anesthetic concentration), observations of movement would be imprecise,<sup>6</sup> leading to incorrect assessment of movement *versus* nonmovement.

Additional material related to this article can be found on the ANESTHESIOLOGY Web site. Go to the following address, click on Enhancements Index, and then scroll down to find the appropriate article and link. <http://www.anesthesiology.org>

\* Assistant Clinical Professor, Department of Anesthesia, University of California San Francisco. † Vice President for Medical Affairs, Durect Corporation, Cupertino, California.

Received from the Department of Anesthesia, University of California San Francisco, San Francisco, California. Submitted for publication March 8, 2001. Accepted for publication July 27, 2001. Support was provided solely from institutional and/or departmental sources. Dr. Fisher has consulted for the following companies that market inhaled anesthetics: Baxter, Madison, New Jersey, and Abbott, Abbott Park, Illinois.

Address reprint requests to Dr. Fisher: Durect Corporation, 10240 Bubb Road, Cupertino, California 95014. Address electronic mail to: [fisher@PlessThan.com](mailto:fisher@PlessThan.com). Reprints will not be available from the authors. Individual article reprints may be purchased through the Journal Web site, [www.anesthesiology.org](http://www.anesthesiology.org).

**Table 1. Results from a Single Simulation**

Patient No.	Concentration			$P_{\text{effect site}}$	Individual MAC	Response	Crossover (n)	Midpoint Concentration
	Target	Machine	Alveolar					
1	1.30	1.296	1.271	1.215	1.16	No	No	
2	1.17	1.165	1.208	1.105	1.05	No	No	
3*	1.04	1.043	0.986	0.901	0.81	No	No	
4	0.91	0.915	0.893	0.827	1.06	Yes	Yes (1)	0.975
5	1.04	1.041	1.018	0.992	0.80	No	†	
6	0.91	0.913	0.983	0.936	0.86	No	No	
7	0.78	0.780	0.744	0.700	0.80	Yes	Yes (2)	0.845
8‡	0.91	0.908	0.843	0.788	0.80	Yes	No	
9‡	1.04	1.036	1.029	0.983	0.97	Yes	No	
10	1.17	1.172	1.240	1.197	1.13	No	Yes (3)	1.105
11*	1.04	1.042	0.959	0.931	0.89	No	No	
12*	0.91	0.906	0.916	0.888	0.83	No	No	
13	0.78	0.783	0.795	0.756	1.14	Yes	Yes (4)	
14*	0.91	0.913	0.988	0.961	0.82	No	†	
15	0.78	0.782	0.747	0.704	1.05	Yes	Yes (5)	0.845
16	0.91	0.907	0.887	0.836	1.02	Yes	No	
17‡	1.04	1.040	1.028	1.001	1.03	Yes	No	
18	1.17	1.170	1.175	1.104	0.82	No	Yes (6)	1.105
19	1.04	1.039	1.025	0.985	1.18	Yes	†	
20‡	1.17	1.173	1.160	1.118	1.11	No	Yes (7)	1.105
21	1.04	1.036	0.945	0.863	1.06	Yes	†	
22	1.17	1.174	1.211	1.181	1.03	No	Yes (8)	1.105
23	1.04	1.038	1.041	1.003	0.82	No	No	
24‡	0.91	0.912	0.852	0.810	0.83	Yes	Yes (9)	0.975
25	1.04	1.039	0.991	0.947	1.12	Yes	No	
26	1.17	1.167	1.138	1.078	1.05	No	Yes (10)	1.105

Estimated minimum alveolar concentration (MAC) from midpoint concentrations for the first 2 crossovers = 0.91; for 4 crossovers = 0.94; for 6 crossovers = 0.95; for 8 crossovers = 0.99; for 10 crossovers = 1.00.

\* Because individual MAC differed  $\leq 2\%$  from the effect site concentration, the response of this patient was assessed as a random event. † Because the previous patient was used in a crossover, this one is not (see text for explanation). ‡ Based on the population value for MAC, this patient's movement response is "misclassified" (*i.e.*, the individual value for MAC, rather than the population value, explained this patient's response).

Population MAC = 1.0%; interindividual variability = 20% of population MAC; starting concentration for this simulation = 1.3% inhaled anesthetic, incremental changes between patients = 10% of the starting concentration (= 0.13% inhaled anesthetic); machine error level = 5%; ventilation/perfusion mismatch = 5%;  $P_{\text{effect site}}$  = partial pressure at the effect site (reported in concentration units).

We recognized that several factors might result in a difference between the concentration reported by the monitoring device and the alveolar concentration, and between the alveolar concentration and the effect site concentration. First, many anesthetic monitors report end-tidal concentrations with no more precision than 0.1%; thus, the concentration reported by the machine is a rounded representation of the actual "internal machine" reading. Second, machines that measure anesthetic concentration in airway gases have measurement errors.‡ Third, end-tidal concentrations differ to a small extent from arterial and effect site concentrations (*e.g.*, ventilation/perfusion [ $\dot{V}/\dot{Q}$ ] mismatch), even at steady state. Assuming "reasonable" values (defined later) for these factors, we simulated the internal machine reading, the end-tidal concentration, and the effect site concentration based on a target end-tidal concentration for each individual.

‡ For example, the accuracy of infrared absorption spectrometers, *e.g.*, Capnomac ULTIMA (Datex Medical Instrumentation Inc., Tewksbury, MA) or Ohmeda 5250 Respiratory Gas Monitor (Ohmeda, Louisville, CO), for measurements within a range of 1–5% of halothane, isoflurane, or sevoflurane is reported in their operating manuals as  $\pm 0.2$  vol%.

Each simulation was performed using the sequence of a MAC study (table 1). For the first simulated patient, we selected a starting value for the target concentration and, using the approach described above, simulated an effect site concentration. For that individual, we also simulated their individual value for MAC ("individual MAC"), selecting this value randomly from a normal distribution with a mean equal to the population MAC and an SD selected for that simulation. We then compared the effect site anesthetic concentration to the individual value for MAC. If the effect site concentration exceeded individual MAC, we assumed that the patient would not respond ("no move"); conversely, if the effect site concentration was less than individual MAC, we assumed that the patient would move. For a patient in whom the difference between effect site concentration and individual MAC was small (*e.g.*, the ratio of these values was between 0.98 and 1.02), we assumed that the observation was ambiguous ("ambiguous response") and assigned that patient randomly the value "move" or "no move."

Once the response of the first patient was determined, a second patient was simulated. If the first patient

“moved,” the target concentration for the second patient was increased by a predetermined increment, (e.g., 10% of the starting concentration value). Conversely, if the first patient did not “move,” the target concentration for the second patient was decreased by the same increment. Target concentration for each subsequent patient was altered in the same manner (*i.e.*, based on the starting concentration). If the movement response of a patient differed from that of the previous patient, those two patients constituted a crossover; the midpoint between the target end-tidal concentrations of these two patients was recorded. However, each patient was permitted to be involved in only a single crossover (see footnote of table 1 for an example). The process of assessing the response of consecutive patients continued until 10 crossovers were attained (note that clinical studies usually terminate after four crossovers).<sup>3,7</sup>

Simulations were performed using the S-PLUS software package (version 5.0, 1998; MathSoft Inc., Seattle, WA) in a Solaris (UNIX) operating environment. For each of the following five issues, simulations were replicated 500 times, and the resulting values were averaged.

#### *Impact of Starting Concentration*

We assumed that the optimal concentration to start a MAC experiment is a value close to the expected result. However, early in clinical development of an inhaled anesthetic (particularly as new populations are being studied<sup>7</sup>), investigators may not be aware of the expected result. We hypothesized that using a starting value that differed markedly from the true value would impact the estimates of MAC minimally but would increase the number of patients required to achieve a predetermined number of crossovers. We also speculated that individuals whose MAC value differed markedly from the population value might produce outlying values for crossover midpoint concentrations, thereby impacting the estimate of MAC in certain instances. We simulated starting concentrations ranging from one half to twice the population MAC value (e.g., assuming that the typical value for MAC is 1%, we simulated starting target concentrations ranging from 0.5 to 2.0%).

#### *Impact of the Number of Crossovers*

Dixon<sup>2</sup> suggested that this type of experiment be continued until at least four crossovers were reached. We hypothesized that continuing a MAC study to a larger number of crossovers may result in more accurate MAC estimates. Therefore, we compared simulations based on 2, 4, 6, 8, or 10 crossovers. This was accomplished by completing each simulation when 10 crossovers were

attained and then analyzing the data for all simulated patients or only through the first 2, 4, 6, or 8 crossovers.

#### *Impact of Increment Size of Concentration Adjustments*

For a clinical MAC experiment, a typical value for the incremental change in target concentration of inhaled anesthetic between two consecutive patients ranges from approximately 10%<sup>3</sup> to approximately 15%<sup>4</sup> of the expected MAC value§ (e.g., 0.10% or 0.15% for an expected MAC of 1%, and 0.6% or 0.9% for an expected MAC of 6%). We hypothesized that a larger increment would decrease the number of patients required to achieve a predetermined number of crossovers but that variability in the estimate of MAC would increase. We simulated increment values as being a fraction of the chosen starting concentration (either 10% or 20%) over a range of starting concentrations from one half to twice the expected MAC value. We further tested whether after achieving the first crossover, decreasing the increment size to half its initial value or one tenth of the first midpoint value would affect the accuracy of MAC estimates and the number of patients needed to achieve accurate results.

#### *Impact of Interindividual Variability*

If interindividual variability is small, then successive individuals differ minimally in their individual MAC, and crossovers should occur in rapid succession once the target concentration is near the true MAC. In turn, the range of crossover midpoint concentration should be small. In contrast, larger interindividual variability should result in a larger range of crossover midpoint concentrations and possibly more outlier (*i.e.*, flawed) estimates for MAC. We simulated interindividual variability ranging from 10 to 30% of the population MAC value.

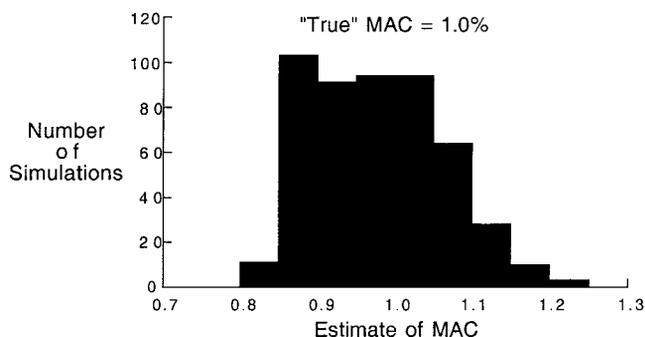
#### *Impact of Measurement Error*

At least three factors (digits displayed by the anesthetic monitor, machine imprecision, and  $\dot{V}/\dot{Q}$  mismatch) create differences between the value reported by the anesthetic monitor and the effect site concentration. We simulated each component of these measurement errors. First, we assumed that precision of reporting end-tidal concentration was either 0.1% (e.g., a reported concentration of 1.1% could represent an internal machine reading of 1.050–1.149%) or 0.01%. Second, we simulated that the effect of  $\dot{V}/\dot{Q}$  mismatch and machine measurement error were each either 5% or 0.1% of scale (the smaller value minimized the influence of these factors on the estimate of MAC).

#### *Assessment of Simulations*

Each simulation continued until 10 crossovers were obtained. The average of the first 2 midpoint values was termed MAC<sub>2</sub>; the averages of the 4, 6, 8, and 10 mid-

§ Note that the investigator is unaware of the true value for MAC in the target population. Therefore, the expected value on which the increment size is based is actually the starting concentration of that experiment.



**Fig. 1.** A sample distribution of minimum alveolar concentration (MAC) values determined in 500 replicate simulated experiments is shown. The range of MAC values in these 500 replicates is 0.63%, the mean is 0.99%, and the SD is 0.084%. Assumptions used in these simulations: population MAC = 1.0%; interindividual variability = 10% of population MAC; incremental changes for inhaled anesthetic = 20% of starting concentration; machine measurement error = 5%; ventilation/perfusion mismatch = 0.1%; starting concentration = 1.3% inhaled anesthetic.

point values were termed MAC<sub>4</sub>, MAC<sub>6</sub>, MAC<sub>8</sub>, and MAC<sub>10</sub>, respectively. SDs (SD<sub>2</sub>, SD<sub>4</sub>, SD<sub>6</sub>, SD<sub>8</sub>, SD<sub>10</sub>) of these crossover values were determined. The number of individuals needed to achieve the desired number of

crossovers (2, 4, 6, 8, and 10) were termed accordingly N<sub>2</sub>, N<sub>4</sub>, N<sub>6</sub>, N<sub>8</sub>, and N<sub>10</sub>, and the number of individuals in whom ambiguous responses were obtained was recorded (Ambig<sub>2</sub>, Ambig<sub>4</sub>, Ambig<sub>6</sub>, Ambig<sub>8</sub>, and Ambig<sub>10</sub>, based on 2–10 crossovers). We further assessed whether MAC<sub>2</sub>, MAC<sub>4</sub>, MAC<sub>6</sub>, MAC<sub>8</sub>, or MAC<sub>10</sub> differed from the true value for MAC by more than 10, 15, 20, or 25%. Each simulation was replicated 500 times, and summary statistics for MAC, SD, Ambig, and N were prepared from these replicates. The maximum (MAX<sub>2</sub>, MAX<sub>4</sub>, MAX<sub>6</sub>, MAX<sub>8</sub>, and MAX<sub>10</sub>, respectively) and minimum (MIN<sub>2</sub>, MIN<sub>4</sub>, MIN<sub>6</sub>, MIN<sub>8</sub>, MIN<sub>10</sub>, respectively) values of MAC with 2, 4, 6, 8, or 10 crossovers were determined from these 500 replicates. Each of the issues described above was examined over a variety of conditions. Only representative examples are reported.

*Comparison of Our Simulations with Published Clinical Studies*

These simulations make various assumptions about the magnitude of interindividual variability and measurement error. If these assumptions are flawed, our simulations would be invalid. Although certain of our assumptions have not been tested, one overall check can be

**Table 2.** Impact of Varying the Starting Concentration on the Resulting Estimates of Population MAC and the Number of Patients Needed to Be Enrolled in a Study

	Starting Concentration (%)								
	0.5	0.7	0.9	1.0	1.1	1.3	1.5	1.7	2.0
MAC <sub>2</sub>	0.958	0.971	0.99	0.998	1.01	1.006	1.014	1.012	1.004
MAC <sub>4</sub>	0.972	0.983	0.995	0.999	1.009	1.005	1.009	1.007	1.003
MAC <sub>6</sub>	0.981	0.99	0.997	0.997	1.007	1.004	1.006	1.005	1.002
MAC <sub>8</sub>	0.986	0.992	0.998	0.998	1.006	1.003	1.004	1.005	1.002
MAC <sub>10</sub>	0.989	0.994	0.998	0.999	1.004	1.003	1.003	1.004	1.003
Max <sub>2</sub>	1.125	1.12	1.17	1.15	1.155	1.17	1.2	1.275	1.2
Max <sub>4</sub>	1.113	1.103	1.125	1.125	1.155	1.138	1.2	1.19	1.15
Max <sub>6</sub>	1.075	1.097	1.11	1.1	1.118	1.105	1.175	1.162	1.133
Max <sub>8</sub>	1.063	1.076	1.103	1.1	1.1	1.105	1.125	1.148	1.15
Max <sub>10</sub>	1.065	1.064	1.107	1.09	1.089	1.092	1.11	1.105	1.1
Min <sub>2</sub>	0.8	0.805	0.81	0.85	0.825	0.78	0.825	0.765	0.8
Min <sub>4</sub>	0.863	0.857	0.878	0.85	0.88	0.878	0.863	0.85	0.85
Min <sub>6</sub>	0.892	0.887	0.9	0.867	0.862	0.867	0.9	0.85	0.867
Min <sub>8</sub>	0.913	0.91	0.9	0.875	0.88	0.877	0.881	0.829	0.875
Min <sub>10</sub>	0.92	0.917	0.918	0.88	0.902	0.897	0.9	0.85	0.9
SD <sub>2</sub>	0.045	0.056	0.057	0.058	0.056	0.069	0.063	0.071	0.085
SD <sub>4</sub>	0.055	0.065	0.07	0.07	0.072	0.082	0.085	0.095	0.103
SD <sub>6</sub>	0.057	0.066	0.074	0.073	0.077	0.087	0.09	0.099	0.109
SD <sub>8</sub>	0.057	0.067	0.074	0.074	0.08	0.087	0.094	0.1	0.109
SD <sub>10</sub>	0.057	0.067	0.075	0.074	0.081	0.088	0.095	0.101	0.109
N <sub>2</sub>	13.4	8.0	5.0	4.5	4.7	6.2	7.6	7.9	8.8
N <sub>4</sub>	18.7	12.9	9.8	8.9	9.3	10.7	11.5	12.2	13.1
N <sub>6</sub>	23.7	17.8	14.5	13.4	13.8	15.2	15.8	16.4	17.2
N <sub>8</sub>	28.6	22.6	19.1	17.9	18.3	19.5	20.2	20.6	21.3
N <sub>10</sub>	33.6	27.4	23.8	22.5	22.8	23.9	24.6	24.9	25.4

Results are average values of 500 replicates for each starting concentration.

Population minimum alveolar concentration (MAC) = 1.0%; interindividual variability = 10% of population MAC; incremental changes for inhaled anesthetic = 10% of starting concentration; machine error level = 0.1%; MAC<sub>2</sub>, MAC<sub>4</sub>, MAC<sub>6</sub>, MAC<sub>8</sub>, MAC<sub>10</sub> = average MAC for 2, 4, 6, 8, 10 crossovers; Max<sub>2</sub>, Max<sub>4</sub>, Max<sub>6</sub>, Max<sub>8</sub>, Max<sub>10</sub> = highest estimated MAC for 2, 4, 6, 8, 10 crossovers; Min<sub>2</sub>, Min<sub>4</sub>, Min<sub>6</sub>, Min<sub>8</sub>, Min<sub>10</sub> = lowest estimated MAC for 2, 4, 6, 8, 10 crossovers; SD<sub>2</sub>, SD<sub>4</sub>, SD<sub>6</sub>, SD<sub>8</sub>, SD<sub>10</sub> = average of standard deviations of 500 replicates for 2, 4, 6, 8, 10 crossovers; N<sub>2</sub>, N<sub>4</sub>, N<sub>6</sub>, N<sub>8</sub>, N<sub>10</sub> = number of patients needed for 2, 4, 6, 8, 10 crossovers.

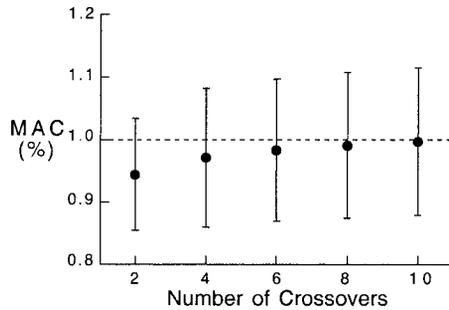


Fig. 2. Values for average minimum alveolar concentration (MAC) and average SD are plotted against the number of crossovers. The dashed line indicates the population MAC used in the simulations. Other assumptions used in these 500 replicates: interindividual variability = 20% of population MAC; starting concentration = 0.5% inhaled anesthetic; incremental changes for inhaled anesthetic = 20% of starting concentration; machine error level = 5%; ventilation/perfusion mismatch = 0.1%. Note that increasing the number of crossovers improves the estimate of MAC and increases the SD toward the true value used in the simulations.

performed. We surveyed the published literature for MAC and found that the anesthetic for which MAC studies were conducted in the largest homogenous group of humans was halothane (97 individuals in five studies).<sup>1,8-11</sup> We used the reported values for end-tidal concentrations and movement response in these individuals in two approaches. First, we assessed how many of these individuals moved at target concentrations greater than the estimated MAC value determined from that study or did not move at concentrations less than that value. Second, we assessed how many of all the patients in these studies moved at end-tidal concentrations greater than 0.75% halothane or did not move at concentrations less than 0.75% halothane. The first approach was also used for the reported data of two studies of sevoflurane MAC.<sup>4,12</sup> If there was no interindividual variability or measurement error, individuals who moved at concentrations greater than MAC or did not move at concentrations less than MAC would be "misclassified." The frequency of these misclassified individuals in the population thereby estimates the sum of interindividual variability and measurement error.

## Results

### Variability in MAC Results

With 500 replicates (assumptions: population MAC = 1.0%; interindividual variability = 10% of population MAC; incremental changes for inhaled anesthetic = 20% of starting concentration; machine measurement error = 5%;  $\dot{V}/\dot{Q}$  mismatch = 0.1%; starting concentration = 1.3% inhaled anesthetic), 23% of replicates yielded MAC estimates more than 10% different from the true value (fig. 1); 2.4% of replicates deviated more than 15% from the true value. The range of MAC values was 0.45%.

With different initial conditions (e.g., incremental changes for inhaled anesthetic = 0.2%, and interindividual variability = 30% of population MAC), MAC values varied from 0.65 to 1.35%. The magnitude of this range depended on the number of crossovers and the starting concentration (data not shown). With interindividual variability of MAC of 10% and minimal error, the average SD increased toward the true value (10%) with either a larger number of crossovers or larger starting concentrations (table 2). The average SD was typically 6-9% with either four or six crossovers and starting concentrations close to the true MAC. With a simulated interindividual variability of MAC of 20%, the average SD was approximately 11% with either four or six crossovers (table A, Web Enhancement; fig. 2).

### Impact of Starting Concentration

The average MAC estimate in 500 replicates was less accurate with small starting concentrations (table 2). Typically, starting concentrations larger than population MAC yielded average values for MAC slightly greater than

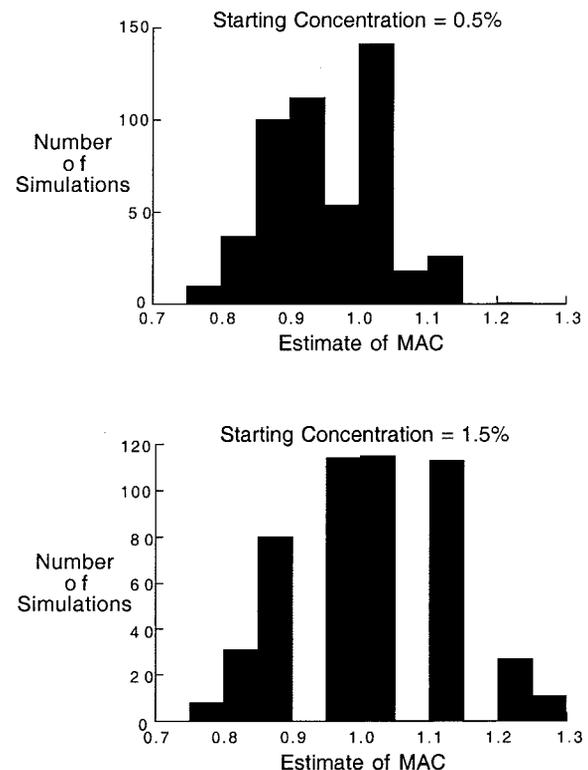


Fig. 3. These histograms display the minimum alveolar concentration (MAC) values obtained in 500 replicates. The increment size between concentration changes was 20% of the starting concentration. In the upper panel, the starting concentration for each replicate was 0.5% (with an increment size of 0.1%); in the lower panel, the starting concentration for each replicate was 1.5% (with an increment size of 0.3%). The lower starting concentration produced a slight bias toward smaller estimates for MAC (60% of values were < 1.0%). Because increment size is a function of the starting concentration, increment size for the lower panel is larger, presumably explaining gaps in the distribution in that panel.

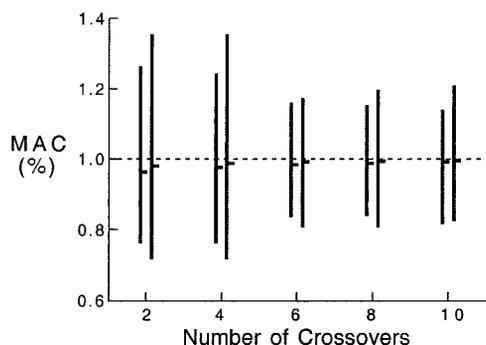


Fig. 4. Values for average minimum alveolar concentration (MAC) and the largest and smallest values obtained in 500 replicates are plotted against the number of crossovers. For each crossover, the left-hand line is based on an increment size of 10%, and the right-hand line is based on an increment size of 20%. The dashed line indicates the population MAC used in the simulations. Other assumptions used in these 500 replicates: interindividual variability = 20% of population MAC; machine error level = 5%; ventilation/perfusion mismatch = 0.01%. Note that increasing the increment size slightly improved the average value for MAC but increased the range.

population MAC, whereas starting concentrations less than population MAC yielded values for MAC less than population MAC (fig. 3 and table 2); this was more pronounced for the smallest starting concentrations. As the starting concentration approached population MAC, the number of individuals needed to attain a desired number of crossovers decreased. With an increment size of one tenth of the starting concentration between consecutive patients, the magnitude of decrease was approximately two individuals per 0.2% change in the starting concentration (based on a true MAC value of 1.0%).

#### Impact of Number of Crossovers

Increasing the number of crossovers from two to six improved MAC estimates (figs. 2, 4, and 5), whereas a further increase in the number of crossovers to 10 affected the average MAC and SD of MAC estimates less (table 2; table A, Web Enhancement). The range of MAC values was smaller with six, compared with four, crossovers (*i.e.*, fewer extreme [outlier] values). Increasing the number of crossovers from 2 to 10 required 2–3 more patients for each additional crossover under many different simulated conditions, *e.g.*, different starting concentrations (table 2).

#### Impact of Increment Size

Increasing the increment size from one tenth to two tenths of the starting concentration slightly improved the average value of MAC estimates but increased the range of MAC estimates; variability (SD) increased and was closer to the true value (table B, Web Enhancement; fig. 4). The number of individuals needed to attain the target crossovers also decreased. The average number of ambiguous decisions was typically small. Decreasing the increment size to half its initial value or one tenth of the

first midpoint value after the first crossover was achieved did not improve MAC estimates (data not shown).

#### Influence of Interindividual Variability

As interindividual variability increased from 10 to 30% of the population MAC, the range for MAC estimates (fig. 5) and the number of ambiguous decisions decreased (table C, Web Enhancement). With interindividual variability of MAC of 10% and a starting concentration of 1.3% (table 3), 4% of 500 replicates yielded MAC estimates that differed more than 10% from the true value; 1% of the 500 replicates yielded a MAC estimate that differed more than 15% from the true value. With interindividual variability of MAC of 30%, approximately one third of replicates yielded MAC estimates more than 10% different from the true value. With a starting concentration of 1.3%, one fifth of replicates differed more than 15% from the true value, one tenth differed more than 20%, and one twentieth differed from the true value more than 25%.

#### Effect of Measurement Error and $\dot{V}/\dot{Q}$ Mismatch

With measurement error and  $\dot{V}/\dot{Q}$  mismatch of 5% each *versus* 0.1% each, the percentage of MAC estimates more than 10, 15, 20, or 25% different from the true value typically doubled (table 4). A starting concentration that deviated more from the true MAC value resulted in more outliers.

#### Comparing Individual Patient Response with the Population MAC

**Published Data for Halothane.** Estimates for halothane MAC in humans ranged from 0.73 to 0.765%. Of the 97 patients, 21 (22%) were misclassified based on

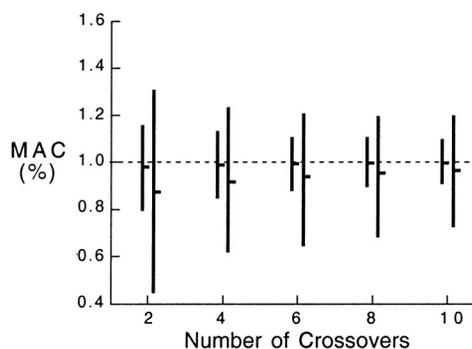


Fig. 5. Values for average minimum alveolar concentration (MAC) and the largest and smallest values obtained in 500 replicates are plotted against the number of crossovers. For each crossover, the left-hand line is based on interindividual variability of 10%, and the right-hand line is based on interindividual variability of 30%. The dashed line indicates the population MAC used in the simulations. Other assumptions used in these 500 replicates: incremental changes for inhaled anesthetic = 20% of starting concentration; machine error level = 5%; ventilation/perfusion mismatch = 0.1%. Note that the larger interindividual variability resulted in a larger range of MAC values.

**Table 3. Impact of Variability of Individual MAC Values on the Accuracy of MAC Estimates (Based on Six Crossovers)**

Fraction of estimates not within this percentage of population MAC	Interindividual Variability			
	0.1%		0.3%	
	Starting Concentration		Starting Concentration	
	0.9%	1.3%	0.9%	1.3%
10%	0.04	0.04	0.26	0.32
15%	0.00	0.01	0.12	0.18
20%	0.00	0.00	0.05	0.09
25%	0.00	0.00	0.01	0.04

Results are average values of 500 replicates; values are shown for two representative starting concentrations.

Population minimum alveolar concentration (MAC) = 1.0%; incremental changes for inhaled anesthetic = 10% of starting concentration; machine error level = 5%; ventilation/perfusion mismatch = 0.1%.

their movement response, their reported target concentration, and the MAC value reported in that manuscript. Applying a single cutoff value for MAC of 0.75% for all five studies, 17% of patients were misclassified.

**Published Data for Sevoflurane.** Five (25%) of 20 patients and five (22%) of 23 patients in two different studies were misclassified by these criteria.

**Our Simulations.** Assuming a population MAC of 1.0, incremental changes for inhaled anesthetic of 0.1%, machine error level of 5%,  $\dot{V}/\dot{Q}$  mismatch of 5%, and a starting concentration of 1.3% inhaled anesthetic, the percentage of patients misclassified was 16% with interindividual variability of 5% of population MAC, 17% with interindividual variability of 10% (table 1), 34% with interindividual variability of 20%, and 36% with interindividual variability of 30%.

## Discussion

To determine the reliability of estimates for MAC in humans, we simulated a typical MAC experiment and

**Table 4. Impact of Magnitude of Error (0.1% vs. 5%) on the Accuracy of MAC Estimates (Based on Six Crossovers)**

Fraction of estimates not within this percentage of population MAC	Magnitude of Error			
	0.1%		5%	
	Starting Concentration		Starting Concentration	
	0.9%	1.3%	0.9%	1.3%
10%	0.14	0.17	0.26	0.29
15%	0.03	0.06	0.08	0.14
20%	0.00	0.02	0.01	0.05
25%	0.00	0.00	0.00	0.02

Results are average values of 500 replicates; values are shown for two representative starting concentrations.

Population minimum alveolar concentration (MAC) = 1.0%; interindividual variability = 20% of population MAC; incremental changes for inhaled anesthetic = 0.2%.

explored a number of issues that may affect these estimates. We also examined the impact of these issues on the number of patients that need to be enrolled to complete an experiment. We found that certain issues impacted the results of the study to a large degree, whereas other factors had less impact.

First, we learned that under many conditions, one or more of 500 replicates yielded estimates for MAC that deviated more than 10% or even more than 25% from the true value. In that an investigator conducts only a single experiment, this suggests that random occurrence may permit inaccurate estimates of MAC. If this occurred in the first experiment with a particular anesthetic in a particular population, there would be no previously published values to suggest an error, and the results might be accepted without criticism. However, if this occurred as the second or later experiment with that anesthetic in that same population, we speculate that the results might be assumed to be flawed (and never submitted for publication), even though the previous, rather than the new, results are actually flawed. Alternatively, the results of both studies might be published, leading the reader to conflicting conclusions as to the true value. For example, Inomata *et al.*<sup>13</sup> reported that sevoflurane MAC in children aged 1–9 yr was 2.03%. In contrast, Lerman *et al.*<sup>14</sup> reported that sevoflurane MAC in children aged 1–12 yr was approximately 2.5%. In that Lerman *et al.*<sup>14</sup> obtained similar estimates of MAC for subgroups aged 1–3, 3–5, and 5–12 yr, it is unlikely that the difference in MAC between studies is a function of the slightly different age groups. One explanation for these different values for sevoflurane MAC is that the populations differ (Japanese *vs.* Americans). However, we propose that the populations may not differ; instead, experimental errors and interindividual variability may result in these findings. Marked differences in MAC were also reported for sevoflurane in healthy adults: values were 1.58,<sup>15</sup> 1.71,<sup>12</sup> and 2.05%.<sup>4</sup> Similarly, our findings may explain why some investigators demonstrate an effect of adjuvant drugs on anesthetic potency (*e.g.*, Forbes *et al.*<sup>9</sup> reported that pancuronium decreased halothane MAC), whereas other investigators do not (*e.g.*, Fahey *et al.*<sup>16</sup> found no effect of pancuronium on halothane MAC).

Second, we learned that the SD determined from the midpoint concentrations underestimates true interindividual variability. A similar observation was suggested by de Jong and Eger<sup>17</sup> in 1975; however, they presented no data to support this claim. The magnitude of the underestimate decreases as the number of crossovers increases, as the starting concentration differs markedly from the true value, and as the increment size increases (compare table C [Web Enhancement] with table 2). We offer the following explanation for this phenomenon. Consider a patient with an individual MAC value (1.45%) that differs markedly from the population MAC (1.0%). If the experiment starts at a high target concentration (*e.g.*,

1.5%), successive iterations of other patients (in whom individual MAC is close to 1.0%) results in successive “no move” responses. If this patient is tested at a target concentration of 1.4%, the patient will move, contributing a crossover value that differs markedly from the population MAC. In contrast, if the starting concentration for the experiment is close to the population value, successive responses will toggle around the population value. When this “aberrant” patient is tested with a target concentration of 1.0%, the “move” response will provide no insight as to whether this patient’s individual MAC value is 1.1% or 1.45% (the actual value). Therefore, a strategy in which MAC is approached from extreme values provides additional information about variability within the population. Note that a design in which the increment size is coupled to the starting concentration results in a larger increment with a larger starting concentration. In that a larger increment size and a starting concentration that differs from the typical value both improve the estimate of variability of MAC, it seems that the optimal strategy to quantify variability in MAC is to start with an initial concentration that exceeds the expected value.

Third, we learned that on average, four crossovers yield nearly the same average MAC as a larger number of crossovers. However, with four crossovers, there are more outlier estimates for MAC than with six or more crossovers (table A, Web Enhancement). In that the investigator can never be sure that the result of a particular experiment is not an outlier, this suggests that the larger number of crossovers might be useful to prevent spurious results. Terminating a study after six or more crossovers requires additional resources; however, our simulations suggest that only 2.0–2.5 additional patients would be needed to attain each additional crossover. The benefit of increasing the number of crossovers beyond six seems to be small.

We also learned that the starting concentration may affect the results of the experiment. First, if the starting concentration deviates markedly from the population MAC, a larger sample size is needed to attain the desired number of crossovers. Second, a starting concentration markedly different from the true MAC value yields slightly biased estimates for MAC and a larger incidence of biased individual estimates.

Finally, a larger increment size of concentration changes between consecutive individuals decreased the number of individuals needed to attain the target crossovers and increased variability of MAC estimates toward the true value.

Several issues of our study design warrant comment. First is our assumption that the response in a MAC experiment can be ambiguous. Antognini *et al.*<sup>6</sup> assessed the movement response in rats repeatedly at different end-tidal anesthetic concentrations. Some animals moved at end-tidal concentration at 110% of their indi-

vidual MAC or did not move at concentrations at 90% of their individual MAC. These findings suggest that there is intraindividual variability in MAC. We assumed that an ambiguous response occurred over only a small range (98–102% of the individual MAC). However, the findings of Antognini *et al.*<sup>6</sup> suggest that the range might be larger. If the range is larger, the opportunity for inaccurate MAC values increases even further than we simulated.

The second issue of study design regards our many assumptions, the validity of several of which can be challenged. If our simulations introduced markedly more interindividual variability or error than exists in nature, then a markedly larger number of patients would be misclassified. For example, interindividual variability may be markedly smaller than the values we tested. If this were true and if the various components of measurement error were trivial, results from MAC studies should show a sharp delineation between movers and nonmovers, *i.e.*, there should be a single end-tidal concentration above which no patient moves and below which all patients move. The compiled results for halothane indicate that a single cutoff value of 0.75% misclassifies approximately 17% of patients; using the MAC estimate from each study results in 22% of patients being misclassified. Similarly, in two sevoflurane MAC studies, 22–25% of patients were misclassified based on single cutoff values.<sup>4,12</sup> Allowing for 5% measurement error, 5%  $\dot{V}/\dot{Q}$  mismatch, a 0.1% precision in machine measurement (*i.e.*, measurements reported to a tenth-percent anesthetic concentration), and 10% interindividual variability yielded a percentage of misclassified patients (17%) similar to published values. This suggests that the combined error and variability used in our simulations is no larger in magnitude than that which exists in nature.

Our simulations assume that each individual’s response to a stimulus is assessed only once. This is the approach typically used in MAC studies in humans. In contrast, a small number of studies in humans and many in animals bracket MAC for each subject. Different statistical issues may apply to studies performed using that design.

These simulations assumed a typical value of MAC of 1%. Although this value is close to that of halothane MAC, our results apply to all anesthetics, regardless of their typical value of MAC. We note that one component of our simulations, measurement error (quantified as a percentage of the true value) may be smaller for anesthetics, such as desflurane, with larger MAC values. The simulations also assumed normal or log-normal distributions. It is possible that certain distributions (*e.g.*, interindividual variability in MAC) may have nonnormal distributions. However, in the absence of information suggesting such distributions, we are unable to evaluate the possible implications of such distributions.

We replicated each simulation 500 times. Initial simulations (not reported) were performed with only 100 replicates. When these simulations were repeated, results sometimes varied, an expected finding with this small sample size. With the larger sample size (500 replicates), repeated simulations yielded similar results. However, to verify that 500 replicates were sufficient to minimize error, we replicated one set of simulations 25,000 times. The results of that markedly larger simulation were quite similar to those of the simulations with 500 replicates. Thus, we report analyses performed with 500 replicates.

Finally, although most studies of anesthetic potency focus on the estimate of MAC rather than on the estimate of interindividual variability in MAC, we focus on both. Clinicians presumably dose anesthetics so that few, if any, of their patients move in response to surgical incision. This is accomplished by dosing to prevent movement in 95% of patients ( $ED_{95}$ ) or 99% of patients rather than administering a dose for which 50% of patients will move ( $ED_{50}$ ). Determining  $ED_{95}$  requires knowledge of both  $ED_{50}$  and its variability.

In summary, we simulated the results of a MAC experiment in humans using an up-down approach. Allowing for reasonable estimates of interindividual variability and error, the value for MAC was estimated accurately, on average, with 500 replicates. However, certain estimates of MAC in individual simulations differed by 25% or more from the true value. In that investigators typically perform MAC studies only once, our findings suggest that the up-down approach for analysis of quantal data may yield incorrect estimates. Our findings suggest that aiming for six or more crossovers, rather than four, decreases the likelihood of reporting an inaccurate estimate and incurs minimal additional costs.

The authors thank Edmond I. Eger II, M.D. (Professor of Anesthesia, Department of Anesthesia, University of California San Francisco, San Francisco, CA), for valuable criticism of their study design.

## References

1. Saidman LJ, Eger EI II: Effect of nitrous oxide and of narcotic premedication on the alveolar concentration of halothane required for anesthesia. *ANESTHESIOLOGY* 1964; 25:302-6
2. Dixon WJ: The up-and-down method for small samples. *J Am Statist Assoc* 1965; 60:967-78
3. Rampil IJ, Lockhart SH, Zwass MS, Peterson N, Yasuda N, Eger EI II, Weiskopf RB, Damask MC: Clinical characteristics of desflurane in surgical patients: minimum alveolar concentration. *ANESTHESIOLOGY* 1991; 74:429-33
4. Scheller MS, Saidman LJ, Partridge BL: MAC of sevoflurane in humans and the New Zealand white rabbit. *Can J Anaesth* 1988; 35:153-6
5. Rampil IJ, King BS: Volatile anesthetics depress spinal motor neurons. *ANESTHESIOLOGY* 1996; 85:129-34
6. Antognini JF, Wang XW, Carstens E: Quantitative and qualitative effects of isoflurane on movement occurring after noxious stimulation. *ANESTHESIOLOGY* 1999; 91:1064-71
7. Fisher DM, Zwass MS: MAC of desflurane in 60% nitrous oxide in infants and children. *ANESTHESIOLOGY* 1992; 76:354-6
8. Bridges BE Jr, Eger EI II: The effect of hypocapnia on the level of halothane anesthesia in man. *ANESTHESIOLOGY* 1966; 27:634-7
9. Forbes AR, Cohen NH, Eger EI II: Pancuronium reduces halothane requirement in man. *Anesth Analg* 1979; 58:497-9
10. Miller RD, Wahrenbrock EA, Schroeder CF, Knipstein TW, Eger EI II, Buechel DR: Ethylene-halothane anesthesia: Addition or synergism? *ANESTHESIOLOGY* 1969; 31:301-4
11. Saidman LJ, Eger EI II, Munson ES, Babad AA, Muallem M: Minimum alveolar concentrations of methoxyflurane, halothane, ether and cyclopropane in man: Correlation with theories of anesthesia. *ANESTHESIOLOGY* 1967; 28:994-1002
12. Katoh T, Ikeda K: The minimum alveolar concentration (MAC) of sevoflurane in humans. *ANESTHESIOLOGY* 1987; 66:301-3
13. Inomata S, Watanabe S, Taguchi M, Okada M: End-tidal sevoflurane concentration for tracheal intubation and minimum alveolar concentration in pediatric patients. *ANESTHESIOLOGY* 1994; 80:93-6
14. Lerman J, Sikich N, Kleinman S, Yentis S: The pharmacology of sevoflurane in infants and children. *ANESTHESIOLOGY* 1994; 80:814-24
15. Kimura T, Watanabe S, Asakura N, Inomata S, Okada M, Taguchi M: Determination of end-tidal sevoflurane concentration for tracheal intubation and minimum alveolar anesthetic concentration in adults. *Anesth Analg* 1994; 79:378-81
16. Fahey MR, Sessler DI, Cannon JE, Brady K, Stoen R, Miller RD: Atracurium, vecuronium, and pancuronium do not alter the minimum alveolar concentration of halothane in humans. *ANESTHESIOLOGY* 1989; 71:53-6
17. de Jong RH, Eger EI II: MAC expanded:  $AD_{50}$  and  $AD_{95}$  values of common inhalation anesthetics in man. *ANESTHESIOLOGY* 1975; 42:384-9