Neuromuscular dose–response studies: determining sample size

A. F. Kopman1*, C. A. Lien1 and M. Naguib2

1 Department of Anesthesiology, Weill Cornell Medical College, 525 East 68th Street, New York City, NY 10065, USA
2 Department of Anesthesiology, Cleveland Clinic, 9500 Euclid Ave., Cleveland, OH 44195, USA

* Corresponding author. E-mail: akopman@nyc.rr.com

In the investigation of any new neuromuscular blocking agent, the drug potency (ED50, the dose required to produce 50% twitch depression) should be estimated early in the development process. Without this information, an estimate of its clinical margin of safety (the ratio of the ED50 to the ED50 dose which produces unacceptable cardiovascular or other adverse side-effects) cannot be ascertained. However, because there is considerable patient-to-patient variability in the sensitivity to relaxants,1 2 the determination of potency with small samples may give misleading results. The development of any new drug is an expensive undertaking, and the recruitment of experimental subjects is time-consuming and costly. When obtaining data, a balance must be struck between scientific rigour and unnecessary expenditure. Surprisingly, there is little consensus in the neuromuscular literature on how large a sample size needs to be before these conflicting demands are both reasonably satisfied. The consensus paper ‘Good Clinical Research Practice in Pharmacodynamic Studies of Neuromuscular Blocking Agents’ (GCRP) which represents the most current thinking for clinical investigators in this area is essentially silent on this subject.3 The comment ‘An adequate sample size should be determined before carrying out the study to prevent any waste of time and resources on an inconclusive investigation’ although obviously correct is not particularly helpful.

When planning a protocol for a new non-depolarizing agent, a pre-requisite for approval by most Institutional Review Boards (IRBs) is an estimate of required sample size as determined by power analysis. In our review of the neuromuscular literature, dose–response studies that include an a priori power analysis are few and far between and the methodology used is usually far from transparent.4 Thus, we had no theoretical guidance as to how to proceed or defend any given sample size we might select. We suspect that the failure of most investigators to use power analysis is because the calculations require some knowledge of the variability around the mean, that is, the standard deviation (SD) of the sample or the coefficient of variation (COV)

Editor’s key points

- Sample size can be difficult to determine in neuromuscular block studies.
- It requires some measure of variability around the mean.
- Retrospective analysis of data was used to calculate a coefficient of variation.
- This gives a starting point for calculating the sample size.
- Size will depend on the allowable error in the estimated ED50: 20%, 15 subjects, and 10%, 50 subjects.

Background. Investigators planning dose–response studies of neuromuscular blockers have rarely used a priori power analysis to determine the minimal sample size their protocols require. Institutional Review Boards and peer-reviewed journals now generally ask for this information. This study outlines a proposed method for meeting these requirements.

Methods. The slopes of the dose–response relationships of eight neuromuscular blocking agents were determined using regression analysis. These values were substituted for γ in the Hill equation. When this is done, the coefficient of variation (COV) around the mean value of the ED50 for each drug is easily calculated. Using these values, we performed an a priori one-sample two-tailed t-test of the means to determine the required sample size when the allowable error in the ED50 was varied from ± 10–20%.

Results. The COV averaged 22% (range 15–27%). We used a COV value of 25% in determining the sample size. If the allowable error in finding the mean ED50 is ± 15%, a sample size of 24 is needed to achieve a power of 80%. Increasing ‘accuracy’ beyond this point requires increasing greater sample sizes (e.g. an ‘n’ of 37 for a ± 12% error).

Conclusions. On the basis of the results of this retrospective analysis, a total sample size of not less than 24 subjects should be adequate for determining a neuromuscular blocking drug’s clinical potency with a reasonable degree of assurance.

Keywords: neuromuscular block; pharmacology, dose response; potency, drug; potency, ED50; power analysis

Accepted for publication: 26 August 2010
often reported, confidence limits are a measure of the reliability of the mean value and not of individual data dispersion. There are very few published data describing the $sd$ or COV for the ED$_{50}$ of commonly used neuromuscular blocking agents.

The method most commonly used to determine neuromuscular potency (least-squares regression analysis) does not lend itself to providing this information. However, once the best-fit slope of the dose–response relationship is determined by regression analysis, this value can be substituted for $\gamma$ in the Hill equation.\textsuperscript{5, 6} If this is done, the ED$_{50}$ for each individual subject can be estimated. Consequently, calculating the average ED$_{50}$ and its $sd$ and COV is an easy next step. With this information, estimating neuromuscular dose–response sample size based on conventional power analysis becomes a relatively simple procedure.

When reviewing data from a recent publication\textsuperscript{7} in which we re-examined the dose–response relationships of succinylcholine, rocuronium, rapacuronium, and cisatracurium, we noticed that the COV of these compounds appeared to be relatively constant from agent to agent. We hypothesized that if this is so, then sample size recommendations for these agents might be generally applicable to other neuromuscular blocking drugs as well. Therefore, we examined data from our files for four additional agents (pipecuronium, pancuronium, atracurium, and vecuronium) to extend our observations on the individual variability in the response to neuromuscular blocking agents. We believe that this information will be of use to investigators in the design of future dose–response studies. In addition, our methodology provides a power analysis-based rationale for determining the sample size when performing dose–response studies of neuromuscular blocking compounds.

**Methods**

The present protocol was approved by the Human Subject Review Committee at Weill Cornell Medical College. All patients consented to participate in the earlier IRB-approved studies from which the data were obtained. The methodology used in obtaining our dose–response data has been previously described in detail.\textsuperscript{2, 3, 8} Briefly, 48 (rapacuronium),\textsuperscript{3} 50 (succinylcholine),\textsuperscript{2} 41 (rocuronium),\textsuperscript{2} 45 (atracurium), and 45 (vecuronium) ASA physical status I–II, adult patients undergoing elective surgical procedures were included in these studies. The atracurium and vecuronium data have not previously been reported. Anaesthesia was induced with alfentanil 40 $\mu$g kg$^{-1}$ and propofol 2.0–2.5 mg kg$^{-1}$ i.v. Tracheal intubation was accomplished without the use of neuromuscular blocking agents. Anaesthesia was maintained with nitrous oxide (65–70% inspired), a propofol infusion, and opioid supplements as needed. Ventilation was controlled, and end-tidal carbon dioxide was maintained at 4.5–5.3 kPa.

The indirectly evoked integrated compound action potential of the first dorsal interosseous muscle to supramaximal stimulation of the ulnar nerve at the wrist was measured and recorded using an NMT 221 monitor (Datex, Tewksbury, MA, USA). Single stimuli at 0.10 Hz were administered during the period of observation, and twitch depression was continuously recorded. Control twitch height was established after a 15–20 min period of baseline stabilization. Immediately after baseline calibration, a single dose of relaxant was administered.

The first subject in each group received a bolus estimated from published peer-reviewed data to approximate an ED$_{50}$. Using the Hill equation with a postulated slope of 4.50–4.75, the subject’s individual ED$_{50}$ was then estimated. The second subject received a dose which equalled the calculated ED$_{50}$ for Patient 1. Patients 3–5 were given a dose that approximated the average estimated value of the ED$_{50}$s of the subjects which preceded them. In a similar manner, Patients 6–10 were administered a dose based on the results of the earlier subjects calculated to achieve 20% twitch depression. The next five subjects received doses which approximated the average estimated value of the ED$_{50}$. In the remaining patients, doses of each blocking drug were selected to provide essentially equally distributed increments in the range estimated to span responses from 15% to 95% twitch depression. We observed 100% block in two patients in the succinylcholine and rapacuronium groups. These data points were recorded as an effect of 99.5% when entered into the Hill equation. We did not encounter any 0% responses. In each group, once all the data points were collected, the slope of the best-fit line of regression was used to recalculate the ED$_{50}$ for each subject and finally the mean ED$_{50}$ and its $sd$ using the Hill equation.

Fifty ASA I adult patients undergoing elective surgery received cisatracurium.\textsuperscript{9} Anaesthesia was induced with midazolam, propofol, fentanyl, and 70% nitrous oxide in oxygen and was maintained with a continuous infusion of propofol and nitrous oxide (70% inspired) with incremental doses of fentanyl as needed. The trachea was intubated after induction of anaesthesia without the aid of a neuromuscular blocking agent. Ventilation was adjusted to maintain end-tidal PC$_{O_{2}}$ at 4.3–5.3 kPa.

The ulnar nerve was stimulated with supramaximal train-of-four (TOF) stimuli every 12 s using a Myotest peripheral nerve stimulator (Biometer International, Odense, Denmark). The contraction of the adductor pollicis muscle was recorded using a force displacement transducer and a neuromuscular function analyzer (Myograph 2000, Biometer International). The first response of the TOF was considered the twitch height. Cisatracurium was administered after stabilization of the response to indirect muscle stimulation. Each patient received a single bolus of cisatracurium. Ten patients each received a pre-determined dose of 20, 25, 30, 40, or 50 $\mu$g kg$^{-1}$. Six of the individuals who received 50 $\mu$g kg$^{-1}$ developed 100% blocks.

Forty subjects received pancuronium and 32 pipecuronium using a methodology similar to that outlined above for cisatracurium. Two patients in the pipecuronium group and four in the pancuronium group had 100% twitch depression from an administered dose of relaxant.
All 100% responses were plotted as a 99.5% response when entered into the Hill equation. The mean ED₅₀ and coefficients of variation were calculated as described above for succinylcholine, rocuronium, rapacuronium, atracurium, and vecuronium.

**Statistical analysis**

**Non-linear regression analysis**

A best-fit non-linear analysis was performed on the data for each of the eight neuromuscular blocking agents using GraphPad Prism version 5.02 for Windows (GraphPad Software, San Diego, CA, USA). The function used to calculate the ED₅₀ was 'log dose vs response—variable slope (four parameters)'. This function, a modification of the Hill equation, is based on the following equation: $Y = \frac{\text{Bottom} + (\text{Top} - \text{Bottom})/(1 + 10(\log\text{ED}_{50} - X) \times \text{Hillslope})}{1}$. We set the upper constraint (top) as 100% and the bottom as zero. Hillslope is synonymous with the Hill coefficient or $g$. The software program also provided the confidence limits for the ED₅₀ and $g$. Complete twitch depression was entered as 100% effect.

**Hill equation**

We calculated the ED₅₀ values ($\text{ED}_{50}$) for all eight drugs by this method using the best-fit values for slope ($\gamma$) determined using non-linear regression. The Hill equation allows the determination of an individual subject's ED₅₀ and the ED₉₅ if three pieces of information are provided—the dose of drug administered, the effect produced by that dose, and the slope ($\gamma$) of the dose–response relationship for that drug. The various forms of this equation may be found in the appendix of Kopman and colleagues.⁵

**Sample size analysis**

If the data from our observations can be generalized to other neuromuscular blocking agents, the ED₅₀ has a COV of 0.15–0.27. If this is so, then an a priori one-sample two-tailed t-test of the means can be used to predict the required sample size. To compute sample size, one must supply a hypothetical ‘true’ ED₅₀ ($H_0$), the degree to which variation ($H_1$) from $H_0$ is allowed, the $\text{sd}$ of the ED₅₀, and values for $\alpha$ and power ($1 - \beta$). Using G*Power3 software, with $\alpha=0.05$ and power equal to 80%, we calculated sample size requirements for a hypothetical drug with an ED₅₀ of 1.0 mg kg⁻¹ at various values for $H_1$ when its $\text{sd}$ was 0.22, 0.25, or 0.28 mg kg⁻¹.

**Results**

As six of the individuals who received cisatracurium 50 μg kg⁻¹ developed 100% block, we also performed non-linear regression analysis with this dose omitted ($n=40$). Calculated COVs for the ED₅₀ ($n=40$ vs $n=50$) did not differ by more than 0.5%.

The ED₅₀ values calculated by non-linear regression or using the Hill equation are essentially interchangeable (Table 1). Both methods produce values which vary on average by less than 5%. The median value for $\gamma$ was 4.15 (range 3.6–6.1) (Table 2), and the median value for the COV was 24.1% (range 15.3–27.1%). A relatively small increase in the COV can result in large increases in the required sample size (Fig. 1 and Table 1).

A major factor in determining the sample size is the magnitude of the allowable error in the estimated ED₅₀ which the investigator is willing to accept. Assuming a COV of 25%, only 15 subjects need to be recruited if an error of ±20% is deemed good enough. For an allowable error of ±15%, a sample size of 24 is required. This value rapidly escalates to 50 subjects if an error of only ±10% is sought (Table 3 and Fig. 2).

**Discussion**

Before power analysis can be implemented, an investigator must have some idea of the variability of the data in his test sample. Most of the information on the degree to which subjects differ in their sensitivity to neuromuscular blocking drugs comes from techniques other than single-dose protocols. When a cumulative dose technique is used to determine drug potency, estimates of patient variability are available since the ED₅₀ for each subject can be determined. Thus, in a comparison of three different population

<table>
<thead>
<tr>
<th>Table 1</th>
<th>ED₅₀ values as determined by non-linear regression analysis (NLR) or by the Hill equation. *The Hill equation confidence limits estimated from the standard error of the mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ED₅₀ (mg kg⁻¹) (95% confidence limits) or (sd)</td>
</tr>
<tr>
<td></td>
<td>Non-linear regression</td>
</tr>
<tr>
<td>Atracurium ($n=45$)</td>
<td>0.130 (0.122–0.139)</td>
</tr>
<tr>
<td>Cisatracurium ($n=50$)</td>
<td>0.0260 (0.0246–0.0275)</td>
</tr>
<tr>
<td>Pancuronium ($n=40$)</td>
<td>0.0329 (0.0314–0.0345)</td>
</tr>
<tr>
<td>Pipecuronium ($n=32$)</td>
<td>0.0268 (0.0258–0.0278)</td>
</tr>
<tr>
<td>Rapacuronium ($n=48$)</td>
<td>0.375 (0.358–0.398)</td>
</tr>
<tr>
<td>Rocuronium ($n=41$)</td>
<td>0.170 (0.158–0.184)</td>
</tr>
<tr>
<td>Succinylcholine ($n=50$)</td>
<td>0.143 (0.135–0.152)</td>
</tr>
<tr>
<td>Vecuronium ($n=45$)</td>
<td>0.0255 (0.0244–0.0267)</td>
</tr>
</tbody>
</table>
groups, a COV for the ED50 of rocuronium of between 26% and 32% was reported.2 Each of their subjects received six incremental doses. A study of 11 subjects receiving vecuronium11 (five doses per subject) estimated its ED50 to be 30.4 (SD 8.8) mg kg⁻¹ (COV = 27%), and a study of 50 individuals estimated the ED50 of atracurium to be 160 (41) mg kg⁻¹ (COV = 25.6%).12 Patients in these three studies 21,1,12 received multiple doses, making comparisons with the single-dose technique we used problematic. In addition, the first two studies cited above looked at only a small number of subjects. When drug potency is measured using a single-dose protocol, estimates of the variability of the ED50 and ED95 are usually not given, and when reported they are hard to interpret. A single-dose method reported an ED90 of 0.30 (0.062) mg kg⁻¹ (COV = 21%) for rocuronium,13 but the method of calculating the value of the SD is not explained.

In contrast to regression methods of determining neuromuscular potency, when using the Hill equation, the calculation of the SD or COV of the ED50 is easily obtained. Although the estimated mean value of ED50 is not greatly altered by significant changes in the γ used, the same is not true for the effect of the slope on the parameter’s COV and thus on the sample size. Any increase in the COV results in larger required sample sizes as determined by power analysis. There are at least theoretical reasons to suggest that traditional linear regression analysis is a flawed approach to determining the neuromuscular dose–response relationship.7 Probit and logit analyses are meant to handle only quantal responses with binomial error distributions, not continuous data such as the percent of maximal response. In addition, 0% and 100% responses are undefined when using logit or probit data transformation. We therefore based our recommendations for the sample size on the COVs which were derived from the best-fit slopes as calculated by non-linear regression.

In our calculations of sample size, we chose a COV of 25% as this approaches the maximum value for COV we observed using non-linear regression analysis. This value produces a more conservative (larger) estimate of the

### Table 2 Slopes of the best-fit lines of regression

<table>
<thead>
<tr>
<th>Drug</th>
<th>Slope (non-linear regression)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atracurium</td>
<td>4.09</td>
</tr>
<tr>
<td>Cisatracurium</td>
<td>5.48</td>
</tr>
<tr>
<td>Pancuronium</td>
<td>4.21</td>
</tr>
<tr>
<td>Pipecuronium</td>
<td>4.53</td>
</tr>
<tr>
<td>Rapacuronium</td>
<td>3.58</td>
</tr>
<tr>
<td>Rocuronium</td>
<td>3.84</td>
</tr>
<tr>
<td>Succinylcholine</td>
<td>3.61</td>
</tr>
<tr>
<td>Vecuronium</td>
<td>6.12</td>
</tr>
</tbody>
</table>

### Table 3 Required sample size (n) as determined by power analysis

<table>
<thead>
<tr>
<th>Difference H₀ vs H₁ (allowable error)</th>
<th>n (COV = 25%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>± 20%</td>
<td>15</td>
</tr>
<tr>
<td>± 18%</td>
<td>18</td>
</tr>
<tr>
<td>± 15%</td>
<td>24</td>
</tr>
<tr>
<td>± 12%</td>
<td>37</td>
</tr>
<tr>
<td>± 10%</td>
<td>52</td>
</tr>
</tbody>
</table>

Fig 1 The effect of allowable error on required sample size; α = 0.05, COV = 25%.

Fig 2 The effect of COV on required sample size. α = 0.05. Allowable error is ± 15%.
required sample size than would be obtained using the slope from traditional log dose logit or probit linear regression analysis.

When determining the required sample size, one key element is determined by the investigator. How large an error is deemed acceptable? As can be seen from Figure 2 and Table 3, as few as 15 subjects may be required if an error of ±20% in the ED_{50} is thought to be tolerable. However, as the allowable error decreases, the required sample size rapidly increases. There is therefore no absolutely ‘correct’ solution to the problem of what the sample size should be. Every sample size determination represents a compromise between scientific precision and practical logistics.

When determining sample size for neuromuscular dose–response studies, we suggest that an allowable error of ±15% constitutes a reasonable middle ground. An error of this magnitude in dosing is unlikely to have major clinical consequences, yet the sample size requirement (\(n=24\)) should not impose an undue logistical burden on the investigator. Once 30 valid observations have been obtained, the point of diminishing returns has to a large extent been reached. This latter value also allows some leeway for protocol violations, outliers, and technical difficulties. Beyond this sample size, increases in accuracy are unlikely to be of clinical importance.

Before this study, single-dose protocol-derived potency data on the variability (\(\sigma\) of ED_{50}) of even commonly used drugs were extremely sparse. Our data and conclusions can be used in two ways. Investigators can use a COV of 25% and Table 3, as few as 15 subjects may be required if an error of ±20% in the ED_{50} is thought to be tolerable. However, as the allowable error decreases, the required sample size rapidly increases. There is therefore no absolutely ‘correct’ solution to the problem of what the sample size should be. Every sample size determination represents a compromise between scientific precision and practical logistics.

Conflict of interest

M.N. has participated in two studies of sugammadex which were funded for Organon Pharmaceuticals.

Funding

The present study was totally unfunded. The cited studies on rapacuronium, rocuronium, succinylcholine, vecuronium, and atracurium were supported in part by funds from the Department of Anesthesiology, St Vincent’s Hospital Manhattan, New York, NY, USA. The cited studies on cisatracurium, pancuronium, and pipercuronium were supported by funds received from the College of Medicine Research Center, College of Medicine, King Saud University.

References

7 Kopman AF, Lien CA, Naguib M. Determining the potency of neuromuscular blockers: are traditional methods flawed? Br J Anaesth 2010; 104: 705–10