Evaluation of methacholine dose-response curves by linear and exponential mathematical models: goodness-of-fit and validity of extrapolation


*Correspondence: G. Verlato
Cattedra di Statistica Medica e Biometria
Istituto di Igiene, Università di Verona
Policlinico di Borgo Roma, Verona, Italy.
**Istituto di Tisiologia e Malattie dell’Apparato Respiratorio, Università di Pavia, Policlinico S. Matteo, Pavia, Italy.
††Departamento de Epidemiologia, State University of Groningen, Groningen, The Netherlands.

Received: March 27 1995
Accepted after revision October 30 1995

The scientific community is still debating how to express bronchial responsiveness [1]. In particular, studies on bronchial reactivity and its relationship with asthma and other chronic obstructive pulmonary diseases have to face the difficult problem of so-called censored data of bronchoprovocation challenge tests. This problem is larger in epidemiological studies than in the clinical context: firstly, the maximal dose of bronchoconstrictor agent which can be administered is limited by safety considerations and side-effects; secondly, most subjects are not asthmatics and, thus, present a low bronchial reactivity. The resulting information loss limits the possibilities for statistical analyses [2, 3]. The problem of censored data becomes even more important in longitudinal studies, having to deal with changes in the estimates of censored data [4–7].

To overcome this problem, several techniques have been proposed: new indices linked to a smaller fall of forced expiratory volume in one second (FEV1) (i.e. 6, 10 and 15%); extrapolations to one doubling-dose beyond the maximum dose administered; area under dose-response curve; dose-response slope [3, 8–14].

Recently Chinn et al. [15] proposed estimating the provocative dose of agonist required to cause a 20% fall in FEV1 (PD20) by fitting an exponential curve to the observed data and by extrapolating by one doubling-dose. They compared extrapolated PD20 with two alternative measures of response, two-point slope [13] and least-squares slope [2], and concluded that PD20 is somewhat more suitable. Indeed this method presents several advantages: its results are repeatable; it allows discrimination between asthmatic and nonasthmatic subjects; and it can be synthesized through variables which satisfy the main requirements for statistical analysis, i.e. normality and stability of variance. However, the validity of extrapolation has not yet been verified through a direct comparison of observed and expected data [16].

The coefficient of determination ($r^2$) was significantly higher with the exponential model (0.81±0.22; mean±SD) than with the linear model (0.69±0.27). With both models, extrapolated values were usually lower than observed values. As a consequence, a 20% fall in FEV1 with respect to postsaline FEV1 was observed in only 24% and 21% of the tests, where a 20% fall had been predicted, respectively, according to the linear and the exponential model.

In conclusion, exponential models are better than linear models with respect to data interpolation of methacholine dose-response curves. However, they are worse with respect to extrapolation to higher doses. With any model, extrapolation of dose-response curves by one doubling-dose should be avoided.
In the three Italian centres participating in the European Community Respiratory Health Survey (ECRHS) [17, 18], the protocol of the methacholine challenge test was slightly modified by administering at least one additional dose of methacholine (4 mg cumulative dose), without impairing safety conditions [19–21]. This gave the opportunity to verify the validity of extrapolation by one doubling-dose in the usual range of epidemiological studies.

The aims of the present work were: 1) to investigate which model (linear [2] vs exponential [15]) and which minimization method (trials and errors [22] vs Levenberg-Marquardt [23]) gives better results in terms of data interpolation (goodness-of-fit); 2) to verify the validity of extrapolation by comparing FEV1 observed after 4 mg methacholine with values extrapolated after truncation of the dose-response curve at 2 mg.

**Material and methods**

**Subjects**

The design of the ECRHS [17] involved two stages. In Stage I, a screening questionnaire on respiratory symptoms was mailed to a probability sample of 20–44 year old males and females resident in the areas. In the second stage, a 20% random sample of responders to the mailed questionnaire was invited to the local chest clinic, in order to undergo a standarized clinical interview, lung function tests, including a bronchial challenge with methacholine, and allergen skin tests. In addition, people who, in the screening questionnaire, had reported asthma attacks, taking medicines for asthma, or awakening due to an attack of shortness of breath, were asked to attend Stage II.

In this way, 1,835 subjects out of 6,031 who had previously returned the questionnaire were invited to the clinical centres. Of these, 899 agreed to participate in stage II, yielding an attendance rate of 49% [24]. An additional 190 subjects who declared asthma-like symptoms in the screening questionnaire underwent clinical examination. Moreover, 15 subjects who had fulfilled the Stage II protocol in Verona without attending Stage I were admitted to the study, yielding an overall number of 1,104 subjects.

Of these 1,104 subjects, 850 performed technically satisfactory methacholine challenge tests. After eliminating 18 subjects receiving just one or two doses of methacholine, the study base comprised 832 subjects. Informed consent was given by each subject participating in the study.

**Methacholine bronchial challenge**

Bronchial responsiveness was measured by methacholine challenge. For the administration of methacholine Mefar MB3 Dosimeters (Mefar spa, Bovezzi, Italy) were used.

Lung volumes (forced vital capacity (FVC) and FEV1) were measured by means of Biomedin Spirometer (Biomedin srl, Padova, Italy) taking the best of five satisfactory manoeuvres. A solution from lyophilized methacholine chloride (Provocoline, Hoffman La Roche, Basel, Switzerland) was used as a provocative agent.

Inclusion criteria, measurement conditions and laboratory protocol are reported in the international protocol [17, 18]. Briefly, all subjects reporting a heart attack in the last 3 months or any other disease, affected by epilepsy requiring medical treatment, pregnant or breast-feeding, or taking beta-blockers for any reason, were excluded from the challenge test. In addition, methacholine challenge was not performed when the subjects had difficulty in complying with instructions or had a best initial FEV1 less than 70% predicted or a best postdiluent FEV1 less than 90% of initial FEV1.

According to the national protocol, the test was stopped when FEV1 had fallen by 20% or more with respect to postsaline FEV1, when a cumulative dose of 6 mg methacholine had been administered (while the European protocol stopped at 2 mg), or when the subject refused to continue or was not able to perform two technically satisfactory manoeuvres out of five attempts. The following cumulative doses were given: 0.0156, 0.0625, 0.25, 1, 2, 4, 6 mg. If the subject had a past history of asthma or wheezing, a more careful protocol was adopted by starting from 0.00185 mg and progressively doubling the cumulative dose. In one centre (Verona), the dose of 6 mg was replaced by a dose of 8 mg.

In order to standardize the administration methods, all the Italian research teams underwent a training session and all the tests were conducted using the same equipment. Nebulizers, of known output, were checked prior and during each challenge.

Bronchoconstriction was readily reversed by administration of salbutamol *via* metered-dose inhaler. No noticeable adverse reactions were observed at a cumulative dose of 6 or 8 mg methacholine.

**Mathematical models**

The two-point slope [13] was computed as the percentage decline in FEV1 from the postsaline value to that of the total cumulative dose administered, divided by the total dose. The least-squares slope was calculated from all data except the postsaline FEV1 [2]. The exponential model [11]:

\[
\text{FEV1} = c - \exp(a+b(\log_{10}\text{dose}))
\]

was fitted to all data (including postsaline FEV1) using two different minimization methods, the trials and errors [22] and Levenberg-Marquardt method [23]. To include postsaline FEV1 in the model, it was necessary to assume that the corresponding dose of methacholine was higher than zero (0.0001 mg). The models are presented in graphic form in figure 1.

In a few cases, the temporary estimate of \(c\) tended to increase enormously during the minimization process,
yielding unphysiological values of \( c \) (up to 100,000 L·s⁻¹) and turning the exponential model into a linear one. To prevent this outcome, \( c \) was not allowed to increase more than 30% with respect to postsaline FEV₁. This problem did not occur with linear models, since the intercept of the least-squares regression never differed from postsaline FEV₁ by more than 18%.

A QBASIC program and a Mathematica program, the former modified (V.G.) the latter developed in our laboratory (V.A. and P.M), were used to fit the exponential equation to the experimental data by the method of trials and errors and by the method by Levenberg-Marquardt, respectively. This software is available from the Correspondence author, address on first page of paper.

### Truncation of the curves

The goodness-of-fit of different methods was studied for all of the 832 subjects, whilst the validity of the extrapolation was investigated for the 639 subjects who were administered a cumulative dose of 2 mg methacholine without presenting a 20% fall in FEV₁, and thus received an additional dose of 4 mg methacholine. Extrapolated values lower than zero were recoded at zero.

### Statistical analysis

Data are presented as mean±SD. The coefficient of determination \( (r²) \) was used to evaluate the goodness-of-fit of linear versus exponential models. Since this measure was not normally distributed according to the Kolmogorov-Smirnov test, Friedman's nonparametric two-way analysis of variance was used to compare coefficients of determination obtained with different models.

The relationship between observed and extrapolated values of FEV₁ at 4 mg methacholine was studied by a simple linear regression. A t-test was used to assess whether the slope of the regression line was significantly different from that of the identity line [25].

### Results

Of the total 832 subjects given at least three doses of methacholine, 131 presented a 20% fall in FEV₁ below 4 mg methacholine, and 35 at 4 mg. The number of doses administered (mean±SD) was 7.0±1.8 (range 3–14).

#### Data interpolation

Both linear and exponential models could be fitted to all 832 curves. The \( r² \) was significantly lower \((p<0.001)\) with the linear model \((0.69±0.27 with the least-squares method)\) than with the exponential model \((0.81±0.22 both with trials and errors and Levenberg-Marquardt algorithm)\).

The linear model slightly underestimated postsaline FEV₁, whilst the exponential models slightly overestimated it. The percentage difference between the intercept of the linear regression and the corresponding postsaline FEV₁ amounted to -2.2±2.9% \((range -18.1 to +7.8\%)\), whilst the percentage difference between the upper asymptote and postsaline FEV₁ was 3.0±7.2% \((range -8.8 to +30.0\%)\) and 1.9±6.1% \((range -8.8 to +30.0\%)\), respectively, for the trials and errors method and the Levenberg-Marquardt algorithm.

When allowing for any \( c \) values, exponential models yielded unphysiological asymptotes, exceeding postsaline FEV₁ by 30% or more, in 65 cases (8%) with trials and errors and in 31 cases (4%) with the Levenberg-Marquardt algorithm.

#### Data extrapolation

Truncation of the curves, by omitting doses equal to or higher than 4 mg methacholine, enabled FEV₁ extrapolation \((n=639)\). Before extrapolating to 4 mg, goodness-of-fit of linear and exponential models on truncated curves was verified through the \( r² \). Again this measure
was significantly lower (p<0.001) with the linear model (0.64±0.29 with the least-squares method) than with the exponential model (0.74±0.26 with trials and errors and 0.73±0.27 with Levenberg-Marquardt algorithm).

Figure 2 shows the FEV1 extrapolated to 4 mg (exp) versus the experimentally determined FEV1 (obs). The regression line was closer to the identity line with linear models (exp=-0.15+1.00*obs for two-point slope (fig. 2a) and exp=-0.16+1.01*obs for least-squares slope (fig. 2b)) than with exponential models (exp=0.24+0.87*obs for trials and errors (fig. 2c) and exp=0.10+0.92*obs with Levenberg-Marquardt algorithm (fig. 2d)). Indeed, the slope of the regression line was significantly different from unit with exponential models but not with linear ones. The linear correlation coefficient indicated a highly significant relationship for linear models (r=0.97 for both two-point slope and least-squares slope) and a much weaker relationship for exponential models (r=0.70 for trials and errors and r=0.76 for Levenberg-Marquardt method).

The deviation of the extrapolated values in percentage of the measured values amounted to -4.5±6.1%, -4.0±6.2%, -5.8±19.2% and -5.1±17.6% (mean±SD) when adopting two-point slope, least-squares slope, trials and errors and Levenberg-Marquardt algorithm, respectively.

Extrapolation was also evaluated by recoding FEV1 into a dichotomous variable, where the response was considered "positive" in the presence of a 20% fall in FEV1 with respect to postsaline FEV1 and "negative" otherwise (table 1). Specificity, defined as the percentage of predicted negative responses among observed negative responses was similar for all models, ranging 85–87% (table 2). On the contrary, sensitivity, defined as the

### Table 1. – Validity of extrapolation by one doubling-dose, verified through comparison of the bronchial response observed at 4 mg with the bronchial response extrapolated after truncation of the curve at 2 mg (n=639)

<table>
<thead>
<tr>
<th></th>
<th>Two-point slope</th>
<th>Least-squares slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed response</td>
<td>Observed nonresponse</td>
</tr>
<tr>
<td>Expected response</td>
<td>31</td>
<td>91</td>
</tr>
<tr>
<td>Expected nonresponse</td>
<td>4</td>
<td>513</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Trials and errors</th>
<th>Levenberg-Marquardt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed response</td>
<td>Observed nonresponse</td>
</tr>
<tr>
<td>Expected response</td>
<td>21</td>
<td>85</td>
</tr>
<tr>
<td>Expected nonresponse</td>
<td>14</td>
<td>519</td>
</tr>
</tbody>
</table>

†: Bronchial response was coded as dichotomous variable, where response was considered a decrease in FEV1 ≥20% with respect to postsaline FEV1, whilst all other outcomes were coded as nonresponse. Linear models (two-point slope and least-square slope) are compared with exponential models (trials and errors and Levenberg-Marquardt). FEV1: forced expiratory volume in one second.
As far as the first point is concerned, $r^2$ was higher with exponential models than with linear models, suggesting that, at least in this dose range, the pattern of dose-response curves can be approximated better by an exponential model on a semilogarithmic scale than by a linear model on a normal scale.

It must be pointed out, however, that the two models are equivalent from a mathematical point of view: a linear model turns into an exponential one after log transformation of $x$ values. Thus, the statement that exponential models are better with respect to data interpolation is relevant from a methodological point of view rather than from a physiological one. Indeed, dose-response curves of methacholine challenges are usually reported on semi-logarithmic scale, as for instance when computing PD_{20} [26, 27].

Both minimization procedures, trials and errors and Levenberg-Marquardt algorithm, allowed the exponential curves to be fitted equally well to the observed data. As far as extrapolation to one doubling-dose is concerned, to our knowledge, the present study is the first attempt to evaluate extrapolation by one doubling-dose on the basis of the criterion of validity: expected values are directly compared with observed values. Up to now, extrapolation by one doubling-dose has been evaluated mainly on statistical considerations, such as repeatability, normality and stability of variance [15].

Linear models were slightly better than exponential models with respect to extrapolation by one doubling-dose (4 mg methacholine). However, in both cases extrapolation underestimated the value of FEV_{1} which was actually observed. As a consequence, when adopting the 20% fall in FEV_{1} as a cut-off, the number of predicted responses was three-times larger than the number of responses actually observed. Moreover, most false positives did not respond even at higher doses (6–8 mg methacholine). Indeed, for this low positive predictive value extrapolation by one doubling-dose is not suited for epidemiological studies dealing with risk factors for bronchial hyperresponsiveness.

A rather straightforward explanation can be given for this result. The complete dose-response curves present a sigmoid shape [14, 16], which is clearly evident in most subjects only when very high doses of methacholine are administered. In epidemiological studies these doses are rarely attained, so that only the first part of the sigmoid curve is recorded, which can be approximated by a straight line or by an exponential curve. Extrapolation by one doubling-dose could have taken place on that part of the curve (2–4 mg methacholine) where the response is progressively decreasing towards the plateau value. This could be the reason why extrapolation by one doubling-dose underestimated FEV_{1} at 4 mg.

These results are further supported by the fact that the study protocol was very strict and rigorous, since it had been specifically designed to prevent possible biases arising from different testing procedures and difference in subject compliance.

As a consequence, if the study design does require that PD_{20} be assessed in as large as possible a number of subjects, the first attempt to diminish censored data should

<table>
<thead>
<tr>
<th>Extrapolation Method</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>+ve</th>
<th>-ve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-point slope</td>
<td>89</td>
<td>85</td>
<td>25</td>
<td>99</td>
</tr>
<tr>
<td>Least-square slope</td>
<td>77</td>
<td>86</td>
<td>24</td>
<td>98</td>
</tr>
<tr>
<td>Trials and errors</td>
<td>60</td>
<td>86</td>
<td>20</td>
<td>97</td>
</tr>
<tr>
<td>Levenberg-Marquardt</td>
<td>63</td>
<td>87</td>
<td>22</td>
<td>98</td>
</tr>
</tbody>
</table>
consist in increasing the maximum dose administered. The present study as well as previous ones [9, 19–21] confirm that, even at very high doses of methacholine (8 mg cumulative dose), when bronchoconstriction does occur, it can be readily reversed by inhalation of salbutamol via metered-dose inhaler. It must be pointed out that a more careful protocol was used for subjects with a past history of asthma or wheezing. Moreover, no noticeable adverse reactions were observed.

In conclusion, according to the methacholine challenge tests, performed during the ECRHS in Northern Italy, exponential models are better than linear models with respect to data interpolation, but are worse with respect to extrapolation to higher doses. However, with any model the use of extrapolation by one doubling-dose to increase the number of estimates should be discouraged.

Members of ECRHS - Italy


