Statistical Modelling of Cardiovascular Data. An Introduction to Linear Mixed Models

Paulo Gonçalves, Christophe Lenoir, Christophe Heymes, Bernard Swynghedauw, Christian Lavergne

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Statistical Modelling of Cardiovascular Data. An Introduction to Linear Mixed Models

Paulo Goncalves, Christophe Lenoir, Christophe Heymes, Bernard Swynghedauw and Christian Lavergne

N° 5787

Décembre 2005

THÈME 1

ISSN 0249-6399

Statistical Modelling of Cardiovascular Data. An Introduction to Linear Mixed Models

Paulo Gonçalves [1], Christophe Lenoir [2], Christophe Heymes [2]
Bernard Swynghedauw [2], Christian Lavergne [1]

Thème 2 – Systèmes Cognitifs
Projet Mistis

Rapport de recherche n° 5787 – Décembre 2005 - 21 pages

Abstract: Most of statistical approaches in cardiovascular research were based on variance analysis (ANOVA). However, most of the time, the assumption that data are independent is violated since several measures are performed on the same subject (repeated measures). In addition, the presence of intra- and inter-observers variability can potentially obscure significant differences.

The linear mixed model (LMM) is an extended multivariate linear regression method of analysis that accounts for both fixed and random effects. LMM allows for addressing incomplete design cases.

In this paper, LMM was applied to two sets of cardiovascular research data and compared to ANOVA. The first example is an analysis of heart rate in mice after atropine and propranolol injections. LMM shows an important mouse random effects that depends on pharmacological treatment and provides with accurate estimates for each significant experimental factors. When randomly suppressing observations from the data sets (20-30%) the time factor of Anova model becomes non significant while LMM still remains significant. The second example is the analysis of isolated coronary-perfused pressure of transgenic mice hearts. LMM evidenced a significant transgenic effect in both male and female animals, while, with ANOVA, the transgenic effects was limited to male mice only.

In both cases, as compared to ANOVA, the LMM separately accounts for fixed and random effects, allowing thus for studying more adequately incomplete designs on repeated measures.

[1] INRIA Rhône-Alpes, 655 avenue de l'Europe, 38334 Saint Ismier, CEDEX, France. {Paulo.Goncalves, Christian.Lavergne}@inria.fr

[2] U572-INSERM. Hôpital Lariboisière. 41 Bd de la Chapelle 75475. Paris Cedex 10, France. {Christophe.Lenoir, Christophe.Heymes, Bernard.Swynghedauw}@larib.inserm.fr

Keywords: linear mixed models, ANOVA, cardiovascular research data, random effect, information criterion, statistical analysis

Modélisation de Données Cardiovasculaires. Une Introduction aux Modèles Linéaires Mixtes

Résumé: La plupart des études statistiques faites dans le domaine cardio-vasculaire sont basées sur l'analyse de la variance (ANOVA). Cependant, dans une grande majorité de cas, plusieurs séries de mesures sont effectuées sur un même individu (mesures répétées), invalidant ainsi l'hypothèse implicite d'indépendance des données. En outre, l'existence de variabilité intra- et inter individus peut masquer des différences potentiellement significatives.

Les modèles linéaires mixtes (LMM) sont une extension des régressions linéaires multi variées qui prennent en compte à la fois les effets fixes et les effets aléatoires. Les LMM permettent également de traiter les cas de plans d'expériences avec données manquantes.

Dans cet article, les LMM sont appliqués à deux séries de données cardiovasculaires, puis les résultats sont comparés à ceux obtenus par ANOVA. Le premier exemple traite de l'analyse du rythme cardiaque de la souris après injection d'atropine et de propranolol. Les LMM révèlent alors un important effet aléatoire individu, dépendant du traitement pharmacologique, et fournit une estimation précise pour chacun des facteurs expérimentaux significatifs. Lorsqu'une partie (20-30%) des observations est aléatoirement supprimée du jeu de données, le facteur « temps » perd toute significativité avec l'ANOVA, alors qu'il demeure un facteur significatif avec les LMM. Le deuxième exemple traite de l'analyse de la pression dans le cas de cœurs de souris transgéniques, lorsque ceux-ci sont isolés et placés en perfusion coronarienne. Les LMM mettent ici en évidence un effet transgénique significatif pour les souris mâles, comme pour les femelles, alors que l'ANOVA ne souligne cet effet que dans le cas des mâles.

Dans les deux exemples traités, les LMM comparés à l'ANOVA, permettent de distinguer les effets fixes des effets aléatoires, offrant ainsi une étude plus rigoureuse dans le cas de plans d'expériences incomplets avec répétitions.

Mots clés: Modèles Linéaire Mixtes, étude de données cardiovasculaires, effet aléatoire, critère d'information, analyse statistique.

1 Introduction

In biological research, the methods generally employed to describe the data and to account for fixed and random effects are based on the analysis of variance (ANOVA). Fixed effects are factors that do not vary between measurements, such as sex or transgenic strain. Random effects vary at each measurement point. ANOVA performs a linear regression to estimate the *fixed effects* parameters of the underlying models, and to reveal significant differences. Nevertheless, lack of efficient algorithms have traditionally prompted some users to work with balanced designs only, and to comply with such constraint, some data were simply withdrawn from the experimental data set. However, from a theoretical viewpoint, ANOVA does not require to necessarily consider balanced designs, and more recent softwares are now free of this constraint. More importantly, linear models assume more stringent hypothesis that do not match all possible applications. For instance, in a host of situations, data are obtained from a fixed set of subjects observed under different conditions and times (longitudinal studies). Such framework, referred to as *repeated measures,* necessarily implies correlation between observations derived from a same subject. In general, studies are not conditioned to the subjects, and these latter induce on their own an unobservable *random effect.* The ANOVA model does not take into account this dimension, but only accounts for fixed effects, viewed as controlled factors such as sex or transgenic manipulations. The linear mixed models (LMM) have been proposed to circumvent these limitations, by adding random effects aimed at modelling the variability due to peculiarity of the observed subjects (reviewed in (10)). Originally, the method has been developed for balanced designs (3), then generalized to arbitrary Gaussian mixed models (6). Its use is now widespread thanks to the joint development of performing software and hardware.

The purpose of this paper is to introduce Linear Mixed Models to cardiovascular research. Using two examples derived from cardiovascular context, we illustrate the extended capabilities of LMM over ANOVA models. The first example deals with atropine and propranolol effects on the mice heart rate which is known to be a highly variable parameter. The second example is an analysis of isolated coronary-perfused pressure of mice hearts, in order to study a transgenic effect.

2 Theory

As analysis of variance models are particular cases of linear models, analysis of variance models with random effects are part of a more general theory on linear mixed models. In this domain, an abundant literature exists: aside from the "classical literature" on linear models, more recently a large number of books have been exclusively dedicated to linear mixed models, including "Variance components" from Searle et al. (12), and "Mixed-Effects models in S and S-Plus" from Pinheiro et al 2000 (11).

Under normal assumption, LMM's rely on Maximum Likelihood (ML) estimating procedures, as it is also the case for ANOVA models, but in contrast to ANOVA, these ML procedures do not admit analytic solutions in general and can only be numerically solved. Today, statistical softwares like R, S-plus, SAS have turned these methods more accessible, favouring their use in many scientific domains, such as genetics, psychophysiology and psychometry (1, 2, 5, 7, 8).

2.1 The Variance Analysis Model (ANOVA)

The usual ANOVA model aims at interpreting the response variable Y as a linear combination of factors, and quantifies the joint effect of each factor level (or state) according to the following model equation:

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \qquad (1)$$

In this equation, indices *i* and *j* denote the *i*-th and the *j*-th level (or state) of factors F_1 and F_2 respectively, whereas index *k* corresponds to the *k*-th repetition of the occurrence (*i*,*j*) (*k*=1,…,n_{ij}). We pose n = ? $_{i,j}$ n_{ij}, the total number of observations. Moreover, Anova model assumes that e_{ijk} are *n* independent random variables, identically distributed with normal law, zero mean and s_e^2 variance. In short, we can write e ~ $N_R^n(0,s_e^2.Id_n)$, where e is a vector from R^n with components $e_{ijk,}$ and Id_n is the identity matrix of R^n.
The unknown parameter μ_{ij} represents the expectation of Y_{ijk}, and expresses the idea that each combination (*i*,*j*) has a specific effect on the response Y, but does not express the possible individual effects, nor it expresses the interaction of F_1 and F_2. Then, it is necessary to expand each parameter μ_{ij} into factor dependent terms as:

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \qquad (2)$$

This new expression of μ_{ij} induces linear constraints on the parameters α , β and γ. In the sequel of this presentation, the following arbitrary parameterization is selected:

$$\alpha_1 = 0, \qquad \beta_1 = 0, \qquad \gamma_{i1} = \gamma_{1i} = 0, \quad \text{for all } i \text{ and } j. \qquad (3)$$

Commonly, software outputs refer to μ as to the *Intercept*, which corresponds to the first state effects of factors F_1 and F_2 simultaneously. Parameters α_i and β_j quantify the mean differential

effects of the i–th state of factor F_1 and the j–th state of factor F_2, with respect to the first states of F_1 and F_2, respectively. The interaction effect between state i of F_1 and state j of F_2 is then isolated and measured with parameter γ_{ij}.

Since the model is linear, equation (1) is often written in its matrix form:

$$Y = X\theta + \varepsilon \qquad (4)$$

where Y is the observation vector from R^n with components Y_{ijk}, θ is the vector of unknown parameters ($\mu, \alpha_2, \ldots, \alpha_I, \beta_2, \ldots, \beta_J, \gamma_{22}, \ldots, \gamma_{IJ}$). Given also the specific parameterization (3), the design matrix X is filled with 0's and 1's depending on which combination (i,j) does the observation belong to. This matrix can take on any structure in general, and the estimated parameter θ, solution of (4) reads

$$\hat{?} = (X'X)^{-1}X'Y . \qquad (5)$$

This estimator, commonly referred to as the "least square estimator", also coincides with the maximum likelihood estimator of θ, since ε satisfies to the normal hypothesis.

2.2 The linear mixed model (LMM)

The framework of LMM's, or more particularly of ANOVA models with random effects, differs from the framework of classical ANOVA by experimental and / or modelling contexts. For instance, recalling the ANOVA formalism with two factors, let us assume that factor F_1 describes a "drug" factor that takes on three different states associated to distinct drugs, and that factor F_2 represents the gender effect with only two states. These two factors are controlled, and in a classical ANOVA experimental set up, each animal is associated to one single observation. Thus, it is necessary to have n distinct animals at our disposal to realistically assume independence of responses. However, experiments are most often performed on a set of individuals, each of them being observed under all different states (i.e. after each corresponding drug injection). Then, it becomes necessary to improve the model to take into account this repetition effect. It is also clear, that this subject effect can not be modelled like factors F_1 and F_2: its levels are not controlled and moreover, the goal here is not to differentiate among individuals. On the opposite, we consider all subjects as a unique sample issued from the same population, hence the introduction of a "subject" random effect. Then, model equation (1) becomes

$$Y_{ijs} = \mu + \alpha_i + \beta_j + \gamma_{ij} + u_s + \varepsilon_{ijs} \qquad (6)$$

where u_s ($s = 1, \ldots, S$, S being the subject sample size) is the "unobserved" random effect due to subject s. These effects are supposed to be independent, identically distributed with a zero mean Gaussian law and variance σ_u^2, reflecting the idea that subjects come from the same population (same law), and have distinct biological parents (independence). The model assumes independence between effect u and errors ε. As for (4), the new model (6) admits a matrix expression of the form:

$$Y = X\theta + ZU + \varepsilon . \qquad (7)$$

The random effect vector U has dimension S, the matrix Z of size n x S is in our model, filled with 0's and 1's. From the various hypotheses, we get for the variance of the observation Y the following non diagonal matrix:

$$Var(Y)=\sigma_u^2 ZZ'+\sigma_\varepsilon^2 Id_n , \qquad (8)$$

which depends on the two unknown variance components σ_u^2 and σ_ε^2. The maximum likelihood estimation of these quantities is a well-known procedure. However, in contrast to the classical linear case, it does not lead to close form estimates in general. Instead, we must turn to numerical implementation of the solutions (see for instance Searle et al. (12) for a detailed study of these classical estimation procedures). As for the matrix X, the component ZU can take on very general forms, and in particular it can be expressed as $ZU=\sum_{k=1}^{K} Z_k U_k$, where each U_k corresponds to a particular random effect.

2.3 Model selection

This approach allows for constructing a rich class of models generated by a variety of fixed and random designs X and Z, respectively. A "good choice" for the pair (X,Z) must result from a selection model procedure based on information criteria to be minimized. Those criteria rely on the log-likelihood, denoted $L(M)$, evaluated at $\hat{\theta}$ and $\hat{\sigma}^2$, where $\hat{\sigma}^2$ is simply the vector made out of the variance components. We designate by $q(M)$ the number of estimated parameters in the model, i.e. $q(M) = \dim \hat{\theta} + \dim \hat{\sigma}^2$.

AIC: Akaike Information Criterion

$$-2L(M) + 2q(M) \qquad (9)$$

BIC: Bayesian Information Criterion

$$-2L(M) + \ln(n) \cdot q(M) \qquad (10)$$

Difference between these two criteria is uniquely in the penalty factor (2 versus $\ln(n)$). In practice, AIC selects models less parsimonious than BIC, whereas this latter promotes robust models with retained effects that are even less questionable. Finally, let us note that the likelihood ratio tests remain valid as far as fixed effects are concerned (test on the θ vector components), but have no theoretical legitimacy at resolving presence versus absence of random effects.

3 Data analysis

The utilization of LMM, as compared to ANOVA, was illustrated in the two following experimental set up. Both include comparative analysis of cardiovascular data obtained after drug injections. Such set up were conventionally analysed with ANOVA in cardiovascular research.

3.1 *In vivo* study of mice heart rate

Most mammalians have a vagal tone, mice could be an exception, as previously suggested in (9). Such a question was addressed by analysing the effects of two antagonists of the autonomous nervous system on the heart rate of non anesthetized mice.
Eighteen males C57bl/6 mice (10 to 14 wk) were housed individually for two weeks at room temperature before recording. A biocompatible transmitters (TA10ETA-F20, DataSciences International, St Paul, MN) was implanted under isofluran mixture with carbogene anaesthesia (1.5 vol %). Three days later, heart rate was investigated using a telemetric process (Physiotel Receiver RLA1020, DataSciences International, St Paul, MN) in non anaesthetized freely moving animals. Recordings were made through the data acquisition system Powerlab 16 SP and Chart 4 (ADInstruments, Australia) at a sampling frequency of 2,000 Hz. At 2:00 pm, basal ECG was recorded after a single IP injection of either saline, or saturating dose of atropine (1 mg/kg), or propranolol (1 mg/kg) or a combination of atropine and propranolol.
The response variable, RR, is the time interval separating two consecutive heart beats (RR interval) which have been averaged on ECG recording during five minutes. This measure was performed after placebo (control – C), atropine (A), propranolol (P), or propranolol + atropine (AP) injections.
Mice were randomly assigned in these four injection levels, but for each state not all mice enter the experience, leading thus to an unbalanced design. For each of these injections, the RR were calculated before injection (T0), then 20, 40 and 60 minutes after injection (T20, T40 and T60, respectively). The working sample contains 171 measures.

3.1.1 Descriptive analysis

The measures of pharmacological autonomic blockade on the baseline heart rate are summarised in Table 1. This table reveals, before any statistical modelling, a strong placebo effect, the lack of a proper atropine tachycardia, and a pronounced slowing down of heart rate in state P. The combined injection of atropine and propranolol (AP) produces similar trends as in state P, but with smaller magnitude. Statistical modelling was performed with LMM, since we are in presence of repeated, and thus not independent measures. The results of the analysis were compared to those of an ANOVA model, although in that case the statistical hypotheses fail to apply to our experimental set up.

3.1.2 Statistical analysis.

The goal of this study is to model the variable to be explained RR, according to the two inputs: state injection (S) and time (T). Each measure is denoted RR_{STM}, with $S \in \{C, A, P, AP\}$, $T \in \{T0, T20, T40, T60\}$ and M is the mouse label. In the course of the process, among all addressed models, the retained one corresponds to the smallest criterion value.

Step 1: random effect characterisation

In a first step, we consider the elementary fixed effect only due to the state injection factor S, $\mu + \alpha_S$, with $\alpha_{S='C'} = 0$. The random effect was then formalized through five possible models.

We start with an ANOVA model that involves no random effect:

$M_0 : RR_{STM} = \mu + \alpha_S + \varepsilon_{STM}$.

Next, our first linear mixed model introduces a mouse random effect, denoted u_M:

$M_M : RR_{STM} = \mu + \alpha_S + u_M + \varepsilon_{STM}$.

Then, we consider a more elaborate model that adds a state dependent mouse random effect $u_{M/S}$, implicitly assumed to be independent with u_M:

$M_{M!S} : RR_{STM} = \mu + \alpha_S + u_M + u_{M/S} + \varepsilon_{STM}$.

At this point, none of the models M_M or $M_{M/S}$ integrate in their random effect the time influence. So, we now propose three extended versions of M_M and $M_{M/S}$, namely M_M^T, $M_{M/S}^T$ and $M_{M,M/S}^T$, that assume isolated or joint time dependent random effects, denoted v_M and $v_{M/S}$:

$M_M^T : RR_{STM} = \mu + \alpha_S + u_M + T.v_M + \varepsilon_{STM}$,

$M_{M/S}^T : RR_{STM} = \mu + \alpha_S + u_M + u_{M/S} + T.v_{M/S} + \varepsilon_{STM}$,

and $M_{M,M/S}^T : RR_{STM} = \mu + \alpha_S + u_M + T.v_M + u_{M/S} + T.v_{M/S} + \varepsilon_{STM}$.

In the last three models, all random effects are supposed to be independent, and T is a functional of the time variable t. For each of these six models, the corresponding criteria were computed (Table 2). From these AIC and BIC values, the first LMM (M_M) does not reveal a very strong random effect and therefore does not really outperform the simple ANOVA model (the two AIC's being close one to the other). However, the $M_{M/S}$ model shows a clear inter-individual mouse variability, provided it is conditioned to the injection type. The same conclusions apply to the specific time-dependent random effect model $M_{M/S}^T$, showing that time and mice random effects are significant only through the state injection factor. On the other hand, the two models M_M^T and $M_{M,M/S}^T$, are not relevant. In the sequel then, only the random effects u_M, $u_{M/S}$ and $T.v_{M/S}$ are retained. This structure implies not only a correlation between all measures coming from the same mouse, but an even stronger correlation between measures coming from the same mouse observed at different times in a given state. If σ_M^2, $\sigma_{M/S}^2$, $\tau_{M/S}^2$ and σ_ε^2, designate the variances of u_M, $u_{M/S}$, $v_{M/S}$ and ε_{MTS}, respectively, then, the following covariances hold (these four variables are supposed to be independent):

$\mathrm{cov}(RR_{STM}, RR_{S'T'M}) = \sigma_M^2$, for $S \neq S'$,

and

$\mathrm{cov}(RR_{STM}, RR_{ST'M}) = \sigma_M^2 + \sigma_{M/S}^2 + TT'\tau_{M/S}^2$, for $T \neq T'$.

Similarly, the variance of each observation is

$\mathrm{var}(RR_{STM}) = \sigma_M^2 + \sigma_{M/S}^2 + T^2\tau_{M/S}^2 + \sigma_\varepsilon^2$.

A simple linear time functional T was initially tried in the model, however in practice this arbitrary choice needs to be confronted with more general forms. In this direction, we also plugged in the model several non-linear functionals of the form $T : t \mapsto t^a$, with a varying. Actually, none of these choices outperformed the simple linear functional ($a = 1$), which was finally retained.
So far, the independence between random effects has been assumed. Very often though, such an assumption is unlikely to hold true and needs further investigation. In this example, no dependence between any two of the retained random effects improved the model.

Finally, data variance fitted random effect is:

$$u_M + u_{M/S} + T.v_{M/S},$$

with no correlation between random effects.

Step 2: Fixed effect characterization
After having determined a sensible random effect, we now focus on the fixed effect modelling. In this experimental set up, the only two explicative variables are the state injection S and the time T. Whereas S is clearly a factor since it is not measurable, T can impact the model either as a quantitative variable, or a factor. One goal of this section is to choose between these two possibilities. Jointly, we have to check existence of interactions between the two explicative variables. In other words, we must select, comparing the corresponding information criteria, among four possible models reflecting the fixed effect:

- $\mu + \alpha_S + \beta T$: T is the quantitative variable, and β its coefficient,
- $\mu + \alpha_S + \beta T + \gamma_S T$: same as previous with γ_S an extra state dependent time coefficient,
- $\mu + \alpha_S + \beta_T$: β_T is the time factor effect with conventionnaly $\beta_{T='T0'} = 0$,
- $\mu + \alpha_S + \beta_T + \gamma_{ST}$: same as previous with γ_{ST} the interaction between state and time factors, and $\gamma_{ST} = 0$ if $S = 'C'$ or $T = 'T0'$.

The selection criteria corresponding to these four models are summarized in Table 3. First, let us notice that all four models outperform the models omitting time as a fixed effect. Furthermore, we infer from Table 3, that the time variable behaves as a factor in the fixed effect, without significant interaction with drug injection effect.
This model generates seven fixed effect parameters. Looking at the estimated values and at their associated p-values (not reported here), $\alpha_{S='A'}$ and $\beta_{T='T60'}$ can be neglected, confirming in that, the experimental results of Table 1. Eventually, the tailored model comprises only five highly relevant parameters to describe the fixed effect (Table 4), and four parameters to describe the variance (Table 5), with a corresponding AIC and BIC equal to 1344.5 and 1372.8 respectively.
At last, for comparison purpose with the observed values of Table 1, it is possible to compute the fitted values and associated standard deviations, using the estimates reported in Tables 4 and 5. We give these results in Table 6.

There is a pronounced placebo effect at time T20 and T40 ($p<0.001$). Atropine injection has the same effect as placebo on the mean RR, as states 'C' and 'A' can be confounded at all

experimental times. Propranolol injection, respectively its association with atropine, produces a strong RR enlargement of 33.78 ms, and 25.20 ms respectively, with no interaction with experimental times. Within the 60 minutes duration of this experiment, no propranolol effect attenuation was significantly evidenced.

Comments

Let us first stress that reproducing analysis of step 2, but ignoring random effects (i.e. withdrawing step 1), leads to similar fixed effect conclusions, except that the time factor $\beta_{T='T60'}$ becomes significant with ANOVA. These resembling results between ANOVA and LMM approaches are quite usual, mainly when the model's fixed effects are highly significant, as it is the case here. It means that fixed effects are robust with respect to the model covariance structure. However, when suppressing randomly observations from the data set (20 to 30 %), the time factor $\beta_{T='T60'}$ of ANOVA model becomes non significant, whereas all other parameters of the fixed effect remain significant. Playing with the same data suppression, LMM remains robust with respect to all its significant parameters.
Regarding the error modelling of this ANOVA model, the dispersion parameter σ_{ε}^{2} is estimated to be 347, and it can be related to a combination of all partial variance components, whose estimates are reported in Table 5. Then, it becomes clear that most of the data variability has been explained by the random effects modelling. As for the fixed effects, each random component estimates deserves a thorough interpretation. In this particular study for instance, the relatively large estimated value for $\sigma_{M/S}$, indicates that the four injection types induce a strong intra-individual variability.

From an AIC viewpoint, such fixed effect model (ANOVA) would increase AIC by 148 (AIC = 1492.4), as compared to the linear mixed model (AIC = 1344.5). On the other hand, a null model (where all data are considered independent and identically distributed) has an AIC equal to 1570, which is 88 larger (and thus worse) than the AIC of the sole fixed effect model (AIC = 1492.4). The above differences (148 versus 88) support the claim that in LMM, the overall gain comes rather from a proper random effect modelling than from the fixed effect modelling. This assertion is naturally sustained when the model accounts for the random effects only. AIC of such a model (equal to 1381.5), conveys an improvement over the null model, that surpasses the improvement induced by the ANOVA model.

To conclude, such a study illustrates that modelling properly the random effects brings valuable information that does not question the results obtained with an accurate fixed effect modelling, but rather complements knowledge about the experimental set up. Note also that in this particular example, ANOVA is basically an erroneous model.

3.2 Isolated coronary-perfused hearts in transgenic mice

A transgenic strain of mouse was generated to study the effects of myocardial aldosterone production on endothelium-independent coronary reactivity. The construct includes a cardiac specific promotor.
Nine Adult transgenic (four males and five females) and ten wild-type (five males and five females) FVB mice (8-10 weeks old,) were anaesthetised (pentobarbital sodium, 0.1ml/100g) and heparinised (heparin sodium, 100 IU/100g) using IP injection. Hearts were rapidly excised, cannulated, and retrocoronary perfused at constant flow with a Krebs Henseleit buffer equilibrated with 95% O_2 and 5% CO_2 (pH 7.4, 37°C). During a 20 minutes equilibration period, coronary flow was adjusted to achieve a coronary perfusion pressure of 60±5 mmHg as used previously. A coronary perfusion pressure of 60mmHg produced optimal vasodilator responses and preliminary experiments demonstrated that left ventricular function was normal at this level. The coronary perfusion pressure (CPP) was measured via a pressure transducer and data recorded on a MacLab/8s system (ADI, Australia). A left ventricular balloon was not used as it impairs vasodilator responses in some animals from all strains and introduces additional variability. Bolus (10µl) of sodium nitroprusside (SNP, 10 nmol/L - 1 mmol/L) were injected just above the aortic cannula into the flow (~1% of the flow causing no change in coronary perfusion pressure), with sufficient time allowed between agents to allow recovery of basal coronary perfusion pressure.
The response variable PP, is the coronary perfusion pressure in response to six SNP concentration doses (in Mol/l): $d_1 = 10^{-8}$, $d_2 = 10^{-7}$, $d_3 = 10^{-6}$, $d_4 = 10^{-5}$, $d_5 = 10^{-4}$, $d_6 = 10^{-3}$. This measure is performed in order to assess both transgenic and gender effects and includes 114 determinations.

3.2.1 Statistical analysis

The effects of SNP on coronary pressure were summarized in table 7. Like for the previous experiment, a satisfactory Linear Mixed Model was elaborated using Information Criteria to orient the selection model. For the sake of concision, details of steps 1 and 2 that led to characterize random and fixed effects, are omitted. Rather, the retained model is commented with emphasis on its improvements over more standard procedures (e.g. ANOVA with repeated measures).
As far as random effect was concerned, only a subject (mouse) random effect u_M, a dose dependent mouse random effect $D.v_M$, and their correlation, were significant. It is worth noticing that neither gender G, nor transgenesis T produce significant random effects on PP.
Regarding fixed effects, the roles of the two factors G and T have to be formalized, and d to be identified as a factor or a quantitative variable in the model. After this, relevance of the interactions between these inputs needs to be established.
After all these steps, the model reads

$$PP_{TGdM} = \mu + \gamma_{TG} + \beta.D + \beta_T.D + u_M + D.v_M + \varepsilon_{TGdM},$$

where the functional

$$D : d \to D(d) = \begin{cases} 0, & d \le 10^{-6}, \\ \log_{10}(d/10^{-9}), & d > 10^{-6} \end{cases}$$

transforms the quantitative variable dose d_j into a new discrete log-dose variable $D_0 = 0$ (for d_1, d_2, d_3), $D_1 = 4$ (for d_4), $D_2 = 5$ (for d_5), and $D_3 = 6$ (for d_6). Note that in this model, all injection doses less than 10^{-6} mol/l, can be confounded and their effect neglected. Furthermore, in this expression:

- μ is the intercept corresponding to wild type female mice with an injection dose less than 10^{-6};
- γ_{TG} is the interaction between gender (G) and transgenesis (T), and in this particular case, only the interaction between Transgenic and Male is significant leading to $\gamma_{T='Transgenic',G='Male'} \neq 0$;
- β is the coefficient that expresses the overall linear effect of D on the model;
- β_T is the coefficient that expresses the transgenesis dependent linear effect of D, with conventionally $\beta_{T='Wild'} = 0$;
- the mouse random effect u_M has variance σ_M^2, and the dose dependent mouse random effect v_M has variance τ_M^2. Given a mouse M, the correlation between u_M and v_M is denoted ρ and is unknown.
- The variance of residual error ε_{TGdM}, is denoted σ_ε

This model comprises only four prevailing parameters to describe the fixed effect (estimates given in Table 8), and four parameters to describe the variance (component estimates in Table 9). To be more explicit, we give for each mice group the corresponding model equation with the associated fitted values. We expand these final results in Table 10.

3.2.2 Comments

When these results were compared to the null model (for which AIC = 844.6), the LMM shows an overall gain of 125, originated from both the effective random effect modelling and from the fixed effect modelling (ANOVA model). A model that only accounts for the random effect shows an AIC gain of 87, whereas ANOVA model shows an AIC gain of 110. In contrast to the first study, it is mainly the fixed effect characterization that improves model performances. From a physiological viewpoint, this is quite coherent with the modus operandi that suppresses the natural mice heterogeneity when isolating the heart. Nevertheless, a random effect exists, and the variance components estimates clearly show that the variability increases with the dose, but does not depend on gender or on transgenesis.
From a statistical viewpoint though, that means that ignoring random effects and retaining only this fixed effect model would have led to the same estimated parameters and the same conclusions, albeit the corresponding fixed model is theoretically not valid in the presence of repeated measures. However and obviously, this a posteriori "fortunous" conclusion can only be inferred after having properly and jointly characterized the fixed and random effect. Disregarding random effects in the modelling may lead to miss significant parameters, or conversely to propose an over-dimensioned model.

4 Discussion

Many statistical methods of analysis have been previously developed under balanced design assumption for the analysis of repeated-measures studies. At present, it's clear that in order to develop more effective and more powerful observations studies, mixed models methods should be more systematically used. The above examples illustrate how ANOVA should be considered inadequate, when the study design includes incomplete or unbalanced observations. Linear models that merge together random effects and systematic errors are at the origin of many significance tests, including the most commonly used test, namely ANOVA. In linear models, the fixed effects were usually presented as curves, means and clusters, while the dispersion parameter was ascribed to randomness. One important limitation of linear models is that they do not theoretically account for repeated measures, even if, when the design is balanced, the fixed effect estimation may be numerically acceptable. But, for unbalanced or incomplete designs with repeated measures, linear models, such as the traditional ANOVA, definitely have to be avoided in favour of mixed models. Mixed models are specifically built for non independent measures like repeated measures, and hold for any design structure. In physiological settings for instance, situations of unreadable recordings, or incomplete data series due to missing samples, are very common and should naturally resort to mixed models. Moreover, mixed models allow to estimate the treatment (or drug) effect under a large range of variance structures, while ANOVA requiring sphericity conditions (an assumption which is rarely valid in physiological settings) necessitates numerous and debatable correction schemes. Finally, mixed models produce regression coefficients and also predictors that, for each subject, may be either fixed or variable-dependent.

Mixed model methods of analysis have been applied in various setting of experimental research, including the evaluation of dissolution profiles (1), statistical measurement of foetal growth (5), the study of fragile X syndrome (7), or the variability of blood pressure (8), to cite but a few. For example, it was also used by Bagella et al. (2) in psychophysiology, to analyze the electromyographic response of the left brow region to various kind of musical stimuli. The experimental set up includes a large number of missing data which were analysed separately. Data were extremely scattered with a high variance that increases with the mean (as usual in physiology), a temporal autocorrelation, and conditions of sphericity that were not satisfied. Using the entire electromyographic data points, a comparative study evidenced the robustness of the mixed model as compared to the ANOVA model, when randomly deleting observations from the data set. The present cardiovascular studies yield similar conclusions with respect to data suppression. Various softwares are currently available, that can efficiently fit most of mixed models (reviewed in (1, 4, 10)).

The studies conducted here were performed with *R Toolbox,* a freeware software that can be dowloaded from the Comprehensive R Archive Network at http://cran.r-project.org.

Table 1 : Mean of RR (in ms) over the mice set, for each state injection and each measurement time: mean (SE) [corresponding unit number]

	T0 (n=50)	T20 (n=40)	T40 (n=41)	T60 (n=40)
C	114.4 (17.7) [n=50]	95.9 (7.8) [n=12]	98.6 (7.4) [n=13]	102.7 (8.7) [n=12]
A		92.8 (6.0) [n=10]	96.9 (9.4) [n=10]	100.0 (9.5) [n=10]
P		131.3 (25.8) [n=9]	139.3 (33.3) [n=9]	146.6(41.3) [n=9]
AP		123.0 (22.0) [n=9]	128.0 (15.4) [n=9]	130.2 (12.8) [n=9]

Table 2. Akaike Information. Criterion corresponding to the five studied random effect models

Model	M_0	M_M	$M_{M/S}$	M_M^T	$M_{M/S}^T$	$M_{M,M/S}^T$
AIC	1500.9	1490.9	1375.1	1486.9	1366.8	1368.0
BIC	1516.6	1509.7	1397.1	1508.9	1392.0	1396.3

Table 3 Akaike Information Criterion corresponding to the four studied fixed effect models

Fixed effect	$\mu + \alpha_S + \beta T$	$\mu + \alpha_S + \beta T + \gamma_S T$	$\mu + \alpha_S + \beta_T$	$\mu + \alpha_S + \beta_T + \gamma_{ST}$
AIC	1361.2	1359.9	1347.1	1356.2
BIC	1389.4	1397.6	1381.6	1409.6

Table 4. Estimated values with corresponding p-values of the final five significant fixed effect parameters. Estimated values are relative to the intercept value

	Estimated value	p
μ	108.06	$< 10^{-4}$
$\beta_{T='T20'}$	-10.02	$< 10^{-4}$
$\beta_{T='T40'}$	-4.95	$2.5\ 10^{-3}$
$\alpha_{S='P'}$	33.78	$< 10^{-4}$
$\alpha_{S='AP'}$	25.20	10^{-3}

Table 5. Estimated variance components associated to the six random parameters retained in the final linear mixed model

Variance components	σ_M	$\sigma_{M/S}$	$\tau_{M/S}$	σ_ε
Estimates	0.54	15.88	4.39	7.34

Table 6: Fitted values (SD) corresponding to all possible crossings between significant states and times

	T0	T20	T40	T60
Control or Atropine	108.06 (2.92)	98.04 (3.35)	103.12 (3.28)	108.06 (2.92)
Propranolol	-	131.82 (6.64)	136.90 (6.72)	141.85 (6.67)
Atropine + Propranolol	-	123.24 (6.64)	128.32 (6.72)	133.27 (6.69)

Table 7: Coronary perfusion pressure (PP) in mm Hg. Mean(SE). Effects of sodium nitroprusside bolus injection (SNP in Mol/l) CPP (PP) averaged over the mice set, at each factor and each measurement dose. Units are in mm Hg.

SNP	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}
Wild type female	4.61 (1.81)	5.95 (1.76)	7.33 (1.04)	20.19 (4.81)	24.13 (4.01)	25.34 (3.00)
Wild type male	7.50 (2.73)	4.43 (1.82)	7.94 (1.99)	17.31 (2.10)	25.95 (2.48)	24.02 (2.48)
Transgenic female	3.71 (1.88)	8.50 (5.04)	8.03 (1.89)	8.03 (3.37)	15.41 (2.62)	14.78 (8.70)
Transgenic male	1.91 (2.08)	-0.77 (2.22)	2.27 (1.71)	6.43 (2.18)	6.66 (3.24)	8.70 (2.62)

Table 8: Estimated values with corresponding p-values of the final four significant fixed effect parameters. Estimated values are relative to the Intercept value μ

	Estimated value	p
μ (control)	6.39	< 1e-4
$\gamma_{T='Transgenic',G='Male'}$	-5.33	2.8e-3
β	3.29	< 1e-4
$\beta_{T='Transgenic'}$	-1.96	2e-4

Table 9: Estimated variance components associated to the four random parameters retained in the final linear mixed model

Variance components	σ_M	τ_M	ρ	σ_ε
Estimates	0.71	0.85	0.7	4.68

Table 10: Model equations and fitted values describing each one of the eight mice groups

Dose D Group	D_0 ($d \leq 10^{-6}$)	D_1 ($d = 10^{-5}$)	D_2 ($d = 10^{-4}$)	D_3 ($d = 10^{-3}$)
Wild type female	6.39	$6.39 + 3.29\ D_j$ 19.54	22.83	26.12
Wild type male	6.39	$6.39 + 3.29\ D_j$ 19.54	22.83	26.12
Transgenic female	6.39	$6.39 + (3.29 - 1.96)\ D_j$ 11.71	13.04	14.37
Transgenic male	6.39 – 5.33 = 1.065	$(6.39 - 5.33) + (3.29 - 1.96)\ D_j$ 6.38	7.71	9.04

References

1. Adams E., Coomans D., Smeyers-Verbeke J. and Massart D.L. Application of linear mixed effects models to the evaluation of dissolution profiles. *Int. J. of Pharmaceutics* 226: 107-125, 2001.

2. Bagella E., Sloan R.P., and D.J. Heitjan. Mixed-effects models in psychophysiology. *Psychophysiology* 37: 13-20, 2000.

3. Crump S.L. The estimation of components of variance in multiple classifications. *PhD. Thesis,* Iowa State University, Ames, 1947.

4. Edwards L.J. Modern techniques for the anlysis of longitudinal data in biomedical research. *Pediatr. Pulmonol.* 30: 330-344, 2000.

5. Gurrin L.C., Blake K.V., Evan S.F. and Newnham J.P. Statistical measures of foetal growth using linear mixed models applied to the foetal origins hypothesis. *Stat. Med.* 30:3391-409, 2001.

6. Hartley H.O, Rao J.N.K. Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika*, 54: 93-108, 1967.

7. Hessl D., Glaser B., Dyer-Friedman J., Blasey C., Hastie T., Gunnar M. and Reiss A.L. Cortisol and behavior in fragile X syndrome. *Psychoneuroendocrinology* 2: 855-872, 2002.

8. Leary A.C., Donnan P.T., MacDonald T.M. and Murphy M.B. The influence of physical activityon the variability of ambulatory blood pressure. *Am. J. Hypertens.* 13:1067-73, 2000.

9. Mansier P., Médigue C., Charlotte N., Vermeiren C., Coraboeuf E., Deroubaix E., Ratner E., Chevalier B., Clairambault J., Carré F., Dahkli T., Bertin B., Briand P., Strosberg D., and Swynghedauw B. Decreased heart rate variability in transgenic mice overexpressing atrial ß1-adrenoceptors. *Am. J. Physiol*. 271, H1465-1472, 1996

10. Mutapi F., and A. Roddam. p values for pathogens: statistical inference from infectious-disease data. *Lancet Infect. Dis.* 2 : 219-230, 2002.

11. Pinheiro, J. C., Bates, D. M., Mixed-effects models in S and S-Plus, Statistics and Computing, Springer-Verlag New York, 2000.

12. Searle, Casella, McCulloch, Variance components. John Wiley & Sons, 1992.