

Table of contents

TABLE OF CONTENTS.....	1
1 INTRODUCTION TO MIXED-EFFECTS MODELS.....	3
Fixed-effects regression ignoring data clustering	5
Fixed-effects regression including data clustering	12
Fixed-effects regression with dummy variables	15
Random intercept model	17
Intraclass correlation	21

List of tables and figures

TABLE 1.1: DATA FOR 10 HEMODIALYZERS FROM VONESH & CARTER DATA.....	4
TABLE 1.1: DATA FOR 10 HEMODIALYZERS FROM VONESH & CARTER DATA (CONTINUED).....	5
FIGURE 1.1: EXPLORATORY GRAPHS OF RATE VERSUS PRESSURE FOR HEMODIALYZERS	8
FIGURE 1.2: EXPLORATORY GRAPHS OF RATE VERSUS PRESSURE FOR 10 HEMODIALYZERS	9
FIGURE 1.3: FIXED-EFFECTS REGRESSION LINE FOR 10 DIALYZERS	12
TABLE 1.2: REGRESSION RESULTS FOR 10 DIALYZERS: TAKING CLUSTERING INTO ACCOUNT.....	13
FIGURE 1.4: INDIVIDUAL FIXED-EFFECTS REGRESSION LINES FOR 10 DIALYZERS.....	14
TABLE 1.3: RESULTS OF DUMMY VARIABLE MODEL	16
TABLE 1.4: RESULTS OF REGRESSION MODEL WITH DUMMY VARIABLES.....	17
TABLE 1.5: RESULTS OF RANDOM-INTERCEPT MODEL.....	19
FIGURE 1.5: LEVEL-1 RESIDUALS PLOTTED AGAINST LEVEL-1 PREDICTED VALUES .	20
FIGURE 1.6: 95% CONFIDENCE INTERVALS FOR 20 DEVICES	21

1 Introduction to mixed-effects models

Hierarchical structures are often encountered in numerous research areas. Consider, for example, the study of the effect of administering medication, such as an antidepressant, over time to a patient diagnosed with depression. For each patient, the effect of the drug over time can be modeled in terms of the time since the start of treatment, and also in terms of any other information obtained at the time of each measurement during the study. Measures of family support at the time of measurement can also be incorporated into such a model. The outcome would be described as a function of the information collected at the measurement level, and could be viewed as a measurement-level model for each individual patient. However, the gender of the patient, and other characteristics that may influence the outcome but that do not change over time, cannot easily be accommodated in the model proposed, as the model is at a measurement, rather than a patient, level. It may also be of interest to compare patients in terms of their improvement trajectories, which is easier when outcomes are described in terms of patients rather than measurements.

To allow us to study all of these areas of interest simultaneously, a model that acknowledges the data's inherent hierarchical structure (measurements nested within individual patients), and allows the study of both measurement- and patient-level models along with the way these models are related to each other, is needed. As patients may drop out during the study period, the model should also be suitable for the analysis of unbalanced longitudinal data where each individual may be measured at a different number of occasions, or even at different time points.

In this chapter, data from a study described in Vonesh & Carter (1992) that focused on the assessment of high-flux hemodialyzers' *in vivo* ultrafiltration are used to illustrate the need for and basic characteristics of a mixed-effects regression model. While the eventual application of these findings will be in a medical field, the testing of the dialyzers discussed here may be of interest to any researcher who intends modeling repeated measures data.

The ultrafiltration rates of 20 high-flux dialyzers were measured at seven different transmembrane pressures. The unit of measurement for transmembrane pressure was dmHg, and the filtration rate was recorded in mL/hr. These data, also analyzed in Littell, Milliken, Stroup & Wolfinger (1996), are perfectly balanced in that all seven measurements are available for each of the hemodialyzers. The hemodialyzers, machines for filtering impurities from the blood, are the units within which the actual measurements are nested. Data for 10 of the dialyzers are shown in Table 1.1.

Table 1.1: Data for 10 hemodialyzers from Vonesh & Carter data

Device ID	Supply	Pressure	Rate	Device ID	Supply	Pressure	Rate
11.000	1.000	28.500	1.500	16.000	1.000	23.500	3.600
11.000	1.000	52.000	15.400	16.000	1.000	48.000	20.490
11.000	1.000	100.500	32.520	16.000	1.000	101.000	41.880
11.000	1.000	150.000	42.440	16.000	1.000	149.000	49.990
11.000	1.000	198.500	48.570	16.000	1.000	199.000	57.670
11.000	1.000	249.000	53.690	16.000	1.000	248.000	62.480
11.000	1.000	299.500	53.660	16.000	1.000	300.500	62.150
12.000	1.000	29.500	6.420	17.000	1.000	23.500	1.170
12.000	1.000	51.500	20.250	17.000	1.000	48.500	17.680
12.000	1.000	101.000	43.050	17.000	1.000	102.500	39.700
12.000	1.000	148.000	58.110	17.000	1.000	151.500	52.680
12.000	1.000	200.000	61.990	17.000	1.000	199.000	61.800
12.000	1.000	248.000	60.910	17.000	1.000	251.000	61.480
12.000	1.000	300.500	63.600	17.000	1.000	302.000	61.420
13.000	1.000	25.500	3.880	18.000	1.000	26.000	1.890
13.000	1.000	50.000	19.160	18.000	1.000	51.500	18.510
13.000	1.000	98.000	37.650	18.000	1.000	97.000	37.220
13.000	1.000	149.000	47.900	18.000	1.000	150.500	52.350
13.000	1.000	201.500	54.490	18.000	1.000	199.000	60.910
13.000	1.000	251.000	53.170	18.000	1.000	250.000	62.980
13.000	1.000	298.000	59.350	18.000	1.000	299.500	64.770
14.000	1.000	40.000	10.940	19.000	1.000	35.500	10.410

Table 1.1: Data for 10 hemodialyzers from Vonesh & Carter data (continued)

14.000	1.000	47.000	13.470	19.000	1.000	48.000	19.320
14.000	1.000	101.000	35.350	19.000	1.000	102.500	43.770
14.000	1.000	151.500	45.340	19.000	1.000	150.000	51.230
14.000	1.000	198.000	49.440	19.000	1.000	199.000	58.090
14.000	1.000	251.000	53.630	19.000	1.000	250.000	54.090
14.000	1.000	300.000	56.430	19.000	1.000	300.500	62.010
15.000	1.000	29.000	4.050	20.000	1.000	28.000	5.710
15.000	1.000	49.500	16.590	20.000	1.000	50.500	20.500
15.000	1.000	101.500	40.520	20.000	1.000	100.000	39.410
15.000	1.000	152.000	52.840	20.000	1.000	149.000	50.100
15.000	1.000	202.000	60.440	20.000	1.000	200.000	55.160
15.000	1.000	250.000	64.830	20.000	1.000	250.500	61.190
15.000	1.000	297.500	63.830	20.000	1.000	302.000	50.720

Of interest here is the relationship between the ultrafiltration rate, denoted as Rate in Table 1.1, and the associated transmembrane pressure, indicated as Pressure in the table. The blood flow rate, as represented by the column with header Supply, is also of potential interest.

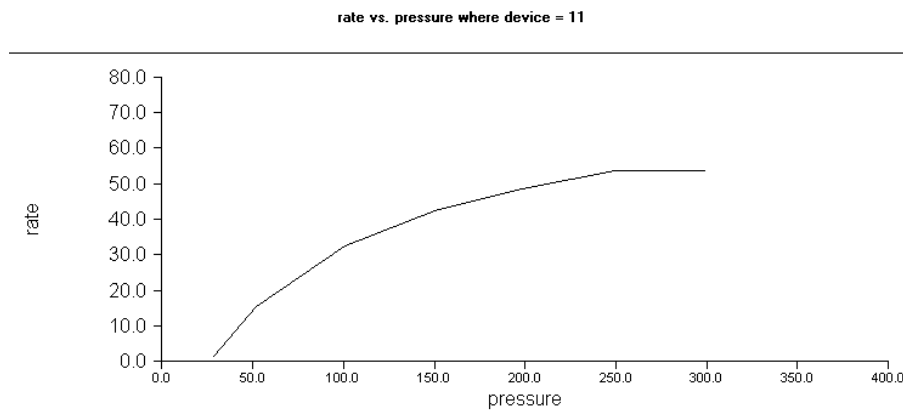
The data as a whole can be viewed as having a hierarchical structure, with measurement-related characteristics of the hemodialyzers at seven measurement occasions; all measurements for each dialyzer are therefore *nested* within that dialyzer. The dialyzers, in turn, form the next level of the hierarchy, and any machine-specific characteristics may be used as potential predictors at this level.

Fixed-effects regression ignoring data clustering

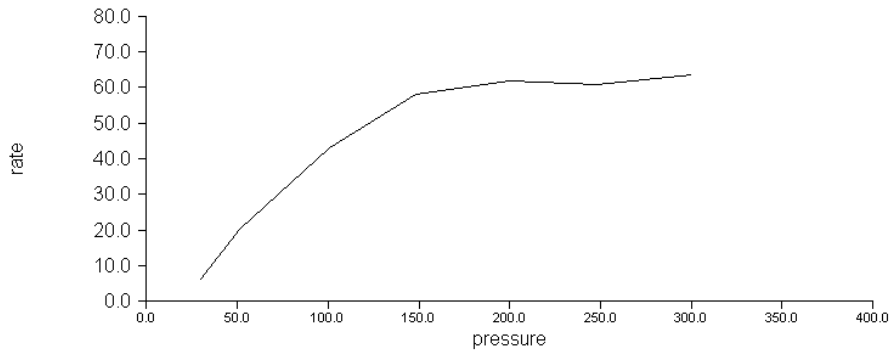
Before proceeding with a mixed-effects analysis of these data, we first look at a fixed-effects analysis that ignores the clustering of measurements within dialyzers. Note that SuperMix can be used for this purpose, and that the analysis is essentially equivalent to performing a traditional multiple linear regression analysis using maximum likelihood, and not least squares, estimation.

Using the information for the second set of 10 dialyzers, for which 70 measurements were available, we now explore the relationship between the Rate of filtration, which serves as our outcome variable, and the transmembrane Pressure at which the measurement was made. Line plots of this relationship for some of the dialyzers are shown in Figure 1.1. These graphs were obtained using SuperMix's exploratory graphs option. Detailed information on how to create such graphs are given elsewhere in the manual.

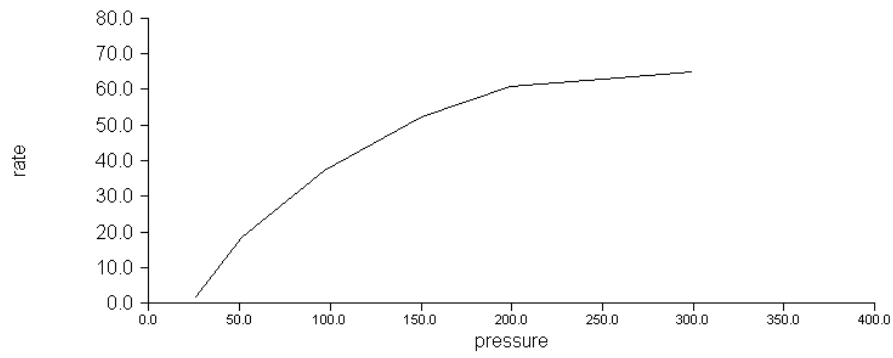
It is clear from these graphs that the relationship between the observed Rate and Pressure at which the measurement was made will be inadequately described by a first-order polynomial. For dialyzer 12 the slope of the line is steep initially, but the curve flattens out at a pressure of about 100 dmHg. This trend is not as clearly observed for the other dialyzers. Also, there seems to be evidence of differences in the rates of dialyzers 18, 19, and 20 towards the higher end of the pressure scale. We conclude that a higher-order polynomial will probably offer a better description of the relationship, and that it may also be wise to make provision for differences between devices (dialyzers).



rate vs. pressure where device = 12



rate vs. pressure where device = 18



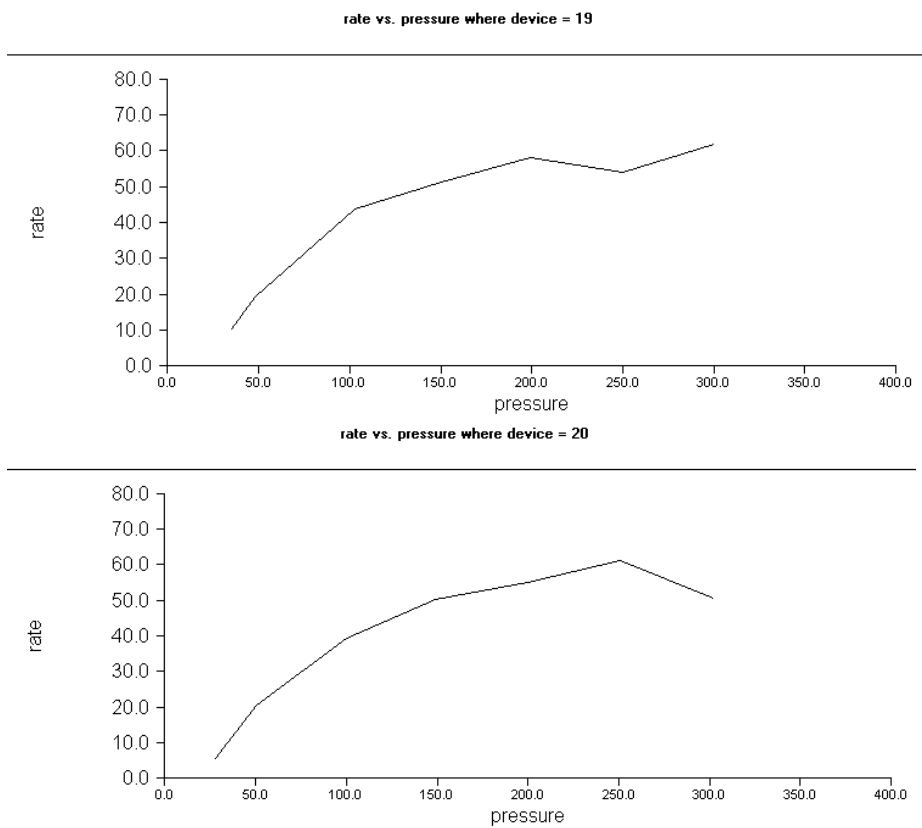


Figure 1.1: Exploratory graphs of rate versus pressure for hemodialyzers

Figure 1.2 represents the same lines for all ten dialyzers simultaneously. While there seems to be little difference in their behavior at the lower level of the pressure scale, the divergence in the plotted lines at higher pressure levels can be seen clearly.

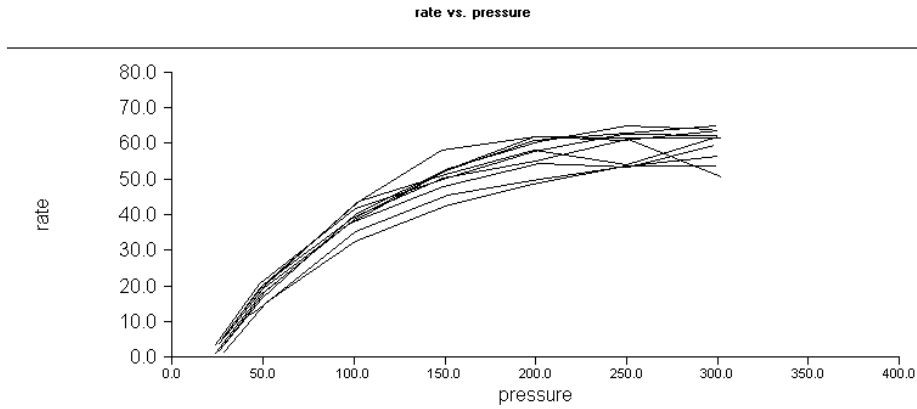


Figure 1.2: Exploratory graphs of rate versus pressure for 10 hemodialyzers

In terms of the variables shown in Table 1.1, we now fit a model of the form

$$y_{ij} = \beta_0 + \beta_1(\text{PRESSURE})_{ij} + \beta_2(\text{PRESSURE})_{ij}^2 + e_{ij} \quad (1.1)$$

where y_{ij} denotes the Rate measurement at time j ($j = 1, 2, 3, \dots, 7$) for hemodialyzer i . $(\text{PRESSURE})_{ij}$ indicates the associated transmembrane pressure, $(\text{PRESSURE})_{ij}^2$ the squared value of the pressure, and e_{ij} measurement error. The coefficients β_0 , β_1 , and β_2 are the fixed, but unknown, parameters to be estimated. The e_{ij} are assumed to have a normal distribution, with mean 0 and variance σ^2 .

For this analysis, we obtain estimates of β_0 , β_1 , and β_2 of -6.5847 , 0.5281 and -0.0011 respectively. The estimated Rate is plotted over time in Figure 1.3. In addition, an estimate of σ^2 of 41.34095 was obtained. The results show that the average predicted Rate, $\hat{\beta}_0$, at a pressure of zero is -6.5847 . However, a value of 0 is outside the range of 23.50 to 303.00 of observed pressure values. As such, the interpretation of the estimate of β_0 in this context is difficult, and we would rather look at the predicted rate for the lowest observed pressure. Another alternative is to transform the values of the variables Pressure and Pressure² in such a way that

interpretation of the intercept estimate is meaningful. Examples of such transformations are given in the chapters to follow.

The coefficient representing the effect of the predictor Pressure, $\hat{\beta}_1$, indicates a predicted increase in Rate with increased pressure: an increase of 0.52807 mL/hr in the Rate is expected for each increase of 1 dmHg in transmembrane pressure. The coefficient β_1 is commonly referred to as a "slope" coefficient, as it indicates both the direction of the relationship between the predictor and the outcome, and the magnitude of the expected change in outcome associated with changes in the predictor.

Similarly, the relationship between the squared values of transmembrane pressure (Pressure²) and the ultrafiltration rate is estimated to be negative: higher values of pressure are predicted to lead to lower predicted rates. The statistical significance of this estimated coefficient indicates that the relationship between pressure and filtration rate is not truly linear, and that the use of a higher-order polynomial may provide a better description of the data. However, while the estimates of β_1 and β_2 are of interest individually, when evaluating the relationship between the transmembrane pressure and the ultrafiltration rate, both estimates should be taken into account. A increase of 1 dmHg in pressure will lead to a change in expected filtration rate of $0.52807(1) - 0.0011(1) = 0.52697$. From this result, we conclude that while the filtration rate and pressure generally shows a positive relationship, this relationship is bound to change with increased pressure. The higher the pressure, the bigger the impact of the estimate of β_2 in the prediction of the rate through use of the formula

$$\hat{y}_{ij} = \hat{\beta}_0 + \hat{\beta}_1(\text{PRESSURE})_{ij} + \hat{\beta}_2(\text{PRESSURE})_{ij}^2$$

The lowest observed pressure is 23.50, and the predicted rate of filtration is thus

$$\begin{aligned}
\hat{y}_{ij} &= \hat{\beta}_0 + 0.52807(\text{PRESSURE})_{ij} - 0.0011(\text{PRESSURE})_{ij}^2 \\
&= -6.5847 + 0.52807(23.5) - 0.0011(23.5)^2 \\
&= -6.5847 + 12.4096 - 0.6075 \\
&= 5.2174.
\end{aligned}$$

For the highest observed pressure of 303, the predicted filtration rate follows as

$$\begin{aligned}
\hat{y}_{ij} &= -6.5847 + 0.52807(303) - 0.0011(303)^2 \\
&= -6.5847 + 160.0052 - 100.9899 \\
&= 52.4306.
\end{aligned}$$

The fixed-effects regression line over all measurements is shown in Figure 1.3 below.

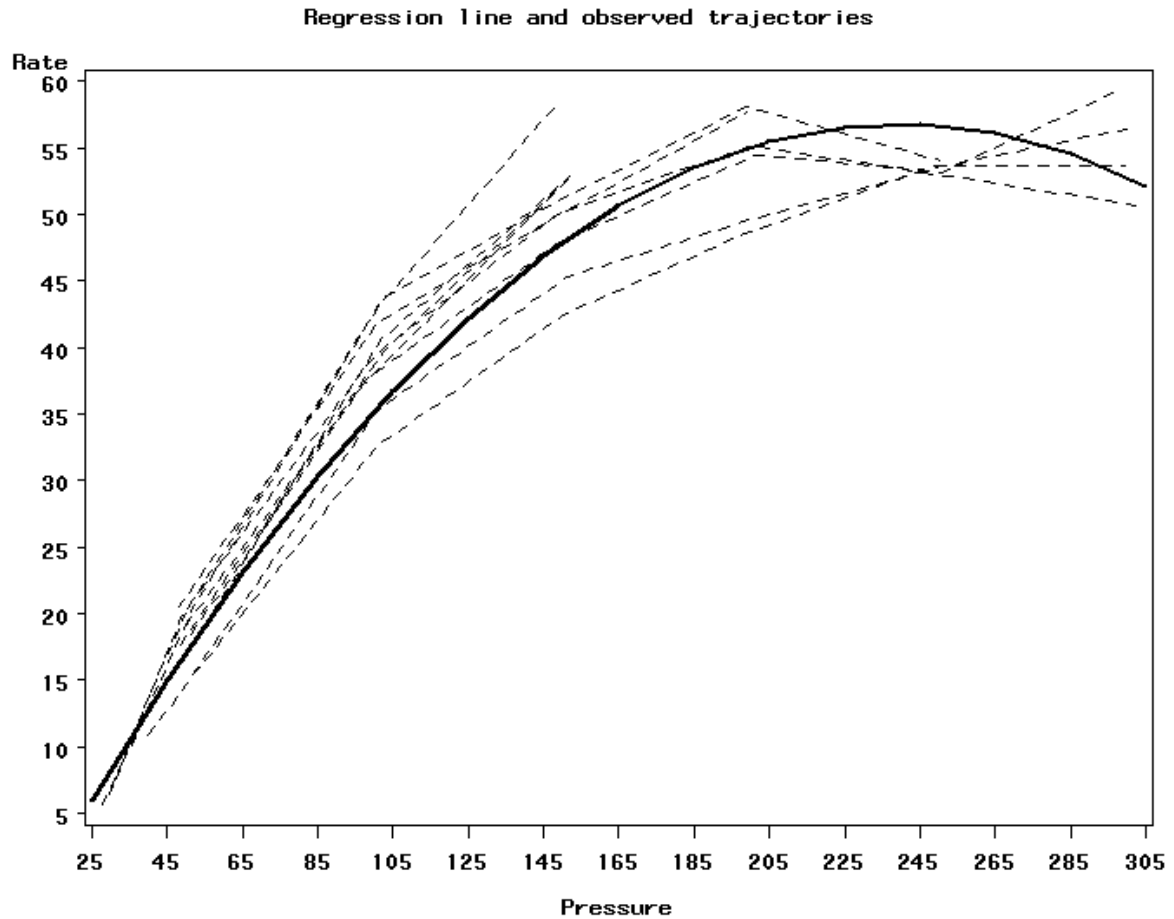


Figure 1.3: Fixed-effects regression line for 10 dialyzers

Fixed-effects regression including data clustering

As noted by Hedeker, Gibbons & Flay (1994) and others, ignoring the data clustering often results in statistical tests which are too liberal, resulting in falsely rejecting the null hypothesis too often. In terms of our data, where multiple measurements "belong" to each dialyzer, it is reasonable to expect that

measurements for a given dialyzer may be more similar to each other than to any other measurement, regardless of the dialyzer it was obtained for. Thus, it may be reasonable to assume that the measurements for a given dialyzer may be correlated. In addition, if it is indeed true that the transmembrane pressure applied impacts on the transfer rate, ignoring the clustering effect may lead to erroneous conclusions concerning the relationship between pressure and transfer rate.

To start addressing these concerns, we modify the previous model to take the clustering of measurements within dialyzers into account. We do so by fitting a line similar to that given in Equation (1.1) for each individual dialyzer. Table 1.2 shows the estimates of β_0 and β_1 for individual dialyzers, and Figure 1.4 a graphical representation of the results.

Table 1.2: Regression results for 10 dialyzers: taking clustering into account

Device	Intercept	Pressure	(Pressure) ²
11	-9.206	0.486	-0.001
12	-9.024	0.629	-0.001
13	-5.115	0.500	-0.001
14	-5.008	0.454	-0.001
15	-10.885	0.602	-0.001
16	-5.255	0.537	-0.001
17	-10.614	0.608	-0.001
18	-10.582	0.590	-0.001
19	-3.645	0.520	-0.001
20	-7.911	0.589	-0.001
overall	-6.585	0.528	-0.001

The estimated coefficients for the intercepts and time slopes of the dialyzers ($\hat{\beta}_0$ and $\hat{\beta}_1$ respectively) in Table 1.2 show that the predicted intercepts of dialyzers differ considerably. Device/dialyzer number 15 has a predicted initial transfer rate of -10.885, which is considerably lower than the predicted initial rate of -3.645 for dialyzer 19. Recall that in the previous analysis, we obtained a value of -6.585 for $\hat{\beta}_0$, which does not provide an adequate description of the initial status of any of the dialyzers except perhaps dialyzers 13, 14, 16, and 20. A "one size fits all" policy for obtaining an estimate of the initial status of patients is clearly inadequate, and does not describe the initial status for individual dialyzers satisfactorily.

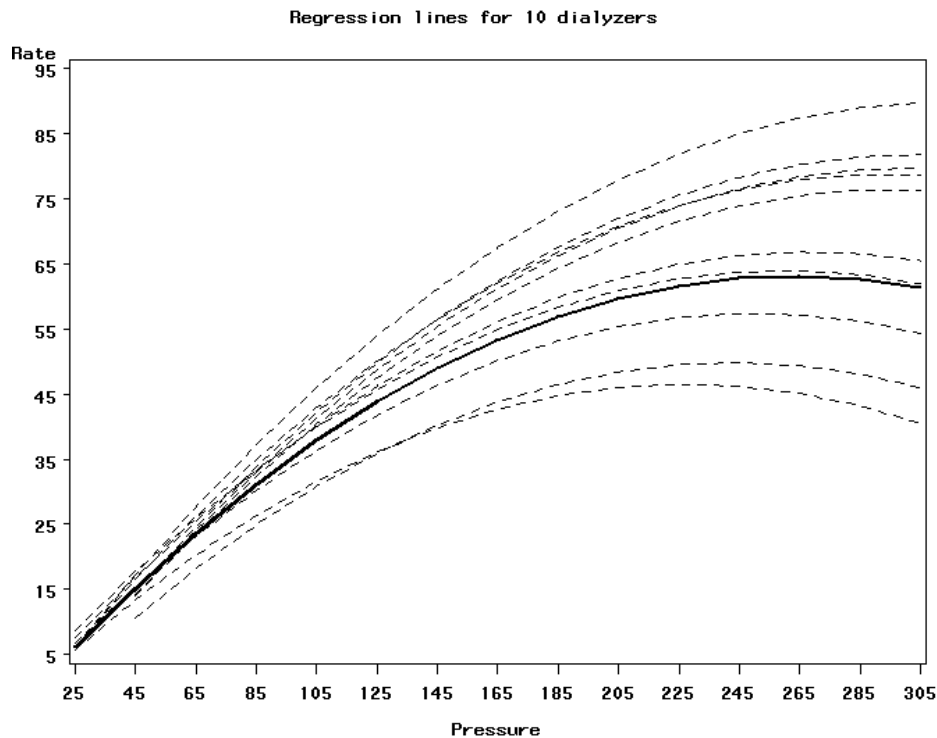


Figure 1.4: Individual fixed-effects regression lines for 10 dialyzers

This conclusion is also apparent from Figure 1.4. While the differences in transfer rates at the lower end of the pressure range are not as clear from the graph as they are in Table 1.2, the graph indicates even larger differences between the dialyzers at high transmembrane pressure. Not only will individual differences in initial transfer rate between devices have to be addressed, but differences in their rates of transfer over the range of applied transmembrane pressure will have to be accommodated in the model.

Fixed-effects regression with dummy variables

Up to this point we have considered two approaches for the modeling of the transfer rates. In the first, all the data were pooled and a common regression model was fitted to the data. In the second approach, a regression line was fitted to each dialyzer's measurements. A summary of the estimated intercepts and slopes showed substantial between-dialyzer variation. The disadvantage of the second approach is that ten separate regression models are fitted. Ideally, a researcher would want to fit a single model that conveys information about between-subject variability.

One approach would be to do a regression analysis with dummy variables. Table 1.3 below shows the data for the first and last dialyzers. We use a dummy variable to represent each dialyzer, coded as follows:

$$D_j = 1 \text{ for dialyzer } j, j = 1, 2, \dots, 10 \\ = 0 \text{ otherwise.}$$

The following regression model is fitted to the data:

$$\text{RATE}_{ij} = \alpha_0 + \alpha_1 (D_1)_{ij} + \alpha_2 (D_2)_{ij} + \dots + \alpha_9 (D_9)_{ij} + \alpha_{10} (\text{PRESSURE})_{ij} + e_{ij}.$$

This model allows for the estimation of individual intercept coefficients, but a common slope parameter α_{10} .

Table 1.3: Results of dummy variable model

device	rate	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Pressure
11	1.50	1	0	0	0	0	0	0	0	0	0	28.5
11	15.40	1	0	0	0	0	0	0	0	0	0	52.0
11	32.52	1	0	0	0	0	0	0	0	0	0	100.5
11	42.44	1	0	0	0	0	0	0	0	0	0	150.0
11	48.57	1	0	0	0	0	0	0	0	0	0	198.5
11	53.69	1	0	0	0	0	0	0	0	0	0	249.0
11	53.66	1	0	0	0	0	0	0	0	0	0	299.5
20	5.71	0	0	0	0	0	0	0	0	0	1	28.0
20	20.50	0	0	0	0	0	0	0	0	0	1	50.5
20	39.41	0	0	0	0	0	0	0	0	0	1	100.0
20	50.10	0	0	0	0	0	0	0	0	0	1	149.0
20	55.16	0	0	0	0	0	0	0	0	0	1	200.0
20	61.19	0	0	0	0	0	0	0	0	0	1	250.5
20	50.72	0	0	0	0	0	0	0	0	0	1	302.0

Table 1.4 contains a summary of the results of this analysis. With the exception of the first and fourth dialyzers (represented by the dummy variables D1 and D4), the estimated coefficients associated with the individual dialyzers are all significantly different from zero at a 5% level of significance. From these results, we expect transfer rates for the second device to be much higher than for the first device, as reflected by the parameter estimates of 14.7153 and 5.2222 respectively. The transmembrane pressure also has a significant and positive relationship to the rate of transfer: for each increase of 1 dmHg in pressure, the rate of transfer is expected to be 0.1959 ml/hr higher.

Table 1.4: Results of regression model with dummy variables

Variable	Parameter estimate	Standard error	t-Value	Pr > t
D1	5.2222	3.9381	1.33	0.1899
D2	14.7153	3.9385	3.74	0.0004
D3	9.3364	3.9343	2.37	0.0209
D4	7.3311	3.9463	1.86	0.0682
D5	13.0271	3.9408	3.31	0.0016
D6	12.6855	3.9312	3.23	0.0020
D7	12.1007	3.9381	3.07	0.0032
D8	12.6124	3.9347	3.21	0.0022
D9	12.3179	3.9439	3.12	0.0028
D10	10.1676	3.9397	2.58	0.0124
Pressure	0.1959	0.0117	16.68	<.0001

Although this model is a compromise between the models for pooled data and separate models for dialyzers' data, the number of parameters to be estimated is proportional to the number of dialyzers and does not allow for the estimation of individual slopes. These issues have led researchers over time to develop mixed-effects models.

Random-intercept model

From the results of the previous models, we concluded that it is not reasonable to assume that the initial transfer rates of dialyzers, or their change in transfer rate with increased transmembrane pressure, can be described adequately by average intercept and slope estimates while the clustering of measurements within individual dialyzers was ignored. While the second of these analyses, where fixed-effects regression lines were fitted for each dialyzer and thus the clustering of measurements was acknowledged, provided better information per dialyzer, neither of these models allows us to obtain average intercept or slope coefficients while simultaneously incorporating the effect of measurements nested within individual devices.

To study differences in the behavior of dialyzers with pressure changes, while acknowledging the clustering of measurements and allowing for differences between devices in initial transfer rate, a random-effects model is needed. From the results obtained thus far, we will have to accommodate not only differences in initial status between dialyzers, but also differences in the slopes of the rates over the range of applied transmembrane pressure.

We start by specifying a model which takes clustering of measurements within dialyzers into account, while allowing the initial transfer rate to vary from device to device. This model, a so-called random-intercept model, contains both fixed and random effects, and can be expressed as

$$y_{ij} = \beta_0 + \beta_1(\text{PRESSURE})_{ij} + \beta_2(\text{PRESSURE})_{ij}^2 + v_{i0} + e_{ij} \quad (1.2)$$

where y_{ij} denotes the Rate measurement at measurement j ($j = 0, 1, 2, 3, 4, 5, 6,$ or 7) for dialyzer i , $(\text{PRESSURE})_{ij}$ the associated transmembrane pressure, $(\text{PRESSURE})_{ij}^2$ the squared value of $(\text{PRESSURE})_{ij}$, and e_{ij} measurement error. The coefficients β_0 , β_1 and β_2 are the fixed, but unknown, parameters to be estimated. The coefficient v_{i0} , in contrast, denotes a random parameter, and represents the amount by which the intercept of dialyzer i differs from the average (fixed) intercept for all devices, as represented by β_0 . By including v_{i0} , we allow intercepts to vary randomly over the dialyzers. We assume that v_{i0} is normally distributed with mean 0 and variance $\phi_{(2)}$ and that the e_{ij} , too, as in the first model, have a normal distribution with mean 0 and variance σ^2 for all dialyzers. In contrast to the model in (1.1), where all unexplained variations in transfer rates were captured by e_{ij} , the current model assumes that there are two potential sources of unexplained variation: variation between measurements as represented by e_{ij} , and variation between dialyzers in terms of their intercepts, as represented by v_{i0} . Viewing the measurements as the lowest level of a nested structure in our data, with measurements nested within devices, we refer to σ^2 as the level-1 (measurement-

level) variance and to $\phi_{(2)}$ as the level-2 (dialyzer-level) variance. In fitting this model, data from all 20 devices are used. The results of the analysis are reported in Table 1.5. All of the estimated coefficients are statistically significant at a 5% level of significance. We see that the rate of transfer is expected to increase with an increase in pressure. However, as pressure increases, the squared value of pressure increases quickly, and the small negative coefficient for this will lead to larger decreases in transfer rate at high pressures. At first glance, these estimates indicate a somewhat nonlinear curve.

Table 1.5: Results of random-intercept model

Parameter	Estimate	Standard error
Intercept	-6.56547	1.56214
Pressure	0.52792	0.01840
Pressure ²	-0.00114	0.00006
var(v_{i0})	16.28786	6.29943
var(e_{ij})	25.05420	3.23435

What is really interesting, and something we have not been able to look at previously, is the amount of variation within and between devices. While most of the variation is at measurement level, i.e. within devices, as indicated by $\text{var}(e_{ij}) = 25.0542$, there is a sizable amount of variation in the intercepts of the devices themselves. As this estimated coefficient is statistically significant, it indicates that it is not adequate to try and describe the intercepts of the devices using a single, common fixed effect as we have done previously. If we had more characteristics of the individual devices, these could have been added to our current model in an attempt to explain away the variation in device intercepts. Likewise, we could have used any other type of measurement made at the measurement occasions to explain more of the residual variation. In Chapter XXX, models with a continuous outcome are described in which the use of additional characteristics at both levels is illustrated.

In addition to these estimates, which describe the average estimated intercept and slope over all devices, we also obtain estimates for the unique deviations from the intercept associated with each of the individual devices. The estimates of the deviations of the predicted from the observed values are depicted graphically in Figure 1.5. The residuals associated with devices 11 and 20 are highlighted: residuals for device 11 are shown as square black boxes, and those for device 20 as asterisks. We see that almost all the residuals are within a (-10,10) interval. For device 11, the residuals are closer to zero in value at lower transfer rates, but vary quite a bit more above a transfer rate of 40. The residuals for device 20, however, vary more over the entire rate of transfer range.

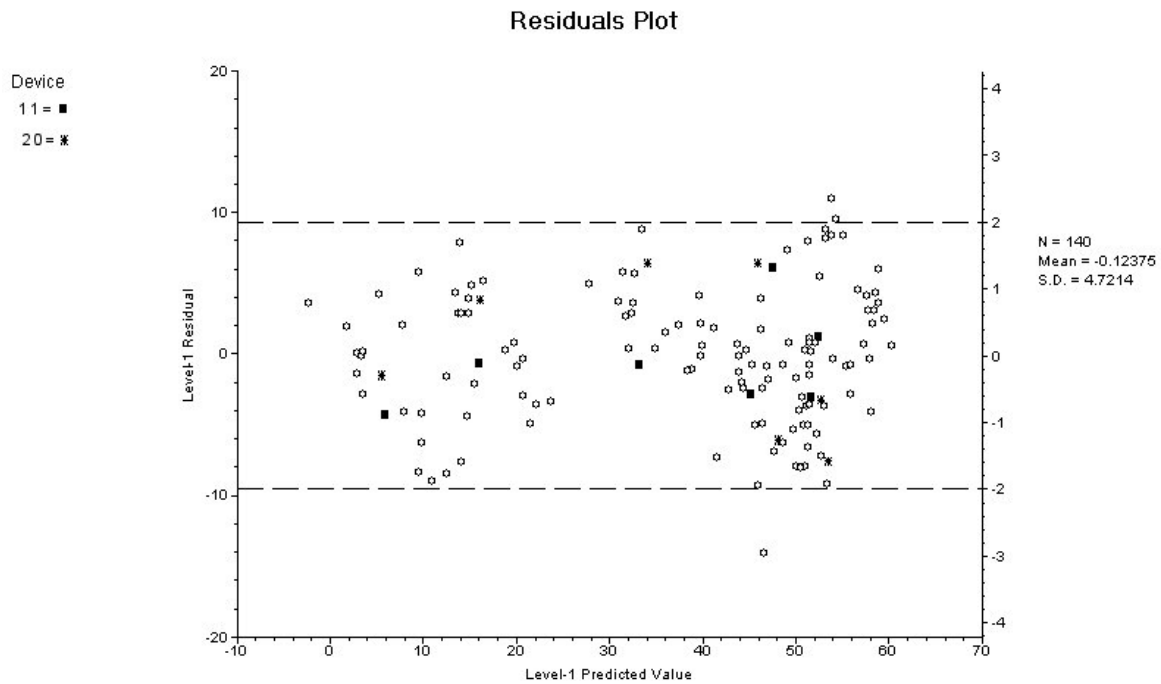


Figure 1.5: Level-1 residuals plotted against level-1 predicted values

Another way to look at these results is to inspect confidence intervals for the deviations of the device intercepts from the estimated value of -6.565 . These are shown in Fig 1.4. The units appear in numerical order, and we can see that the 95% confidence interval for the intercept of device 20 is approximately centered above 0, while that of device 11 is centered below zero. Looking at the confidence intervals for devices 1 and 12, our result that there is significant variation in the device intercepts makes sense.

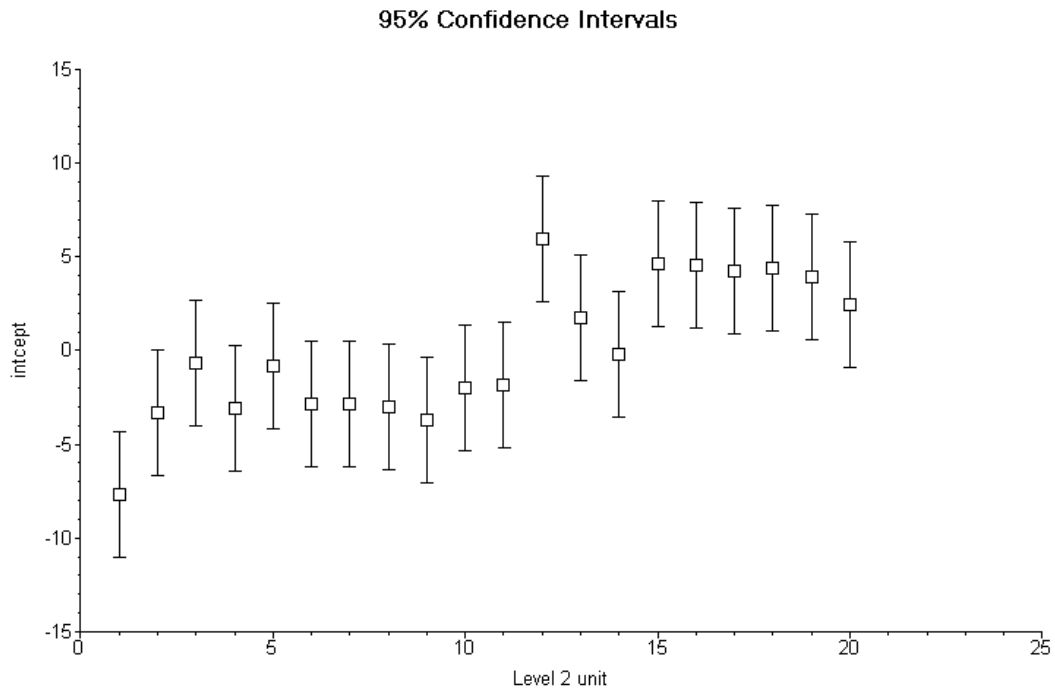


Figure 1.6: 95% confidence intervals for 20 devices

Intraclass correlation

The intraclass correlation is a measure of the degree of dependence of the higher-level units, in this case the devices. It is realistic to assume that measurements from the same device are more alike with respect to certain traits than measurements from different devices.

For data having a two-level hierarchical structure, the intraclass correlation ρ is defined as the proportion of the variance in the outcome variable that is between the second-level units:

$$\rho = \frac{\text{between group variability}}{\text{between group variability} + \text{within group variability}}$$

In the current example, we obtain $\hat{\rho}$ as

$$\hat{\rho} = \frac{16.28786}{16.28786 + 25.05420} = 0.39398.$$

As pointed out by Kreft and de Leeuw (1998), if intraclass correlation is present, as is usually the case when we are dealing with clustered data, the assumption of independent observations in the traditional linear model is violated. They also pointed out that tests of significance lean heavily on the number of independent observations and that the existence of intraclass correlation makes the test of significance in traditional linear models too liberal. Barcikowski (1981) shows that in most applications of analysis of variance, the standard errors of the parameter estimates will be underestimated and that even a small intraclass correlation can inflate the alpha level substantially.

While the random-intercept model has allowed us to accommodate some of our modeling concerns for an unbalanced data set such as the nesting of measurements within devices and allowing intercepts to vary over devices, other concerns remain. From the results shown in Table 1.5, we know that there is a sizable amount of variation between devices, variation that may be explained by the inclusion of additional device characteristics in the model. To address these concerns, extended models are required. Examples of such models, based on the Reisby data, are shown in detail in Section XX.X.