SOME CAPABILITIES AND LIMITATIONS OF
MULTIPLE REGRESSION ANALYSIS:
APPLICATION TO CANINE CORONARY BLOOD FLOW
A. M. S. BLACK AND P. FOEX

SUMMARY
Changes in inspired oxygen and carbon dioxide concentration produced concurrent changes in systemic and coronary haemodynamics in 39 anaesthetized dogs. These changes provided the data base for an exercise in multiple regression analysis using SPSS packages. The physiological problem was to identify influences on coronary blood flow additional to those associated with cardiac work or oxygen consumption. Such influences were identifiable to an extent which was probably limited as much by experimental design as by the underlying physiology. Multiple regression relationships described the integrated response of coronary blood flow to the experimental challenge more completely than any univariate approach. The exercise provided useful insights into a multifactorial problem and the need for deliberate experimental design in attempting to answer specific physiological questions.

Multiple regression analysis is one small facet in a large field which has now acquired the name of "Knowledge Engineering". This embraces the use of computational methods for data processing, problem solving and decision making, and has applications in the medical field (Shortcliffe, Buchanan and Feigenbaum, 1979). Multiple regression analysis is a form of pattern recognition technique in which a data set is examined for similarities to a multiple regression model. The more flexible the model, the easier will the recognition be.

Although the technique is a relatively modest extension from single regression analysis, it is probably under-used, considering its applicability to multifactorial clinical problems (Black and Harris, 1982; Black, Harris and Judson, 1982). This paper uses multiple regression analysis as a relatively straightforward exercise to account for changes in experimentally measured coronary blood flow in terms of changes in a number of concurrently measured variables. It is always reassuring if the results of applying an unfamiliar technique can be "calibrated against" existing experience.

Foex and his co-workers have already published some of their observations on the relationship between changes in canine coronary blood flow and changes in haemodynamics, myocardial oxygen consumption and arterial and venous blood-gas composition (Foex et al., 1976; Foex and Ryder, 1979, 1981; Burt and Foex, 1979; Foex, Ryder and Bennett, 1980). Their experimental design kept the heart in situ, did not separate the coronary and systemic circulations and did not exercise independent control over cardiac work and the composition of the blood supplying the myocardium. One of their observations was of an increase in coronary blood flow with hypercapnia which was disproportionate to the accompanying increase in myocardial oxygen consumption. In this respect, their conclusions, drawn from conventional univariate analysis, were consistent with those of Vance and others (1979), and different from those of Van den Bos, Drake and Noblè, (1979). Their data set is an ideal basis on which to examine the capabilities and limitations of a more demanding form of analysis.

METHODS
Experimental techniques
Thirty-nine dogs were studied under halothane anaesthesia. Inspired $PO_2$ and $PCO_2$ were changed against a background of artificial hyperventilation of the lungs. Measurements were made of arterial, left atrial and left ventricular pressures, aortic and left circumflex coronary artery blood flows and arterial and coronary venous blood-gases and pH. Foex and colleagues have described the details of surgery and instrumentation in the publications quoted above.

Pooling four experimental series, there were 334 experimental impositions of varying combinations of inspired oxygen and carbon dioxide (table I).
These gave rise to corresponding combinations of high, normal or low $P_{O_2}$ with high, normal or low $P_{CO_2}$ in the arterial blood. One hundred and twenty-one of the observations were made in the presence of beta blockade induced either by metoprolol $1 \text{ mg kg}^{-1}$ or oxprenolol $0.3 \text{ mg kg}^{-1}$.

When the responses to each of the imposed conditions in table I had stabilized, measurements of coronary blood flow were made, and were accompanied by measurements of systemic haemodynamics and of arterial and coronary venous blood-gas compositions.

### Data analysis: primary and derived variables

The aortic pressure trace gave systolic, diastolic and mean values. The left ventricular trace yielded left ventricular end diastolic pressure (LVEDP) and left ventricular $dP/dt$. The aortic blood flow trace yielded stroke volume ($SV$) and aortic blood acceleration ($dQ/dt$). Simple products of pressure and flow terms gave conventional measures of cardiac work and power ($SV \times$ mean arterial pressure = left ventricular stroke work ($LVSW$); peak flow $\times$ systolic arterial pressure = peak left ventricular power (PLVP)).

Coronary arterial and venous $P_{O_2}$ and pH and the equations of Rossg and Cain (1966) were used to derive arterial and venous oxygen content (ml dl$^{-1}$), and the product of arteriovenous oxygen content difference and the measured coronary blood flow gave the oxygen consumption of that part of the myocardium supplied by the left circumflex coronary artery.

Acid–base exchange between blood and myocardium was crudely derived in an attempt to allow any effects of changing lactate metabolism to manifest themselves. The net rate of “base uptake” or “acid output” by the myocardium was calculated as the product of coronary blood flow with that part of the arteriovenous difference in standard bicarbonate which was not caused by the Haldane effect. The standard bicarbonates were estimated using the $P_{CO_2}$ and pH values, and (by default of any better), the “normal” buffer line for human blood (Siggaard-Andersen, 1966). The contribution of the Haldane effect was estimated, in mequiv/litre of base (Siggaard-Andersen, 1966), as:

$$0.27 \times \text{Hb}(S_{O_2} - S_{V_2})$$

Beta blockade was indicated by a dummy variable (BK) with a value of unity in the presence, and zero in the absence, of beta blockade (Armitage, 1971). No distinction was made between the effects of metoprolol and oxprenolol. The set of observations from each dog was accompanied by the heart weight and body weight of the dog.

A set of 19 primary and derived variables was taken from these measurements. These are listed with their symbols and reference numbers in table II.

### The regression process: definition of terms

Multiple regressions were carried out using the SPSS package (Nie et al., 1975), available on the Oxford University ICL 1906A computer.

The total corrected sum of squares of the flow observations is the sum of squares of the differences between the individual flow observations and the mean of the 334 observations. This will be abbreviated to TCSS for the purpose of this paper. A regression of flow observations, ($F_0$), with a single independent variable, ($X$) determines a relationship,

$$F_t = C + M.X,$$

such that the sum of squares of the quantities ($F_0 - F_t$) is reduced to a minimal “residual sum of squares”. The difference between this and the TCSS is the sum of squares of the regression. The correlation coefficient ($r$) is the square root of the ratio of regression sum of squares to the TCSS.

The multiple regression of observed flow ($F_0$) determines a relationship between flow and more than one other variable, ($X_1, X_2, X_3, \ldots, X_n$).

$$F_t = C + M_1.X_1 + M_2.X_2 + M_3.X_3 + \ldots + M_n.X_n$$

A combination of values of $C, M_1, M_2, M_3, \ldots, M_n$ is found such that the sum of the quantities ($F_0 - F_t$)$^2$ (residual sum of squares) is again minimized. The

### Table I. The combinations of oxygen and carbon dioxide conditions considered in the study. The figures in parentheses indicate the numbers of observations obtained in the presence of beta blockade.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Hypoxemia</th>
<th>Normoxemia</th>
<th>Hypoxemia</th>
<th>$P_{CO_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypercapnia ($7-9% CO_2$)</td>
<td>9</td>
<td>10</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>Hypercapnia ($3-4% CO_2$)</td>
<td>7 (6)</td>
<td>34 (8)</td>
<td>19 (9)</td>
<td>4</td>
</tr>
<tr>
<td>Hypoxemia ($0% CO_2$)</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Downloaded from https://academic.oup.com/bja/article-abstract/54/12/1319/257159 by guest on 25 December 2018
CANINE CORONARY BLOOD FLOW

Table II. The 77 variables generated from the measured variables

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Symbol</th>
<th>Variable</th>
<th>Reciprocal</th>
<th>Square</th>
<th>Product terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body weight</td>
<td>BW</td>
<td>1</td>
<td></td>
<td></td>
<td>47 BW.MVO2</td>
</tr>
<tr>
<td>Heart weight</td>
<td>HW</td>
<td>2</td>
<td></td>
<td></td>
<td>48 HW.MVO2</td>
</tr>
<tr>
<td>Arterial blood-gases</td>
<td>PBO2</td>
<td>3</td>
<td>20</td>
<td>34</td>
<td>49 PBO2.PACO2</td>
</tr>
<tr>
<td></td>
<td>PaCO2</td>
<td>4</td>
<td>21</td>
<td>35</td>
<td>50 PaCO2.MVO2</td>
</tr>
<tr>
<td></td>
<td>pHa</td>
<td>5</td>
<td></td>
<td></td>
<td>51 PaCO2.MVO2</td>
</tr>
<tr>
<td>Venous blood-gases</td>
<td>PvO2</td>
<td>6</td>
<td>22</td>
<td>36</td>
<td>52 PvO2.BK</td>
</tr>
<tr>
<td></td>
<td>pVCO2</td>
<td>7</td>
<td>23</td>
<td>37</td>
<td>53 pVCO2.BK</td>
</tr>
<tr>
<td></td>
<td>pHv</td>
<td>8</td>
<td>24</td>
<td></td>
<td>54 pVCO2.PACO2</td>
</tr>
<tr>
<td>Oxygen consumption</td>
<td>MVO2</td>
<td>9</td>
<td>25</td>
<td>38</td>
<td>55 MVO2.BK</td>
</tr>
<tr>
<td>Acid-base exchange</td>
<td>Ac-Bas</td>
<td>10</td>
<td>26</td>
<td>39</td>
<td>56 MVO2.Ac-Bas</td>
</tr>
<tr>
<td></td>
<td>BK</td>
<td>11</td>
<td></td>
<td></td>
<td>57 MVO2.MVO2</td>
</tr>
<tr>
<td>Heart rate</td>
<td>HR</td>
<td>12</td>
<td>27</td>
<td>40</td>
<td>58 MVO2.Ac-Bas</td>
</tr>
<tr>
<td>LV end-diastolic</td>
<td>LVEDP</td>
<td>13</td>
<td>28</td>
<td>41</td>
<td>59 MVO2.Ac-Bas</td>
</tr>
<tr>
<td>LV dP/dt</td>
<td>dP/dt</td>
<td>14</td>
<td>29</td>
<td>42</td>
<td>60 MVO2.BK</td>
</tr>
<tr>
<td>Cardiac output</td>
<td>Q</td>
<td>15</td>
<td></td>
<td></td>
<td>61 MVO2.BK</td>
</tr>
<tr>
<td>Aortic acceleration</td>
<td>dQ/dt</td>
<td>16</td>
<td>30</td>
<td>43</td>
<td>62 MVO2.PVO2</td>
</tr>
<tr>
<td>LV stroke work</td>
<td>LVSW</td>
<td>17</td>
<td>31</td>
<td>44</td>
<td>63 MVO2.PvO2</td>
</tr>
<tr>
<td>Peak LV Power</td>
<td>PLVP</td>
<td>18</td>
<td>32</td>
<td>45</td>
<td>64 MVO2.PvO2</td>
</tr>
<tr>
<td>Mean arterial pressure</td>
<td>MAP</td>
<td>19</td>
<td>33</td>
<td>46</td>
<td>65 MVO2.pHv</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>66 HR.MAP</td>
</tr>
</tbody>
</table>

The difference between the residual and total corrected sum of squares is the multiple regression sum of squares, and the ratio of the latter to the TCSS is the square of the multiple correlation coefficient, multiple $r^2$. The multiple $r^2$ is one measure of completeness of the description provided by the multiple regression equation, although other measures may give different impressions of completeness (see Results).

The coefficients $M_1, M_2, M_3, \ldots, M_n$ are the regression coefficients. The terms $X_1, X_2, X_3, \ldots, X_n,$ contribute significantly to the multiple regression to the extent that their regression coefficients are significantly different from zero. This will not occur if a variable $X_i$ is totally unrelated to flow, nor if its descriptive effect is completely contained in the descriptive effect of another variable to which $X_i$ is closely related. Merely offering a large number of variables to the regression process does not mean that every variable will be incorporated with significant effect into the multiple regression relationship. It does not create relationships which do not exist. Unimportant terms are effectively ignored by allocating them regression coefficients which are not significantly different from zero.

It may be, however, that, if variables $X_p$ and $X_q$ are very closely related to each other, the regression process may include $X_p$ in the regression and "ignore" $X_q$, even though the latter is more directly related in a causal sense to changes in flow. A reason
for incorporating the term $X_i$ instead of $X_q$ would be if it had been measured with greater precision. Nonetheless, the descriptive relationship between $X_p$ and flow would be a valid one.

**The expansion of the variable set**

The ability of the regression process to recognize similarities of pattern between a data set and a regression model depends somewhat on the flexibility of the regression model. If, for instance, the relationship between coronary blood flow and coronary venous $P_{O_2}$ were markedly non-linear and the regression model allowed only for a linear relationship, pattern recognition would be less successful than if the model did allow for non-linearities. The inclusion of square and reciprocal terms in coronary venous $P_{O_2}$ would make this allowance sufficiently well for many practical purposes. Thus the data could be examined for similarity to an equation of the form:

$$F_i = C + M_1.P_{O_2} + M_2.(1/P_{O_2}) + M_3.(P_{O_2})^2.$$

If the relationship was predominantly linear, the inclusion of the non-linear terms would not improve the description, but in the presence of non-linearities, it would.

By extension of this reasoning, square and reciprocal terms were added to most of the primary and derived variables in table II. The reciprocal and square terms are included in the same table.

The product terms listed in table II are an attempt to confer additional flexibility on the regression model by relaxing the restriction that the effect of every variable should be independent of the values assumed by other variables. If, for instance, the regression process was constrained to examine the data set for similarity to a relationship between flow, $P_{O_2}$ and $P_{CO_2}$ which had the form:

$$F_i = C + M_1.P_{O_2} + M_2.P_{CO_2}$$

then the effect of a change of $P_{CO_2}$ ($\Delta P_{CO_2}$) would be $M_2.\Delta P_{CO_2}$ irrespective of the prevailing $P_{O_2}$. Provision for a term $M_3.(P_{O_2}.P_{CO_2})$ to be included in the above equation would allow the regression process to search for a similarity (if it existed) to a regression model in which the effect of a change in $P_{CO_2}$ could vary with $P_{O_2}$. A similar logic applies to the inclusion of the product terms with the dummy variable BK. A regression model which allows for a relationship:

$$F_i = C + M_1.P_{O_2} + M_2.(BK \times P_{O_2}) + M_3.BK$$

allows both the slope and the intercept of the flow to $P_{O_2}$ relationship to be varied by beta-blockade.

Anomalies in the inclusion of reciprocal, square and product terms in table II reflect the empirical development of the study as new possibilities for flexibility were successively realized. The addition of new variables ceased when it became clear that many variables could be omitted from the 77-variable regression without noticeable detriment to the explanation of the TCSS of the flow observations. The description of the changes in flow was as complete as could be achieved in terms of the associated variables which had been measured.

**Analysis of TCSS by “backwards regression”**

If, in order to allow for non-linear and interactive descriptive effects, a given “primary” variable is represented additionally by its reciprocal, square and product terms, its overall contribution cannot readily be examined by the contributions of its individual terms. What is required is to compare the regression sum of squares (or multiple $r^2$) for a regression which includes all the terms in that “primary” variable with that for a regression which is similar except that it contains none of the terms. If the term “backwards regression” (Nie et al., 1975) is applied to the procedure in which regression sums of squares are compared for relationships with and without individual terms, “backwards regression by groups” might be used to describe the modification proposed above.

“Backwards regression by groups” lies behind much of this study. The “groups” or categories used were: Myocardial oxygen consumption ($M_v$), haemodynamics (Haem), arterial blood-gases (Pa), coronary venous blood-gases ($P_v$), arterial and venous pH terms ($pH$), acid–base transfer (Ac-Bas), beta-blockade (BK), heart weight (HW) and body weight (BW). The variables in each group or category are indicated in table III by reference to the numbers listed in table II.

The SPSS multiple regression package was used to determine a selection of regression sums of squares whose interrelationships allow an analysis of the observed TCSS of coronary blood flow.

**RESULTS AND DISCUSSION**

**Control of coronary blood flow**

The particular selection of results has been determined by consideration of the background physiology of the control of coronary blood flow. The considerable complexities of coronary vascular control at the cellular level are beyond the scope of this study, and are well reviewed elsewhere (Berne, 1964; Rubio...
CANINE CORONARY BLOOD FLOW

TABLE III. The contents of the variable categories by reference to the variables listed and numbered in table II

<table>
<thead>
<tr>
<th>Category name</th>
<th>Symbols</th>
<th>Variable numbers contained</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myocardial O2 consumption</td>
<td>Mv</td>
<td>9, 25, 38, 47, 48, 50, 51, 56, 57, 60</td>
<td>10</td>
</tr>
<tr>
<td>Haemodynamics</td>
<td>Haem</td>
<td>12–19, 27–33, 40–46, 66–77</td>
<td>34</td>
</tr>
<tr>
<td>Arterial blood-gases</td>
<td>Pa</td>
<td>3, 4, 20, 21, 34, 35, 49–53</td>
<td>11</td>
</tr>
<tr>
<td>Coronary venous blood-gases</td>
<td>Pv</td>
<td>6, 7, 22, 23, 36, 37, 54–59, 62–64</td>
<td>15</td>
</tr>
<tr>
<td>Arterial Po2</td>
<td>PAO2</td>
<td>3, 20, 34, 49, 50, 52</td>
<td>6</td>
</tr>
<tr>
<td>Arterial PCO2</td>
<td>PACO2</td>
<td>4, 21, 35, 49, 51, 53</td>
<td>6</td>
</tr>
<tr>
<td>Coronary venous Po2</td>
<td>PVo2</td>
<td>6, 22, 36, 54, 55, 56, 58, 62, 64</td>
<td>9</td>
</tr>
<tr>
<td>Coronary venous PCO2</td>
<td>PVCO2</td>
<td>7, 23, 37, 54, 57, 59, 63, 64</td>
<td>8</td>
</tr>
<tr>
<td>pH</td>
<td>pH</td>
<td>5, 8, 24, 55, 65</td>
<td>5</td>
</tr>
<tr>
<td>Acid–base exchange</td>
<td>Ac-Bas</td>
<td>10, 26, 39, 58, 59, 61</td>
<td>6</td>
</tr>
<tr>
<td>Beta blockade</td>
<td>BK</td>
<td>11, 52, 53, 60–65</td>
<td>9</td>
</tr>
<tr>
<td>Heart weight</td>
<td>HW</td>
<td>1, 47</td>
<td>2</td>
</tr>
<tr>
<td>Body weight</td>
<td>BW</td>
<td>2, 48</td>
<td>2</td>
</tr>
</tbody>
</table>

and Berner, 1975; Belloni, 1979). This data set can only address the problem at a more simplistic level. The close and long recognized association between changes in coronary blood flow and changes in myocardial oxygen consumption (Eckenhoff et al., 1947), raises the question of whether the latter, or the cardiac work rate which it represents, is also a direct cause of the former. A finding that coronary blood flow is uniquely dependent on myocardial oxygen consumption would suggest that this may be so — that is that some process very closely related to cellular energy utilization or oxidative energy production is on the final common pathway in the mechanisms of coronary vascular control.

If the relationship between coronary blood flow and myocardial oxygen consumption is not unique, but can be modified by factors unrelated to cardiac work, alternative hypotheses can be entertained. At constant rates of oxygen supply (coronary blood flow times arterial oxygen content) changes in myocardial oxygen consumption will alter the amount of oxygen in the coronary venous blood, and the Po2 in the myocardial tissue. The associated changes in carbon dioxide production would also change myocardial tissue PCO2 and pH. Therefore myocardial chemosensitivity to Po2, PCO2 and pH could conceivably be partly or wholly responsible for the observed changes in coronary blood flow with changes in myocardial oxygen consumption.

With observations from this data set, it can be shown that the relationship between coronary blood flow and oxygen consumption is not unique. On two-dimensional plots of coronary blood flow against oxygen consumption, combinations of coronary venous hypoxia and hypercapnia can be found in which the blood flow/oxygen consumption relationship is displaced. However, such plots cannot distinguish if the displacement is attributable uniquely to the changes in the coronary venous blood, if changes in the arterial blood are more closely related, or if, even, they can be a result of changes in cardiac work which fail, in some unforeseen way, to be represented by the oxygen consumption. The multiple regression approach was adopted in the hope of achieving a different insight on the problem.

The most complete regression description

The coronary blood flow observations covered a wide range. The mean was 28.34 ml min⁻¹ (SD 20.56) and the TCSS of flow was 140 737. The regression of coronary blood flow on all 77 variables in table III had a regression sum of squares of 136 740, and thus accounted for 97.16% of the TCSS of flow. This appears satisfactorily complete, as judged by a multiple r² of 0.9716, and even more complete as judged by a multiple correlation coefficient (multiple r) of 0.9857 (the square root of 0.9716). However, these figures correspond to a residual standard deviation of 4 ml min⁻¹. Thus only 80.8% of the standard deviation of the flow observations is accounted for. The analysis is carried out in terms of sums of squares, since these have additive properties, but residual standard deviations are also displayed in table IV to provide an alternative impression of the completeness of the explanations of the variability of flow.

Although the distribution of the flow observations is clearly not normal, their distribution about the regressions appears to be tolerably so, although
Although the regression of all 77 variables in Table III on coronary blood flow satisfactorily explains most of the TCSS of flow, it is extremely cumbersome. The next step was to reduce it by removing redundant variables.

The first backwards regression was performed to eliminate the category of pH variables. As no deliberate attempt was made to change pH independently of PCO₂, the explanatory effect of the pH category was expected to be largely contained in that of the PCO₂ terms. The regression sums of squares from this pH category appeared to be statistically significant (P<0.05), as indeed was the contribution of the pH category to the 77-variable regression. Statistical significance indicates consistency of observation and the small contributions of the PCO₂ terms were expected to be largely contained in that of the pH category. The next step was to reduce it by comparing away redundant variables.

Table IV compares the sum of squares for this 72-variable regression with the sums of squares for the regressions in which the other variable categories have been omitted, one at a time. The top line of solid circles indicates that every one of the overlying variables was represented (72 variables in all). The next line has solid circles under every category except BW, indicating that the two body weight variables have been removed, leaving a regression of 70 variables. In the next line, only the two heart weight variables have been removed; in the next, only the eight beta blockade variables, and so on. The changes in regression sums of squares from that of the 72-variable regression are a measure of the portion of the explanation of TCSS which is provided uniquely by each variable category, without overlap from any other variable category, from that of the 72-variable regression sums of squares.

Table IV also shows the probability, P, that its unique contribution is not a significant improvement to the explanation of the TCSS. By these measures, even the small contributions of BK, Ac-Bas, and Haem appeared to be statistically significant. Statistically significant categories indicate consistency of observation and that the small contributions of BK, Ac-Bas, and Haem appeared to be statistically significant.
rather than physiological importance. For the purposes of later discussion, the categories HW, BW, pH, BK and Ac-Bas were declared redundant, although some interest was retained in the haemodynamic variables.

The pH variables were "redundant" in regressions containing PCO\textsubscript{2} variables, but their contribution was seven-fold larger in regressions which did not. As expected, the explanatory effect of pH variables was largely overlapped by that of the PCO\textsubscript{2} variables. In similar fashion, the explanatory effects of the beta blockade variables and the HW and BW variables were shown to be largely covered by those of the oxygen consumption variables.

After excluding the categories pH, HW, BW, Ac-Bas and Haem, 21 variables were left in the categories Mv, Pa and Pv. These are the variables listed for Mv, Pa, and Pv in table III, without the product terms with variables in the other categories. These 21 variables accounted for 94.5% of the explanation of TCSS, or 75% of the standard deviation of flow.

"Extended backwards regression" — Venn diagrams

The "backwards regression by groups" in table IV was adequate to delineate the unique explanatory contributions of the different variable categories. It was "extended" to describe also the areas of overlap between the explanatory contributions.

Table V illustrates an "extended backwards regression" applied to the explanatory contributions from the three variable categories — Mv, Pa and Pv. The analysis is detailed in the appendix. It requires the sums of squares of seven regressions (the regression containing all three variable categories, the three regressions containing two out of the three, and the three containing only one category each). From these, it calculates seven portions of the explanation (three portions of unique contribution, three portions of overlap between explanatory contributions from two of the three categories and one portion of overlap between all three categories). The interrelationships of these areas are displayed in figure 1 using a Venn diagram.

**TABLE V.** The relationships of the descriptive contributions of Mv, Pa and Pv categories. Seven regression sums of squares (first column) are required to define the percentage of the TCSS of flow explained by the overlap between all three categories (last column, row 1) by the three two category overlaps (last column, rows 2, 3 and 4) and the three unique contributions (last column, rows 5, 6 and 7). The required calculations are given in the Appendix.

<table>
<thead>
<tr>
<th>Regression sum of squares</th>
<th>Multiple $r^2$</th>
<th>Variable categories</th>
<th>Percentage of flow TCSS explained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mv</td>
<td>Pa</td>
</tr>
<tr>
<td>132,996</td>
<td>94.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>120,974</td>
<td>86.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>108,121</td>
<td>76.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>71,366</td>
<td>50.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>38,405</td>
<td>27.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>62,013</td>
<td>44.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>94,488</td>
<td>67.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
An even more extended backwards regression, requiring 15 regressions and defining 15 portions of explanation, allows the explanatory effects of the haemodynamic variables to be located in relation to those of the Mv, Pa and Pv categories. In figure 2, the explanatory effects of the latter three categories are portrayed similarly to their portrayal in figure 1. The superimposed explanatory effect of the haemodynamic terms encroaches on, and so reduces the unique explanatory contributions of the Mv, Pa and Pv terms in comparison with figure 1.

It would have been interesting to separate the effects of PaO₂ from those of PaCO₂, and Pvo₂ from PvcO₂, in a six-category Venn diagram. However, this would require the sums of squares of 63 regressions to delineate 63 portions of the explanation in order to draw a Venn diagram which would be indecipherable. In any case figures 1 and 2 are sufficient for the present discussion.

Inferences from the Venn diagrams

If the explanatory effects of total cardiac work on changes in coronary blood flow are encompassed within those of oxygen consumption, any explanatory effects outside those of oxygen consumption might be regarded as being unrelated to cardiac work. On this basis, one would argue from figure 1 that 26% of the TCSS of flow is unrelated to cardiac work. Twenty-six percent is the sum of the unique contributions of Pa terms, the Pv terms and the Pa/Pv overlap (see table V). However, figure 2 shows that a portion of the explanatory effects of the haemodynamic terms are also outside those of the oxygen consumption terms. Their encroachment on the effects of the Pa and Pv categories is such as to reduce the sum of the Pa, Pv and Pa/Pv effects to only 14% of the TCSS of flow. The extensive areas of overlap between the explanatory contributions of terms related (Mv and Haem) and unrelated (Pa and Pv) to cardiac work are unhelpful in this context, since they cannot distinguish which terms might be more closely related to the flow changes.

The extension of the explanatory contribution of the haemodynamic terms beyond those of oxygen consumption requires further thought. Although it is perfectly possible to calculate the rate at which the heart performs external work on the blood which it pumps, that external work is a poor predictor of oxygen consumption of intact hearts in situ (Braunwald, 1969; Gibbs, 1978). The oxygen consumption is affected also by the internal work required to maintain, amongst other things, given ventricular pressures at different degrees of dilatation. Happily, the approach in this study is unaffected by these difficulties. A conglomerate of haemodynamic variables was offered to the regression so as to determine an empirical relationship which accounted for as much as possible of the TCSS of flow. A more informed choice of variables might have been more effective, but subsidiary investigation has not proved this. In any case, the problem is that the explanatory effects of the haemodynamic terms seem already too extensive.

If cardiac work and the associated oxygen consumption affect myocardial conductance, the additional explanatory effects of the haemodynamic terms may be attributable to their representation of myocardial perfusion pressure. As myocardial vascular conductance and perfusion pressure both vary during the cardiac cycle, this explanation cannot be satisfactorily tested. All that can be argued is that the inclusion of the haemodynamic terms reduces the percentage of the TCSS of flow which cannot be explained in terms of cardiac work, oxygen consumption or myocardial perfusion pressure.

The unique explanatory contributions of Pvo₂, PvcO₂, PaO₂ and PaCO₂ are very highly significant, but relatively small. Their smallness allows conflicting arguments as to the importance of factors unrelated to cardiac work in the control of coronary blood flow.
One may dismiss the explanatory effects, arguing that the deficiencies in the account given by oxygen consumption for changes in coronary blood flow are the result of systematic errors in estimation of arteriovenous difference in oxygen content; these errors are, in turn, systematically linked to the Pa and Pv variables by consistent inaccuracies in blood-gas analysis or the equation of Rossing and Cain (1966).

Alternatively, one might argue that the changes in Pa and Pv terms do have effects in their own right, but that these have been largely obscured by the experimental design in this study; the concurrence of changes related and unrelated to cardiac work has been too extensive to enable their effects to be adequately separated. In a different experimental design, Case and Greenberg (1976) and Case and colleagues (1978), perfused the coronary vasculature separately from the systemic circulation so that myocardial PO2 and PCO2 could be changed without change in cardiac work or oxygen consumption. Changes in coronary vascular conductance were observed which could now be related only to blood-gas changes.

Subsidiary considerations

"Negative overlap". Table V shows that the Mv/Pv overlap has negative sign. For want of better means, such "negative overlaps" are depicted as "holes" in the Venn diagrams (figs 1, 2).

The explanation of this needs a simpler example. For the single categories, PaO2 and PvO2, the respective regression sums of squares are 43062 and 23410. If there were no overlap between the explanatory effects of PaO2 and PvO2, one would expect the two category regression with PaO2 and PvO2 to have a sum of squares of 43062 plus 23410, which equals 67472. Since the regression actually has a smaller sum of squares (55090), the expected-minus-observed difference of 12382 has positive sign and is attributed to positive overlap between the explanatory effects of PaO2 and PvO2.

The regression with the PaCO2 variables has a sum of squares of 11545. The two category regression with PaO2 and PaCO2 terms has an observed sum of squares of 62013. This is considerably more than the expected value of 54607 (43062 plus 11545) for the absence of overlap. The negative sign of the expected-minus-observed difference is negative overlap, arising because of supra-additive explanatory interaction between the variable categories PaO2 and PaCO2.

More complex overlaps are hard to explain and depict, but they do not affect the main inferences based (above) on the unique explanatory contributions of the Pa and Pv terms.

Inevitable correlation. Myocardial oxygen consumption was calculated as the product of a varying coronary blood flow and an arteriovenous oxygen content difference. If the latter had been constant the correlation of oxygen consumption with flow would have been perfect. Variation in arteriovenous oxygen content difference rendered the correlation less than perfect, although still strong. Some of the strength will have arisen from the inevitable correlation of flow, as dependent variable, with the flow term contained in the oxygen consumption. Any variations in flow will affect both sides of the regression equation and such variations will include inter-animal variation, random variation and measurement error.

Inter-animal variation can be allowed for by a variety of approaches based on the examination, for each animal, of the discrepancies between the observations and the expectations from the regression relationships. When such allowance is made, the TCSS of the flow observations is reduced, but the percentages explained by corresponding regressions remain largely similar, except as expected, for a particular reduction of the explanatory contribution which is unique to myocardial oxygen consumption.

Even after allowance for inter-animal variation, random and error variation in flow will still artificially strengthen the relationship between flow and oxygen consumption, but they are not necessarily major factors. It will be noted the AB terms also consisted of the product of coronary blood flow with an arteriovenous difference. Although inevitable correlation may be held responsible for the significance of the unique explanatory contribution of the AB terms, this contribution constituted a very small proportion of the TCSS.

CONCLUSIONS

The application of multiple regression analysis to the data available for this study has satisfied one requirement of an exercise with an unfamiliar technique — in not contradicting the conclusions reached in previous publications. The conclusions are broadly consistent with concepts derived from more disruptive experimentation. Multiple regressions have afforded an extraordinarily complete account of the observed variation in flow, compared with univariate approaches, and have yielded penet-
rater insights into the overlapping accounts rendered by interdependent explanatory variables. The relatively small proportion of the flow variation which seems not to be related to cardiac work reaffirms that no approach to data analysis can obviate the need for experimental design aimed directly at the problems to be solved. The design in this study was aimed at more general aspects of cardiovascular control.

There are many clinical problems which do not allow the option of corroboration by more invasive and disruptive approaches. For such problems, a more penetrating approach to data analysis may be the only way to do justice to the data set.

**APPENDIX**

**UNIQUE AND OVERLAPPING EXPLANATORY CONTRIBUTIONS OF CATEGORIES MV, PA AND PV (TABLE V)**

The explanation provided by the Pa, Pv and Mv categories in combination is the sum of squares of the regression containing the terms in all three categories (that is 132 996, or 94.5% of the TCSS of flow).

The unique contribution of a given variable category is the difference between 132 996 and the sum of squares for the two-category regression which excludes that variable. The regression using Pa and Mv terms excludes Pv terms and has a sum of squares of 120 794. Thus the unique contribution of the Pv category is 132 996 - 120 794 = 12 202 (or 8.7% of the TCSS).

Similarly, the unique contributions of Pa and Mv are, respectively, 24 875 and 61 630 (or 17.7% and 43.8% of the TCSS).

The entire explanatory contribution made by a single variable category is the sum of squares for the regression containing only that category. For the Pa, Pv and Mv categories these are, respectively, 38 405, 94 488 and 60 013 (or 27.3%, 44.1% and 41.8% of the TCSS).

Thus (38 405 - 12 202 = 26 203) of the Pv contribution is not unique, but is overlapped by at least one of the other two categories.

The amount by which Mv, alone and in combination with Pa, overlaps Pv is given by the difference between the sum of the total Pv and Mv contributions (38 405 + 94 488), and the sum of squares (108 121) of the two-category regression using Pv and Mv categories together: 38 405 + 94 488 - 108 121 = 24 772

Since 24 772 of the overlapped contribution of Pv (26 023), is overlapped by Mv, the remainder, 26 023 - 24 772 = 1 251 (1% of TCSS), is the amount by which Pa alone overlaps Pv.

By similar reasoning, the amount by which Mv alone overlaps Pa is 8 086 (or 5.7% of TCSS), and the amount by which Mv alone overlaps Pv is 2 849, (or -2% of TCSS).

The above reasoning defines six of the seven portions of the explanation of TCSS. These are three unique contributions (Pv, Pa, and Mv) and three explanatory overlaps between two categories (Pa and Pv, Mv and Pa, and Mv and Pv). The remaining portion is the three-category explanatory overlap (Mv and Pa and Pv), which is the difference between the sum of squares for the three-category regression and the sum of the squares of the six portions of the description already defined: 132 996 - 105 375 = 27 621 (or 19.6% of TCSS).

A more extensive process, using similar reasoning, defines the 15 portions of explanation obtainable from a four-category regression.

**ACKNOWLEDGEMENTS**

Thanks are due to Dr Foëx's collaborators and the departmental technical staff involved in the experimental work; to Professor R. B. Fisher (University Laboratory of Physiology, Oxford) and Professor P. Armitage (University Department of Biomathematics, Oxford), for helpful comments on statistical method; and to the assessor for the British Journal of Anaesthesia for suggestions as to content and presentation.

**REFERENCES**


CANINE CORONARY BLOOD FLOW


EINIGE FÄHIGKEITEN UND GRENZEN DER MULTIPLEN REGRESSIONSANALYSE.
ANWENDUNG AUF DEN BLUTSTROM IN DEN KORONARIEN DES HUNDES

ZUSAMMENFASSUNG


QUELQUES POSSIBILITES ET LIMITATIONS DE L’ANALYSE DE REGRESSION MULTIPLE.
APPLICATION AU DEBIT SANGUIN CORONAIRE CHEZ LE CHIEN

RESUME

Les modifications des concentrations d’oxygène et de gaz carbonique inspirées ont entraîné des modifications des paramètres hémodynamiques systémiques et coronaires chez 39 chiens anesthésiés. Ces modifications ont constitué la base de données pour un exercice d’analyse de régression multiple utilisant les conditionnements SPSS. Le problème physiologique était d’identifier des facteurs agissant sur le débit sanguin coronarien en plus de ceux associés au travail cardiaque ou à la consommation d’oxygène. De tels facteurs ont pu être retrouvés dans une mesure probablement limitée autant par le schéma expérimental que par la physiologie sous-jacente. Les relations de régression multiple décrivaient la réponse intégrée du débit sanguin coronarien au protocole expérimental de façon plus complète que n’importe quelle approche à une seule variable. L’exercice a fourni des données utiles sur un problème multifactoriel et a montré la nécessité de schémas expérimentaux réfléchis pour tenter de répondre à des questions spécifiques en physiologie.

QUELQUES POSSIBILITES ET LIMITATIONS DE L’ANALYSE DE REGRESSION MULTIPLE.
APPLICATION AU DEBIT SANGUIN CORONAIRE CHEZ LE CHIEN

RESUME

Les modifications des concentrations d’oxygène et de gaz carbonique inspirées ont entraîné des modifications des paramètres hémodynamiques systémiques et coronaires chez 39 chiens anesthésiés. Ces modifications ont constitué la base de données pour un exercice d’analyse de régression multiple utilisant les conditionnements SPSS. Le problème physiologique était d’identifier des facteurs agissant sur le débit sanguin coronarien en plus de ceux associés au travail cardiaque ou à la consommation d’oxygène. De tels facteurs ont pu être retrouvés dans une mesure probablement limitée autant par le schéma expérimental que par la physiologie sous-jacente. Les relations de régression multiple décrivaient la réponse intégrée du débit sanguin coronarien au protocole expérimental de façon plus complète que n’importe quelle approche à une seule variable. L’exercice a fourni des données utiles sur un problème multifactoriel et a montré la nécessité de schémas expérimentaux réfléchis pour tenter de répondre à des questions spécifiques en physiologie.

ALGUNAS POSIBILIDADES Y LIMITACIONES DEL ANÁLISIS DE REGRESIÓN MÚLTIPLE.
APLICACIÓN A LA CORRIENTE SANGUíNEA CORONARIA DE LOS CANES

SUMARIO

En 39 perros anestesiados, los cambios de concentración de anhídrido carbónico y oxígeno inspirados tuvieron por efecto modificaciones de la hemodinámica sistémica y coronaria. Estos cambios suministraron la base de datos de un ejercicio en el marco de un análisis de regresión múltiple en el que se usó los conjuntos SPSS. El problema fisiológico consistía en identificar las influencias sobre la corriente sanguínea coronaria que se superponían a las asociadas con el consumo de oxígeno o la actividad cardíaca. Dichas influencias pudieron ser identificadas en una medida probablemente limitada tanto por el diseño experimental como por la fisiología subyacente. Las relaciones de regresión múltiple describieron la respuesta integrada de la corriente sanguínea coronaria al estímulo experimental de manera más cabal que la de cualquier enfoque univariante. El ejercicio suministró unos conocimientos útiles de un problema multifactorial y la necesidad de un diseño experimental deliberado con miras a dar una respuesta a unos interrogantes fisiológicos específicos.