MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

Statistical Department

Director: Prof. Dr D.van Dantzig

Chief of Statistical Consultation: Prof. Dr J.Hemelrijk

S 110

The effect of a prolonged intake of phosphoric acid and citric acid in rats
Statistical analysis of the data

by

R.Doornbos and J.Hemelrijk

1953

1. Introduction.

S.L.BONTING [1] [1]) executed a number of experiments about the influence of phosphoric acid and citric acid in the diets of rats. A number of groups of albino rats were given diets with and without these acids for different periods of time and tissue analyses comprising analyses of the blood serum, liver, muscle, kidneys and tibiae were carried out afterwards. The dietary groups, with details about the experiment, are summarized in table 1.

Table 1

Dietary groups.

nr. of series	nr. of rats total	nr. of rats p.group	sex	age	dietary groups	diet during	generation
1	21	3	♂	9 weeks	PA, PB, PC, PD, PE, PF, PG	6 weeks	1st
2	21	3	♀	9 weeks		6 weeks	1st
3	21	3	♀	15 weeks		12 weeks	1st
4	21	3	♂	15 weeks		12 weeks	1st
5	21	3	♂	26 weeks		23 weeks	1st
6	36	12	♀	15 months	PA	14 months	1st
7	36	14,15,7	♂	6 months	PB, PC	5 months	2nd

The dietary groups mentioned above are:

PA: basic diet
PB: the same + 0.05 % phosphoric acid
PC: the same + 0.15 % phosphoric acid
PD: the same + 0.40 % phosphoric acid
PE: the same + 0.15 % citric acid
PF: the same + 0.45 % citric acid
PG: the same + 1.20 % citric acid

The statistical analysis of the results of the experiments was carried out at the Statistical Department of the Mathematical Centre at Amsterdam by J.van KLINKEN, mainly by means of WILCOXON's two sample test (cf. [1] pp. 40-42). The first 5 series of rats, with an intake of acids over a relativity short period, were taken together and the series 6 and 7, with a longer period of intake, were analysed separately. This was done, because different periods of intake may give different results. Furthermore the effect of the acids may be more pronounced for the second

---

1) Numbers in square brackets refer to the references at the end of this report.

generation than for the first. The dietary groups PB, PC and PD were treated as one groups, and the same was done for the diets PE, PF and PG. These groups will be denoted by PD and PG respectively.

During VAN KLINKEN's investigation a serious difficulty arose, owing to the order in which the chemical analyses had been carries out. This order was: PA, PD, PG and VAN KLINKEN found indications of a trend in the observations, within the groups PA, PD and PG separately in the order of observation. There was, at the moment of his investigations, no time available for analysing these difficulties completely or for trying to correct the data for trend. As a result of this only those conclusions could be given, which could not have been the result of a trend instead of being the consequence of the difference between the diets and thus the conclusions given in BONTING [1] were only preliminary. In this report a further analysis of BONTING's data will be given.

2. Indications of a trend in the observations.

The test against trend, which has been used, is described by BONTING [1] pp. 43-44.

In series 1,...,5 no indications of trend of any importance were found, perhaps owing to the small number of observations per group (i.e. 3 for each tissue analysis per diet group). In series 6 however, 4 rather small tail probabilities [2]) were found and in series 7 this number was 3.

The smallest tail probabilities all indicated the presence of a negative trend. The total number of tissue analyses was 19 for series 6 as well as for 7 and the number of small tail probabilities found was rather large in comparison with this total number of 38 tests. Combining the results of all 19 tests for series 6 and 7 separately resulted in a small tail probability for series 7, but not for series 6.

These results are summarized in table II, (see next page).

To illustrate the dangers of a trend in the observations, two cases have been sketched in fig. 1 and 2, for Ash content of the tibia in series 7 and for k in liver in series 6. In fig. 2 we have a case where WILCOXON's test would indicate differences, which may in reality be due to the trend only. In such cases we would therefore easily reach wrong conclusions.

---

2) BONTING used the term "tail error" where we use "tail probability".

Table II

Indications of a trend in the observations.

tissue analysis	k	sign of trend
series 6		
Alkaline phosphatase in kidney	0.03	-
Na in muscle	0.02	-
K in liver	0.03	-
Na in liver	0.09	+
series 7		
Weight of kidney	0.08	-
Ash content of the tibia	0.09	-
Total base in blood serum	0.08	-
series 6 in total	0.46	-
series 7 in total	0.03	-

In fig. 1 WILCOXON's test indicates a difference between the groups PA and PD, which might again be the consequence of a trend, but does not indicate a difference between e.g. group PD and PG. It is, however, quite possible that a really existing difference between these two groups has been obliterated by a negative trend.

Conclusion: The statistical analysis of the data in hand not be satisfactorily executed without applying some kind of correction for trend. The tests applied do give indications of the presence of a negative trend and even for those cases, where no trend has been found, this may be due to the small number of observations for each analysis separately. The dangers of attributing the effect of a trend to the difference between the diets makes the application of a correction imperative.

3. Two methods of correction for trend.

We shall use the following notation.

C_0: No correction for trend has yet been applied. The tail probabilities are in this case the same as given by BONTING. The results are not reliable without comparison with the results after correction.

C_I: A parameterfree correction for trend, making no suppositions about the form of the probability distributions of the observed quantities, but assuming a linear trend with time (including the absence of trend as a special case). This methode is described in some detail below.

C_{II}: Analysis of covariance. Here normality of the probability distributions and equality of the variances has to be assumed in additions to the assumption of a linear trend (as for C_I). The reason of using this method in addition to method C_I is explained below.

A short description of method C_I can most easily be given by means of illustrating its application to the case of fig. 2. In the first place we estimate the trend by computing, for each of the three dietary groups PA, PD and PG separately, the slope of each line, connecting two points of one group. We this find $3 \times \frac{11 \times 12}{2} = 198$ slopes and we take the median of these slopes as our estimate of the trend.

Then all points of the three groups are projected parallel to this estimated trend on one veritcal line, and WILCOXON's test is applied to these projections. In fig. 3 this procedure has been carried out for the groups PD and PG, the trend always being estimated from all three groups together.

This method, which has been applied to 27 of the 38 tissue analyses of series 6 and 7, has the disadvantage of introducing stochastical dependence in the data because all projections for one tissue analysis in a series are obtained by means of one common estimate of trend. This interdependence violates the suppositions necessary for applying WILCOXON's test and we have not been able to find a method of taking this fact into account. Therefore the results of this correction alone are not foolproof, the real tail probabilities usually being a little higher (but it is not known how much) than the values found in this way. This was the reason of applying correction C_{II} also. The disadvantage of the interdependence vanishes in that case, but this is only possible at the cost of the assumption of normality and equal variances.

A detailed description of C_{II} is given by W.J.DIXON and F.J.MASSEY [2] pp. 173-183. For the comparison of two dietary groups test (1) of p. 181 was used, not using the data of the third group. This has the advantage that a possibly different trend in the third group cannot cause disturbances for this test.

For both C_I and C_{II} the supposition has been made, that the time interval between consecutive observations has been constant for each tissue analysis separately and that the trend is linear and the same for the two groups compared. We may be sure, that this is not quite true and this makes the corrections hazardous but not applying them would be more hazardous still.

The analysis of covariance enables us, to test the hypo-

thesis, that the trends for two (or more) groups are the same (test (2a) of [2], p. 181). This test has always been applied, before C_{II} was executed. To get an overall picture of the trend for the three groups PA, PD and PG together, this test was first of all applied to these three groups together. If the data of a tissue analysis in series 6 or 7 failed on this test it does not seem advisable to draw conclusions from these data at all and no further tests have been applied. If the data pass this test a second overall test, denoted by C^{*}_{II} was applied, which compares the three groups PA, PD and PG with one another, taking the trend into account. This is the same test as C_{II} itself, but for three groups instead of one. If C^{*}_{II} did not give any indication of a difference, C_{II} has usually been omitted. Only if C_0 and C_I indicated differences between two of the three groups, C_{II} was still applied to these two. Sometimes even C^{*}_{II} could be omitted if previous results were of such a nature, that no result could possibly be expected.

Further details about the reasons, why some tests have not been applied are given in section 4.

The data of series 1,...,5 did not admit the applications of either C_I or C_{II}, because the number of observations per dietgroup were too small. This was caused by the fact, that the diets PD and PG were in reality groups of 3 diets each, thus reducing the number of observations in one real dietgroup to only 3.

4. Results of tests.

As a first step of the statistical analysis graphs were made of all cases to be tested. These graphs did not furnish more than a superficial judgment of the case in hand, but they were an aid in avoiding mistakes. These graphs were all of the same form as figure 1 and 2 and only drafts have been made of them, which have not been added to this report.

The results of the statistical analysis are summarized in table 3. The meaning of the notes $^{1)}$, $^{2)}$,... is as follows.

1) These data failed on the test of the hypothesis, that the trend is the same for the three groups PA, PD and PG (the value of k for this test has been entered in brackets).

2) The groups of observations are too small (only 3 observations in each group).

3) The graphs indicated clearly, that the correction in question would not change the results, which had already been obtained.

4) C^{*}_{II} does not indicate a difference, but C_0 and C_I do. Therefore the test for equality of trend has been applied to these two groups and, if the data passed this test, C_{II} has been executed. In the latter case a entry for k is given, in the former case the value of k for the test for equality of trend is given in brackets.

5) C_{II} has not been applied, because neither C^{*}_{II} nor C_I indicate a difference.

A "+"sign after an entry for k indicates that the second dietary group gives higher values of the investigated quantity than the first one, a "-"sign the opposite.

5. Conclusions and remarks.

In drawing conclusions the following rules have been observed.

1. No conclusions have been drawn from data of series 1,...,5 alone. These series have in general been considered as less reliable than those of series 6 and 7 (the reason for this is the small number of observations per group, making correction for trend impossible).

2. Columns C_I and C_{II} for series 6 and 7 are considered to contain the most important results. These columns have carries most of the weight for our conclusions.

3. If the test for equality of trend for the three groups PA, PD and PG indicated a difference between the trends, the data concerned have been omitted in forming the conclusion.

4. If series 6 and 7, or corrections C_I and C_{II}, gave contradictory results (i.e. small values of with opposite signs for the difference between two groups), no conclusion has been drawn. In this context a comparatively large value of is not to be considered as contradicting a small value of in another series or test.

5. "No conclusion" is not to be read as "no difference", but only as "no difference has been proved". A difference may nevertheless be present and might have been found with more observations or with an experimental design, which eliminates the possibility of wrong results as a consequence of a trend in the observations.

Conclusions.

According to these rules the following conclusions may be found from table 3.

1. Total phosphorus in muscle and Na in liver were systematically smaller for PG than for PA and PD (between PA and PD no systematic difference could be established).

2. Na in muscle increased systematically in the direction PA - PD - PG. This result was the most unambiguously established one of the whole analysis.

3. The water content of the kidney was systematically larger for PA than for PD.

4. The ash content of the tibia was systematically smaller for PD than for PG.

5. Calcium in bone ash was systematically smaller for PA than for PD and also smaller for PD than for PG (no difference could be proved to exist between PA and PG, although this follows logically from the two conclusions mentioned).

6. All further results were negative, i.e. did not give small enough values of k to warrant the drawing of a conclusion.

Remarks.

The fact that no more conclusions can be drawn safely from the data is largely due to the trend in the observations. The dangers of such a trend are clearly shown by table 3, e.g. in the case of the inorganic phosphorus in blood serum (second line of table 3), where an uncorrected value $k = 0.0003$ (-) was found, which is almost certainly due to trend; all other values of k are much larger and point in the opposite direction. The chloride in blood serum is another example of this danger. Here, in series 7, C_O and C_I gave small values of k, but these are probably due to different trends in the PA and PG series, as was shown by the analysis of covariance test. A number of such cases could be enumerated. On the other hand Na in liver in series 6 and 7 without correction did not show at all, that PD gave systematically larger values than PG; this showed clearly after correction and corroborated the result of series 1,...,5 for total base in liver, which of itself could not be deemed trustworthy enough for drawing the conclusion concerned.

Altogether we have tried to draw only those conclusions, which may be considered as reliable in spite of the difficulties of the statistical analysis, for which no really adequate method was available. If further experimental research is planned, some more indication may be read from table 3 and it

may be expected that, with due precautions against a trend in the observations (a trend, which could scarcely have been foreseen before the experiment started, but which may be avoided in the future) on repitition of the experiment, or of parts thereof, more positive results would be found.

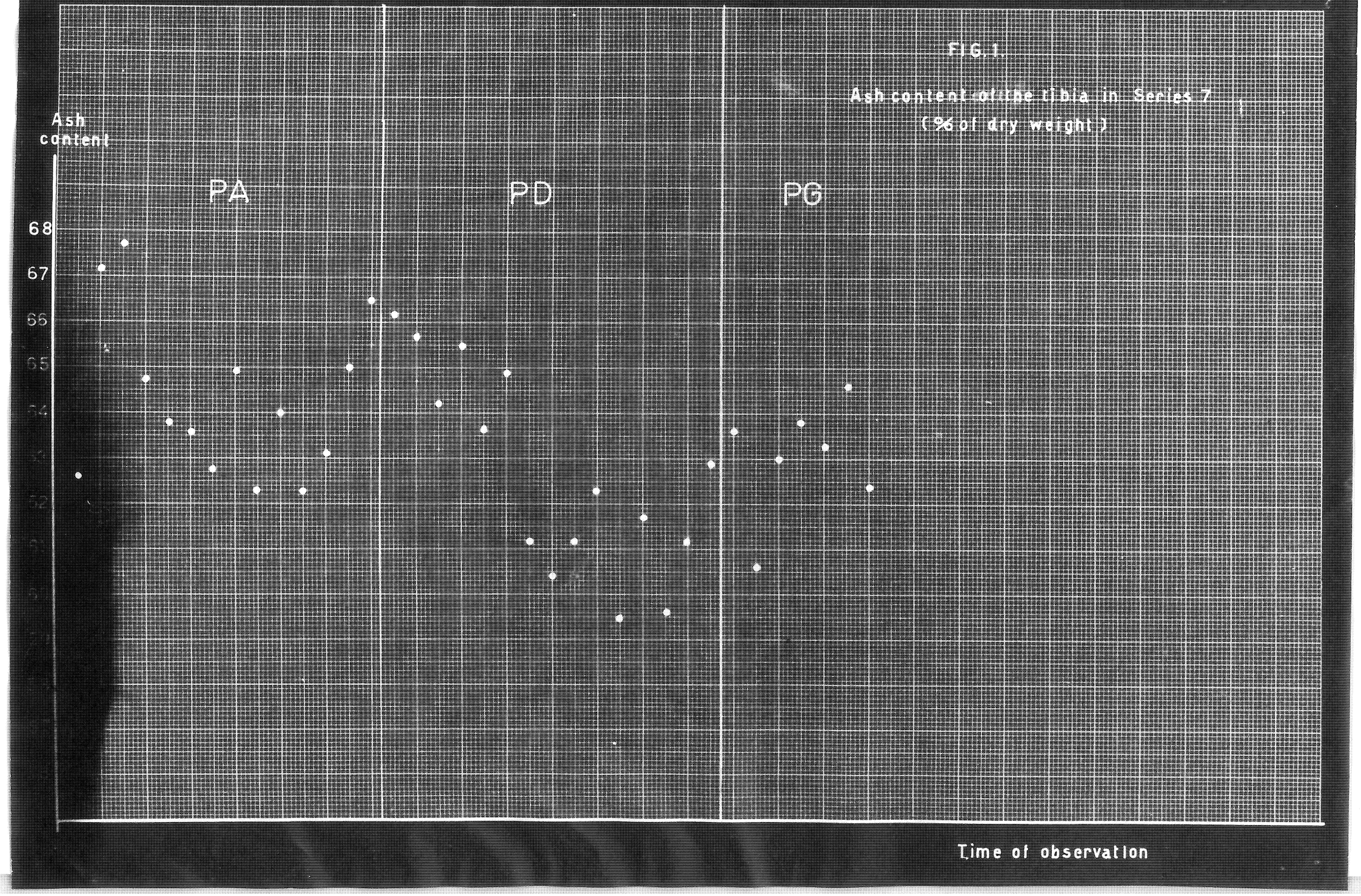
References.

[1] S.L.BONTING, The effect of a prolonged intake of phosphoric acid and citric acid in rats; Haarlem, 1952.

[2] W.J.DIXON and F.J.MASSEY, Introduction to staticical analysis; New York, Toronto, London, 1951.

FIG. 1.

Ash content of the tibia in Series 7
(% of dry weight)



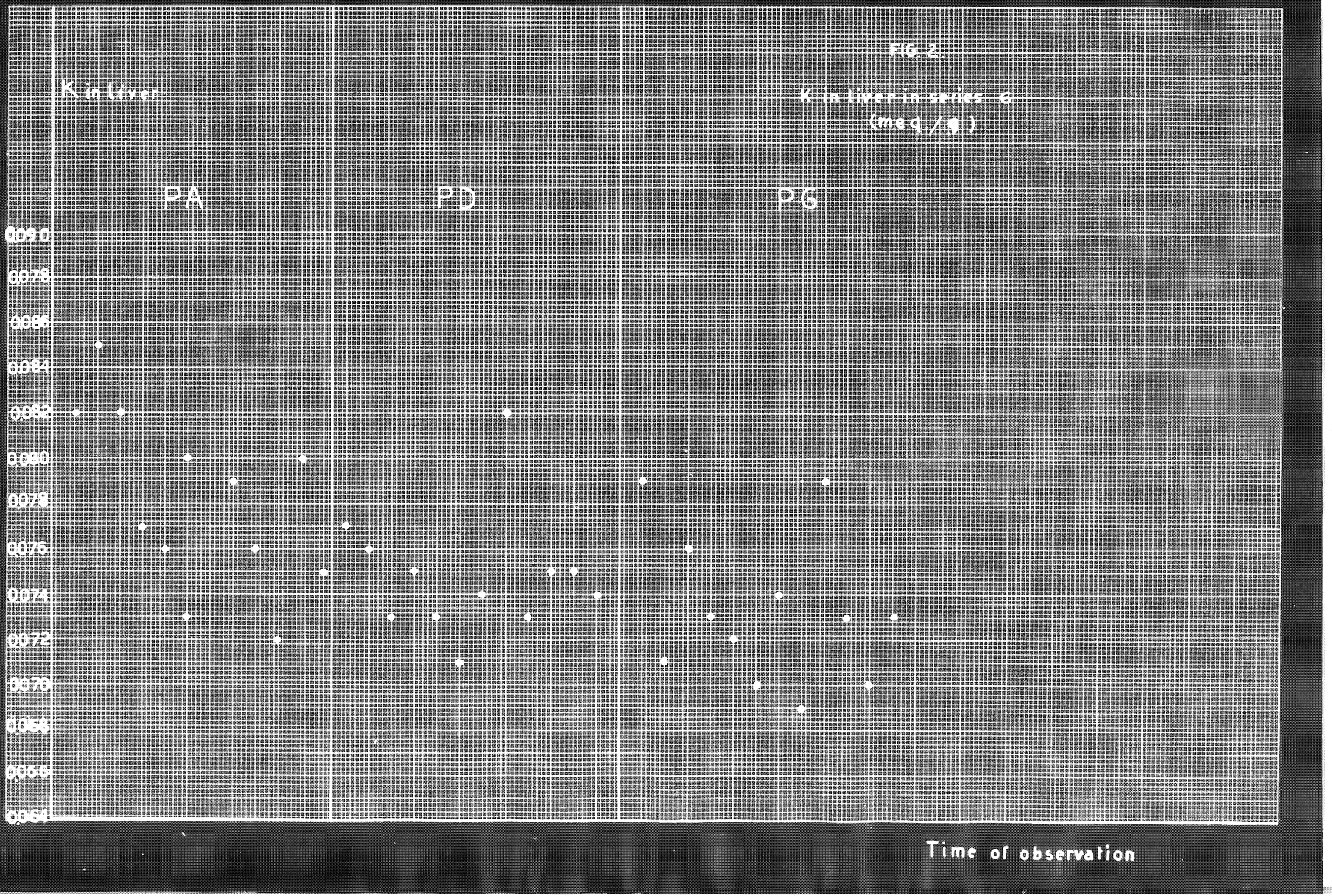


Fig-2
K in Liver
K in Liver in arbitrary
(meq/gm)
PA
PD
PG
Time of observation

FIG. 3.

Ash content of tibia in Series 7
(% of dry weight)