**NOTES**

Test of Peto’s paradox → Vincze 2021 Nature (***How to deal with the large number of 0s measurements for the cancer respective trait?)***

Owing to the large number of zero cancer mortality risk estimates and

thus non-Gaussian distributions, cancer risks were analysed using

zero-inflated phylogenetic models (Methods), as a function of sample

size, body mass and life expectancy. The probability of detecting at least

one individual with cancer in a species increased steeply with increasing

number of individuals with available postmortem pathological records

(Extended Data Table 3). In fact, cancer was detected in at least one indi-

vidual in almost all species with more than 82 individual pathological

records available (Extended Data Figs. 4 and 5). Exceptions were the

blackbuck (Antilope cervicapra) and the Patagonian mara (Dolichotis

patagonum), where no cancer was detected despite postmortem patho-

logical records being available for 196 and 213 individuals, respectively.

**Zero-inflated phylogenetic methods listed**

Therefore, the

first part of this consisted of a phylogenetic binomial regression (using

the function binaryPGLMM, in the R package ape44), where the depend-

ent variable explained the presence of zeros and non-zeros in CMR or

ICM. This model contained the log number of deceased individuals with

available postmortem pathological records as an explanatory variable,

due to the higher probability of detecting cancer with a higher number

of dead individuals inspected. Additionally, the model contained body

mass and adult life expectancy as covariates. The second part of the

model consisted of a PGLS regression that investigated variance only

in non-zero cancer risks. ICM and CMR were logit-transformed in all

PGLS models as recommended when analysing proportions45. These

models were weighted by log number of deceased individuals with

available postmortem pathological records, as the precision of can-

cer mortality risk estimates is expected to increase with the number of

dead individuals subject to postmortem examination, but it is not

expected to explain bias in the estimation of the dependent variable in

any particular direction (as in the case of the binomial models). These

models also contained body mass and adult life expectancy as explana-

tory variables. Given the expected additive effect of body mass and

longevity, the interaction between body mass and longevity metrics

was also tested in all four models (binomial and logistic regressions

for CMR and ICM), but these interactions did not increase model fit in

any case and are therefore not presented. Both models were controlled

for phylogenetic relatedness among species, where the scaling param-

eter of phylogenetic dependence (that is, s2/Pagel’s λ in PGLMMs and

PGLSs respectively) was set to the most appropriate values assessed by

likelihood ratio statics in each model separately. PGLS models in which

Pagel’s λ converged to negative values were refitted with Pagel’s λ fixed

at 0. Three species (Lagurus lagurus,Cricetus cricetus andDasyuroides

byrnei) had been removed from the latter models, due to their high

leverage caused by their very low adult life expectancy compared to

the rest of the species and therefore concerns of strong influence of

these points over model fit. Nonetheless, all models were performed

using the entire dataset, and the results were highly consistent with

and without the exclusions (Supplementary Table 4 and Extended

Data Fig. 7).

**SUMMARY**

They study two different models:

1. Non-cancer (0s) vs Cancer (ALL that is above 0) -> CMR (cancer mortality risk) y ICM (corrected cancer mortality risk) (BINARY PGLMM) podrian haber usado un phylo binary
2. For all those which are no 0 (0,1 or 0.7 or 0.9), VARIANCE between them with covariated (PGLS)

They control for the presence of 0s in these 2 different ways!!!!