



HAL
open science

Reassessing Animal-to-Human Concordance: Flawed Methodologies and the Overestimation of Predictive Value in Preclinical Research

Cédric Sueur

► **To cite this version:**

Cédric Sueur. Reassessing Animal-to-Human Concordance: Flawed Methodologies and the Overestimation of Predictive Value in Preclinical Research. 2024. hal-04815761

HAL Id: hal-04815761

<https://hal.science/hal-04815761v1>

Preprint submitted on 3 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reassessing Animal-to-Human Concordance: Flawed Methodologies and the Overestimation of Predictive Value in Preclinical Research

Cédric Sueur

Ineichen et al. (2024) investigate the concordance between animal and human studies in the development of new therapies, specifically examining whether successes observed in animals translate into clinical successes in humans. It seeks to identify the limitations of transposing animal data to clinical practice and explores the factors influencing this concordance. The goal is to provide a critical perspective on the effectiveness of animal models in the therapeutic development process and their relevance in predicting human outcomes.

The authors of the study claim that ‘only 5% of therapeutic interventions tested on animals obtain regulatory approval for use in humans’, while reporting an 86% concordance between positive results in animal and clinical studies in their meta-analysis. However, this figure of 86% is misleading and invalid due to flawed reasoning in its calculation.

The study data include a total of A_t animal studies, of which A_p yielded positive results, and H_t clinical studies, of which H_p yielded positive results. The authors calculate concordance as $(H_p/H_t)/(A_p/A_t)$, or $(317/515)/(1181/1496) = 0.61/0.79 = 0.86$. This calculation is incorrect because it does not reflect the true concordance between the two datasets.

A fictitious example highlights this inconsistency: If $A_t=10$, $A_p=5$, $H_t=4$, and $H_p=2$, their formula would yield $(H_p/H_t)/(A_p/A_t) = (2/4)/(5/10) = 1$, or 100% concordance. However, in reality, only 2 clinical trials correspond to the 5 positive animal trials, giving a true concordance of $H_p/A_p=2/5 = 0.4$ (40%). This demonstrates that achieving 100% concordance would require $A_p=H_t=H_{pp}$, which is far from the case in this study.

In the study, the actual concordance between positive results is given by $H_p/A_p=317/1181=0.268$ (26.8%). This figure is already significantly lower than the claimed 86%. Even this value, however, cannot be considered valid without a direct correspondence between individual animal and clinical studies. The fact that $A_p=1181$ and $H_t=515$ indicates that only 43.6% of positive animal studies are represented in clinical trials, without explanation. Moreover, it is not specified whether the data compared pertain to the same therapies; if they do not, the comparison is invalid, and the true concordance would be zero.

For a valid comparison, each animal study must be associated with a corresponding clinical trial, and their results directly compared. Positive clinical trials could also emerge from negative animal studies or vice versa, a possibility not considered in this analysis. Even adjusting the ratio to account for this probability ($0.268 \times 0.86 = 0.23$), the concordance would drop to approximately 23%, far from the claimed 86%.

Additionally, the sample analysed is biased, excluding therapies with fewer than five animal studies. This selective approach skews the conclusions, akin to evaluating road safety only using SUVs or asserting that all vegetables are green by only examining salads. By including all available studies (4,443 animal studies and 1,516 clinical studies), recalculated success rates under both high and low hypotheses reveal a relative risk (probability of clinical success given animal success) dropping from 0.86 to a range of 0.71–0.75. The adjusted concordance falls to 19% – 21%, confirming the authors’ approach as biased and inaccurate. Moreover, this

range represents the upper limit, with the actual concordance likely being even lower in reality.

In conclusion, the study claiming 86% concordance between positive results in animal and human studies demonstrates the methodological biases often present in analyses of animal-to-human data transposability. Animal testing in preclinical drug development suffers from fundamental limitations, with unreliable predictive value for human outcomes, particularly for toxicity, underscoring the urgent need for alternative methods as highlighted by low success rates and selection biases in clinical trial data (Bailey et al., 2014; Wong et al., 2019).

References

- Bailey, J., Thew, M., & Balls, M. (2014). An Analysis of the Use of Animal Models in Predicting Human Toxicology and Drug Safety. *Alternatives to Laboratory Animals*, 42(3), 181 – 199. <https://doi.org/10.1177/026119291404200306>
- Ineichen, B. V., Furrer, E., Grüninger, S. L., Zürcher, W. E., & Macleod, M. R. (2024). Analysis of animal-to-human translation shows that only 5% of animal-tested therapeutic interventions obtain regulatory approval for human applications. *PLOS Biology*, 22(6), e3002667. <https://doi.org/10.1371/journal.pbio.3002667>
- Wong, C. H., Siah, K. W., & Lo, A. W. (2019). Estimation of clinical trial success rates and related parameters. *Biostatistics*, 20(2), 273–286. <https://doi.org/10.1093/biostatistics/kxx069>