Title: ***In silico, In vitro*** **and nonhuman primate research models within Major Depressive Disorder research**

17/09/19 accepted for publication in *Altern Lab Anim*

Carvalho, Constança; Universidade de Lisboa Centro de Filosofia das Ciências. constanca.carvalho@sapo.pt Address: Campo Grande 016, 1749-016 Lisboa, Portugal.

Varela, Susana A. M.; cE3c—Centre for Ecology, Evolution and Environmental Changes. Address: Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal. Current address: Instituto Gulbenkian de Ciência, R. Q.ta Grande 6, 2780-156 Oeiras, Portugal.

Bastos, Luísa Ferreira; Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal; INEB – Instituto de Engenharia Biomédica, Porto, Portugal. l.bastos@fe.up.pt Address: R. Alfredo Allen 208, 4200-135 Porto, Portugal.

Orfão, Inês; cE3c—Centre for Ecology, Evolution and Environmental Changes; Centro de Filosofia das Ciências da Faculdade de Ciências da Universidade de Lisboa. iodias@fc.ul.pt Address: Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

Beja, Vanda; Clinical Psychology Private Practice. vanda.as.beja@gmail.com

Sapage, Manuel; cE3c—Centre for Ecology, Evolution and Environmental Changes. masapage@fc.ul.pt Address: Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

Marques, Tiago A., Centre for Research into Ecological and Environmental Modelling. Departamento de Biologia Animal, Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa tiago.marques@st-andrews.ac.uk Address: The Observatory, University of St Andrews, St Andrews, KY16 9LZ, Scotland.

Knight, Andrew; Centre for Animal Welfare, University of Winchester, UK. andrew.knight@winchester.ac.uk. University of Winchester, Sparkford Road, Winchester Hampshire, SO22 4NR, UK.

Vicente, Luís; Universidade de Lisboa Centro de Filosofia das Ciências. lmvicente@fc.ul.pt Address: Faculdade de Ciências, Universidade de Lisboa, Campo Grande 016, 1749-016 Lisboa, Portugal.

Summary

Major Depressive Disorder (MDD) is the most severe form of Depression, which is the leading cause of disability worldwide. When considering research approaches aimed at understanding MDD it is important to evaluate their effectiveness.

Here, we studied the effectiveness of original studies addressing this disorder using nonhuman primate (NHP), and human-based *in vitro* and *in silico* research approaches, and compared their respective contributions to subsequent medical publications. For each publication we conducted a quantitative citation analysis, and a systematic qualitative analysis of citations.

In the majority of cases, human-based research approaches (both *in silico* and *in vitro*) received more citations by subsequent human medical publications than NHP studies, and were considered to be more relevant to the hypothesis and/or to the methods in the subsequent human medical publications.

The results of this study suggest that *in silico* and *in vitro* approaches are taken into account by medical researchers more often than NHP-based approaches. In addition, these human-based approaches are also cheaper and less ethically contentious than NHP studies. Therefore, we suggest that the standard animal-based procedure for testing medical hypotheses should be revised, and that more opportunities to further develop human-relevant innovative techniques should be created.

Key-words: major depressive disorder, *in silico*, nonhuman primate, animal use alternatives, *in vitro,* 3Rs

Introduction

According to the World Health Organization, depression is the leading cause of morbidity worldwide. It affects more than 300 million people of all ages and is a major contributor to the overall global burden of disease (1). Furthermore, people who suffer from depression are more prone to an early death either by suicide or by the development of other conditions such as cancer, heart disease or stroke (2, 3). These patients are also more prone to develop other disorders like osteoporosis (4), which – even if not life threatening – significantly impacts quality of life, public health and national economies.

Accordingly, there is major investment in research aiming to improve the understanding of depression in all its eight forms (5). Major depressive disorder (MDD) is the most severe depression type and the third leading cause of years lived with disability (6). Likewise, the only available global European studies (from 2004 and 2010) show that MDD was the costliest brain disorder in Europe, accounting for at least 1% of the total European economy (7, 8). In the United States of America (USA) the economic burden of MDD alone was $210.5 billion in 2010 (9).

Clinical research is expensive, time-consuming, and potentially ethically contentious. Therefore, nonclinical research approaches using non-human animals, and human-based *in vitro* and *in silico* approaches are seen as valuable tools within the early steps of biomedical research, as they might simplify and accelerate drug and treatment discovery. However, to optimize the efficiency of nonclinical research it is crucial to evaluate which research approaches can potentially deliver better results for patients.

Since the second half of the twentieth century, animal-based research has been accepted as the 'gold standard' approach for pre-clinical biomedical research and testing (10). Within this approach, nonhuman primate (NHP) research has been considered particularly relevant due to the similarity to humans of NHPs. However, this same similarity has led to legal protections, albeit with important differences, in various regions of the world. For example, Europe (11), USA

(12), and New Zealand (13) applied considerable restrictions to the use of NHPs for scientific purposes. These restrictions are due to the understanding that laboratory confinement alone, as well as the use of invasive or intrusive techniques, has resulted in psychosomatic injuries, mutilations, and even physiology traits that have been compared to those of people with Post-Traumatic Stress Disorder (13-20). Moreover, NHPs are expensive to acquire (21) and are the most expensive animals to maintain (22).

The legislation of several countries (*e.g.* Directive 2010/63/EU) require for a cost-benefit assessment prior conducting a procedure on a non-human animal. The likely harm to the animal should be balanced against the potential benefits of each project and the project should only go further if the expected benefits outweighs the harms inflicted to the animals involved.

Considering all the above, it is assumed that when research is conducted on NHPs, it should provide highly relevant data that leads to concrete improvements in patient outcomes, due to the ethical and economic concerns surrounding this practice. While some authors assert that animal research approaches, and NHP approaches in particular, are crucial for biomedical progress (23), more evidence-based papers have increasingly shown that the contribution of animal-based research for the advancement of human healthcare has been poor (24), namely in understanding MDD (25). However, it is yet to be established whether this poor contribution is due to the intrinsic limitations that all nonclinical research approaches unavoidably have, or if human-based nonclinical research approaches are more effective in helping biomedical progress, at least when seeking to understand complex multifactorial origin disorders such as MDD.

In vitro and *in silico* methods are thought to potentially allow for faster development of medical treatments (26, 27) when they rely on human-based knowledge and/or material directly as a reference. Usually, they are also more cost effective. However, on one hand these relatively recent methods are still judged against the standard biomedical research paradigm: they are

considered to be preliminary steps prior to animal-based testing (28), instead of yielding data of sufficient value to be used without additional and, many times contradictory, animal testing, . On the other hand, combinations of human-relevant methods like *in vitro* and *in silico* are thought to enable sufficient understanding of a disease in humans, and providing the means to test new therapies for specific patients (29).

To shed light on this debate, this study examines and compares the contribution of NHP, *in silico* and *in vitro* approaches to human medical publications addressing MDD. This comparison allows us to: a) evaluate whether the low transferability of knowledge to clinical research is a common trait of all indirect research approaches, and b) evaluate the relevance of each approach to human medical studies.

Considering its dominance within the current pre-clinical research paradigm, we expect NHP studies to have a higher contribution to human medical research than *in silico* and *in vitro* studies. A similar or lower contribution of NHP papers would suggest that clinical research is becoming less reliant on more expensive and ethically questionable NHP research, thus suggesting that the time for a paradigm shift has come.

Methods

The design of this study is based on a previously developed method consisting of a quantitative citation analysis and systematic qualitative analysis of citations (30).

Quantitative citation analysis

Bibliographic search

Our citation analysis was performed between September 2016 and June 2017. The bibliographic database 'PubMed' was searched for papers using NHP, *in vitro* and *in silico* research approaches to investigate MDD. We used Medical Subject Heading (MeSH) search terms: "Depressive disorder, major" AND title/abstract: "primate" OR "ape" OR "macaque" OR "macaca" OR "rhesus" OR "chimpanzee" OR "bonobo" OR "gorilla" OR "gorila" OR "Pan" OR "orangutang" OR "orang-utan" OR "Orang utan" OR "orangutan" OR "ourang-outang" OR "Pongo" OR "gibbon" OR "Hylobates" OR "Colobus" OR "Baboon" OR "Papio" OR "Mandrillus" OR "Mandrill" OR "Cebus" OR "Cebuella" OR " Brachyteles" OR "Loris" OR "Nycticebus" OR "lemur " OR "Callithrix" OR "*in silico*" OR "computer model" OR "mathematical model" OR "computer simulation" OR "*in vitro*" OR "cell culture" Or "culture technique" OR "cell line" OR "organ culture" OR "tissue culture". MeSH terms are a comprehensive list of key terms related to each human disorder, designed to identify all relevant studies in an area (31). So, searching for "Depressive Disorder, Major" retrieves other nomenclatures for the same disorder (e.g. Melancholia). There were no exclusive MeSH terms for non-human primates, so our search retrieved additional non-human animals' papers that we excluded by hand. We also excluded all *in vitro* and *in silico* located papers that resorted to animal data (*e.g.* rat cell line data).

We included papers from scientific journals, books, research reports and conference proceedings written in English, Portuguese or Italian, which are within our linguistic fluencies. We used PubMed filters to exclude review papers ("review", "systematic review", "meta-analysis", "bibliography") as well as editorials an other types of non-research papers ("biography", "auto¬biography", "comment", "opinion paper", "interview"), since our aim was to evaluate the impact of original data. We restricted our search to publications prior to 31 December 2011, to allow adequate time for subsequent citation of papers (32). We retrieved 19 NHP papers, 29 *in silico* papers and 38 *in vitro* papers describing data from original MDD research.

Citation data

We performed a citation analysis on the retrieved papers using the cited reference search facility within the 'Web of Science' bibliographic database. For each retrieved paper, we listed the papers that cited it and recorded three types of citation data: the total number of times the retrieved paper was cited, the total number of times the retrieved paper was cited per research category, and the total number of times the retrieved paper was cited per research subject *i.e.* on MDD or other subjects, as detailed bellow.

We ascribed each citing paper to one or more of the following eight research categories: "invasive animal research papers", "human research papers", "review papers", "editorials", "*in vitro* papers", "*in silico* papers", "non-invasive animal research papers" (e.g. observational studies with wild animals) and "other human papers" (e.g. on social perceptions). By "human research papers" we mean any human-based research that might involve, among other things, analysis of biological samples, epidemiological and behavioural studies, medical case studies, and clinical studies. Citing papers could be allocated to several categories if they described different research approaches. Whenever it was not possible to define the category of the citing paper (due to language barriers or absence of the abstract), the paper was labelled as "not available" and removed from further analysis.

Amongst the categories "human research papers", "*in silico* papers", "*in vitro* papers" and "invasive animal research papers", we also recorded which papers focused on MDD and which focused on other subjects.

Statistical analysis

To test for differences between the numbers of citations across research approaches we implemented three generalized linear models (GLM), each with a Poisson response and a log link function. Each model tested one of the following response variables: total number of citations, total number of citations by human research papers, and total number of citations by

human research papers on MDD. In each model, the type of research approach was the only explanatory variable, with three levels: NHP, *in silico* and *in vitro* research approaches. The GLM's goodness of fit was evaluated by visual inspection of diagnostic plots. Additionally, we used a Gaussian GLM to evaluate if the proportions of citations made by human medical papers, and human medical papers on MDD were different across research approaches. The analyses were performed in R 3.6.1 (33) using the function *glm*. Results were considered significant when $P < 0.05$.

Systematic qualitative analysis of citations

Papers under the category "human research papers on MDD" were systematically analysed by two independent raters to qualitatively evaluate the contribution of each citation of NHP, *in vitro* or *in silico* research paper to the respective human medical study.

Each study was rated according to the following classes, defined prospectively, and as in Carvalho *et al.* (30):

– Redundant: when the cited study was only mentioned amongst other studies as an example. When there were multiple studies used as an example of one or more points, the raters were instructed to rate the study as redundant only if there were older or human studies stating exactly the same points.

– Minor relevance: when the cited study was cited in the discussion or introduction providing information not directly related to the hypothesis explored in the human medical publication.

– Relevant to the hypothesis: when the cited study was cited in the introduction, providing information relevant for the hypothesis explored in the human medical publication.

– Relevant to the methods: when the human medical publication used the same methodology as the cited paper, with the exception of species in the case of NHP papers.

A paper considered relevant could be both relevant for the hypothesis and the methods. Other options in the classes are mutually exclusive. In all cases, disagreement between the raters was resolved via detailed discussions until a consensus was reached.

Whenever it was not possible to assess the contribution of a cited paper to the citing human medical publication due to unavailability of the latter, the paper was labelled as "not available" and removed from further analysis.

We used a statistical test for comparing proportions (Pearson's chi-squared test implemented via via R's *prop.test* function) to assess differences between the three types of citations: "NHP papers", "*in vitro* papers", and "*in silico* papers". Since even for the pair with the largest difference the null hypothesis of equal proportions could not be rejected under the usual significance levels, we did not attempt corrections for multiple comparisons.

Results

Citation analysis — NHP Results

We retrieved 19 papers describing NHP data in the field of MDD research, which were cited, in total, 841 times. Of these 19 papers, five described both human and NHP data.

Citing publications belonged to the following categories: animal research papers (312); followed by review papers (245); human research papers (152); *in vitro* papers (81); *in silico* papers (14); non-invasive animal papers (six); and opinion papers (including editorials, comments or replies to comments) (four). Eighty-five citing papers were not categorized due to being unavailable or written in a language other than English, Portuguese or Italian. Of the 312 citations by animal research papers, 63 were by papers focusing on MDD. Of the 152 citations by human research papers, 71 were by papers focusing on MDD.

Citation analysis – *in silico* results

We retrieved 29 papers describing *in silico* data regarding MDD research, which were cited, in total, 806 times. Of these 29 papers, seven described both patient data and computer simulations.

Citing publications belonged to the following categories: human research papers (317); followed by *in silico* papers (193); review papers (193); animal research papers (44); *in vitro* papers and editorials (17 on each category). Fifty-eight citing papers were not categorized due to being unavailable or written in a language other than English, Portuguese or Italian.

Of the 317 citations by human research papers, 94 were by papers focusing on MDD. Of the 193 citations by *in silico* research papers, 36 were by papers focusing on MDD.

Citation analysis – *in vitro* results

We retrieved 38 papers describing *in vitro* data regarding MDD research, which were cited, in total, 2,574 times. All *in vitro* papers used samples of human biological material, most of them (34 papers) obtained from MDD patients.

Citing publications belonged to the following categories: *in vitro* papers (1,239) resorting to human (789), laboratory animals (373) or both (12) biological materials. Nine hundred and seventy-eight citations were made by human medical papers (189 of which were solely with human participants, without concurrent use of in vitro research approaches'), 844 citations were made by review papers, followed by 464 animal papers (79 of which were solely with live animals without concurrent use of in vitro research approaches) *in vitro* methodologies), 27 editorials and 16 *in silico* papers. One hundred and fifty-four citing papers were not categorized due to being unavailable or written in a language other than English, Portuguese or Italian.

Of the 978 citations by human research papers, 482 were by papers focusing on MDD. Of the 1239 citations by *in vitro* research papers, 487 were by papers focusing on MDD.

Comparison of citations of NHP, *in vitro* and *in silico* papers

Inspection of diagnostic plots showed no reasons for concern regarding the GLM fit. Amongst *in vitro* papers, there was one highly cited (711 citations). We implemented the analysis with and without this potential outlier value and found no significant differences between the two.

GLM showed the total number of citations received, on average. Each NHP paper received 3.73 (standard error (se): 0.03) citations. *In silico* papers received less citations (3.29, *i.e.* -0.44, se: 0.05) and *in vitro* papers received more citations (4.23, *i.e.* +0.5, se: 0.04). Both differences were statistically significant ($P << 0.0001$).

Regarding average citations by human research papers, each NHP paper was cited 2.03 (se: 0.08) times. In comparison to NHPs, both *in vitro* and *in silico* papers received higher number of citations (+1.09, se: 0.09 and +0.33, se: 0.10, respectively). The differences are statistically significant ($P < 0.001$).

Concerning citations by human research papers on MDD, each NHP paper was cited, on average, 1.27 (se: 0.12) times which is not statistically different from *in silico* citations (-0.12; se: 0.16). *In vitro* papers received, on average, more citations (+1.3, se: 0.13) and the difference is statistically significant from NHP ($P < 0.001$).

The proportion of citations of NHP papers by human research papers was 0.13 (se: 0.05). This proportion is significantly higher in *in silico* papers (+0.20, se: 0.07, $P = 0.004$) and in *in vitro* papers (+0.30, se: 0.07, $P << 0.0001$). The proportion of citations of NHP papers by human research papers on MDD was 0.06 (se: 0.03). The proportion of citations by *in silico* papers was not statistically different (+0.06, se: 0.04, $P = 0.1389$). The proportion of citations by *in vitro* papers (+0.14, se: 0.04) was significantly different from the same proportion in NHP papers ($P = 0.001$).

Figure 1 illustrates these comparisons.

Systematic qualitative analysis of citations

Fifty of the 71 (70%) human research papers on MDD that cited NHP papers were available for further analysis, along with 401 of the 482 (83%) human research papers on MDD that cited *in vitro* papers, and 58 of the 94 (62%) human research papers on MDD that cited *in silico* papers. Around 16%, 25% and 25% of citations of, respectively, NHP, *in vitro* and *in silico* papers were of relevance for the hypothesis and/or the methods in the citing human research paper on MDD.

The statistical test used to compare the proportions did not reveal any significant differences between the proportions of citations with relevance between NHP-*in vitro*, NHP-*in silico*, and *in vitro*-*in silico* (P = 0.31, 0.20 and 1). The contingency table is presented in Table 1.

Discussion

We quantitatively and qualitatively analysed the contribution of NHP, *in vitro* and *in silico*-based research to the contemporary understanding of MDD. Of the three approaches studied, NHP-based research seems to be the one that provides the lowest likelihood of contributing to human medical research. Amongst the three research approaches, human-based *in vitro* seems to be the one that influences clinical research the most. However, all approaches seem to be equally relevant in informing the hypothesis and/or methods of human medical studies.

Overall, our results suggest that less funded research approaches (34) are more or equally effective in reaching their final goal — informing clinical research to improve human healthcare. Our quantitative results show that *in silico* and *in vitro* papers are more successful than NHP papers in providing contribution to researchers publishing in human medical papers, as the proportion of their total citations by human medical papers is higher than the proportion of NHP papers cited by human medical papers. NHP papers are mainly cited by subsequent animal

experimentation papers, which suggest that they are mainly contributing to subsequent animal research rather than contributing to advances in healthcare. *In vitro* seems to be the most effective approach since it was the one receiving significantly more citations by human medical papers on MDD or other subjects.

Of the five NHP papers on MDD we analysed that were relevant to the hypothesis, method or both of the citing human research papers on MDD, one described both NHP and human research, and its two citing papers referred to the human research data contained within this NHP paper. Another one of these five NHP papers on MDD was considered relevant to the methods and was cited once. The citing paper described both human and rhesus monkey data, and the citation was relevant to the methods used with the rhesus monkeys. After removing these cases, only three out of 19 NHP studies were relevant for the hypothesis and/or methods of subsequent human medical studies on MDD.

The results of our citation analysis also suggest that the widely accepted approach to testing medical hypotheses — which relies on *in vitro* and *in silico* research approaches as a step prior to animal testing - is not actually working as intended, since clinical papers tend to cite *in silico* and *in vitro* papers directly too. However, the citations within human medical publications on MDD constituted a low percentage (50% or less) of the total citations received in all three categories. This may be explained by the complexity of MDD, which shares genes, phenotypic traits and possible neurologic pathways with different disorders. Hence, a human study on anorexia might cite a nonclinical study on MDD focused on weight loss, since changes in weight is one of the symptoms of MDD.

As to the qualitative results**,** the relevance of the retrieved papers for the publications citing them seems insufficient in all analysed research approaches. Even though *in silico* and *in vitro* papers showed higher percentages of cited papers that were relevant for the hypothesis and/or methods used by the citing clinical studies, there were no statistical differences between the

three approaches. In this case, the effect size observed, where the proportion of citations by NHP is much lesser than that of *in silico* or *in vitro* suggesting that while not statistically significant, due to lack of statistical power, this might be a relevant practical difference.

Recent advancements in *in vitro* technologies, such as organs-on-chips (35) or *in silico* technologies, such as advanced artificial Intelligence based on sophisticated machine learning tools (36) have been published after 2011. These studies have been excluded from our analysis to guarantee sufficient time for citations. However, it is reasonable to expect that these cutting edge technologies are currently being widely used to generate new hypothesis in human medicine (27). Similarly, induced pluripotent stem cells, though known for more than a decade (37), have only recently been recommended for MDD research (38). In light of the above, it would be interesting to replicate this study a decade from now to verify (a) if there is an increase in the number of citations that *in vitro* and *in silico* papers on MDD receive within MDD human publications, and (b) if the number of times in which *in silico* and *in vitro* papers are cited with relevance for the hypothesis and/or methods, within subsequent human publications, has also increased.

We recognize that our study has certain limitations. Due to resource constraints, we were unable to use a greater number of search engines (e.g. CAB Abstracts) to increase the likelihood of retrieving all *in silico, in vitro* and NHP papers investigating MDD, which might have increased our sample, making it more comprehensive. Similarly, we were unable to examine the reference lists of retrieved papers to locate additional relevant papers. This inevitably means that some relevant publications may not have been located. Because our sample was small, our results should be interpreted with caution. Finally, we are aware of the difficulty in objectively determining the relevance of a cited paper to the publication citing it. We have used two different raters to decrease errors in assessment. The initial assessment by the raters was sometimes divergent, indicating that different raters using the same criteria might have rated

some papers differently. However, our experience suggests that these would comprise only a small proportion of papers assessed. Despite the limitations of citation analysis and systematic qualitative analyses of citations, we consider that the method we followed is useful when evaluating the effectiveness of different research approaches, and we hope that similar studies will arise investigating different disorders.

Our results suggest that the contribution of NHPs to contemporary understanding of MDD is poor, and that other approaches with potentially superior relevance to humans should be used. Our results also shed light on the controversy around the efficacy of NHP based research for investigating human disorders. This controversy is longstanding, with some authors claiming that their use is crucial for medical advancement (23), while others asserting the opposite (39, 40). However, ongoing scientific advances in non-animal methods for the acquisition of knowledge and the development of new treatments, may provide alternative solutions to help sidestep the dilemmas and concerns surrounding NHP use.

Conclusion

To our knowledge, this is the first study to compare the effectiveness of original studies involving the use of NHP, *in vitro* and *in silico* research approaches to inform the medical research community, at least within the MDD field. Our results suggest that within this field, compared to NHP studies, human-based research approaches are more promising in generating new hypothesis and methods in clinical research.

Given the scientific advances in human-based research methods, we suggest that future research using our methodology should examine the impact of more recent approaches in informing human medical research. Such research could examine if and how the standard paradigm for testing medical hypotheses is still being used, from applied research, through animal use in preclinical testing, and on to clinical research and development. Such review could

provide further insight into how the 'gold standard' that considers *in vitro* and *in silico* research approaches as merely steps prior to animal testing, could be challenged and revised. Given the scientific and ethical solutions that innovative human-based approaches are providing, with relatively little investment when compared to the investments in animal-based research, a reallocation of Research & Development resources is clearly warranted, in favour of MDD nonclinical research using human-based approaches.

Acknowledgements

The authors would like to acknowledge the University of Lisbon EcoComp team members for their valuable insights during the conception and design of this study. This study was financed by Animalfree Research – Switzerland, and by Portuguese National Funds through FCT – Fundação para a Ciência e a Tecnologia, within the CFCUL Unit funding UID/FIL/00678/2013 and through MS PhD grant (PD/BD/128349/2017) and SAMV Post-Doctoral Research Grant (SFRH/BPD/66042/2009). TAM thanks partial support by CEAUL (funded by FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the project UID/MAT/00006/2019). None of these funders played any role in the study conceptualisation, design, conduct, data collection or analysis, authorship of the resultant manuscript, or the decision to submit it for publication.

References

1. Depression – Fact sheet. World Health Organization, 2018. Available at: https://www.who.int/news-room/fact-sheets/detail/depression (Accessed 29.06.19).
2. Cuijpers, P., Vogelzangs, N., Twisk, J., Kleiboer, A., Li, J. & Penninx, B. W. (2014). Comprehensive meta-analysis of excess mortality in depression in the general

community versus patients with specific illnesses. *American Journal of Psychiatry* **171**, 4, 453–462.

3. Li, C.T., Bai, Y.M., Tu, P.C., et al. (2012). Major depressive disorder and stroke risks: a 9-year follow-up population-based, matched cohort study. *PLoS One* **7**, 10, e46818.
4. Rauma, P.H., Pasco, J.A., Berk, M., et al. (2015). The association between major depressive disorder, use of antidepressants and bone mineral density (BMD) in men. *Journal of Musculoskeletal and Neuronal Interactions* **15**, 2, 177–185.
5. American Psychiatric Association. (2014). *DSM-5: Manual diagnóstico e estatístico de transtornos mentais.* São Paulo, Brasil: Artmed Editora.
6. Vos, T., Allen, C., Arora, M., Barber, R.M., Bhutta, Z.A., Brown, A., et al. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* **388**, 10053, 1545-1602.
7. Sobocki, P., Jonsson, B., Angst, J. & Rehnberg C. (2006). Cost of depression in Europe. *J. Ment. Health Policy Econ* **9**, 87–98.
8. Wittchen, H.U., Jacobi, F., Rehm, J. et al. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology* **21**, 9, 655–679.
9. Greenberg, P.E., Fournier, A.A., Sisitsky, T., Pike, C.T. & Kessler, R.C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of Clinical Psychiatry* **76**, 2, 155–162.
10. Maurer, K. & Quimby, F. (2015). Animal Models in Biomedical Research. In Laboratory Animal Medicine (Third Edition) (ed. J. G. Fox, L. C. Anderson, G. M. Otto, K. R. Pritchett-Corning & M. T. Whary.), 1497–1534. Academic Press. [doi: https://doi.org/10.1016/C2012-0-03109-2].

11. Special Eurobarometer 340/Wave 73.1: Science and Technology Report. European Comission, 2010. Available at: http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_340_en.pdf. (Accessed 29.06.19).
12. Lankau, E.W., Turner, P.V., Mullan, R.J. & Galland, G.G. (2014). Use of nonhuman primates in research in North America. *Journal of the American Association for Laboratory Animal Science* **53**, 3, 278–282.
13. Animal welfare Act, New Zealand, 1999. Available at http://www.legislation.govt.nz/act/public/1999/0142/59.0/DLM49664.html. (Accessed 19.08.2019).
14. Kraskov, A., Prabhu, G., Quallo, M.M., Lemon, R.N. & Brochier, T. (2011). Ventral premotor–motor cortex interactions in the macaque monkey during grasp: response of single neurons to intracortical microstimulation. *The Journal of Neuroscience* **31**, 24, 8812–8821.
15. Vigneswaran, G., Philipp, R., Lemon, R.N. & Kraskov, A. (2013). M1 corticospinal mirror neurons and their role in movement suppression during action observation. *Current Biology* **23**, 3, 236–243.
16. Maninger, N., Capitanio, J.P., Mason, W.A., Ruys, J. D. & Mendoza, S.P. (2010). Acute and chronic stress increase DHEAS concentrations in *Rhesus* monkeys. *Psychoneuroendocrinology* **35**, 7, 1055–1062.
17. Knight, A. (2010). Laboratory Animal Statistics for Great Britain: Implications for Animal Welfare. *Japanese Society for Alternatives to Animal Experiments*, 17, 2, 51–52.
18. Gottlieb, D.H., Capitanio, J.P. & McCowan, B. (2013). Risk factors for stereotypic behavior and self-biting in rhesus macaques (*Macaca mulatta*): Animal's history, current environment, and personality. *American journal of primatology* **75**,10, 995–1008.

19. Morgan, K.N. & Tromborg, C.T. (2007). Sources of stress in captivity. *Applied Animal Behaviour Science* **102**, 3, 262–302.

20. Buller, T. (2014). Animal Minds and Neuroimaging. *Cambridge Quarterly of Healthcare Ethics* **23**, 02, 173–181.

21. Keen, J. (2019). Wasted money in United States biomedical and agricultural animal research. In *Animal Experimentation: Working Towards a Paradigm Change* (ed. K. Herrmann & K. Jayne). 244–272. Brill. [doi: 10.1163/9789004391192_002].

22. Fitzgerald, T.A. (1982). Comparison of research cost: man--primate animal--other animal models. *Journal of Medical Primatology* **12**, 3, 138–145.

23. Worlein, J.M. (2014). Nonhuman primate models of depression: effects of early experience and stress. *ILAR Journal* **55**, 2, 259–273.

24. Pound, P. & Bracken, M.B. (2014). Is animal research sufficiently evidence based to be a cornerstone of biomedical research? *BMJ* **348**, g3387.

25. Carvalho, C., Alves, D., Knight, A. & Vicente, L. (2019). Is animal-based biomedical research being used in its original context? In *Animal Experimentation: Working Towards a Paradigm Change* (ed. K. Herrmann & K. Jayne), 376–390. Brill. [doi: 10.1163/9789004391192_002].

26. Benam, K.H., Gilchrist, S., Kleensang, A., Satz, A. B., Willett, C. & Zhang, Q. (2019). Exploring new technologies in biomedical research. *Drug Discovery Today* **24**, 6, 1242–1247.

27. Archibald, K., Tsaioun, K., Kenna, J. G. & Pound, P. (2018). Better science for safer medicines: the human imperative. *Journal of the Royal Society of Medicine* **111**, 12, 433–438.

28. Taylor, K. (2019). Recent developments in alternatives to animal testing. In *Animal Experimentation: Working Towards a Paradigm Change* (ed. K. Herrmann & K. Jayne). 583–609). Brill. [doi: 10.1163/9789004391192_002].

29. Langley, G.R., Adcock, I.M., Busquet, F., et al. (2017). Towards a 21st-century roadmap for biomedical research and drug discovery: consensus report and recommendations. *Drug Discovery Today* **22**, 327–339. https://doi.org/10.1016/j.drudis.2016.10.011.

30. Carvalho, C., Crespo, M.V., Bastos, L.F., Knight, A. & Vicente, L. (2016). Contribution of animal models to contemporary understanding of Attention Deficit Hyperactivity Disorder. *ALTEX* **33**, 3, 243–249.

31. Uman, L. (2011). Systematic reviews and meta-analyses. *J Can Acad Child Adolesc Psychiatry* **20**, 57–59.

32. Wooding, S., Pollitt, A., Castle-Clarke, S., et al. (2014). Mental Health Retrosight: Understanding the returns from research (lessons from schizophrenia): Policy Report. Rand Health Quarterly 4. Available at https://www.rand.org/pubs/research_reports/RR325.html (Accessed 21.08.2019).

33. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

34. Frank, J. (2005). Technological lock-in, positive institutional feedback, and research on laboratory animals. *Structural Change and Economic Dynamics* **16**, 4, 557–575.

35. Bhatia, S.N., & Ingber, D.E. (2014). Microfluidic organs-on-chips. *Nature biotechnology* **32**, 8, 760-772.

36. Batool, M., Ahmad, B., & Choi, S. (2019). A Structure-Based Drug Discovery Paradigm. *International journal of molecular sciences* **20**, 11, 2783.

37. Romito, A., & Cobellis, G. (2016). Pluripotent stem cells: current understanding and future directions. *Stem cells international*, p. 9451492.

38. Licinio, J., & Wong, M. L. (2016). Serotonergic neurons derived from induced pluripotent stem cells (iPSCs): a new pathway for research on the biology and pharmacology of major depression. *Molecular Psychiatry* **21**, 1–2.

39. Knight, A. (2007). The poor contribution of chimpanzee experiments to biomedical progress. *J Appl Anim Welf Sci* **10**, 281–308.

40. Bailey, J. & Taylor, K. (2016). Non-human primates in neuroscience research: The case against its scientific necessity. *ATLA* **44**, 1, 43–69.

Citations of	NHP papers	In silico papers	In vitro papers
Redundant or Minor Relevance	42	43	301
Relevant for the Hypothesis or for the Methods	8	15	100

Table 1. Number of citations of redundant or minor relevance and relevant to the hypothesis or for the methods by human research papers on MDD.

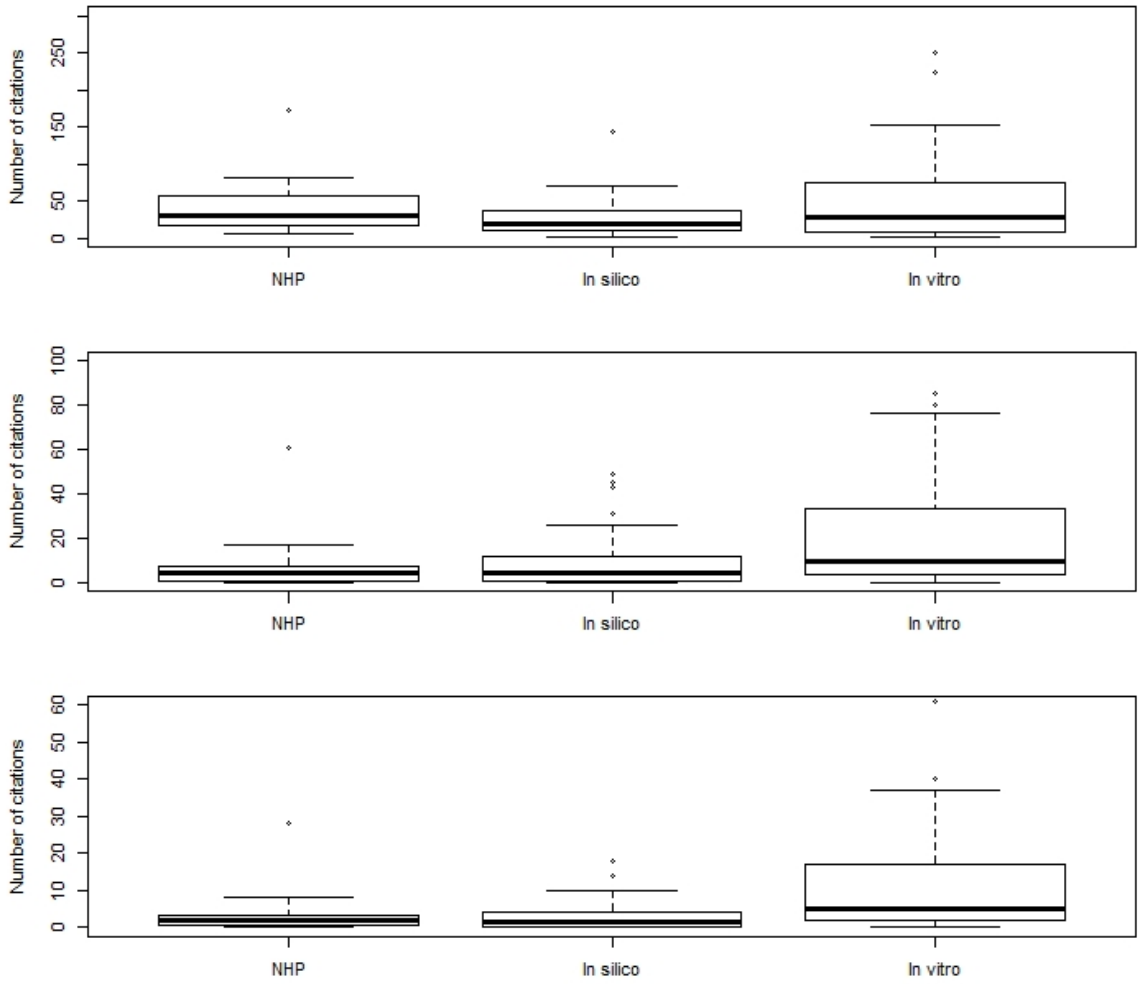

Fig. 1 Number of citations received by research method. The first row illustrates the total number of citations by research method; the second row illustrates the number of citations by human medical papers and the third row illustrates the number of citations by human medical paper on MDD. For visualization purposes, the largest observation in *In vitro*, included in the analysis was removed from the plots.