PLOS BIOLOGY

Analysis of animal-to-human translation shows that only 5% of animal-tested therapeutic interventions obtain regulatory approval for human applications

Benjamin V. Ineichen [1,2]*, **Eva Furrer**[1], **Servan L. Grüninger**[1,3], **Wolfgang E. Zürrer**[1], **Malcolm R. Macleod**[4]

1 Centre for Reproducible Science, University of Zurich, Zurich, Switzerland, **2** Clinical Neuroscience Center, University of Zurich, Zurich, Switzerland, **3** Department of Mathematics, University of Zurich, Zurich, Switzerland, **4** Centre for Clinical Brain Sciences, The University of Edinburgh, Edinburgh, United Kingdom

* benjamin.ineichen@uzh.ch

Abstract

There is an ongoing debate about the value of animal experiments to inform medical practice, yet there are limited data on how well therapies developed in animal studies translate to humans. We aimed to assess 2 measures of translation across various biomedical fields: (1) The proportion of therapies which transition from animal studies to human application, including involved timeframes; and (2) the consistency between animal and human study results. Thus, we conducted an umbrella review, including English systematic reviews that evaluated the translation of therapies from animals to humans. Medline, Embase, and Web of Science Core Collection were searched from inception until August 1, 2023. We assessed the proportion of therapeutic interventions advancing to any human study, a randomized controlled trial (RCT), and regulatory approval. We meta-analyzed the concordance between animal and human studies. The risk of bias was probed using a 10-item checklist for systematic reviews. We included 122 articles, describing 54 distinct human diseases and 367 therapeutic interventions. Neurological diseases were the focus of 32% of reviews. The overall proportion of therapies progressing from animal studies was 50% to human studies, 40% to RCTs, and 5% to regulatory approval. Notably, our meta-analysis showed an 86% concordance between positive results in animal and clinical studies. The median transition times from animal studies were 5, 7, and 10 years to reach any human study, an RCT, and regulatory approval, respectively. We conclude that, contrary to widespread assertions, the rate of successful animal-to-human translation may be higher than previously reported. Nonetheless, the low rate of final approval indicates potential deficiencies in the design of both animal studies and early clinical trials. To ameliorate the efficacy of translating therapies from bench to bedside, we advocate for enhanced study design robustness and the reinforcement of generalizability.

Introduction

Animal studies remain foundational in basic research, accounting for a substantial share of global biomedical research investment. These experiments have provided insight on aspects of human diseases and have paved the way for therapeutic innovations. For example, mitoxantrone and glatiramer acetate, FDA-approved drugs for multiple sclerosis, owe their inception at least partly to animal studies [1]. Yet, in recent years, concerns have grown about the low translatability of findings from animal experiments to humans, a concern that certain drugs with beneficial findings in animal experiments did not show similar effects in humans [2–4]. For example, while NXY-059 showed substantial promise in animal stroke studies, it failed in human trials [5]. Natalizumab displayed considerable efficacy in both animal and human multiple sclerosis trials, but the animal studies did not detect a severe side effect caused by a virus not present in rodents [6,7]. Opicinumab demonstrated significant promise in multiple sclerosis animal studies but failed its primary endpoint in human trials for multiple sclerosis [8,9].

The concerns of low translatability of animal research are particularly relevant to the debate on the ethical use of animals in research because clinical translation is one of the primary justifications for such research [10]. Discussions around the usefulness of animal experiments persist, but much of the current debate relies on anecdotal findings from discrete research areas [11]. High-level evidence—spanning various biomedical sectors and assessing translational success rates—is scarce. Hence, here we aimed to (1) offer a quantitative perspective on animal-to-human translation across diverse biomedical fields; (2) scrutinize the agreement between findings in animal and human drug development studies; and (3) probe the time intervals separating animal and human trials during drug development.

Materials and methods

Study registration

We registered the study protocol on the Open Science Framework platform (OSF, https://osf.io/jh2d8) and used the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines for reporting [12].

Approach to identification of therapeutic interventions

Our method to identify therapeutic interventions assessed for bench-to-bedside translation was identified via a two-stage process: in a first step, we identified systematic/scoping reviews assessing animal-to-human translation of specific therapeutic interventions (described in the Methods section "Search strategy for systematic reviews"). In the second step, these therapeutic interventions were included in our analysis of translational proportions, consistency of animal and human findings, and development times (described in the Methods section "Assessment of translational proportions and development times").

Search strategy for systematic reviews

We searched for studies published from inception up to August 01, 2023, in Medline (Ovid), Embase, and Web of Science Core Collection (Clarivate). We created the search string in Medline and translated to the other databases. The exact search strings are provided in the S1 Data. In brief, the search string comprised one block with terms for translation and one block for systematic/scoping reviews, and was limited to animal studies by employing the SYRCLE animal filter [13]. To probe the sensitivity of this approach, we also tested a broad search string comprising only the systematic review block and the SYRCLE animal filter (i.e., without the block for translation). Our protocol stated that we would use this broader strategy if we

identified in a subgroup of search returns additional studies equivalent to 5% of the total. However, we did not identify any additional eligible studies using the broad search strategy (i.e., 0%) and so we used the narrower search string for this systematic review (see protocol for details). We deduplicated the references in Endnote using the Bramer method [14].

Inclusion and exclusion criteria

Inclusion criteria. Systematic reviews or scoping reviews with or without meta-analysis which investigate translation of interventions in animal models of human diseases, i.e., the study must have the goal of assessing animal-to-human translation of therapies. Any type of intervention with the goal of improving at least 1 disease outcome was eligible (e.g., drugs, surgical interventions, neuromodulation, diets, behavioral therapy). The minimum requirement to be eligible as systematic review, scoping review, and/or meta-analysis was at abstract level: (1) Having at least 2 authors; (2) mentioning a systematic literature search; and (3) at full-text level having a paper section describing methodology of the systematic review.

Exclusion criteria. Original studies and/or studies not assessing bench-to-bedside translation, non-English articles, and gray literature (conference abstracts, book chapters). We excluded non-systematic reviews but retained them to find potential additional references.

Study selection and data extraction

Three independent reviewers (BVI, SG, and EF) screened titles and abstracts of studies for their relevance in the web-based application SyRF (RRID: SCR_018907) in duplicate [15]. We resolved discrepancies by discussion. Subsequently to full-text screening, we extracted the following data: bibliographic data (author names, journal, title, publication year, digital object identifier), number of included animal studies and clinical trials, disease classes, any data related to translation, any provided definition on translation as well as data on year/outcome of clinical studies (see below). We extracted all data from text/tables if possible, and if not extracted from figures using Universal Desktop Ruler [16].

Critical appraisal of included studies

We assessed the quality of each included study against predefined criteria by 3 independent reviewers in duplicate (BVI, WEZ, and EF), based on a checklist proposed by Sena and colleagues [17], and extended with additional items for a more granular critical appraisal. Concretely: (1) Was an a priori study protocol defined? (2) Was a flowchart for study selection provided? (3) Was a conflict-of-interest statement provided? (4) Was screening and/or extraction conducted by 2 or more reviewers? (5) Was a clear research question defined? (6) Were in- and exclusion criteria reported? (7) Were 2 or more literature databases searched? (8) Was a search date provided? (9) Was a search string provided? (10) Was a critical appraisal conducted? (11) Did the study mention alignment with relevant guidelines, e.g., SYRCLE, CAMARADES, or PRISMA? We resolved discrepancies by discussion. This appraisal method was initially tested in a separate umbrella review. Detailed application guidelines are available in the S2 Data. The inter-rater agreement was calculated using Cohen's Kappa. Of note, we included all systematic reviews into our final analysis, regardless of their risk of bias.

Assessment of translational proportions and development times

Definition of translation. We used the following working definition of translation: the process of turning observations from animal experiments into interventions that improve the health of human individuals and the public [18].

Study year and outcomes.   For each individual intervention identified in the eligible systematic reviews, 2 independent reviewers extracted: (1) the first published animal study testing the respective intervention; and (2) the first clinical study testing the respective intervention. We included any type of clinical studies including pilot studies or case series, and (3) the first randomized controlled trial (RCT) testing the respective intervention. In addition, for all clinical studies/RCTs, we extracted the main study outcome as defined by the authors of the respective study, grouped into 4 classes, i.e., whether the intervention had a positive, negative, mixed, or neutral effect on the corresponding disease outcome [19], any outcome was considered, e.g., primary or secondary outcome. We extracted these data in first priority from the respective systematic reviews. If these data were not available in the systematic reviews, we searched Medline (Ovid) and Embase for corresponding clinical studies/RCTs. For this, we used a search string comprising the intervention and disease name, including synonyms. For clinical approval of an intervention, we considered FDA approval (by searching the FDA webpage for respective therapies) as well as UK and Swiss medical guidelines for recommended use of respective interventions. We resolved discrepancies by discussion.

Data synthesis and analysis

Narrative synthesis and descriptive statistics.   We provide a narrative summary of the extent of bench-to-bedside translation, supported by descriptive statistics of extracted parameters.

Meta-analysis on relative risks.   To assess the concordance between animal and human studies, we conducted a meta-analysis on relative risks, i.e., the ratio of the proportion of positive animal studies to the proportion of positive clinical studies. We conducted this only in case of specific interventions for specific diseases entities (e.g., atorvastatin for glioblastoma). To reduce noise of the dataset, we restricted our analysis to therapies that were the subject of 5 or more published animal studies. As primary outcome, we pooled relative risks to obtain an overall relative risk and 95% confidence intervals. We fitted a random-effects model to the data [20] and estimated the amount of heterogeneity, i.e., τ^2, using the DerSimonian—Laird estimator. We calculated the Q-test for heterogeneity and the I^2 statistic. We used the R package *meta* for the meta-analysis [21] (RRID: SCR_019055).

Kaplan–Meier survival analysis for development times.   To estimate the time-to-event from first animal study to clinical studies/RCTs and eventually official endorsement (i.e., the lag time), we conducted a Kaplan—Meier analysis. We used the packages *survminer* (RRID: SCR_021094) and *survival* (RRID: SCR_021137) for this survival analysis.

All statistical analyses were conducted in the R programming environment (version 4.2.2). We considered a two-tailed P value < 0.05 statistically significant.

Results

Eligible publications and general study characteristics

Eligible studies.   In total, 5,227 original publications were retrieved from our database search, and 1 publication from reference lists of reviews on related topics. After abstract and title screening, 656 publications were eligible for full-text search. After screening the full text of these studies, 122 articles (2% of deduplicated references) were included for qualitative synthesis (Table 1) and a subset of 62 for quantitative synthesis (Fig 1).

The eligible systematic reviews comprised a total of 4,443 animal studies and 1,516 clinical studies (median 21 animal and 8 human studies per systematic review).

Most studies have been published in the last 5 years (88 since 2018). Most studies were from the United States of America (27 studies, 22%), Canada (19, 16%), the Netherlands, (16,

Table 1. Number of included translational systematic reviews, diseases/conditions, and therapeutic interventions.

Number of systematic reviews included		122		
Number of unique diseases/conditions		54		
Number of unique therapeutic interventions (including drug and non-drug interventions)	367	No. in any human study		165 (45%)
		No. in an RCT		132 (36%)
		No. with (FDA) approval		14 (4%)

The percentage in brackets refers to the relative numbers of therapeutic interventions entering the respective therapy development stage.

The data underlying this table can be found on https://osf.io/frjm4 (Sheet: *Mastersheet_including_RoB*).

FDA, food and drug administration; No., number; RCT, randomized controlled trial.

https://doi.org/10.1371/journal.pbio.3002667.t001

13%), Australia (13, 11%), Italy (13, 11%), and the United Kingdom (10, 8%). Thirty-six studies were collaborative efforts between 2 or more countries.

Diseases and therapies. The studies covered 54 unique different human diseases/conditions (Table 1). The most common ICD-11 disease classes addressed by the systematic reviews
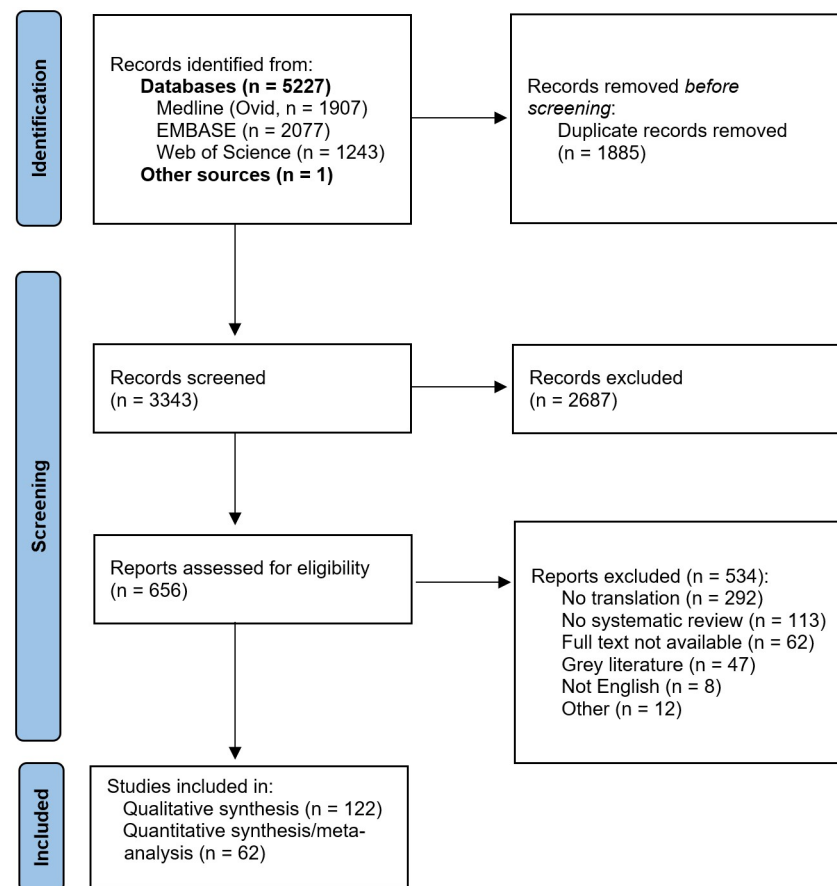
Fig 1. Flow chart for study inclusion. Flow chart for inclusion of systematic reviews or scoping reviews with or without meta-analysis which investigate translation of interventions in animal models of human diseases.

https://doi.org/10.1371/journal.pbio.3002667.g001
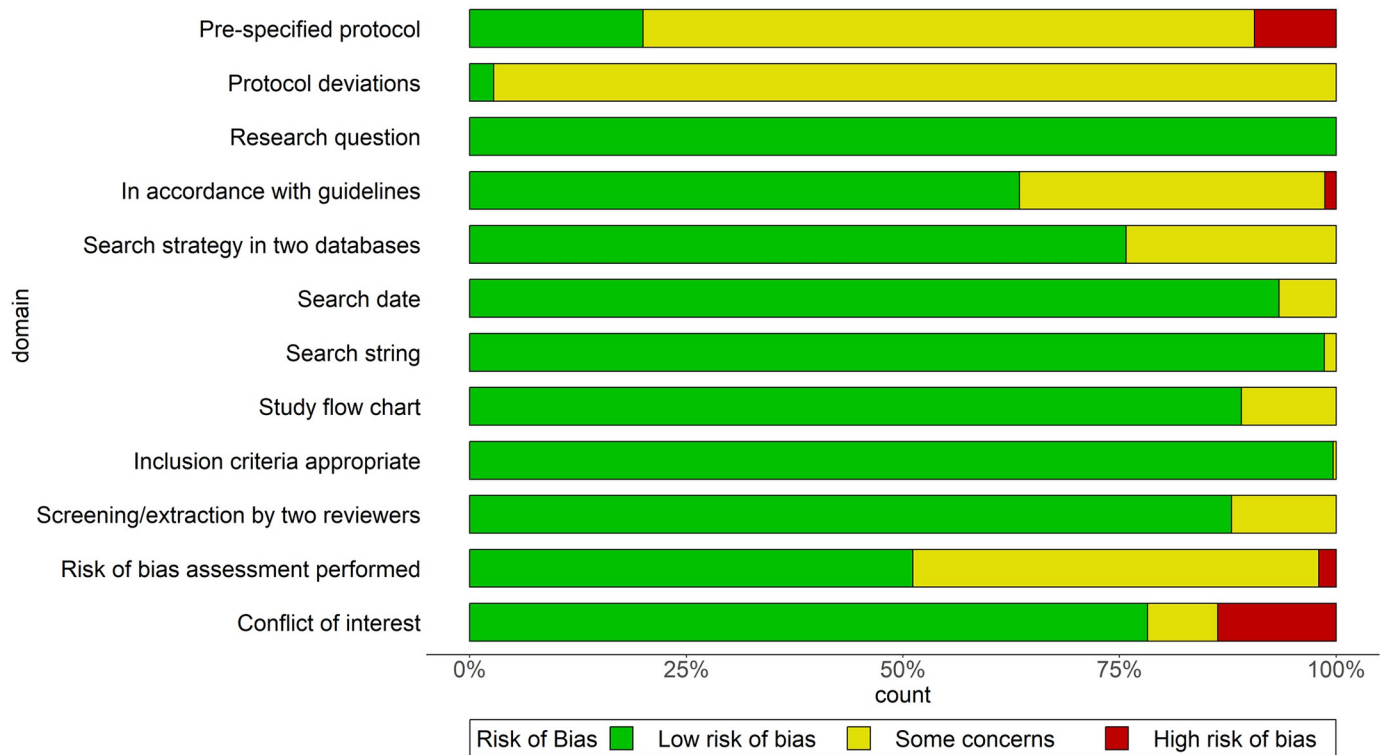
Fig 2. Risk of bias assessment for included systematic reviews. The data underlying this figure can be found on https://osf.io/frjm4 (Sheet: *Mastersheet_including_RoB*). The code underlying this figure can be found S1 Code (https://osf.io/9fgru).

https://doi.org/10.1371/journal.pbio.3002667.g002

were "diseases of the nervous system" (32% of studies), "diseases of the musculoskeletal system and connective tissue" (11%), "mental and behavioral disorders" (9%), "diseases of the circulatory system" (9%), "diseases of the digestive system" (8%), and "neoplasms" (8%).

These reviews included a total of 367 unique therapeutic interventions including drug and non-drug treatments (Table 1). The median interval between the first animal experiment and publication of the respective systematic review was 15 years (range: 3 to 63 years, total observation time 6,736 years). All therapies and diseases are listed in S1 Table.

Risk of bias assessment. The inter-rater agreement Kappa was 0.76. We considered most of the included reviews to be at low risk of bias for providing a search date, a search string, a study flow chart, and reporting screening and data extraction by 2 reviewers. However, few studies published a study protocol, and a substantial number of systematic reviews did not perform a risk of bias assessment, thus posing high risk of bias in these domains (Fig 2). Given that most included systematic reviews were deemed at low risk of bias, we anticipated no relevant impact on our overall conclusions.

Qualitative summary of translation in different biomedical fields

Therapies were tested in a variety of diseases, including neurological, musculoskeletal, psychiatric, circulatory system, digestive, skin, lung, and metabolic diseases (S2–S8 Tables).

Most studies discussed potential hurdles for translation of findings from animal to human studies. A prevalent observation was the disparity in methodological approaches between animal experiments and human studies. Specifically, several studies highlighted that experimental conditions in animal research often do not mirror clinically relevant scenarios [22–33]. For

example, treatments for stroke were frequently tested on young, healthy animals, which contrast the typically multimorbid elderly patient population in clinical settings [22]. Poor study quality and inadequate reporting, predominantly in animal studies, were also recurrent concerns [34–36]. A noticeable reduction in effect size from animal to human studies was documented by several reviews [24,37–40]. This trend was substantiated by a systematic review and meta-analysis which examined the preclinical-to-clinical development trajectory of 37 treatments for acute ischemic stroke, encompassing 50 phase 3 clinical trials, 75 early clinical trials, and 209 animal studies [36]: It observed a progressive reduction in efficacy from animal research to early clinical trials and then to Phase 3 clinical trials. This decline was attributed to shortcomings in preclinical study rigor (such as the absence of randomization and blinding), differences in study design, including the use of differing outcomes, and insufficient statistical power in both animal studies and preliminary clinical studies.

Another comprehensive systematic review covering therapies for cardiac arrest encompassed 415 animal and 43 clinical studies. This review, which evaluated 190 pharmacological interventions, found a limited number of positive outcomes in clinical studies [41]. In addition, many animal studies were conducted subsequent to the publication of a corresponding clinical study [41]. And similar to stroke, there were substantial variation in experimental methodologies between animal and human studies. For example, drugs were typically administered approximately 9.5 min post-cardiac arrest in animal studies, compared to roughly 19.4 min in human studies [42].

Finally, one study traced the developmental pathway of an oncolytic virus used in cancer therapy [43]. While animal experiments exhibited 80% to 100% regression rates in tumours, human studies only demonstrated a range of 0% to 24%. Intriguingly, more rigorously conducted studies showed smaller effect sizes. And the authors emphasize that even successful biotherapeutic interventions might not present a straight-forward translational journey.

Quantitative overview of therapy translation overall and by discipline

In these systematic reviews, and only accounting for therapies where more than 10 years has elapsed since the initial animal experiment, 50%, 40%, and 5% of therapies entered any human study, an RCT, or have been (FDA) approved (281 therapies, Fig 3A). Diseases of the circulatory system (166 therapies, including stroke: 34%, 29%, and 1%, respectively) and mental health disorders (16 therapies: 50%, 31%, and 0%, respectively) showed particularly low translational proportions (Fig 3B and 3D). Diseases of the musculoskeletal system (13 therapies, 100%, 62%, and 15%, respectively) and cancer (15 therapies, 73%, 47%, and 20%, respectively) showed relatively high translational proportions (Fig 3E and 3F). Translational proportions considering all therapies independent of time gap between animal and clinical studies were slightly lower (Table 1 and S1 Fig). Translational proportions per disease are illustrated in Fig 4 with cardiac arrest, multiple sclerosis, and stroke assessing the most therapeutic interventions in animals (S9 Table).

Chronology of animal-to-human translation

Because less than half of the therapies transitioned from animal experiments to clinical studies, we were not able to estimate the overall median durations (Fig 5). However, if only considering therapeutic interventions entering any clinical study, an RCT, or obtaining (FDA) approval, the median lag times were 5 years [95% CI: 5 to 6], 7 years [95% CI: 6 to 8], and 10 years [95% CI: 4-not estimable], respectively (S2 Fig). The maximum time from first animal study to any clinical trial, and RCT, or FDA approval was 44 years, 58 years, and 34 years, respectively.
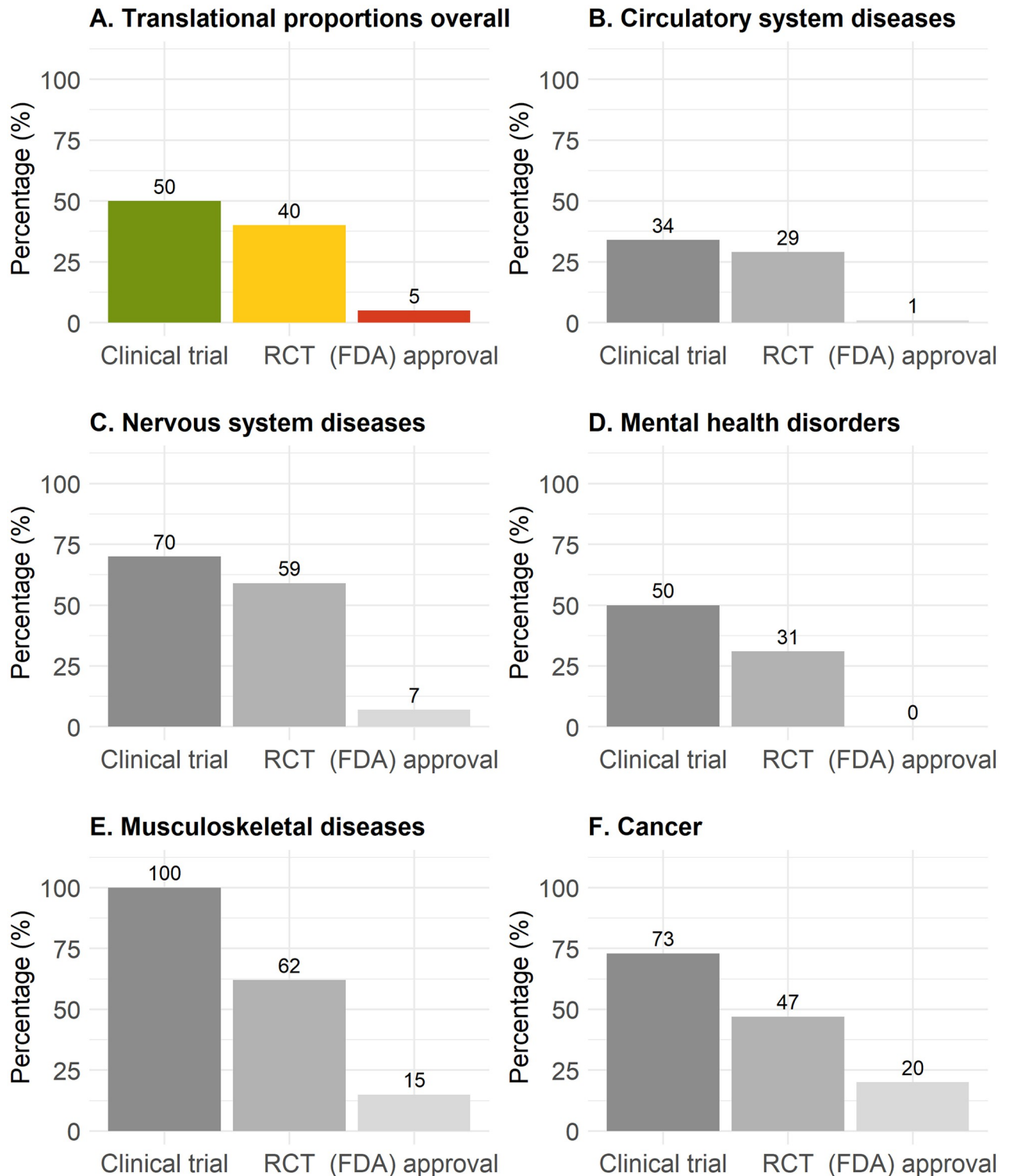
Fig 3. Translational proportions for all therapies with ≥10 years development time since the first published animal study. Proportions of translation from animals to any clinical study (green), to an RCT (yellow), or to (FDA) approval (red) overall (A), for circulatory system diseases (B), for neurological diseases (C), for mental health disorders (D), for musculoskeletal diseases (E), and for cancer (F). The data underlying this figure can be found on https://osf.io/frjm4 (Sheet: *Translation*). The code underlying this figure can be found in S1 Code (https://osf.io/9fgru). RCT, randomized controlled trial.

https://doi.org/10.1371/journal.pbio.3002667.g003
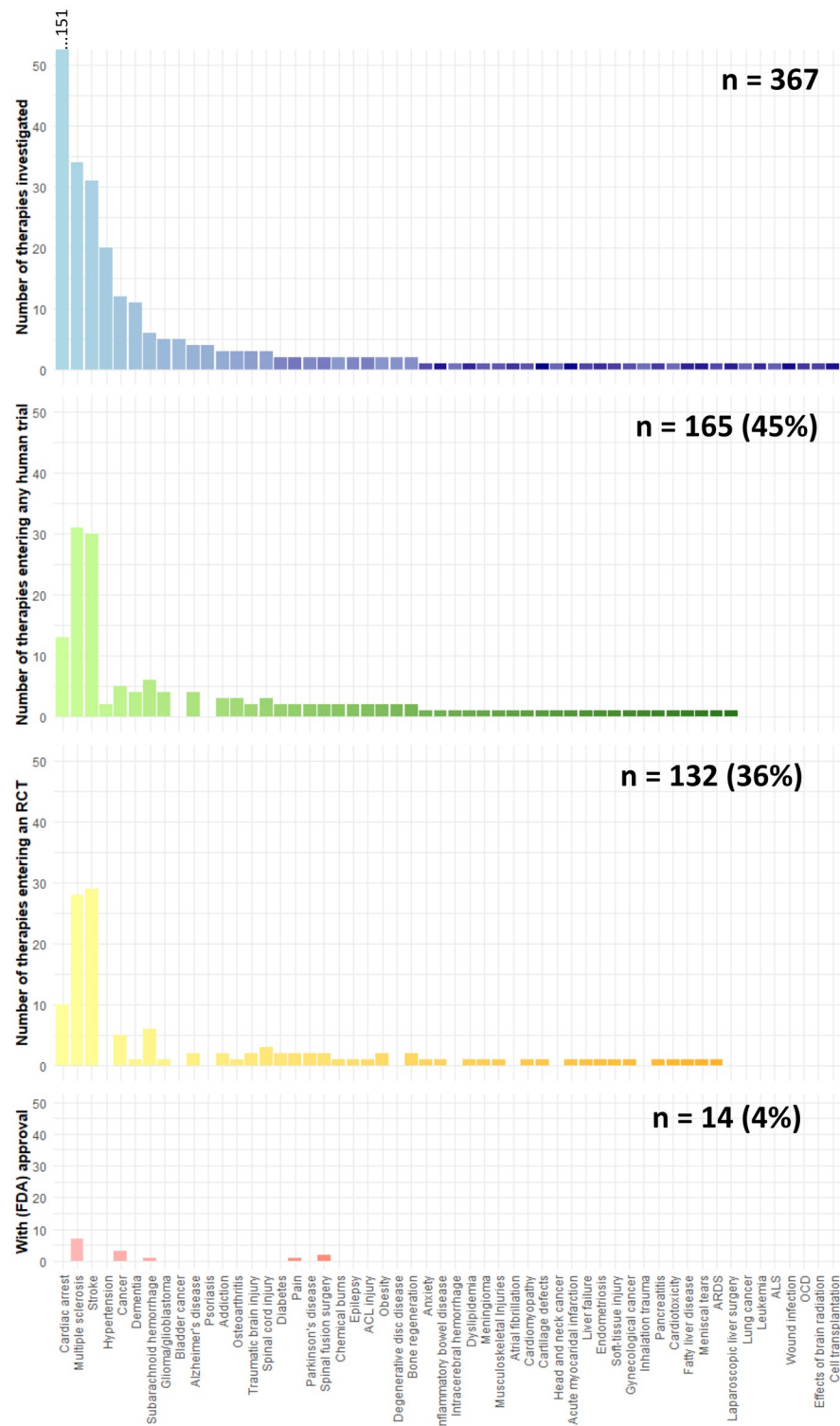
Fig 4. Progression of therapies through different stages of clinical development, categorized by disease or condition. The quantity of therapies that are initially tested in animal studies (blue), subsequently entering any clinical trial (green), advance to a randomized controlled trial (yellow), and eventually achieve regulatory approval (red). The total number of therapies for each category is annotated at the upper right corner of the respective graph. The data underlying this figure can be found on https://osf.io/frjm4 (Sheet: *Diseases*). The code underlying this figure can be found S1 Code (https://osf.io/9fgru). ACL, anterior cruciate ligament; ALS, amyotrophic lateral sclerosis; ARDS, acute respiratory distress syndrome; OCD, obsessive compulsive disorder.

https://doi.org/10.1371/journal.pbio.3002667.g004
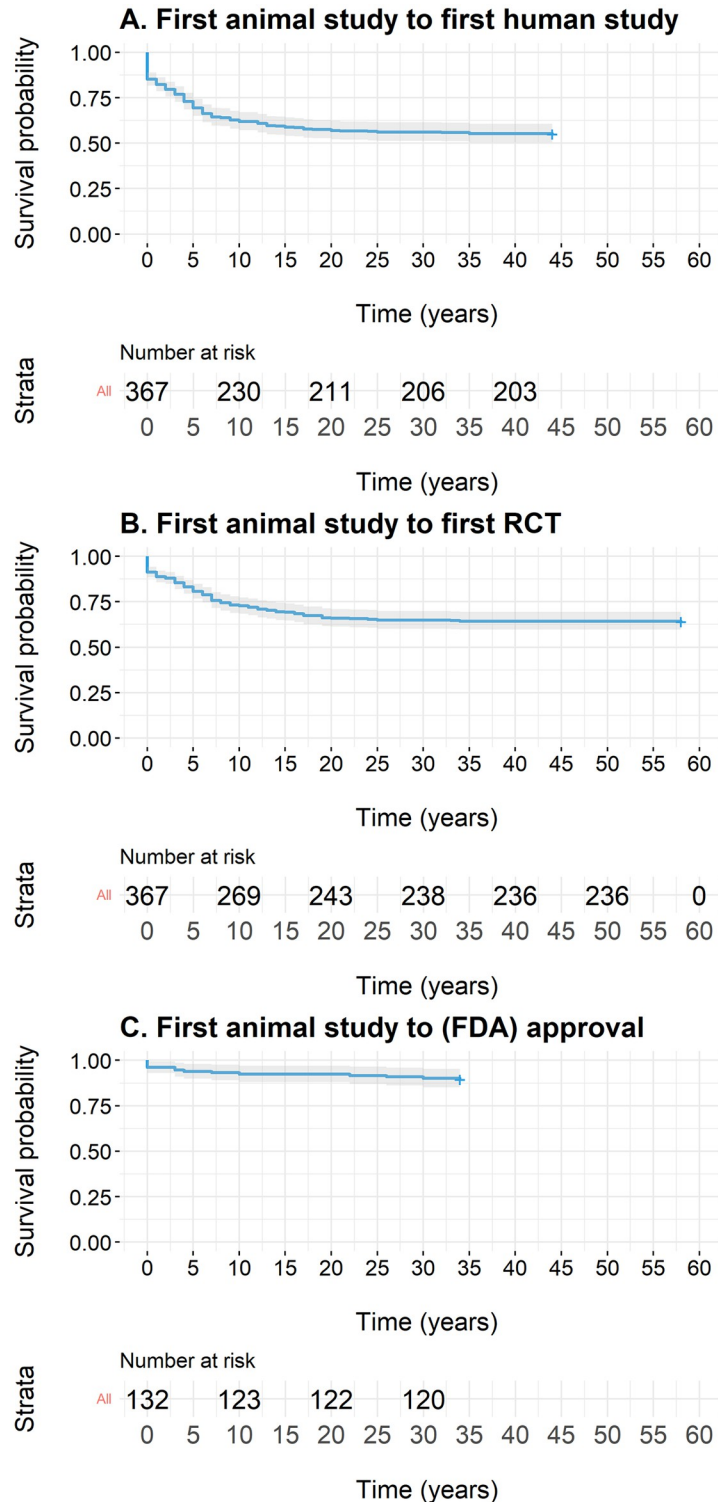
Fig 5. Lag times for clinical therapy development from first animal study. Lag times for therapies from first animal study to any clinical study (A), to a randomized controlled trial (RCT, B), or to (FDA) approval (C). The data underlying this figure can be found on https://osf.io/frjm4 (Sheet: *Translation*). The code underlying this figure can be found S1 Code (https://osf.io/9fgru). RCT, randomized controlled trial.

https://doi.org/10.1371/journal.pbio.3002667.g005

Table 2. Outcome of animal studies, clinical studies, and RCTs.

	Positive (%)	Neutral (%)	Mixed (%)	Negative (%)	Total
Animal studies	1,181 (79%)	278 (19%)	10 (<1%)	27 (<2%)	1,496
Clinical studies	317 (62%)	178 (35%)	5 (<1%)	15 (<3%)	515
RCTs	111 (50%)	103 (47%)	2 (<1%)	4 (2%)	220

We restricted our analysis to therapies that were the subject of 5 or more published animal studies, encompassing 62 therapeutic interventions.

The data underlying this table can be found on https://osf.io/frjm4 (Sheet: *Mastersheet_including_RoB*).

RCT, randomized controlled trial.

https://doi.org/10.1371/journal.pbio.3002667.t002

Notably, in several instances, the first animal experiment was documented after the first clinical trial (49 therapies, representing 31%) or RCT (28 therapies, accounting for 22%).

Concordance between animal and human studies

In evaluating concordance, we restricted our analysis to therapies that were the subject of 5 or more published animal studies, encompassing 62 therapies. These therapies were examined across 1,496 animal studies, 515 clinical studies, and 220 RCTs. Out of these, positive outcomes were identified in 1,181 (79%) animal studies, 317 (61%) clinical studies, and 111 (50%) RCTs (Table 2). The overall alignment between positive outcomes in animal and clinical studies, represented by the relative risk, was 0.86 [95% CI: 0.80 to 0.92] (Fig 6). This alignment was especially pronounced in therapies for neurological diseases at 0.99 [0.66 to 1.06], involving 23 therapies, and circulatory system diseases (inclusive of stroke) at 0.88 [0.76 to 1.01], involving 12 therapies. Conversely, lower concordance was observed for digestive system diseases (0.86 [0.75 to 0.99], 7 therapies), musculoskeletal diseases (0.67 [0.515 to 0.862], 6 therapies), cancer (0.66 [0.49 to 0.88], 3 therapies), and mental health disorders (0.60 [0.37 to 0.97], 7 therapies) (S2–S8 Figs).

Discussion

Main findings

Our umbrella systematic review evaluated (1) the proportion of therapies which translate from animal studies to human application, including timeframes; and (2) the consistency between animal and human study results. We observe a notable consistency between results from animal and human studies including a relatively large proportion of therapeutic interventions entering early clinical trials. However, only a minority of therapeutic intervention achieved regulatory approval.

Findings in the context of existing evidence

Translation across biomedical fields. Our review shows a high consistency between findings from animal and human studies, similar to studies outside of therapy translation [44]. In addition, a surprisingly high proportions of therapies entered early clinical development: half of the therapeutic interventions made the transition from animal studies to early human clinical studies (34% to 100% across different biomedical fields). Furthermore, 40% of these therapies progressed to the more rigorous RCT stage (29% to 62% for different biomedical fields). However, a strikingly low proportion—only 5%—of therapies achieved official approval (0% to 20% across the biomedical spectrum).

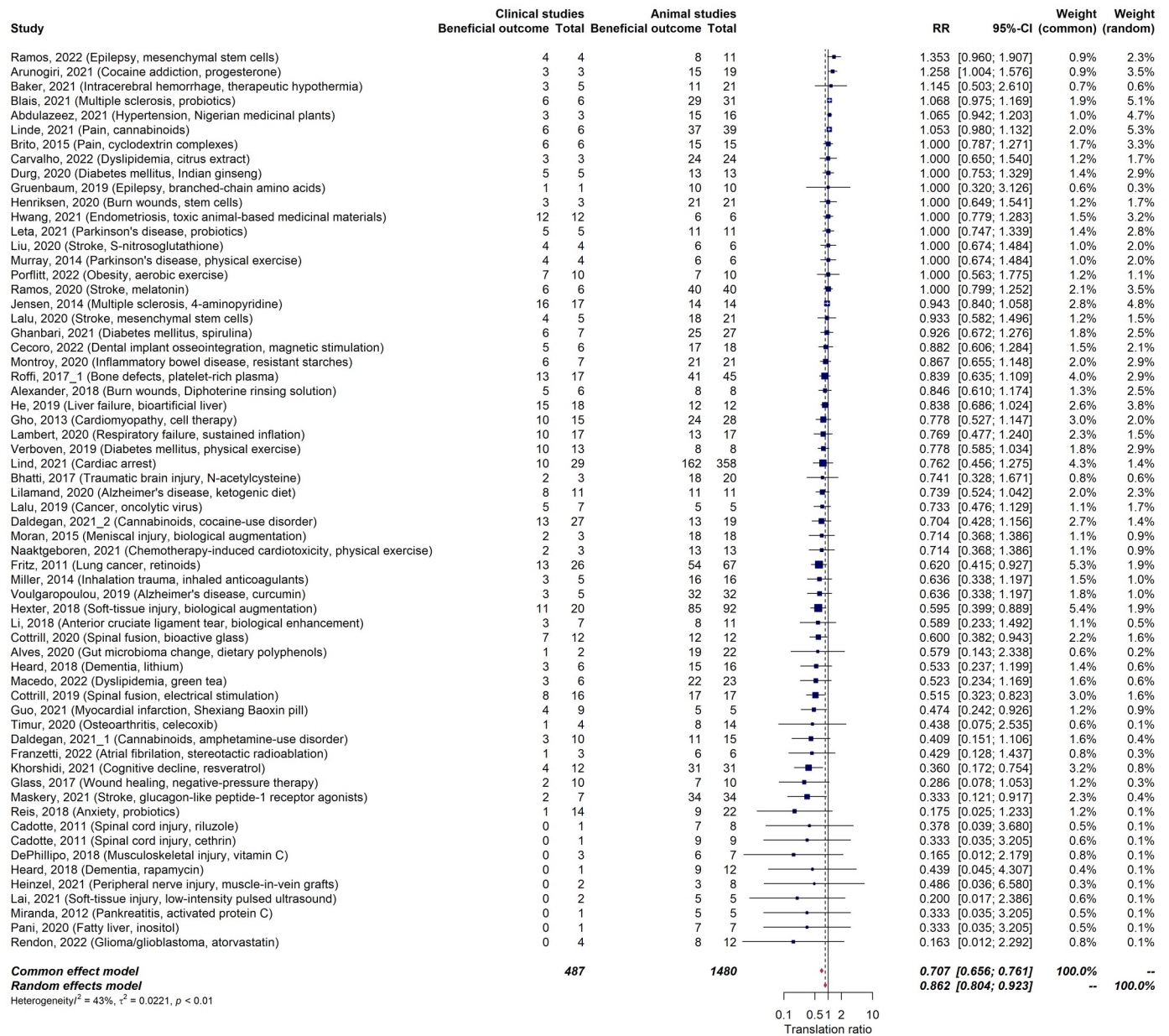| Study	Clinical studies Beneficial outcome	Total	Animal studies Beneficial outcome	Total		RR	95%-CI	Weight (common)	Weight (random)
Ramos, 2022 (Epilepsy, mesenchymal stem cells)	4	4	8	11		1.353	[0.960; 1.907]	0.9%	2.3%
Arunogiri, 2021 (Cocaine addiction, progesterone)	3	3	15	19		1.258	[1.004; 1.576]	0.9%	3.5%
Baker, 2021 (Intracerebral hemorrhage, therapeutic hypothermia)	3	5	11	21		1.145	[0.503; 2.610]	0.7%	0.6%
Blais, 2021 (Multiple sclerosis, probiotics)	6	6	29	31		1.068	[0.975; 1.169]	1.9%	5.1%
Abdulazeez, 2021 (Hypertension, Nigerian medicinal plants)	3	3	15	16		1.065	[0.942; 1.203]	1.0%	4.7%
Linde, 2021 (Pain, cannabinoids)	6	6	37	39		1.053	[0.980; 1.132]	2.0%	5.3%
Brito, 2015 (Pain, cyclodextrin complexes)	6	6	15	15		1.000	[0.787; 1.271]	1.7%	3.3%
Carvalho, 2022 (Dyslipidemia, citrus extract)	3	3	24	24		1.000	[0.650; 1.540]	1.2%	1.7%
Durg, 2020 (Diabetes mellitus, Indian ginseng)	5	5	13	13		1.000	[0.753; 1.329]	1.4%	2.9%
Gruenbaum, 2019 (Epilepsy, branched-chain amino acids)	1	1	10	10		1.000	[0.320; 3.126]	0.6%	0.3%
Henriksen, 2020 (Burn wounds, stem cells)	3	3	21	21		1.000	[0.649; 1.541]	1.2%	1.7%
Hwang, 2021 (Endometriosis, toxic animal-based medicinal materials)	12	12	6	6		1.000	[0.779; 1.283]	1.5%	3.2%
Leta, 2021 (Parkinson's disease, probiotics)	5	5	11	11		1.000	[0.747; 1.339]	1.4%	2.8%
Liu, 2020 (Stroke, S-nitrosoglutathione)	4	4	6	6		1.000	[0.674; 1.484]	1.0%	2.0%
Murray, 2014 (Parkinson's disease, physical exercise)	4	4	6	6		1.000	[0.674; 1.484]	1.0%	2.0%
Porflitt, 2022 (Obesity, aerobic exercise)	7	10	7	10		1.000	[0.563; 1.775]	1.2%	1.1%
Ramos, 2020 (Stroke, melatonin)	6	6	40	40		1.000	[0.799; 1.252]	2.1%	3.5%
Jensen, 2014 (Multiple sclerosis, 4-aminopyridine)	16	17	14	14		0.943	[0.840; 1.058]	2.8%	4.8%
Lalu, 2020 (Stroke, mesenchymal stem cells)	4	5	18	21		0.933	[0.582; 1.496]	1.2%	1.5%
Ghanbari, 2021 (Diabetes mellitus, spirulina)	6	7	25	27		0.926	[0.672; 1.276]	1.8%	2.5%
Cecoro, 2022 (Dental implant osseointegration, magnetic stimulation)	5	6	17	18		0.882	[0.606; 1.284]	1.5%	2.1%
Montroy, 2020 (Inflammatory bowel disease, resistant starches)	6	7	21	21		0.867	[0.655; 1.148]	2.0%	2.9%
Roffi, 2017_1 (Bone defects, platelet-rich plasma)	13	17	41	45		0.839	[0.635; 1.109]	4.0%	2.9%
Alexander, 2018 (Burn wounds, Diphoterine rinsing solution)	5	6	8	8		0.846	[0.610; 1.174]	1.3%	2.5%
He, 2019 (Liver failure, bioartificial liver)	15	18	12	12		0.838	[0.686; 1.024]	2.6%	3.8%
Gho, 2013 (Cardiomyopathy, cell therapy)	10	15	24	28		0.778	[0.527; 1.147]	3.0%	2.0%
Lambert, 2020 (Respiratory failure, sustained inflation)	10	17	13	17		0.769	[0.477; 1.240]	2.3%	1.5%
Verboven, 2019 (Diabetes mellitus, physical exercise)	10	13	8	8		0.778	[0.585; 1.034]	1.8%	2.9%
Lind, 2021 (Cardiac arrest)	10	29	162	358		0.762	[0.456; 1.275]	4.3%	1.4%
Bhatti, 2017 (Traumatic brain injury, N-acetylcysteine)	2	3	18	20		0.741	[0.328; 1.671]	0.8%	0.6%
Lilamand, 2020 (Alzheimer's disease, ketogenic diet)	8	11	11	11		0.739	[0.524; 1.042]	2.0%	2.3%
Lalu, 2019 (Cancer, oncolytic virus)	5	7	5	5		0.733	[0.476; 1.129]	1.1%	1.7%
Daldegan, 2021_2 (Cannabinoids, cocaine-use disorder)	13	27	13	19		0.704	[0.428; 1.156]	2.7%	1.4%
Moran, 2015 (Meniscal injury, biological augmentation)	2	3	18	18		0.714	[0.368; 1.386]	1.1%	0.9%
Naaktgeboren, 2021 (Chemotherapy-induced cardiotoxicity, physical exercise)	2	3	13	13		0.714	[0.368; 1.386]	1.1%	0.9%
Fritz, 2011 (Lung cancer, retinoids)	13	26	54	67		0.620	[0.415; 0.927]	5.3%	1.9%
Miller, 2014 (Inhalation trauma, inhaled anticoagulants)	3	5	16	16		0.636	[0.338; 1.197]	1.5%	1.0%
Voulgaropoulou, 2019 (Alzheimer's disease, curcumin)	3	5	32	32		0.636	[0.338; 1.197]	1.8%	1.0%
Hexter, 2018 (Soft-tissue injury, biological augmentation)	11	20	85	92		0.595	[0.399; 0.889]	5.4%	1.9%
Li, 2018 (Anterior cruciate ligament tear, biological enhancement)	3	7	8	11		0.589	[0.233; 1.492]	1.1%	0.5%
Cottrill, 2020 (Spinal fusion, bioactive glass)	7	12	12	12		0.600	[0.382; 0.943]	2.2%	1.6%
Alves, 2020 (Gut microbioma change, dietary polyphenols)	1	2	19	22		0.579	[0.143; 2.338]	0.6%	0.2%
Heard, 2018 (Dementia, lithium)	3	6	15	16		0.533	[0.237; 1.199]	1.4%	0.6%
Macedo, 2022 (Dyslipidemia, green tea)	3	6	22	23		0.523	[0.234; 1.169]	1.6%	0.6%
Cottrill, 2019 (Spinal fusion, electrical stimulation)	8	16	17	17		0.515	[0.323; 0.823]	3.0%	1.6%
Guo, 2021 (Myocardial infarction, Shexiang Baoxin pill)	4	9	5	5		0.474	[0.242; 0.926]	1.2%	0.9%
Timur, 2020 (Osteoarthritis, celecoxib)	1	4	8	14		0.438	[0.075; 2.535]	0.6%	0.1%
Daldegan, 2021_1 (Cannabinoids, amphetamine-use disorder)	3	10	11	15		0.409	[0.151; 1.106]	1.6%	0.4%
Franzetti, 2022 (Atrial fibrillation, stereotactic radioablation)	1	3	6	6		0.429	[0.128; 1.437]	0.8%	0.3%
Khorshidi, 2021 (Cognitive decline, resveratrol)	4	12	31	31		0.360	[0.172; 0.754]	3.2%	0.8%
Glass, 2017 (Wound healing, negative-pressure therapy)	2	10	7	10		0.286	[0.078; 1.053]	1.2%	0.3%
Maskery, 2021 (Stroke, glucagon-like peptide-1 receptor agonists)	2	7	34	34		0.333	[0.121; 0.917]	2.3%	0.4%
Reis, 2018 (Anxiety, probiotics)	1	14	9	22		0.175	[0.025; 1.233]	1.2%	0.1%
Cadotte, 2011 (Spinal cord injury, riluzole)	0	1	7	8		0.378	[0.039; 3.680]	0.5%	0.1%
Cadotte, 2011 (Spinal cord injury, cethrin)	0	1	9	9		0.333	[0.035; 3.205]	0.6%	0.1%
DePhillipo, 2018 (Musculoskeletal injury, vitamin C)	0	3	6	7		0.165	[0.012; 2.179]	0.8%	0.1%
Heard, 2018 (Dementia, rapamycin)	0	1	9	12		0.439	[0.045; 4.307]	0.4%	0.1%
Heinzel, 2021 (Peripheral nerve injury, muscle-in-vein grafts)	0	2	3	8		0.486	[0.036; 6.580]	0.3%	0.1%
Lai, 2021 (Soft-tissue injury, low-intensity pulsed ultrasound)	0	2	5	5		0.200	[0.017; 2.386]	0.6%	0.1%
Miranda, 2012 (Pankreatitis, activated protein C)	0	1	5	5		0.333	[0.035; 3.205]	0.5%	0.1%
Pani, 2020 (Fatty liver, inositol)	0	1	7	7		0.333	[0.035; 3.205]	0.5%	0.1%
Rendon, 2022 (Glioma/glioblastoma, atorvastatin)	0	4	8	12		0.163	[0.012; 2.292]	0.8%	0.1%
Common effect model		**487**		**1480**		**0.707**	**[0.656; 0.761]**	**100.0%**	**--**
Random effects model						**0.862**	**[0.804; 0.923]**	**--**	**100.0%**

Heterogeneity: $I^2 = 43\%$, $\tau^2 = 0.0221$, $p < 0.01$

0.1   0.5 1 2   10
Translation ratio

Fig 6. Meta-analysis on concordance rate between animal and human studies (relative risk). Forest plot of concordance rates (relative risks) between animal and human studies. A random-effects model was fitted to the data. The data underlying this figure can be found on https://osf.io/frjm4 (Sheet: *Mastersheet_including_RoB*). The code underlying this figure can be found S1 Code (https://osf.io/9fgru). CI, confidence interval; RR, relative risk.

How can we make sense of the fact that animal studies and early clinical trials seem to show promise, yet there is very limited official approval for these therapies? There are 2 possible explanations: One scenario is that the strict requirements of RCTs and regulatory approval are causing many potentially valuable treatments to be left behind. The other scenario is that both animal studies and early clinical trials may have limitations in their design, such as a lack of proper randomization and blinding, which affects their internal validity [45]. This could lead to unreliable findings in both domains, ultimately resulting in the exclusion of these therapies in more rigorous clinical trial settings like RCTs. This line of reasoning also includes the so

called efficacy-effectiveness gap, i.e., the differences in outcomes between patients treated in ideal and controlled circumstances of clinical trials versus real-world scenarios [46,47].

We lean towards the second scenario for 2 reasons. First, as therapies progress to more rigorous study designs, their numbers do decrease as shown by our data, which contrasts to the mostly small and uncontrolled early clinical trials where these therapies were initially tested. Second, drawing from the field of clinical neurology, many therapies that have shown promise in animal studies and early trials reported as successful candidates herein, such as melatonin and mesenchymal stem cells for stroke, have not yet become standard clinical practice [48]. A similar pattern can be seen in other neurological diseases like Alzheimer's disease and spinal cord injury, where there are several therapies with promising preclinical results but limited practical translation [49,50].

Potential hurdles for successful translation. Several factors contribute to the challenges in translating therapies from animal models to human application, as discussed by many of the included reviews. First, there is a notable discrepancy in the contexts of animal testing versus human application. For example, treatment strategies tested on young, healthy animals, such as those for stroke, may not directly apply to the more complex scenarios of elderly patients with multiple health conditions. Second, there is an overall poor quality of many animal studies. These studies often have inherent design flaws, lacking critical elements like blinding or randomization. This absence can bias the results and affect their applicability to the human case, i.e., their external validity. Third, there seems to be a disconnect between animal and human research [51]. This warrants a stronger focus on educating a new generation of translational scientists [52]. Fourth, when it comes to human studies, they can suffer from being underpowered or relying on outcome measures that do not capture the efficacy of a treatment [53]. For example, early phase clinical trials testing interventions for neurological diseases are commonly underpowered [54]. Similarly, clinical trials may use trial outcomes not genuinely reflecting real-world patient settings such as complex composite outcomes, commonly seen in trials of cardiovascular diseases [55] or assessing cognitive domains in dementia trials not relevant for patients [56]. Lastly, animal and human studies commonly address different questions: whereas animal studies tend to focus on mechanisms, human studies tend to focus on effectiveness of an intervention.

How can we improve translation? While finding a straightforward solution is challenging, we emphasize 2 crucial elements to enhance the transition from preclinical studies to clinical applications: the robustness and generalizability of data. ALS research illustrates data robustness issues, with treatments effective in animal models often failing in human studies [57]. On the other hand, the variability in experimental protocols and tools affects how generalizable the results are [58,59]: drugs effective across diverse laboratory settings tend to promise better outcomes in human studies [60]. In addition, outcomes from animal and early clinical studies must align with actual clinical needs [59]. Incorporating these measures could not only streamline drug development but also positively impact animal welfare by reducing research waste [61,62].

Could low translation be an innate characteristic of translational animal research? Instead of disposing animal research due to low translation, a more telling comparison might be drawing parallels between translational rates in animal research and sectors like medical device approvals, where development largely bypasses animal use [63]. It might emerge that translation proportions could be similarly modest in these animal-free sectors.

We did not systematically assess differences in animal models or therapy parameters such as therapy type (drug versus non-drug) for different disease categories—for example, comparing conditions where animal and human therapy outcomes show high consistency (such as neurological or circulatory system diseases) against those with lower consistency (such as

mental health issues or cancer). Future research could explore these factors as potential moderators on the alignment of results between animal studies and human applications. Additionally, future studies should examine how the source of trial sponsorship, comparing investigator-initiated versus industry-sponsored trials, influences translational success, given that industry-sponsored trials appear to exert a greater influence on clinical guidelines [64].

Limitations

The findings of this study come with notable limitations:

First and most importantly, the therapies herein included were identified through systematic reviews specifically focusing on translation which will likely bias the factual translation. Such reviews may be more commonly conducted in fields where at least some therapies have already achieved clinical translation.

Second, the journey of transitioning a therapy from animal experiments to human application is intricate, bearing a multidimensional nature [43]. Our methodology, however, simplified this process by categorizing clinical studies as positive, neutral, or negative.

Third, the outcomes of clinical studies were classified based on the authors' conclusions, which may lead to biased interpretations if the authors frame their findings to support a beneficial therapeutic outcome [65]. This phenomenon, known as "spin," could result in an overestimation of clinical trials with beneficial outcomes.

Fourth, it is prudent to recognize that animal studies can have indirect benefits in the translation process. For example, they might enhance our mechanistic comprehension of diseases, even if not directly leading to a successful therapeutic application in humans.

Fifth, this umbrella review is confined to studies published in English, which may lead to the omission of relevant data published in other languages.

Strengths

Our umbrella review also has strengths:

First, we employed a rigorous systematic review methodology, which aids in mitigating potential biases.

Second, it offers a comprehensive perspective by spanning a variety of biomedical fields, therapeutic modalities, and human diseases, and including a relatively large number of therapies and diseases.

Third, we have undertaken both qualitative and quantitative assessments of translational rates, constituting sensitivity analyses.

Conclusions

Our umbrella review presents translational proportions across various biomedical fields, detailing the progression times from animal studies to clinical development. Although the consistency between animal and early clinical studies was high, only a minority of therapeutic interventions achieved regulatory approval. To enhance development of therapies for clinical application, it is imperative to emphasize the robustness and generalizability of experimental approaches, ensuring rigorous animal and human research.

Supporting information

S1 Data. Supplementary data.
(PDF)

S2 Data. Supplementary data.
(DOCX)

S1 Fig. Translational proportions for all therapies (neglecting development time).
(DOCX)

S2 Fig. Lag times for clinical therapy development from first animal study only considering therapies which transitioned to a clinical trial.
(DOCX)

S3 Fig. Meta-analysis on concordance rate for neurological diseases (relative risk).
(DOCX)

S4 Fig. Meta-analysis on concordance rate for circulatory system diseases (relative risk).
(DOCX)

S5 Fig. Meta-analysis on concordance rate for diseases of the digestive system (relative risk).
(DOCX)

S6 Fig. Meta-analysis on concordance rate for musculoskeletal diseases (relative risk).
(DOCX)

S7 Fig. Meta-analysis on concordance rate for cancer (relative risk).
(DOCX)

S8 Fig. Meta-analysis on concordance rate for mental health disorders (relative risk).
(DOCX)

S1 Table. Studies, interventions, and development status of included systematic reviews/ therapeutic interventions.
(DOCX)

S2 Table. Translational assessment of interventions for neurological diseases.
(DOCX)

S3 Table. Translational assessment of interventions for diseases of the musculoskeletal system and connective tissue.
(DOCX)

S4 Table. Translational assessment of interventions for psychiatric disorders.
(DOCX)

S5 Table. Translational assessment of interventions for diseases of the circulatory system.
(DOCX)

S6 Table. Translational assessment of interventions for diseases of the digestive system.
(DOCX)

S7 Table. Translational assessment of interventions for cancer.
(DOCX)

S8 Table. Translational assessment of interventions for other disorders/conditions.
(DOCX)

S9 Table. Number of therapeutic interventions tested in animals, entering any clinical trial, any RCT, and eventually obtaining regulatory approval per disease/condition

(n = 54).
(DOCX)

S1 Code. Analysis code.
(R)

Acknowledgments

Author Contributions

Conceptualization: Benjamin V. Ineichen, Eva Furrer, Servan L. Grüninger, Malcolm R. Macleod.

Data curation: Benjamin V. Ineichen, Eva Furrer, Servan L. Grüninger, Wolfgang E. Zürrer.

Formal analysis: Benjamin V. Ineichen.

Methodology: Benjamin V. Ineichen, Eva Furrer, Servan L. Grüninger, Malcolm R. Macleod.

Supervision: Malcolm R. Macleod.

Visualization: Benjamin V. Ineichen.

Writing – original draft: Benjamin V. Ineichen, Malcolm R. Macleod.

Writing – review & editing: Eva Furrer, Servan L. Grüninger, Wolfgang E. Zürrer.

References

1. Steinman L, Zamvil SS. How to successfully apply animal studies in experimental allergic encephalomyelitis to research on multiple sclerosis. Ann Neurol. 2006; 60 (1):12–21. https://doi.org/10.1002/ana.20913 PMID: 16802293.

2. Howells DW, Sena ES, Macleod MR. Bringing rigour to translational medicine. Nat Rev Neurol. 2014; 10(1):37. https://doi.org/10.1038/nrneurol.2013.232 PMID: 24247324

3. Ioannidis JP. Extrapolating from animals to humans. Sci Transl Med. 2012; 4(151):151ps15. Epub 2012/09/14. https://doi.org/10.1126/scitranslmed.3004631 PMID: 22972841.

4. Seyhan AA. Lost in translation: the valley of death across preclinical and clinical divide—identification of problems and overcoming obstacles. Transl Med Commun. 2019; 4(1):1–19.

5. Dirnagl U, Macleod MR. Stroke research at a road block: the streets from adversity should be paved with meta-analysis and good laboratory practice. Br J Pharmacol. 2009; 157(7):1154–1156. https://doi.org/10.1111/j.1476-5381.2009.00211.x PMID: 19664136

6. Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L, Miller DH, et al. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med. 2006; 354(9):899–910. Epub 2006/03/03. https://doi.org/10.1056/NEJMoa044397 PMID: 16510744.

7. Friese MA, Montalban X, Willcox N, Bell JI, Martin R, Fugger L. The value of animal models for drug development in multiple sclerosis. Brain J Neurol. 2006; 129(Pt 8):1940–1952. https://doi.org/10.1093/brain/awl083 PMID: 16636022.

8. Mi S, Hu B, Hahm K, Luo Y, Kam Hui ES, Yuan Q, et al. LINGO-1 antagonist promotes spinal cord remyelination and axonal integrity in MOG-induced experimental autoimmune encephalomyelitis. Nat Med. 2007; 13(10):1228–33. Epub 2007/10/02. https://doi.org/10.1038/nm1664 PMID: 17906634.

9. Ahmed Z, Fulton D, Douglas MR. Opicinumab: is it a potential treatment for multiple sclerosis? Ann Transl Med. 2020; 8(14). https://doi.org/10.21037/atm.2020.03.131 PMID: 32793736

10. Ruan Y, Robinson NB, Khan FM, Hameed I, Rahouma M, Naik A, et al. The translation of surgical animal models to human clinical research: A cross-sectional study. Int J Surg. 2020; 77:25–29. https://doi.org/10.1016/j.ijsu.2020.03.023 PMID: 32198098

11. Ormandy EH, Schuppli CA. Public attitudes toward animal research: a review. Animals. 2014; 4 (3):391–408. https://doi.org/10.3390/ani4030391 PMID: 26480314

12. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015; 4(1):1. https://doi.org/10.1186/2046-4053-4-1 PMID: 25554246

13. de Vries RB, Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M. Updated version of the Embase search filter for animal studies. Lab Anim. 2014; 48(1):88. Epub 2013/07/10. https://doi.org/10.1177/0023677213494374 PMID: 23836850.

14. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc. 2016; 104(3):240. https://doi.org/10.3163/1536-5050.104.3.014 PMID: 27366130

15. Bahor Z, Liao J, Currie G, Ayder C, Macleod M, McCann SK, et al. Development and uptake of an online systematic review platform: the early years of the CAMARADES Systematic Review Facility (SyRF). BMJ Open Sci. 2021; 5(1):e100103. https://doi.org/10.1136/bmjos-2020-100103 PMID: 35047698

16. AVPSoft. Universal Desktop Ruler. https://universal-desktop-rulerensoftoniccom/?ex=RAMP-176802020.

17. Sena ES, Currie GL, McCann SK, Macleod MR, Howells DW. Systematic reviews and meta-analysis of preclinical studies: why perform them and how to appraise them critically. J Cereb Blood Flow Metab. 2014; 34(5):737–742. https://doi.org/10.1038/jcbfm.2014.28 PMID: 24549183

18. Austin CP. Translating translation. Nat Rev Drug Discov. 2018; 17(7):455–6. Epub 2018/04/21. https://doi.org/10.1038/nrd.2018.27 PMID: 29674698.

19. Cwikel J, Behar L, Rabson-Hare J. A comparison of a vote count and a meta-analysis review of intervention research with adult cancer patients. Res Soc Work Pract. 2000; 10(1):139–158.

20. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc. 2009; 172(1):137–159. https://doi.org/10.1111/j.1467-985X.2008.00552.x PMID: 19381330

21. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010; 36 (3):1–48.

22. Hansen LF, Nielsen NSK, Christoffersen LC, Kruuse C. Translational challenges of remote ischemic conditioning in ischemic stroke—a systematic review. Ann Clin Transl Neurol. 2021; 8(8):1720–1729. https://doi.org/10.1002/acn3.51405 PMID: 34133841.

23. Ramos E, Farre-Alins V, Egea J, Lopez-Munoz F, Reiter RJ, Romero A. Melatonin's efficacy in stroke patients; a matter of dose? A systematic review. Toxicol Appl Pharmacol. 2020; 392:114933. https://doi.org/10.1016/j.taap.2020.114933 PMID: 32112789.

24. Blais LL, Montgomery TL, Amiel E, Deming PB, Krementsov DN. Probiotic and commensal gut microbial therapies in multiple sclerosis and its animal models: a comprehensive review. Gut Microbes. 2021; 13(1):1943289. https://doi.org/10.1080/19490976.2021.1943289 PMID: 34264791.

25. Plemel JR, Juzwik CA, Benson CA, Monks M, Harris C, Ploughman M. Over-the-counter anti-oxidant therapies for use in multiple sclerosis: A systematic review. Mult Scler. 2015; 21(12):1485–1495. https://doi.org/10.1177/1352458515601513 PMID: 26286700.

26. Voulgaropoulou SD, van Amelsvoort T, Prickaerts J, Vingerhoets C. The effect of curcumin on cognition in Alzheimer's disease and healthy aging: A systematic review of pre-clinical and clinical studies. Brain Res. 2019; 1725:146476. https://doi.org/10.1016/j.brainres.2019.146476 PMID: 31560864.

27. Regner GG, Pereira P, Leffa DT, de Oliveira C, Vercelino R, Fregni F, et al. Preclinical to Clinical Translation of Studies of Transcranial Direct-Current Stimulation in the Treatment of Epilepsy: A Systematic Review. Front Neurosci. 2018; 12:189. https://doi.org/10.3389/fnins.2018.00189 PMID: 29623027.

28. Heinzel JC, Quyen Nguyen M, Kefalianakis L, Prahm C, Daigeler A, Hercher D, et al. A systematic review and meta-analysis of studies comparing muscle-in-vein conduits with autologous nerve grafts for nerve reconstruction. Sci Rep. 2021; 11(1):11691. https://doi.org/10.1038/s41598-021-90956-3 PMID: 34083605.

29. Daldegan-Bueno D, Maia LO, Glass M, Jutras-Aswad D, Fischer B. Co-exposure of cocaine and cannabinoids and its association with select biological, behavioural and health outcomes: A systematic scoping review of multi-disciplinary studies. Eur Neuropsychopharmacol. 2021; 51:106–131. https://doi.org/10.1016/j.euroneuro.2021.06.002 PMID: 34273801.

30. Daldegan-Bueno D, Maia LO, Glass M, Jutras-Aswad D, Fischer B. Co-exposure of cannabinoids with amphetamines and biological, behavioural and health outcomes: a scoping review of animal and human studies. Psychopharmacology. 2022; 239(5):1211–1230. https://doi.org/10.1007/s00213-021-05960-2 PMID: 34613429.

31. Reis DJ, Ilardi SS, Punt SEW. The anxiolytic effect of probiotics: A systematic review and meta-analysis of the clinical and preclinical literature. PLoS ONE [Electronic Resource]. 2018; 13(6):e0199041. https://doi.org/10.1371/journal.pone.0199041 PMID: 29924822.

32. Li Y, Fu SC, Cheuk YC, Song G, Feng H, Yung SH. The non-reconstructive treatment of complete ACL tear with biological enhancement in clinical and preclinical studies: A systematic review. Asia Pac J Sports Med Arthrosc Rehabil Technol. 2018; 14:10–16. https://doi.org/10.1016/j.asmart.2018.04.002 PMID: 30302314

33. Goldberg A, Mitchell K, Soans J, Kim L, Zaidi R. The use of mesenchymal stem cells for cartilage repair and regeneration: a systematic review. J Orthop Surg. 2017; 12(1):39. https://doi.org/10.1186/s13018-017-0534-y PMID: 28279182.

34. Dalton EJ, Churilov L, Lannin NA, Corbett D, Campbell BCV, Hayward KS. Early-phase dose articulation trials are underutilized for post-stroke motor recovery: A systematic scoping review. Ann Phys Rehabil Med. 2022; 65(1):101487. https://doi.org/10.1016/j.rehab.2021.101487 PMID: 33429089.

35. Zoerle T, Ilodigwe DC, Wan H, Lakovic K, Sabri M, Ai J, et al. Pharmacologic reduction of angiographic vasospasm in experimental subarachnoid hemorrhage: systematic review and meta-analysis. J Cereb Blood Flow Metab. 2012; 32(9):1645–1658. https://doi.org/10.1038/jcbfm.2012.57 PMID: 22534672.

36. Schmidt-Pogoda A, Bonberg N, Koecke MHM, Strecker JK, Wellmann J, Bruckmann NM, et al. Why Most Acute Stroke Studies Are Positive in Animals but Not in Patients: A Systematic Comparison of Preclinical, Early Phase, and Phase 3 Clinical Trials of Neuroprotective Agents. Ann Neurol. 2020; 87 (1):40–51. https://doi.org/10.1002/ana.25643 PMID: 31714631.

37. Baker TS, Durbin J, Troiani Z, Ascanio-Cortez L, Baron R, Costa A, et al. Therapeutic hypothermia for intracerebral hemorrhage: Systematic review and meta-analysis of the experimental and clinical literature. Int J Stroke. 2022; 17(5):506–516. https://doi.org/10.1177/17474930211044870 PMID: 34427479

38. Lalu MM, Montroy J, Dowlatshahi D, Hutton B, Juneau P, Wesch N, et al. From the Lab to Patients: a Systematic Review and Meta-Analysis of Mesenchymal Stem Cell Therapy for Stroke. Transl Stroke Res. 2020; 11(3):345–364. https://doi.org/10.1007/s12975-019-00736-5 PMID: 31654281.

39. Cottrill E, Pennington Z, Ahmed AK, Lubelski D, Goodwin ML, Perdomo-Pantoja A, et al. The effect of electrical stimulation therapies on spinal fusion: a cross-disciplinary systematic review and meta-analysis of the preclinical and clinical data. J Neurosurg Spine. 2019:1–21. https://doi.org/10.3171/2019.5.SPINE19465 PMID: 31593923.

40. Cottrill E, Pennington Z, Lankipalle N, Ehresman J, Valencia C, Schilling A, et al. The effect of bioactive glasses on spinal fusion: A cross-disciplinary systematic review and meta-analysis of the preclinical and clinical data. J Clin Neurosci. 2020; 78:34–46. https://doi.org/10.1016/j.jocn.2020.04.035 PMID: 32331941.

41. Lind PC, Johannsen CM, Vammen L, Magnussen A, Andersen LW, Granfeldt A. Translation from animal studies of novel pharmacological therapies to clinical trials in cardiac arrest: A systematic review. Resuscitation. 2021; 158:258–269. https://doi.org/10.1016/j.resuscitation.2020.10.028 PMID: 33147523.

42. Reynolds JC, Rittenberger JC, Menegazzi JJ. Drug administration in animal studies of cardiac arrest does not reflect human clinical experience. Resuscitation. 2007; 74(1):13–26. https://doi.org/10.1016/j.resuscitation.2006.10.032 PMID: 17360097

43. Lalu M, Leung GJ, Dong YY, Montroy J, Butler C, Auer RC, et al. Mapping the preclinical to clinical evidence and development trajectory of the oncolytic virus talimogene laherparepvec (T-VEC): a systematic review. BMJ Open. 2019; 9(12):e029475. https://doi.org/10.1136/bmjopen-2019-029475 PMID: 31796474.

44. Beuker C, Strecker JK, Rawal R, Schmidt-Pogoda A, Ruck T, Wiendl H, et al. Immune Cell Infiltration into the Brain After Ischemic Stroke in Humans Compared to Mice and Rats: a Systematic Review and Meta-Analysis. Transl Stroke Res. 2021; 12(6):976–90. Epub 20210126. https://doi.org/10.1007/s12975-021-00887-4 PMID: 33496918.

45. Friedman LM, Furberg CD, DeMets DL, Reboussin DM, Granger CB. Fundamentals of clinical trials. Springer; 2015.

46. Eichler H-G, Abadie E, Breckenridge A, Flamion B, Gustafsson LL, Leufkens H, et al. Bridging the efficacy—effectiveness gap: a regulator's perspective on addressing variability of drug response. Nat Rev Drug Discov. 2011; 10(7):495–506. https://doi.org/10.1038/nrd3501 PMID: 21720406

47. Djulbegovic B, Paul A. From efficacy to effectiveness in the face of uncertainty: indication creep and prevention creep. JAMA. 2011; 305(19):2005–2006. https://doi.org/10.1001/jama.2011.650 PMID: 21586716

48. Campbell BC, De Silva DA, Macleod MR, Coutts SB, Schwamm LH, Davis SM, et al. Ischaemic stroke. Nat Rev Dis Primers. 2019; 5(1):1–22.

49. Knopman DS, Amieva H, Petersen RC, Chételat G, Holtzman DM, Hyman BT, et al. Alzheimer disease. Nat Rev Dis Primers. 2021; 7(1):33. Epub 2021/05/15. https://doi.org/10.1038/s41572-021-00269-y PMID: 33986301.

50. Ahuja CS, Wilson JR, Nori S, Kotter M, Druschel C, Curt A, et al. Traumatic spinal cord injury. Nat Rev Dis Primers. 2017; 3(1):1–21.

51. Dalton E, Churilov L, Lannin NA, Corbett D, Hayward KS. Dose Articulation in Preclinical and Clinical Stroke Recovery: Refining a Discovery Research Pipeline and Presenting a Scoping Review Protocol. Front Neurol. 2019; 10:1148. Epub 20191106. https://doi.org/10.3389/fneur.2019.01148 PMID: 31781018.

52. Ogier R, Knecht W, Schwab ME. Academic leadership: (with)holding the keys to translational medicine? Nat Med. 2019; 25(12):1812–3. Epub 2019/12/07. https://doi.org/10.1038/s41591-019-0670-5 PMID: 31806891.

53. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials. 2017; 18:1–7.

54. Khan A, Fahl Mar K, Brown WA. The Impact of Underpowered Studies on Clinical Trial Results. Am J Psychiatry. 2018; 175(2):188. https://doi.org/10.1176/appi.ajp.2017.17091016 PMID: 29385825.

55. Ferreira-González I, Permanyer-Miralda G, Domingo-Salvany A, Busse JW, Heels-Ansdell D, Montori VM, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. BMJ. 2007; 334(7597):786. https://doi.org/10.1136/bmj.39136.682083.AE PMID: 17403713

56. Molnar FJ, Man-Son-Hing M, Fergusson D. Systematic review of measures of clinical significance employed in randomized controlled trials of drugs for dementia. J Am Geriatr Soc. 2009; 57(3):536–546. https://doi.org/10.1111/j.1532-5415.2008.02122.x PMID: 19187414

57. Scott S, Kranz JE, Cole J, Lincecum JM, Thompson K, Kelly N, et al. Design, power, and interpretation of studies in the standard murine model of ALS. Amyotroph Lateral Scler. 2008; 9(1):4–15. https://doi.org/10.1080/17482960701856300 PMID: 18273714

58. Voelkl B, Würbel H, Krzywinski M, Altman N. The standardization fallacy. Nat Methods. 2021; 18(1):5–7. https://doi.org/10.1038/s41592-020-01036-9 PMID: 33408399

59. Lourbopoulos A, Mourouzis I, Xinaris C, Zerva N, Filippakis K, Pavlopoulos A, et al. Translational Block in Stroke: A Constructive and "Out-of-the-Box" Reappraisal. Front Neurosci. 2021; 15:652403. Epub 20210514. https://doi.org/10.3389/fnins.2021.652403 PMID: 34054413.

60. Lapchak PA, Zhang JH, Noble-Haeusslein LJ. RIGOR guidelines: escalating STAIR and STEPS for effective translational research. Transl Stroke Res. 2013; 4:279–285. https://doi.org/10.1007/s12975-012-0209-2 PMID: 23658596

61. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet. 2014; 383(9912):166–175. https://doi.org/10.1016/S0140-6736(13)62227-8 PMID: 24411645

62. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, et al. Biomedical research: increasing value, reducing waste. Lancet. 2014; 383(9912):101–104. https://doi.org/10.1016/S0140-6736(13)62329-6 PMID: 24411643

63. Bretthauer M, Gerke S, Hassan C, Ahmad OF, Mori Y. The new European Medical Device Regulation: balancing innovation and patient safety. Ann Intern Med. 2023. https://doi.org/10.7326/M23-0454 PMID: 37068279

64. Blümle A, Wollmann K, Bischoff K, Kapp P, Lohner S, Nury E, et al. Investigator initiated trials versus industry sponsored trials-translation of randomized controlled trials into clinical practice (IMPACT). BMC Med Res Methodol. 2021; 21(1):182. https://doi.org/10.1186/s12874-021-01359-x PMID: 34465296

65. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA. 2010; 303(20):2058–2064. https://doi.org/10.1001/jama.2010.651 PMID: 20501928