An analysis of citations used in the For Life on Earth paper "Pharmaceutical Companies Acknowledge the Failure of Animal Models in their Drug Development Process, and Write about this Openly in the Scientific Literature"

Location of paper: https://www.forlifeonearth.org/wp-content/uploads/2013/05/Pharmaceutical-Company-Quotes2.pdf

	Comments	Type	Peer-reviewed paper?	
The success rate for new drugs in all areas of development is dismal. Out of 5,000- 10,000 chemicals that enter the drug development pipeline only one will enter the market. (European Commission 2008; [1] Hughes et al. 2011 [2])	Vaguely referenced, but the final report of the Innovative Medicines Initiative (ref 1) barely mentions animals. Drug attrition is mentioned in relation only to late-stage human trials. https://ec.europa.eu/research/health/pdf/imi_final_evaluation.pdf The Hughes paper also makes no reference to drug attrition rates or their causes https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3058157/ **The claim is unsupported by the references.**	Claim not supported by the reference	1 N 2 Y	3
Moreover, the major cost of drug development occurs during the clinical trials and the attrition rate during this stage is equally dreadful. (Unknown 2002 [3]; Shaffer 2012 [4]; Paul et al. 2010 [5]; Schachter 2007 [6])	No. Candidate drugs are knocked out at every stage of development, but in diminishing absolute numbers. Much of the cost does come after phase 1 and 2 human trials, but the **greatest spend is during discovery**, as Greek's reference (7) attests, in disagreement with Greek. The references at best support late stage attrition, not attrition following the animal phase. https://www.rdmag.com/article/2012/01/safety-through-sequencing	Claim not supported by the references	3 N 4 N 4 N 6 N	3
Drugs entering Phase I trials have approximately a 9% chance of coming to market. (FDA 2004 [7]; Sarkar 2009 [8]; Editorial 2007 [9]; Paul et al. 2010 [10])	The references say 8%. There are many reasons for failure - including lack of efficacy once the drug has proven safe in animal tests - and commercial reasons. This does not implicate the animal model.	The claim is true but does not support the hypothesis being advanced.	7 N 8 N 9 N 10 Y	4
Of the drugs that advance to Phase III,	True, but incomplete. The author goes on to say that lack	The claim is true but does not	11 N	4

less than 50% are marketed.	of efficacy, at 66%, was by far the biggest cause of failure for the 83 drugs studied, notes that the most common failures were in difficult-to-treat diseases and speculates the reasons for failure were progressing things to phase 3 after they showed only marginal efficacy in phase 2 human trials. https://www.nature.com/articles/nrd3375.epdf?no_publisher_access=1&r3_referer=nature	support the hypothesis being advanced.		
*The failure rate for oncology drugs is even higher. (Editorial 2011 [12]; Caponigro & Sellers 2011 [13]; Arrowsmith 2011 [14]; Begley & Ellis 2012 [15]) * Only 5% of cancer drugs that have an Investigational New Drug Application (IND) eventually go to market. (Kummar et al. 2007[16])	True. Some cancers can be hard to treat. Ref 14 is the same paper as ref 11. The Arrowsmith paper is used 6 times as a reference. Ref 13 is a comment article.	True.	12 N 13 Y 14 N 15 N 16 Y	1
* Lack of safety or efficacy accounts for approximately 90% of drug failures during clinical trials. (Kola & Landis 2004 [17]; Arrowsmith 2011 [18]).	True but incomplete, which changes its meaning. Efficacy constitutes 66% of the failure, safety 21% but problems with the latter manifest in equal percentages across phase 2 and 3 trials, which means it tested safely in both animals and humans before later failing, perhaps as the dose was increased.	The claim is true but does not support the hypothesis being advanced.	17 N 18 N	4
Both safety and efficacy determinations rely on animal models. To complicate matters further, the pipeline in Pharma is drying up and fewer drugs, especially new chemical entities (NCEs) are being marketed. (Editorial 2008 [19]; GBI Research 2011 [20]).	Incorrect first sentence. Animals are used in regulatory testing for safety rather than efficacy. The second sentence is the part supported by the reference and is uncontroversial.	The first claim is incorrect. The second claim is true but does not support the hypothesis being advanced.	19 N 20 N	2, 4
Björquist and Sartipy state: "Furthermore, the compound attrition rate is negatively affected by	This isn't a scientific paper. It's an article promoting stem cell assays and their claims are in turn unreferenced. The authors both work for Cellartis, which sells non-animal	The quote is accurately reproduced but is itself untrue, unreferenced and derives from	21 N	5

the inability to predict toxicity and efficacy in humans. These shortcomings are in turn caused by the use of experimental pre-clinical model systems that have a limited human clinical relevance..." (Björquist & Sartipy 2007 [21])	experimental tools. https://www.ddw-online.com/therapeutics/p92860-human-es-cell-derived-functional-cells-as-tools-in-drug-discoverywinter-2007.html	a business magazine not a peer-reviewed paper.		
* Then-U.S. Secretary of Health and Human Services Mike Leavitt stated in 2006: "Currently, nine out of ten experimental drugs fail in clinical studies because we cannot accurately predict how they will behave in people based on laboratory and animal studies." (FDA 2006 [22])	True but incomplete, which changes its meaning. Note *laboratory* and animal studies, so mainly non-human methods. Again, this is not a criticism of safety testing.	The quote is accurately reproduced but is talking about efficacy failures following primarily non-animal safety testing.	22 N	6
Johnson et al. found that out of 39 anticancer drugs tested on xenograft mice, only one mimicked the response in humans. (Johnson et al. 2001 [23])	True but incomplete, which changes its meaning. The authors go on "However, for compounds with *in vivo* activity in at least one-third of tested xenograft models, there was correlation with ultimate activity in at least some Phase II trials. Thus, an efficient means of predicting activity *in vivo* models remains desirable for compounds with anti-proliferative activity in vitro." https://www.ncbi.nlm.nih.gov/pubmed/11355958	True, but the model being discussed (in 2001) was outdated and atypical of animal experiments.	23 Y	6
Oncology drugs fail more frequently in clinical trials than most other categories. (DiMasi & Grabowski 2007[24]; DiMasi et al. 2010 [25])	Repetition of point 5. Some cancers are hard to treat.	True, but the quote doesn't support the hypothesis being advanced since the main reason for failure is advancing drugs to phase 3 human trials despite a poor showing in phase 2 .	24 Y 25 Y	4
There have been many attempts to reproduce human cancers in mice. The nude mouse lacked the FOX1 gene, the SCID mouse was created	Again, incomplete which changes its meaning. This is an article not a scientific paper. The article is about genetically engineering mice to make them better predictors of a drug's efficacy and does not refer to all mouse models. The	The reference does not back up the claim nor the wider hypothesis.	26 N	3, 4

with a very deficient immune system, and there have been many more models. All have failed to predict human response and have misled researchers. Zielinska discusses mouse models of cancer stating they: "rarely predict how a human will respond to the same treatment." Zielinska then quotes Marks of the NCI, and who is also head of the Mouse Models of Human Cancers Consortium, as saying: "we had loads of models that were not predictive, that were [in fact] seriously misleading."(Zielinska 2010 [26])	author goes on: "In this trial, however, principal investigator Pier Paolo Pandolfi and others have engineered the mice to develop cancers that carry mutations similar to those seen in cancer patients—mutations scientists suspect may explain why some patients respond to a particular treatment and some don't." https://www.the-scientist.com/uncategorized/building-a-better-mouse-43400 The quote, from 9 years ago, "…we had loads of models that were not predictive, that were [in fact] seriously misleading…" **was referring to the situation before the new animal models were available**.			
* The NCI had previously tested mice with 12 anti-cancer drugs being successfully used to treat humans. The mice were growing 48 different kinds of human cancers. The study revealed that 30 out of 48 times (63%) the drugs that were effective against human cancers were ineffective in the mice that were growing the human cancers. The NCI believes efficacious treatments for human cancers have been lost because of animal testing. (Gura 1997 [27])	Incomplete, which changes its meaning. This is an article not a scientific paper. Similar to the above, the article is referring to an older way of doing things. In this case xenograph mice. http://science.sciencemag.org/content/278/5340/1041.long The author goes on "And attempts to use human cells in culture don't seem to be faring any better, partly because cell culture provides no information about whether a drug will make it to the tumor site. To create better models of cancer development in humans, investigators are now drawing on knowledge of human cancer-related gene mutations to **genetically alter mice so that they carry the same kinds of changes that**	Incorrect application of a historic claim to refer to the present situation. The NCI believes efficacious treatments for human cancers have been (temporarily) lost because of lots of reasons including development being halted because it did not treat the target system, yet was useful for other diseases like cancers. An example is Thalidomide, which is today used a cancer drug.	27 N	6

	lead to cancer in humans"			
The problem of animal models is well known to the drug development community. Cook et al state: "Over many years now there has been a poor correlation between preclinical therapeutic findings and the eventual efficacy of these [anti-cancer] compounds in clinical trials (Johnson et al. 2001; [28] Suggitt & Bibby 2005 [29]).	Sentence one falsely equates animal trials with being the totality of preclinical testing. Sentence two refers again to all non-human methods of conducting research, **does not affect animals' record** with regards to safety testing and is narrowly focussed on cancer.	The first sentence falsely stylises preclinical testing as animal research and uses a quote about efficacy testing in the context of safety testing.	28 Y 29 Y	4
* The development of antineoplastics is a large investment by the private and public sectors, however, the limited availability of predictive preclinical systems obscures our ability to select the therapeutics that might succeed or fail during clinical investigation."(Cook, Jodrell, and Tuveson 2012 [30])	True but incomplete, which changes its meaning. http://csmres.co.uk/cs.public.upd/article-downloads/Cook_2012_Drug-Discovery-Today.pdf The authors go on "**Selecting the most appropriate in vivo model is essential** during the drug development process to enable accurate modelling of therapeutic efficacy. By developing innovative preclinical trials using sophisticated animal models that recapitulate the human malignancies in question, we might be able to advance the field of drug discovery, and improve success rates for potential novel therapeutics in clinical trials."	Equates preclinical testing with animal testing, then ignores the recommendation to use animal models.	30 Y	6
Singh and Ferrara echo this, stating: "Over 90% of phase 3 clinical trials in oncology fail to meet their primary endpoints despite encouraging preclinical and even early-stage clinical data. This staggering and sobering figure underscores the limitations of existing animal models for the evaluation of potential	Half true, but this is not an indictment of the animal model alone – preclinical is all non-human methods and early-stage clinical is human studies. These areas all share the 'failure' rate. The 'paucity' of models means the lack of them in the areas listed. The authors add "In addition, technological and logistical advances in mouse models of human cancer over the past five years have the potential to increase the clinical	The reference doesn't indict animal models alone, but what part is down to animals supports the hypothesis.	31 Y	1

anticancer agents. The paucity of models is especially apparent with the advent of drugs that target the tumor milieu, or microenvironment, such as anti-angiogenics . . . immunotherapies and compounds directed against tumor-associated fibroblasts."(Singh & Ferrara 2012 [31])	translatability of animal studies", which is an acknowledgement that their analysis may not reflect the current situation.			
Wittenburg and Gustafson agree, stating: "The current drug development pathway in oncology research has led to a large attrition rate for new drugs, in part due to a general lack of appropriate preclinical studies that are capable of accurately predicting efficacy and/or toxicity in the target population. . . . One of the most serious challenges currently facing pharmaceutical research of novel anti-cancer therapeutics is the lack of translation of efficacy and safety from preclinical models to human clinical trials, leading to a large attrition rate of investigational compounds. For new oncology drugs, only about 5% of investigational new drug applications submitted progress beyond the investigational phase due to a general lack of preclinical systems that can accurately predict efficacy and toxicity of new agents."(Wittenburg & Gustafson 2011 [32])	Dr Greek has cut a key sentence out of the middle. Following the first sentence should be "Because of an obvious need for novel therapeutics in many types of cancer, new compounds are being investigated in human Phase I and Phase II clinical trials before a complete understanding of their toxicity and efficacy profiles is obtained." Therefore, a significant part of the failure rate is hurrying them through human trials in case they work, because the disease is cancer. More compounds could be eliminated before human trials, thus lowering the 'failure rate', but this would not yield more drugs.	Wittenburg and Gustafson agree with the previous reference, but neither agrees with Dr Greek.	32 Y	6, 4

Animal models fail to predict safety as well as efficacy. Reviewers of Phase I trials conducted by the National Cancer Institute (NCI) from 1991-2002 discovered that 15% of participants undergoing single agent chemotherapy agents suffered serious side effects. (Horstmann et al. 2005 [33])	A 15% failure rate means an 85% success rate, which cannot be described as a failure overall. This is also an example from a niche area. The paper is looking at experimental chemotherapy and side-effects of any chemotherapy are inevitable, let alone with experimental doses. Dr Greek's use of single-agent chemotherapy is particularly unfortunate since they write: "In our view, it is inaccurate to refer to phase 1 oncology studies as if they are all similar to one another." And "The response rates of 4 to 6 percent and the toxicity-related death rate of 0.5 percent continue to be viewed as representative of phase 1 oncology trials, but these rates are based on reviews of single-agent trials. They do not take into full account the development of new types of anticancer agents, trials of combinations of agents, new trial designs, or improvements in supportive care, and they do not present a comprehensive picture of the benefits and risks associated with phase 1 trials." https://www.nejm.org/doi/full/10.1056/NEJMsa042220	The reference does not support the claim.	33 Y	3
Richard Klausner, then-director of the NCI said: "The history of cancer research has been a history of curing cancer in the mouse. . . . We have cured mice of cancer for decades—and it simply didn't work in humans."(Cimons et al. 1998 [34])	A fair use of an inaccurate and outdated comment. This is not a paper, although is referenced like a paper, but an article in the LA Times from 1998. In it, Klausner uses an inaccurate rhetorical flourish to both exaggerate 'cures' in mice and underplay advances in human medicine.	An inaccurate throwaway quote cited as fact.	34 N	5
In an editorial to two articles, Nature Medicine stated: "The complexity of	True. The quote is a perfectly fine use of the first reference. The second referenced paper (Van Dyke 2010)	Incomplete quote describing an earlier model that was	35 N 36 N	1

human metastatic cancer is difficult to mimic in mouse models. As a consequence, seemingly successful studies in murine models do not translate into success in late phases of clinical trials, pouring money, time and people's hope down the drain."(Ellis & Fidler 2010; [35] Van Dyke 2010 [36])	praises the discoveries that came from investigating the reasons behind the historic lack of translation from mouse to man, which led to better translation. It goes on: "First, if the model is, in fact, representative of the human disease, the observation that pancreatic cancers are resistant to drug uptake may explain why virtually every therapy tested for this disease has failed and represents, therefore, a breakthrough in advancing therapeutic effectiveness. Second, the use of the GEM model facilitated testing for therapeutic agents that promote effective drug delivery, resulting in the development of a protocol for combined therapy that targets both the microenvironment, with the Smoothened inhibitor, and the tumor cells, with gemcitabine."	superseded.		
Caponigro and Sellers of the Novartis Institutes For BioMedical Research, Oncology Research and Oncology Translational Medicine stated in 2011: "Despite an improved understanding of the biology of cancer, and an unprecedented volume of new molecules in clinical trials, the number of highly efficacious drugs approved by the regulatory authorities remains disappointingly low. The significant attrition rate of drugs entering clinical trials comes at a high price. This price is paid primarily by the underserved patient and secondarily by the pharmaceutical and biotechnology community, which invests enormous resources perfecting a molecule only	True, but the paper doesn't implicate animal models. The paper states "The often empirical treatment of cancer--which was initially based on inhibiting DNA synthesis and cellular division--while having led to a number of remarkable successes, remains prone to a high rate of clinical failure that **results partly from a lack of understanding of how best to implement drugs in the clinic.**"	The reference does not support the hypothesis.	37 Y	4

to watch it fail in humans.				
Cancer researcher Robert Weinberg, of Massachusetts Institute of Technology, was quoted by Leaf in Fortune magazine as saying: “And it’s been well known for more than a decade, maybe two decades, that many of these preclinical human cancer models have very little predictive power in terms of how actual human beings—actual human tumors inside patients—will respond . . . preclinical models of human cancer, in large part, stink . . . hundreds of millions of dollars are being wasted every year by drug companies using these [animal] models.”(Leaf 2004 [38]) Leaf also quotes Homer Pearce, “who once ran cancer research and clinical investigation at Eli Lilly and is now research fellow at the drug company” as saying: “. . . that mouse models are ‘woefully inadequate’ for determining whether a drug will work in humans. ‘If you look at the millions and millions and millions of mice that have been cured, and you compare that to the relative success, or lack thereof, that we’ve achieved in the treatment of metastatic disease clinically,’ he says, ‘you realize that there just has to be something wrong with those models.’”(Leaf 2004 [39])	True only for one mouse model. This is another magazine article, not a scientific paper and it’s talking about a specific form of mouse model (hence ‘these mouse models’ in the quote), not all preclinical mouse models. http://fortune.com/2004/03/22/cancer-medicines-drugs-health/ “One of the most frequently used experimental models of human cancer is to take human cancer cells that are grown in a petri dish, put them in a mouse—in an immunocompromised mouse—allow them to form a tumor, and then expose the resulting xenograft to different kinds of drugs that might be useful in treating people.”	This is a magazine article commenting on a specific, outdated method of research.	38 N	1

Others have also pointed out the inadequacy of animal models of cancer, including genetically modified animal models (Frese & Tuveson 2007; [40] Kerbel 2003; [41] Singh et al. 2010;[42] Talmadge et al. 2007; [43] Peterson & Houghton 2004;[44] Francia & Kerbel 2010; [45] Johnson et al. 2001; [46] Zielinska 2010; [47] Wade 2009 [48])	Various misrepresentations. Not all papers. The six papers are all about tweaking models to get better results, such as moving away from xenografts. Some are critical of particular models at the time of writing some years ago and the reference doesn't support a contemporary criticism. The only mention of genetic models is reference 42, which concludes that they model human responses well so that part of the statement is not supported by the reference. Frese & Tuveson 2007 (40) write "Animal models of cancer provide an alternative means to determine the causes of and treatments for malignancy, **thus representing a resource of immense potential for cancer medicine**. The sophistication of modelling cancer in mice has increased to the extent that investigators can both observe and manipulate a complex disease process in a manner impossible to perform in patients." Kerbel (41) writes "Close inspection of retrospective and prospective studies in the literature, however, reveals that human tumor xenografts-even non metastatic ectopic/subcutaneous "primary" tumor transplants-can be **remarkably predictive of cytotoxic chemotherapeutic drugs that have activity in humans**" Singh et al (42) write " Comparisons with corresponding clinical trials indicate that these GEMMs **model human responses well."** Talmadge et al (43) Collectively, murine models are **critical in drug development**	Misrepresentation of several papers, selective quoting leaving out key context and examples of animal models succeeding.	40 Y 41 Y 42 Y 43 Y 44 Y 45 N 46 Y 47 N 48 N	6, 4, 7

	Peterson & Houghton (44) are sceptical in a similar way suggested by Dr Greek. They express ' reasonable scepticism' over the value specifically of xenograft rodent tumour models and do not pass judgement on the efficacy of newer techniques such as transgenic mouse models. Francia & Kerbel (45) is not a paper but a comment on a paper behind a paywall. Johnson et al. (46) are talking only about a specific type of mouse model, writing in 2001. Zielinska (47) is a magazine article not a paper. In it, new animal models are being tried. They write, "Using the mice to screen for more effective treatment combinations, they found that APL 15;17 mice could be cured of their leukemia if given a combination of RA and arsenic trioxide, another chemotherapy drug. The APL 11;17 mice, in contrast, responded to RA combined with a newer drug, phenylbutyrate, a histone deacetylase. Again, **both predictions bore out in the clinic, turning a fatal form of leukemia into one with a 70–90 percent cure rate.**" Wade 2009 is an article in the New York Times. Its title "New Treatment for Cancer Shows Promise in Testing".			
Tamoxifen is a good example of the shortcomings of animal models in general. Tamoxifen was originally touted as a birth control pill based on rat studies and was only later found	Incorrect. The reference is in regard to human tumours transplanted into mice only. Tamoxifen was developed as a contraceptive and anti-cancer drug simultaneously and was tested in "pre-menopausal patients with mammary carcinoma, which was justified on the grounds that it might	The reference doesn't support the claim.	49 Y	2, 3

to be an anticancer chemical. Moreover, it was ineffective as an oral contraceptive as it actually increased a woman's likelihood of becoming pregnant. (Jordan & Robinson 1987 [49])	have a therapeutic as well as an anti-fertility effect." So, women with cancer. "Walpole wrote that the compound not only provided an interesting lead in oral contraception, but also in hormone-dependent cancers of the prostate and breast" https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5600945/ The reference doesn't support the claim.			
Tamoxifen acts by binding to the protein known as tubulin thus inhibiting cell division. After discovered to be effective against cancer, Tamoxifen was shown to causes liver tumors in some strains of rat, but not in mice or hamsters.(Powles 1992 [50])	True. The reference is a letter to The Lancet https://www.thelancet.com/pdfs/journals/lancet/PII0140-6736(92)93162-G.pdf and its contents are a reasonable reference in support of the statement.	The reference is correctly used, but doesn't support the hypothesis.	50 N	4
If this had been discovered in preclinical trials, the drug would not have come to market.(Editorial 2003 [51])	The reference is not a paper (https://www.nature.com/articles/nrd1057) but supports the statement, although is itself flawed, since it suggests that animal data from rats was the only data used to declare the drug's safety. The fact it was different in mice and hamsters would have been enough to keep investigations going. However the reference is fairly used.	The reference is correctly used, but is itself incorrect.	51 N	5
According to D. N. Richardson of the Imperial Chemistries Industries PLC: "No laboratory tests for anti-tumour activity were carried out for Nolvadex [tamoxifen] until after the activity in human patients had been confirmed."(Richardson 1988 [52])	True but irrelevant. The reference supports this point, but Tamoxifen is an unusual case in that it was tested in humans early because they had cancer.	The reference is correctly used, but doesn't support the hypothesis.	52 Y	4
The most common side effect of Tamoxifen is nausea and vomiting,	Not entirely true, but this 40 year old text book may have asserted that. As this states	Not a paper but the reference is largely fairly used.	53 N	1

which was not seen in dogs, which are touted as the best species to use when looking for that side effect.(Tucker et al. 1984 [53]	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2737646/ "The question can only be addressed by asking a supplementary question 'In response to which stimulus?'" As the reference is very out of date it's not salient to a discussion of modern knowledge or techniques.			
Sadly, even the drugs that do come to market are too frequently not very effective against cancer. In the case of breast cancer, for instance, most women do not benefit from chemotherapy. As a general rule, one-third of women diagnosed with breast cancer would have improved without the chemotherapy and one-third would have died with or without it. Only one-third actually benefit from the treatment. Along the same lines, chemotherapies for cancer have decreased the size of the tumors but at the expense of an increase in frequency of secondary tumors and a very adversely affected lifestyle. Furthermore, most chemotherapy does not prolong life or result in a longer, high quality life. (Bear 2003; [54] Savage 2008; [55] Mittra 2007 [56])	True but not more broadly representative. These cherrypicked examples are not sufficient to back up such a huge claim. The failure rates cited here are all connected with niche forms of cancer or describe later attempts at treatment. 50–80% of early breast cancers are cured by surgery alone, with the survival rate boosted by various treatments. The first paper it cites is very clear that the study it's reporting on is poorly designed. The first reference is called "Earlier chemotherapy for breast cancer: perhaps too late but still useful." It says of the data behind this statement "This trial was started in 1988, and because of slow accrual, was stopped short of the number of patients that were really needed for adequate statistical power. Nevertheless, the addition of perioperative chemotherapy in this trial did not significantly improve overall or disease-free survival. Unfortunately, despite a valiant effort, the design of this trial probably doomed it to reach this result from the beginning, for a number of reasons." The second reference is very specific – referring only to higher intensity chemo in small-cell lung cancer vs normal chemo. High-Intensity Chemotherapy Does Not Improve Survival in	Use of niche, unrepresentative examples.	54 N 55 Y 56 N	7

	Small Cell Lung Cancer. Journal of the National Cancer Institute 100 (8):519. The third reference is concerned with epithelial cancers, and supports the spirit of the statement that chemotherapy can shrink tumours, but lead them to grow back faster if the cancer cells are not eradicated, but equally notes that doctors have therefore shifted to other drugs and treatments for these types of cancer https://www.nature.com/articles/ncponc0772			
Enna and Williams, in 2009, state: Success in federally funded drug discovery initiatives has had a checkered history. As one example, while the 1971 National Cancer Act gave the National Cancer Institute a charter to cure cancer, the incidence of this disease in the United States remains the highest in the world, with a death rate that has remained unchanged for over 50 years (193.9 per 100,000 in 1950 vs. 193.4 per 100,000 in 2002). This lack of progress is both surprising and disappointing given the billions of dollars spent over the past 40 years on improving treatment options, reducing cancer-related behaviors, such as smoking, and increasing efforts in early detection (Aggarwal, Danda, Shan Gupta, & Gehlot, 2009). Many are now coming to the realization that, as	This reference, from a book not a paper, is fairly used but out of date and no longer true. The overall rate of cancer deaths in the U.S. has declined by 27% during the past 25 years, using the same report that gave us the figures of 193.9 per 100,000. “The cancer death rate reached its peak in 1991, with 215.1 deaths per 100,000 population, but dropped steadily by about 1.5% per year to 156 per 100,000 population in 2016”. https://www.healio.com/hematology-oncology/lung-cancer/news/in-the-journals/%7B49a68303-a995-4a28-961e-b17ebb69d95a%7D/us-cancer-deaths-down-27-in-25-years-but-socioeconomic-gaps-widening With regard to this paper, the authors don’t say that animals *cannot* predict human effects, but that poor experimental modelling using animals leads to poor results. “The poor translation record of animal models to humans	Both the claim and the Enna and Williams reference are no longer reflective of modern cancer survival rates. The papers are misused and do not support the claims, with one pointedly highlighting the many non-animal causes for study failure such as inadequate sample sizes, which Dr Greek has wrongly labelled ‘deficiencies in the animal models’, when the deficiency is in the study design.	57 N	5, 3

in other therapeutic areas, the greatest limitation for identifying new drugs for treating cancer are the deficiencies in the animal models used for testing NCEs [new chemical entities, also referred to as new molecular entities or NMEs] (Aggarwal et al., 2009) . . .A major hurdle in the translational medicine undertaking is the fact that most preclinical animal models of disease generally lack predictive value with respect to the human condition under study. Indeed, the false positives that result from the present generation of animal assays are a major cause of NCE attrition in the clinic either because of lack of efficacy or the appearance of unacceptable side effects that were not detected preclinically [in animals]. While there are notable, albeit retrospective, exceptions (Zambrowicz & Sands, 2003), this weakness in the conventional drug discovery process has not been resolved with the use of transgenic animals which themselves contribute additional confounds that further complicate data interpretation. [57]	has been **attributed to poor preclinical methodologies** (Green, 2008, Hackam, 2001, Perel et al ., 2007) which include a lack of blinding and randomisation, adequate power/size (animal numbers), and an "optimization bias" in that very often only positive results are reported."(page 12) With the second point they are recommending more animals are used to strengthen the statistical power of the study. https://books.google.co.uk/books?id=jW763NvA_jAC&pg=PA9&dq=Aggarwal,+Danda,+Shan+Gupta,+%26+Gehlot,+2009&hl=en&sa=X&ved=0ahUKEwii2920w8DhAhWSTRUIHZyIDCwQ6AEIKjAA#v=onepage&q=false%20positives&f=false+			
Schreiber et al., in 2010, state: The ability of recombinant DNA to provide nearly unlimited access to human	The quote as separated from the paper is fairly used, but the paper is recommending using 'humanized' mice and stage 5 of their recommended way forward involves	The paper's claims are outdated and no longer true.	58 Y	5

proteins resulted in a second approach that is also common today—target-based drug discovery. Here, therapeutic targets are selected using insights gained most often from biochemistry, cell biology and model organisms. Small molecules are identified that modulate the targets (often by small-molecule screening) followed by optimization and clinical testing. Although this is a robust process, the common failure of candidate drugs in late-stage clinical testing, owing to unforeseen toxicity or lack of efficacy, reveals limits in our ability to select targets using surrogates of human physiology, such as in vitro assays and animal models. [58]	animal tests. Their meaning therefore was 'current' animal models at the time, not animal models that were being brought on-line in around 2010. They write: "Transplantable mouse models offer the advantage of speed since genetic lesions are introduced into stem or progenitor cells that are then transplanted into recipient animals. Such models exist for a number of cancer types, including lymphoma, glioblastoma, and carcinomas of the liver[19–21]. These models can be used to screen large numbers of genes for oncogenicity and acquired dependencies[22] and to determine the efficacy of small-molecule probes that have been optimized for animal testing." https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2939009/			
Markou, Chiamulera, Geyer, Tricklebank (of Eli Lilly), and Steckler (of Johnson and Johnson) state in 2009: Despite great advances in basic neuroscience knowledge, the improved understanding of brain functioning has not yet led to the introduction of truly novel pharmacological approaches to the treatment of central nervous system disorders. This situation has been partly attributed to the difficulty of predicting efficacy in patients based on results from preclinical studies. . . .	The text removed by Dr Greek is important. The net sentence is: "To address these issues, this review critically discusses the traditional role of animal models in drug discovery, the difficulties encountered, and the reasons why this approach has led to suboptimal utilization of the information animal models provide. The discussion focuses on how animal models can contribute most effectively to translational medicine and drug discovery and the changes needed to increase the probability of achieving clinical benefit." "Despite the extensive criticism of animal models	Papers selectively quoted to remove positive examples of animal efficacy.	59 Y	6

Few would dispute the need to move away from the concept of modeling CNS diseases in their entirety using animals. However, the current emphasis on specific dimensions of psychopathology that can be objectively assessed in both clinical populations and animal models has not yet provided concrete examples of successful preclinical-clinical translation in CNS drug discovery. . . . Since the founding of the American College of Neuropsychopharmacology (ACNP) in December 1961, there have been tremendous advances in neuroscience knowledge that have greatly improved our understanding of brain functioning in normal and diseased individuals. Unfortunately, however, these scientific advancements have not yet led to the introduction of truly novel pharmacological approaches to the treatment of central nervous system (CNS) disorders in general, and psychiatric disorders in particular (Hyman and Fenton, 2003; Fenton et al., 2003; Pangalos et al., 2007). . . .	(e.g., Horrobin, 2003), they continue to play a major role in drug discovery because of the need to calculate parameters, such as margin of safety referred to above, as well as for the primary purpose of target validation." They conclude: "In summary, the current translational approach recognizes that accurate predictions are based on the quality, reliability, and relevance to the disorder of *both* the preclinical and clinical measures. Although this requirement increases the burden on the animal models because extensive refinement and revalidation are required, the improved predictability of the models is expected to outweigh the effort required. Additionally, the requirement of extensive validation is not only an issue for animal studies; the same applies to challenge studies in healthy volunteers or sophisticated neurobiologically informed tests in patient trials that need to prove their validity to regulatory authorities." Once again, the authors are pointing to the value of animal models for safety testing while suggesting changes to older protocols to improve efficacy translation. http://www.ncbi.nlm.nih.gov/pubmed/18830240 .			
Neuzil et al., states in 2012: Animal testing is not ideal either, as the predictive value of such tests is limited owing to metabolic differences between humans and	Quote is incomplete, changing its meaning. The preceding sentence to this reads "Currently, however, the results obtained with new *in vitro* systems cannot replace animal testing because they do not take into account the complex interactions between different tissues and organs." It goes	The paper is discussing the prospects for technologies like organs-on-chips to replace animal testing one day.	60 Y	6

animals, and many ethical issues are raised by the testing.[60]	on "Despite their physiological differences to humans, whole-organism-based screens can provide deep insights into the effects of drug candidates on developmental processes, tissue-tissue interactions and metabolism." https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6493334/			
Björquist et al., in Drug Discovery World 2007: Furthermore, the compound attrition rate is negatively affected by the inability to predict toxicity and efficacy in humans. These shortcomings are in turn caused by the use of experimental pre-clinical model systems that have a limited human clinical relevance . . . Animal models are today important tools to detect adverse effects of compounds but are costly and their clinical relevance is widely debated. In fact, animal models are about 50% effective in predicting human toxicity to the liver, heart and during development.[61]	This is an article not a scientific paper, a repeat of reference 21 by a company, Cellartis, selling the competition to animal assays. The claims are unsupported by references and are factually incorrect in terms of the situation today. https://www.ddw-online.com/therapeutics/p92860-human-es-cell-derived-functional-cells-as-tools-in-drug-discoverywinter-2007.html	The repetition of an unsupported claim in an article by people who have a commercial interest in undermining animal models.	61 N	5, 8
Sharp and Langer write in 2011: The next challenge for biomedical research will be to solve problems of highly complex and integrated biological systems within the human body. Predictive models of these systems in either normal or disease states are beyond the capability of current knowledge and technology [62].	An article in science magazine, it doesn't mention animals at all. Instead it talks about building multi-disciplinary teams from different fields to address complex problems, for instance using big data, engineering and traditional methods together https://science.sciencemag.org/content/333/6042/527	Reference doesn't support the hypothesis	62 N	4
* Zhang et al., state in 2010: [The	This quote is fairly used.	The quote is fairly used and is	63 Y	1

publication of the report Toxicity Testing in the 21st Century: A Vision and a Strategy by the National Research Council of the National Academies of Science (NAS)] is a long-due response to the call by many for alternatives to the currently standard, whole-animal-based methodologies, which are inefficient, costly, and have had only limited success in making informative connections to human health risk associated with environmental chemical exposures. [63]		reflective of the general desire to move away from animal models if and when possible.		
Elias Zerhouni, former director of NIH and current head of R&D at Sanofi was quoted in the June 25, 2012 issue of Forbes as saying: "R&D in pharma has been isolating itself for 20 years, thinking that animal models would be enough and highly predictive, and I think I want to just bring back the discipline of outstanding translational science, which means understand the disease in humans before I even touch a patient."	This does not implicate the animal model. The animal model is intended to be instructive, not precisely predictive, and it is indeed madness to rely solely upon it. The role of the animal model is still required in Zerhouni's suggested amendment to the drug development process i.e. after animal models comes further specialisation of drugs for factors like genetic predispositions. Full article here https://www.forbes.com/sites/matthewherper/2012/06/06/can-bushs-nih-chief-fix-the-drug-industry/	The reference does not support the hypothesis. The animal model has not failed in its intended purpose, but the charge is that some people thought it could exceed its purpose.		4
* Raven wrote in 2012: " 'The mouse models really don't reflect the human condition,' says Shaw Warren, an infectious disease specialist at the Massachusetts General Hospital in Boston. 'Clearly, current animal models seem to be incapable of	https://www.nature.com/nm/articles?type=news&year=2012 News article, not a paper, talking about one particular disease. This does not implicate all animal models, but is an example of a situation where mice make a poor model for the particular human disease being investigated. In this	Dr Greek has selectively quoted from a news article to imply that a narrow criticism (of mouse models of human sepsis) are all mouse models, the referenced this editorial article as is it were a paper.	64 N	7

predicting results in human trials of new agents,' says Mitchell Fink, a surgeon at the University of California– Los Angeles." [64]	case, sepsis. Mice remain very good models for other diseases, for instance Familial ALS. The issue is discussed well here https://sciencebasedmedicine.org/mouse-model-of-sepsis-challenged/			
Mullane and Williams [65] state in 2012: "The difficulties in predicting drug efficacy from preclinical models have been of concern for more than two decades . . . Thus, novel findings apparently related to the systems and targets involved in disease causality; the delineation of the efficacy, selectivity and safety of NCEs; and the predictive relevance of biomarkers and animal model data to the human disease state, even when there is evidence for target engagement in humans, all frequently fail to enhance the success rate for new drug applications (NDAs)." They continue stating that one reason for the problems Pharma is facing is: "(i) An over-reliance on animal models of diseases that are poorly validated in the manner they are applied."	True. Another example of poor translational work and over-expecting the animal model to precisely model efficacy. The authors note that "over-reliance on animal models of diseases that are poorly validated in the manner they are applied" is the problem, not that animal models are useless when used for the correct applications.	A fair use of the quote, although preclinical models do not solely mean animal data, but also computers, cell cultures, organs on chips etc. The predictive relevance of human biomarkers is also criticised.	65 N	1
Clearly, scientists, not just animal advocates, do link the failure rate of new drugs to animal models. This is mainly due to the inability of animal models to predict efficacy and safety—the very things they are	Not a conclusion supported by the preceding references. Where the references are legitimate, the authors have criticised the animal model for some disease, like sepsis, using a particular species, like mice. Animals are not unable to predict safety at all, and achieve this end with a 90%+ success rate. Efficacy is a more patchy picture, very much	Unsupported, inaccurate summing-up.		2

supposed to predict. While there are many other problems with Pharma, reliance on the animal model is well recognized and discussed. Peruse just about issue of a drug development journal and you will find an article discussing the problems with animal models and why early human testing is the key to solving the pipeline problem as well as the efficacy and safety problems.	reliant on which species and which disease is being investigated. Good experimental design can mean choosing the right species. The claim that "early human testing is the key to solving the pipeline problem as well as the efficacy and safety problems" will not solve the pipeline problem, but probably is a good idea as part of addressing efficacy. Safety is not a major concern nor one leading to many drugs at all being withdrawn.			