


RESEARCH PAPER

Preclinical efficacy in therapeutic area guidelines from the U.S. Food and Drug Administration and the European Medicines Agency: a cross-sectional study

Correspondence Daniel Strech, Charité - University Medicine Berlin, QUEST - Center for Transforming Biomedical Research, Berlin Institute of Health (BIH), Berlin, Germany. E-mail: strech.daniel@mh-hannover.de

Received 6 March 2018; **Revised** 28 July 2018; **Accepted** 6 August 2018

Holger Langhof^{1,2} , William Wei Lim Chin², Susanne Wieschowski², Carole Federico³, Jonathan Kimmelman³ and Daniel Strech^{1,2}

¹Charité - University Medicine Berlin, QUEST - Center for Transforming Biomedical Research, Berlin Institute of Health (BIH), Berlin, Germany,

²Institute for History, Ethics and Philosophy of Medicine, Hannover Medical School (MHH), Hannover, Germany, and ³STREAM (Studies of Translation, Ethics and Medicine), Biomedical Ethics Unit, McGill University, Montreal, QC, Canada

BACKGROUND AND PURPOSE

Therapeutic area guidelines (TAGs) published by the EMA and the FDA offer guidance in planning the launch of a trial in a certain indication. We assessed and compared the guidance on preclinical efficacy of all available TAGs from EMA and FDA.

EXPERIMENTAL APPROACH

EMA and FDA websites and databases were searched for all TAGs. A mixed deductive and inductive approach was applied to analyse and cluster content for preclinical efficacy.

KEY RESULTS

A total of 114 EMA and 120 FDA TAGs were identified, covering 126 indications. Our core finding is that 75% of EMA TAGs and 58% from the FDA TAGs do not offer any guidance on preclinical efficacy. TAGs varied widely on the extent, nature and detail of guidance.

CONCLUSIONS AND IMPLICATIONS

Guidance on preclinical efficacy in a consistent, comprehensive and explicit way that still allows for justified deviations is an important but neglected aspect of transparency for drug development. This transparency would help sponsors in designing preclinical studies and in negotiating more efficiently with regulators.

Abbreviations

EMA, European Medicines Agency; FDA, U.S. Food and Drug Administration; ICH, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; IRBs, Institutional Review Boards; STAIR, Stroke Therapy Academic Industry Roundtable; TAGs, Therapeutic Area Guidelines

Introduction

Preclinical studies are intended to provide a key resource for risk–benefit assessment prior to early phase clinical trials. In general, the preclinical safety studies (mainly pharmacokinetics and toxicology) inform judgments about risk, and the preclinical efficacy studies (as part of the pharmacodynamics) inform judgments about benefit ('clinical promise'). Those designing a particular early human study need to provide a set of preclinical studies (most often animal studies) that is sufficient to demonstrate a favourable balance of safety and clinical promise.

Whether a set of preclinical animal studies is 'sufficient' to demonstrate efficacy and safety or not depends on the questions these studies addressed as well as on their study design features that aim to reduce validity threats. Over the past 10 years, many commentators have raised concerns about the design and reporting of preclinical reports (Kilkenny *et al.*, 2009; Prinz *et al.*, 2011; Begley and Ellis, 2012; Howells *et al.*, 2014), and some linked these concerns to the high attrition rates in clinical research (Prinz *et al.*, 2011). Some have further argued that regulatory bodies do not routinely assess preclinical efficacy studies when authorizing early phase studies (Kimmelman and Federico, 2017). Recent analyses on how preclinical efficacy data are reported within investigator brochures for early phase studies intensified this concern (Wieschowski *et al.*, 2018; Yasinski, 2018). One opportunity for probing this claim more deeply is to look at the recommendation on preclinical efficacy studies contained in regulatory guidelines.

There are no overarching regulatory guidelines that specify guidance on preclinical efficacy. For preclinical safety studies, in contrast, extensive regulatory guidance exists. The European Medicines Agency (EMA) provides over 50 guidelines for non-clinical toxicology that "help medicine developers prepare marketing authorisation applications for human medicines" (European Medicines Agency, 2017a). Another five guidelines specify non-clinical guidance on pharmacokinetics and toxicokinetics as well as safety pharmacology (European Medicines Agency, 2017a). Most of these safety guidelines have an overarching character and address non-clinical guidance that applies to several drug classes or therapeutic areas.

This lack of 'overarching' efficacy guidelines, however, might not be surprising because judgements about whether a certain set of preclinical efficacy studies is sufficient for clinical translation is much more dependent on specific disease models, therapeutic areas or drug class. Therefore, regulatory guidelines for specific therapeutic areas could provide a more suitable approach to navigate the complexity of preclinical efficacy recommendations. Within the European Union, common market applications for marketing authorization are required to fully comply with the EMA 'guidelines on clinical efficacy and safety' and to justify any deviation (European Medicines Agency, 2009). These guidelines are "intended to give guidance (...) in planning the overall pharmaceutical product development, as well as the non-clinical and clinical tests and studies of a compound intended to be used as human or veterinary medicinal products" (European Medicines Agency, 2009). The U.S. Food and Drug Administration (FDA) also recommends compliance with similar

guidelines, and deviations must be discussed with the agency (Food and Drug Administration, 2017b). We will refer to these EMA and FDA guidelines as Therapeutic Area Guidelines (TAGs).

This study assessed and compared the content for preclinical efficacy in all accessible TAGs from the EMA and the FDA. In addition, we compared the recommendations in TAGs with those presented in a systematically derived sample of experts guidelines for *in vivo* animal experiments (Henderson *et al.*, 2013).

Methods

TAGs search

EMA and FDA websites were searched for TAGs in December 2016. The EMA lists all guidelines on 'clinical efficacy and safety' under a common heading (European Medicines Agency, 2017b). A subsequent search using the site-wide search on the EMA's website was conducted to retrieve potentially relevant TAGs not listed under the common heading. Because the FDA does not group scientific guidelines under a common heading, a search was conducted using the database for FDA Guidance Documents (Food and Drug Administration, 2017a). All EMA and FDA guidelines were checked for relevance by two authors (W.C. and S.W.), and all documents relating to clinical development in specific therapeutic areas were identified and downloaded.

Comparative content analysis of EMA and FDA guidelines

All included EMA and FDA TAGs were categorized under one of 12 therapeutic areas as proposed by the EMA (European Medicines Agency, n.y.). We then applied a mixed deductive and inductive approach to analyse the TAGs' recommendations on preclinical efficacy. First, for the deductive approach, a matrix of different preclinical efficacy study design elements was derived from the International Council for Harmonization (ICH) Guideline on Good Clinical Practice (International Conference on Harmonization Working Group, 1996). We referred to the ICH Guideline as it is recognized by the EMA and FDA, and the preclinical provisions outlined in this guideline apply to all clinical trials, though ICH does not itself provide any TAGs. Second, to increase sensitivity and specificity in our assessment, the matrix also allowed the (inductive) inclusion of 'Other preclinical efficacy study design elements' not addressed in the ICH guideline.

Two researchers (W.C. and S.W.) independently analysed all included TAGs. For each TAG, all relevant text passages that gave guidance on preclinical efficacy were extracted. Thematic text analysis guided by the assessment matrix was applied to all text passages to assess preclinical study design recommendations presented in each TAG. The researchers (W.C. and S.W.) compared their extractions in order to check for differences in rating and extracting. All differences were resolved by discussion. H.L. and D.S. further double-checked the results of the thematic text analysis for validity and consistency, each taking a subsample of 20 guidelines.

Comparison with scientific expert guidelines for *in vivo* animal experiments

For the identification of scientific expert guidelines, we referred to a systematic review of guidelines addressing design and conduct of preclinical efficacy studies (Henderson *et al.*, 2013). The 26 guidelines identified in this review cover a diverse range of therapeutic areas with a total of 18 distinct indications. For those expert guidelines covering an indication that matched with a TAG indication, we checked if the corresponding EMA and FDA TAGs contained guidance on preclinical efficacy.

As this study was literature based and did not involve human subjects, no ethical approval was necessary.

Results

We identified 114 unique TAGs from EMA, and 120 from FDA that met our eligibility criteria (see flow diagram in Figure 1). These 234 TAGs covered 126 distinct indications (Table 1; also Supporting Information Table S1), which fitted into 11 distinct therapeutic areas. Indications ranged from common (e.g. diabetes mellitus) to orphan diseases (e.g. Duchenne muscular dystrophy). For 34 indications, both the FDA and the EMA provided at least one guideline. For the remaining 92 indications, we found guidance from one agency only.

In total, 25% ($n = 29$) of all 114 EMA and 42% ($n = 50$) of all 120 FDA TAGs included at least one recommendation on a particular aspect of preclinical efficacy studies or at least one remark on the need to test efficacy in preclinical studies (see Table 2). More specifically, 9% ($n = 10$) of all EMA and 25% ($n = 30$) of all FDA TAGs described a suitable preclinical efficacy model. Efficacy characteristics such as the desired nature, intensity or duration of effects were mentioned in 12% ($n = 14$) of all EMA and 16% ($n = 19$) of all FDA TAGs.

We further differentiated guidance on preclinical efficacy based on whether it addressed general versus specific study design elements, with the underlying assumption that general guidance is less helpful for practical needs in risk–benefit assessment (see Table 2). Guidance on choice of animal models tended to be specific (in 7 of 10 EMA TAGs and in 18 of the 30 FDA TAGs).

An example for specific guidance on efficacy models was found in the EMA TAG on ‘Non-clinical and clinical development of similar biological medicinal products containing low-molecular-weight heparins (LMWH)’:

“(…) *in vivo* pharmacodynamic activity of the similar and the reference LMWH should be quantitatively compared in an appropriate *in vivo* pharmacodynamic model, which takes into account state of the art knowledge about clinically relevant pharmacodynamic effects of LMWH and includes, at least, an evaluation of anti-FXa, and

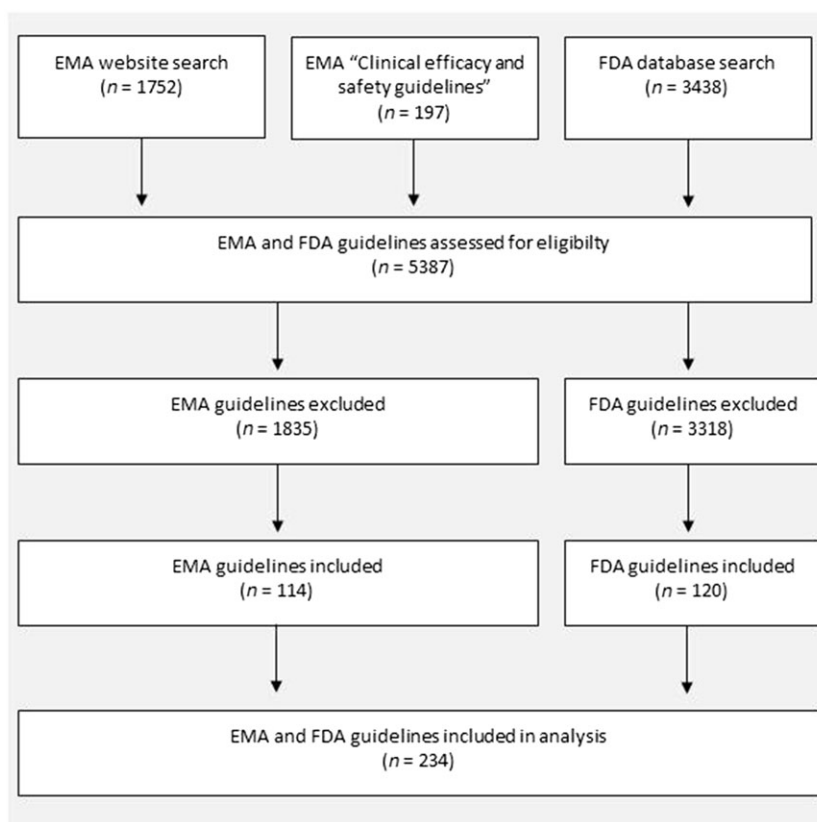


Figure 1

Flow chart: Search for EMA and FDA TAGs.

Table 1

Comparison of EMA and FDA TAGs

Therapeutic area	Guideline topic	EMA guidelines	FDA guidelines
Alimentary tract & metabolism	Weight control	2	1
	Irritable bowel syndrome	1	1
	Diabetes mellitus	2	2
	Nausea and vomiting	1	–
	Crohn's disease	1	–
	Gastroesophageal reflux disease	1	–
	Chronic constipation	1	–
	Ulcerative colitis	1	1
	Exocrine pancreatic insufficiency	–	2
	Gastroparesis	–	1
	Duodenal ulcer disease	–	1
	Enzyme replacement therapy	–	1
	Endocrine-related drug toxicity	–	1
	Growth	2	1
	Liver injury	–	1
Allergy & immunology	Allergic rhino-conjunctivitis	1	2
	Transplantation	1	1
	Allergic diseases	1	–
	Disorders affecting the haematopoietic system	–	2
	Primary humoral immunodeficiency	–	1
Anti-infectives for systemic use	Infectious disease	2	11
	HIV	2	10
	Fungal disease	1	1
	Hepatitis C	1	2
	Hepatitis B	1	1
	Sepsis	1	–
	Acute bacterial otitis media	–	1
	Acute bacterial sinusitis	–	2
	Influenza	2	4
	Clostridium difficile infection	–	1
	Gingivitis	–	3
	Sexually transmitted diseases	–	2
	Recurrent herpes labialis	–	1
	Sinusitis	–	1
	Smallpox	1	1
	Intra-abdominal infection	–	1
Head lice infestation	–	1	
Antineoplastic & immune-modulating agents	Cancer	5	7
	Chronic primary immune thrombocytopenia	1	–
	Infection	1	–
	Chronic lymphocytic leukaemia	1	–
	Non-small cell lung cancer	–	2
	Breast cancer	–	1
Blood product & biotech	Colon and rectal cancer	–	1
	Hepatitis B	1	–
	Haemostasis	1	1

continues

Table 1

(Continued)

Therapeutic area	Guideline topic	EMA guidelines	FDA guidelines
	Haemophilia A	1	–
	Haemophilia B	1	–
	Immuno-deficiency syndromes	2	–
	Bleeding disorder	3	–
	Blood	3	3
Cardiovascular system	Hypertension	3	1
	Lipid disorders	2	–
	Antiarrhythmics	1	–
	Restenosis	1	2
	Pulmonary arterial hypertension	1	–
	Stroke and systemic embolic events	1	–
	Venous thromboembolism	3	–
	Acute heart failure	2	–
	Chronic heart failure	1	–
	Angina pectoris	1	–
	Acute myocardial infarction	1	–
	Acute coronary syndrome	1	–
	Peripheral-arterial occlusive disease	1	–
	Cardiovascular disease	2	1
Dermatologicals	Skin conditions	1	1
	Psoriasis	1	–
	Acne	–	1
	Acute bacterial skin and skin structure infection	–	1
	Chronic cutaneous ulcer and burn wounds	–	1
	Upper facial lines	–	1
Genito-urinary system & sex hormones	Postmenopausal & vasomotor symptoms and vulvar and vaginal atrophy symptoms	1	2
	Urinary incontinence	1	–
	Contraceptives	1	1
	Chronic renal insufficiency	1	–
	Autosomal dominant polycystic kidney disease	–	1
	Bacterial vaginosis	–	1
	Urinary tract infection	–	1
	Vulvovaginal candidiasis	–	1
Nervous system	Anxiety disorder	2	1
	Epileptic disorders	1	1
	Duchenne and Becker muscular dystrophy	1	1
	Alzheimer's disease and other dementias	1	1
	Pain	1	–
	Alcoholism	1	1
	Migraine	1	1
	Depression	1	1
	Premenstrual dysphoric disorder	1	–
	Attention-deficit/hyperactivity disorder	1	–
	Insomnia	1	–
	Nociceptive pain	1	–

continues

Table 1

(Continued)

Therapeutic area	Guideline topic	EMA guidelines	FDA guidelines
	Post-traumatic stress disorder	1	–
	Panic disorder	1	–
	Obsessive compulsive disorder	1	–
	Smoking	1	–
	Acute stroke	1	–
	Bipolar disorder	1	–
	Schizophrenia	3	–
	Multiple sclerosis	2	–
	Neuropathic pain	1	–
	Amyotrophic lateral sclerosis	1	–
	Autism spectrum disorder	1	–
	Parkinson's disease	1	–
	Chronic fatigue syndrome/myalgic encephalomyelitis	–	1
	Hypnotic	–	1
	Psychoactive	–	1
Respiratory system	Asthma and chronic obstructive pulmonary disease	2	–
	Chronic obstructive pulmonary disease	1	4
	Acute respiratory distress syndrome	1	–
	Cystic fibrosis	1	–
	Asthma	1	–
	Severe acute respiratory syndrome	–	1
	Pneumonia	–	2
	Tuberculosis	–	1
	Bronchitis	–	1
Rheumatology	Osteoporosis	1	1
	Juvenile idiopathic arthritis	1	–
	Osteoarthritis	1	1
	Ankylosing spondylitis	1	–
	Psoriatic arthritis	1	–
	Rheumatoid arthritis	–	2
	Systemic lupus erythematosus and lupus nephritis	1	1
	Chronic disorders	1	–
Others	Anaesthesiology	–	2
	Life-threatening diseases	–	1
	Rare disease	–	2
Total		114	120

anti-FIIa activity and of release of tissue factor pathway inhibitor. (European Medicines Agency, 2016)”

An example of a general guidance for efficacy models can be found in the EMA TAG ‘Evaluation of anticancer medicinal products in man – addendum on paediatric oncology’:

“Sponsors should include a comprehensive overview on any testing of the agent for activity against pre-clinical model systems of paediatric tumours. (European Medicines Agency, 2003)”

Other guidance, such as on outcome measurements, was more often of general nature (in 9 of the 14 EMA TAGs and in 13 of the 19 FDA TAGs). Please refer to Supporting Information Table S2 for further text examples of general and specific guidance. None of the 234 TAGs addressed general internal validity issues like randomization, blinding of outcome assessment or sample size calculation.

Of all 234 TAGs, 18 addressed indications that were also addressed by scientific expert guidelines on preclinical efficacy issues reviewed in the Henderson *et al.* (2013) study. While these 18 TAGs could have drawn on the already

Table 2

Guidance on preclinical efficacy in EMA and FDA TAGs: distribution and specification

Therapeutic area guidelines (TAGs)		EMA (N = 114)		FDA (N = 120)		
Any guidance on preclinical efficacy		29 (25%)		50 (42%)		
Types of preclinical efficacy study design elements	No. of TAGs	Rating of guidance		No. of TAGs	Rating of guidance	
		General ^a	Specific ^a		General ^a	Specific ^a
Efficacy models	10 (9%)	3 (30%)	7 (70%)	30 (25%)	12 (40%)	18 (60%)
Effect characteristics ^b	14 (12%)	9 (64%)	5 (36%)	19 (16%)	13 (68%)	6 (32%)
Dosing characteristics AND/OR administration ^c	12 (11%)	5 (42%)	7 (58%)	29 (24%)	15 (52%)	14 (48%)
Receptor binding and specificity	0 (0%)	0 (0%)	0 (0%)	5 (4%)	2 (40%)	3 (60%)
Other ^d	6 (5%)	2 (33%)	4 (67%)	15 (13%)	5 (33%)	10 (67%)

^aSupporting Information Table S2 gives examples for how guidelines' guidance on preclinical efficacy was rated as general or specific.

^bFor example, nature, frequency, intensity of pharmacological effects, time to onset or duration of effects.

^cFor example, dose duration, dose interval, administration route, dose response or comparison of nontoxic dose findings.

^dOther preclinical efficacy studies, for example, special studies to assess pharmacological actions other than the intended therapeutic effects.

existing set of preclinical recommendations, half of them ($n = 9$) did not address any of these preclinical efficacy recommendations (see Table 3 for details). Only one TAG explicitly referred to the existing expert guidelines.

Discussion

In this study, we reviewed the full sample of TAGs (Therapeutic Area Guidelines) from EMA and FDA and assessed to what extent they provide advice on preferred design measures or other characteristics of preclinical efficacy studies. A core finding is that most TAGs (75% of EMA and 58% of FDA TAGs) do not include any remarks on preclinical efficacy. Among those TAGs that do contain at least some advice, most recommendations are of general nature, addressing neither specific design measures nor specific characteristics of animal models.

Whereas regulatory agencies provide extensive guidance on the design and objectives of preclinical safety studies, they do not do so for preclinical efficacy studies. However, regulatory agencies might have good reason for more recommendations on preclinical study design. Though the majority of new investigational drugs survive the phase I trials evaluating safety, 70% of phase II trials fail to demonstrate efficacy (Hay *et al.*, 2014). Such failures are costly and expose patients to ineffective therapeutic regimens. Many commentators have questioned design and reporting standards for preclinical efficacy studies (Prinz *et al.*, 2011; Begley and Ellis, 2012; Kimmelman and Federico, 2017; Mattina *et al.*, 2017). Better regulatory guidance on the design of preclinical efficacy studies might help reduce the burdens and costs associated with attrition during drug development. Just recently, the need for more robust study designs in preclinical efficacy studies was highlighted (Mattina *et al.*, 2016). Regulatory guidelines could strongly support these activities (Begley and Ellis, 2012).

The limited guidance on the design of preclinical efficacy studies in TAGs contrasts with the abundance of

recommendations in the scientific literature (Henderson *et al.*, 2013). For instance, the Stroke Therapy Academic Industry Roundtable (STAIR) publishes and updates recommendations for preclinical standards in the development of drugs for acute ischaemic stroke (Fisher *et al.*, 2009). STAIR provides a section with 'Recommendations to Clinicians on the Evaluation of Preclinical Data With Neuroprotective Drugs'. These recommendations include, for example, (i) demonstrating dose response, (ii) the window of opportunity or therapeutic time window, (iii) physiological monitoring performed in randomized and blinded studies, (iv) outcome measures including infarct volume and functional assessment in acute as well as long-term phase studies and (v) initial studies performed in rodents, with subsequent studies in larger species, such as cats. However, the EMA TAG for stroke does not refer to any of these preclinical recommendations (European Medicines Agency, 2001) and the FDA does not provide a stroke-specific TAG. Indeed, we found only one intervention – drug eluting stent – where regulatory guidelines appeared to draw on state-of-the-art thinking about design of preclinical efficacy studies (Schwartz *et al.*, 2008).

Opponents of more explicit guidance on preclinical design standards in TAGs might argue that the determination of the sufficient set of preclinical efficacy studies to inform phase I/II trials is too case specific. Vestergaard *et al.* claim that for innovative medicines with high complexity, classical regulatory approaches based on guidelines are too inflexible and thus insufficient (Vestergaard *et al.*, 2013; Narayanan *et al.*, 2014). Instead, they argue that regulators should establish preclinical study recommendations on a case-by-case approach. Their conclusion is based on an analysis of the scientific advice provided by the Committee for Medicinal Products for Human Use at the EMA and is thus limited to Advanced Therapy Medicinal Products. However, the authors further argue that such tailored approaches might also be suitable for other medicinal products with high levels of specificity (Vestergaard *et al.*, 2013).

Nevertheless, tailored approaches without any underlying (specific) standards have drawbacks. First, decisions on

Table 3Comparison of EMA and FDA TAGs with scientific guidelines for *in vivo* animal experiments

Guideline topic	Guidelines/authors	EMA		FDA	
		No. of TAGs on topic	Preclinical guidance contained?	No. of TAGs on topic	Preclinical guidance contained?
Acute renal failure	Bellomo <i>et al.</i> (2004)	0	n.a.	0	n.a.
ALS	Ludolph <i>et al.</i> (2010) and Scott <i>et al.</i> (2008)	0	n.a.	0	n.a.
Alzheimer	Shineman <i>et al.</i> (2011)	1	n.a.	1	n.a.
Analgesic drugs	Rice <i>et al.</i> (2008)	1	n.a.	0	n.a.
Antiarthritic molecules	Bolon <i>et al.</i> (2010)	2	General	2	Specific ^a
Arrhythmia	Curtis <i>et al.</i> (2013)	1	Specific	0	n.a.
Brain injury	Margulies and Hicks (2009)	0	n.a.	0	n.a.
Chemoprevention	Verhagen <i>et al.</i> (2003) and Kelloff <i>et al.</i> (1994)	0	n.a.	0	n.a.
Drug-eluting stents	Schwartz <i>et al.</i> (2008)	1	Specific	2	Specific
Endometriosis	Pullen <i>et al.</i> (2011)	0	n.a.	0	n.a.
Morbus Duchenne	Willmann <i>et al.</i> (2012) and Grounds <i>et al.</i> (2008)	1	Specific	1	General
Multiple sclerosis	Moreno <i>et al.</i> (2012)	1	n.a.	0	n.a.
Myocardial infarction	Bolli <i>et al.</i> (2004)	1	n.a.	0	n.a.
Rett syndrome	Katz <i>et al.</i> (2012)	0	n.a.	0	n.a.
Sepsis	Piper <i>et al.</i> (1996)	1	n.a.	0	n.a.
Stroke	Stroke Therapy Academic Industry Roundtable (STAIR) (1999), Macleod <i>et al.</i> (2009), Liu <i>et al.</i> (2009), Garcia-Bonilla <i>et al.</i> (2011), Savitz <i>et al.</i> (2011) and US National Institutes of Health National Institute of Neurological Disorders and Stroke (2011)	1	n.a.	0	n.a.
TBC	Kamath <i>et al.</i> (2005)	0	n.a.	1	General
Total		11	General: 2 Specific: 3	7	General: 2 Specific: 2

^aOnly one of the two FDA TAG on antiarthritic molecules gives specific guidance.

preclinical efficacy studies that are based solely on expert judgments made in individual negotiations between sponsors and regulatory agencies will lack transparency. Second, it seems highly inefficient to 'tailor' negotiations for preclinical efficacy study design if there are recurring features; made-to-measure clothes also start with at least basic models and body plans. The more transparent and specific regulatory standards for preclinical studies are, the easier it is to justify deviations from such standards.

Making guidance on preclinical efficacy studies for drug development more explicit in TAGs would also provide a measure of transparency on how regulatory agencies assess risk–benefit in early phase trials. It would also help Institutional Review Boards (IRBs), Data and Safety Monitoring Boards or investigators to clarify and judge the risk–benefit profile for early human studies (Kimmelman and Federico, 2017). Furthermore, explicit guidance on preclinical efficacy

studies might help to reduce the number of animals used in research, especially when accompanied by other means to contribute to this reduction, such as animal study registries (Wieschowski *et al.*, 2016).

We recognise that our analysis has several limitations. We only reviewed TAGs from the EMA and FDA. TAGs from other regulatory agencies might complement a sample for future studies. Moreover, further research is needed to better understand the rationales for why regulatory TAGs from FDA and EMA do not include recommendation on how to design preclinical efficacy studies more often and in more detail.

The EMA, FDA and other regulatory bodies are accountable for scientific oversight in drug development including risk–benefit assessment for clinical trials. TAGs provide a basis to support structure and transparency in judgements on whether the evidence for preclinical efficacy is appropriate for starting early phase clinical trials on new investigational

drugs. Our findings indicate, however, that this basis could in many cases be improved and extended by giving more explicit guidance on how to appropriately design preclinical efficacy studies. Further research, expert discussion and pilot tests are needed to determine the characteristics of efficient, but at the same time sufficient, guidance on preclinical efficacy studies that should be included in TAGs.

Acknowledgement

This study was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG).

Author contributions

The research study was conceived and designed by D.S., J.K., H.L. and C.F. The data were acquired by W.C., S.W. and H.L. The data were analysed by W.C., S.W., H.L. and D.S. H.L., D.S., C.F. and J.K. wrote the manuscript

Conflict of interest

The authors declare no conflicts of interest.

Declaration of transparency and scientific rigour

This **Declaration** acknowledges that this paper adheres to the principles for transparent reporting and scientific rigour of preclinical research recommended by funding agencies, publishers and other organisations engaged with supporting research.

References

- Begley CG, Ellis LM (2012). Drug development: raise standards for preclinical cancer research. *Nature* 483: 531–533.
- Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P (2004). Acute renal failure – definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* 8: R204–R212.
- Bolli R, Becker L, Gross G, Mentzer R Jr, Balshaw D, Lathrop DA (2004). Myocardial protection at a crossroads: the need for translation into clinical therapy. *Circ Res* 95: 125–134.
- Bolon B, Stolina M, King C, Middleton S, Gasser J, Zack D *et al.* (2010). Rodent preclinical models for developing novel antiarthritic molecules: comparative biology and preferred methods for evaluating efficacy. *Biomed Res Int* 2011; 2011: 569068.
- Curtis MJ, Hancox JC, Farkas A, Wainwright CL, Stables CL, Saint DA *et al.* (2013). The Lambeth Conventions (II): guidelines for the study of animal and human ventricular and supraventricular arrhythmias. *Pharmacol Ther* 139: 213–248.
- European Medicines Agency (2001). Clinical investigation of medicinal products for the treatment of acute stroke.
- European Medicines Agency (2003). Evaluation of anticancer medicinal products in man – addendum on paediatric oncology.
- European Medicines Agency (2009). Procedure for European Union guidelines and related documents within the pharmaceutical legislative framework.
- European Medicines Agency (2016). Guideline on non-clinical and clinical development of similar biological medicinal products containing low-molecular-weight-heparins.
- European Medicines Agency (2017a). Non-clinical guidelines. Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000083.jsp&mid=WC0b01ac0580027548 (accessed 3/15/2017)
- European Medicines Agency (2017b). Clinical efficacy and safety guidelines. [Online] Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000085.jsp&mid=WC0b01ac0580027549 (accessed 3/15/2017).
- Fisher M, Feuerstein G, Howells DW, Hurn PD, Kent TA, Savitz SI *et al.* (2009). Update of the stroke therapy academic industry roundtable preclinical recommendations. *Stroke* 40: 2244–2250.
- Food and Drug Administration (2017a). Search for FDA Guidance Documents (database). [Online] Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000085.jsp&mid=WC0b01ac0580027549 (accessed 3/15/2017).
- Food and Drug Administration (2017b). Guidance for Industry. [Online] Available at: <https://www.fda.gov/RegulatoryInformation/Guidances/default.htm> (accessed 3/15/2017).
- Garcia-Bonilla L, Rosell A, Torregrosa G, Salom JB, Alborch E, Gutierrez M *et al.* (2011). Recommendations guide for experimental animal models in stroke research. *Neurologia* 26: 105–110.
- Grounds MD, Radley HG, Lynch GS, Nagaraju K, De Luca A (2008). Towards developing standard operating procedures for pre-clinical testing in the mdx mouse model of Duchenne muscular dystrophy. *Neurobiol Dis* 31: 1–19.
- Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J (2014). Clinical development success rates for investigational drugs. *Nat Biotechnol* 32: 40–51.
- Henderson VC, Kimmelman J, Fergusson D, Grimshaw JM, Hackam DG (2013). Threats to validity in the design and conduct of preclinical efficacy studies: a systematic review of guidelines for in vivo animal experiments. *PLoS Med* 10: e1001489.
- Howells DW, Sena ES, Macleod MR (2014). Bringing rigour to translational medicine. *Nat Rev Neurol* 10: 37–43.
- International Conference on Harmonisation Working Group (1996). ICH harmonised tripartite guideline: guideline for good clinical practice E6 (R1) 1996. ICH: Geneva.
- Kamath AT, Fruth U, Brennan MJ, Dobbelaer R, Hubrechts P, Ho MM *et al.* (2005). New live mycobacterial vaccines: the Geneva consensus on essential steps towards clinical development. *Vaccine* 23: 3753–3761.
- Katz DM, Berger-Sweeney JE, Eubanks JH, Justice MJ, Neul JL, Pozzo-Miller L *et al.* (2012). Preclinical research in Rett syndrome: setting the foundation for translational success. *Dis Model Mech* 5: 733–745.
- Kelloff GJ, Johnson JR, Crowell JA, Boone CW, DeGeorge JJ, Steele VE *et al.* (1994). Guidance for development of chemopreventive agents. *J Cell Biochem Suppl* 20: 25–31.

- Kilkenny C, Parsons N, Kadyszewski E, Festing MF, Cuthill IC, Fry D *et al.* (2009). Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS One* 4: e7824.
- Kimmelman J, Federico C (2017). Consider drug efficacy before first-in-human trials. *Nature* 542: 25–27.
- Liu S, Zhen G, Meloni BP, Campbell K, Winn HR (2009). Rodent stroke model guidelines for preclinical stroke trials (1st edition). *J Exp Stroke Transl Med* 2: 2–27.
- Ludolph AC, Bendotti C, Blaugrund E, Chio A, Greensmith L, Loeffler JP *et al.* (2010). Guidelines for preclinical animal research in ALS/MND: a consensus meeting. *Amyotroph Lateral Scler* 11: 38–45.
- Macleod MR, Fisher M, O'Collins V, Sena ES, Dirnagl U, Bath PM *et al.* (2009). Good laboratory practice: preventing introduction of bias at the bench. *Stroke* 40: e50–e52.
- Margulies S, Hicks R (2009). Combination therapies for traumatic brain injury: prospective considerations. *J Neurotrauma* 26: 925–939.
- Mattina J, Carlisle B, Hachem Y, Fergusson D, Kimmelman J (2017). Inefficiencies and patient burdens in the development of the targeted cancer drug sorafenib: a systematic review. *PLoS Biol* 15: e2000487.
- Mattina J, MacKinnon N, Henderson VC, Fergusson D, Kimmelman J (2016). Design and reporting of targeted anticancer preclinical studies: a meta-analysis of animal studies investigating sorafenib antitumor efficacy. *Cancer Res* 76: 4627–4636.
- Moreno B, Espejo C, Mestre L, Suardiaz M, Clemente D, de Castro F *et al.* (2012). Guidelines on the appropriate use of animal models for developing therapies in multiple sclerosis. *Rev Neurol* 54: 114–124.
- Narayanan G, Cossu G, Galli MC, Flory E, Ovelgonne H, Salmikangas P *et al.* (2014). Clinical development of gene therapy needs a tailored approach: a regulatory perspective from the European Union. *Hum Gene Ther Clin Dev* 25: 1–6.
- Piper RD, Cook DJ, Bone RC, Sibbald WJ (1996). Introducing critical appraisal to studies of animal models investigating novel therapies in sepsis. *Crit Care Med* 24: 2059–2070.
- Prinz F, Schlange T, Asadullah K (2011). Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10: 712.
- Pullen N, Birch CL, Douglas GJ, Hussain Q, Pruijboom-Brees I, Walley RJ (2011). The translational challenge in the development of new and effective therapies for endometriosis: a review of confidence from published preclinical efficacy studies. *Hum Reprod Update* 17: 791–802.
- Rice AS, Cimino-Brown D, Eisenach JC, Kontinen VK, Lacroix-Fralish ML, Machin I *et al.* (2008). Animal models and the prediction of efficacy in clinical trials of analgesic drugs: a critical appraisal and call for uniform reporting standards. *Pain* 139: 243–247.
- Savitz SI, Chopp M, Deans R, Carmichael T, Phinney D, Wechsler L (2011). Stem cell therapy as an emerging paradigm for stroke (STEPS) II. *Stroke* 42: 825–829.
- Schwartz RS, Edelman E, Virmani R, Carter A, Granada JF, Kaluza GL *et al.* (2008). Drug-eluting stents in preclinical studies: updated consensus recommendations for preclinical evaluation. *Circ Cardiovasc Interv* 1: 143–153.
- Scott S, Kranz JE, Cole J, Lincecum JM, Thompson K, Kelly N *et al.* (2008). Design, power, and interpretation of studies in the standard murine model of ALS. *Amyotroph Lateral Scler* 9: 4–15.
- Shineman DW, Basi GS, Bizon JL, Colton CA, Greenberg BD, Hollister BA *et al.* (2011). Accelerating drug discovery for Alzheimer's disease: best practices for preclinical animal studies. *Alzheimers Res Ther* 3: 28.
- Stroke Therapy Academic Industry Roundtable (STAIR) (1999). Recommendations for standards regarding preclinical neuroprotective and restorative drug development. *Stroke* 30: 2752–2758.
- US National Institutes of Health National Institute of Neurological Disorders and Stroke (2011) Improving the quality of NINDS-supported preclinical and clinical research through rigorous study design and transparent reporting. US National Institutes of Health National Institute of Neurological Disorders and Stroke: Bethesda (Maryland).
- Verhagen H, Aruoma OI, van Delft JH, Dragsted LO, Ferguson LR, Knasmuller S *et al.* (2003). The 10 basic requirements for a scientific paper reporting antioxidant, antimutagenic or anticarcinogenic potential of test substances in in vitro experiments and animal studies in vivo. *Food Chem Toxicol* 41: 603–610.
- Vestergaard HT, Apote LD, Schneider CK, Herberts C (2013). The evolution of nonclinical regulatory science: advanced therapy medicinal products as a paradigm. *Mol Ther* 21: 1294–1296.
- Wieschowski S, Chin WWL, Federico C, Sievers S, Kimmelman J, Strech D (2018). Preclinical efficacy studies in investigator brochures: do they enable risk–benefit assessment? *PLoS Biol* 16: e2004879.
- Wieschowski S, Silva DS, Strech D (2016). Animal study registries: results from a stakeholder analysis on potential strengths, weaknesses, facilitators, and barriers. *PLoS Biol* 14: e2000391.
- Willmann R, De Luca A, Benatar M, Grounds M, Dubach J, Raymackers JM *et al.* (2012). Enhancing translation: guidelines for standard pre-clinical experiments in mdx mice. *Neuromuscul Disord* 22: 43–49.
- Yasinski E (2018). Study questions animal efficacy data behind trials. *Science* (New York, NY) 360: 142.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

<https://doi.org/10.1111/bph.14485>

Table S1 Comparison and matching of EMA and FDA therapeutic area guidelines (TAG) – Full list. All TAG offering any guidance on preclinical efficacy (specific or general) underlined.

Table S2 Rating of preclinical guidance in EMA and FDA TAGs: text examples illustrating general versus specific guidance.