Enhancing translation: Guidelines for standard pre-clinical experiments in \textit{mdx} mice

Raffaella Willmann\textsuperscript{a,*}, Annamaria De Luca\textsuperscript{b}, Michael Benatar\textsuperscript{c}, Miranda Grounds\textsuperscript{d}, Judith Dubach\textsuperscript{e}, Jean-Marc Raymackers\textsuperscript{f}, Kanneboyina Nagaraju\textsuperscript{g}, on behalf of the TREAT-NMD Neuromuscular Network

\textsuperscript{a} Biozentrum, University of Basel, Switzerland
\textsuperscript{b} Sezione di Farmacologia, Dipartimento Farmacobiologico, Facoltà di Farmacia, Università di Bari, Italy
\textsuperscript{c} Department of Neurology, Miller School of Medicine, Miami, FL, USA
\textsuperscript{d} School of Anatomy and Human Biology, The University of Western Australia, Perth, Australia
\textsuperscript{e} Santhera Pharmaceuticals, Liestal, Switzerland
\textsuperscript{f} Institute of Neuroscience, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
\textsuperscript{g} Research Center for Genetic Medicine, Children’s National Medical Center, Washington, DC, USA

Abstract

Duchenne Muscular Dystrophy is an X-linked disorder that affects boys and leads to muscle wasting and death due to cardiac involvement and respiratory complications. The cause is the absence of dystrophin, a large structural protein indispensable for muscle cell function and viability. The \textit{mdx} mouse has become the standard animal model for pre-clinical evaluation of potential therapeutic treatments. Recent years have seen a rapid increase in the number of experimental compounds being evaluated in the \textit{mdx} mouse. There is, however, much variability in the design of these pre-clinical experimental studies. This has made it difficult to interpret and compare published data from different laboratories and to evaluate the potential of a treatment for application to patients. The authors therefore propose the introduction of a standard study design for the \textit{mdx} mouse model. Several aspects, including animal care, sampling times and choice of tissues, as well as recommended endpoints and methodologies are addressed and, for each aspect, a standard procedure is proposed. Testing of all new molecules/drugs using a widely accepted and agreed upon standard experimental protocol would greatly improve the power of pre-clinical experimentations and help identifying promising therapies for the translation into clinical trials for boys with Duchenne Muscular Dystrophy.

© 2011 Elsevier B.V. All rights reserved.

Keywords: DMD; \textit{mdx}; Methods; Standard operating procedures; Pre-clinical experiments

Introduction

The last few years have witnessed increasing efforts to evaluate potential therapeutic compounds for Duchenne Muscular Dystrophy (DMD) in the \textit{mdx} mouse model. A plethora of data was published reporting important effects of various strategies, making it difficult to select the best candidates for human studies and often facing less enthusiastic or virtually no positive results in patients. The risk of false positive or false negative is still very high, due to intrinsic problem in animal models as well as improper definition of either study-design or readout parameters. Indeed, these can vary considerably with respect to important methodological features such as the selection of the outcome measures and the choice of sampling times for assessment (see Table 1 in [1] and in [2]), confounding comparison of results and hindering the prioritization of human clinical trials in already rare and restricted patient populations.
populations. One way to avoid this is to do animal experiments in the same manner as human clinical trials, using a standardized and methodologically rigorous approach to pre-clinical experimental studies in mdx [3–5], similar to those published recently for pre-clinical trials in ALS/MND and stroke [6–8].

The formulation of these recommendations was initiated by a workshop held in Brazil in 2006 [9] and continued in two workshops held in 2007 and 2008 [10]. In an effort to attain a consensus on the methods used in mouse pre-clinical studies, a group of scientists has come together under the auspices of the multi-national TREAT-NMD network (http://www.treat-nmd.eu/). This manuscript summarizes the deliberations of this group with respect to standardization of methods and proposes a number of procedures that relatively easily could be implemented in pre-clinical therapeutic studies.

“Proof of concept” versus “pre-clinical therapeutic” studies

At the outset, we wish to make a distinction between what has been termed “proof-of-concept” and “pre-clinical therapeutic” studies. The former are early stage exploratory studies that raise treatment hypotheses, which then require formal evaluation in pre-clinical “therapeutic” studies. The recommendations of this working group regarding experimental standardization should not be construed as an inhibition of scientific creativity in such initial proof-of-concept studies. However, pre-clinical therapeutic studies, meant to demonstrate the efficacy of a potential therapy, need to be designed with appropriate methodological rigor. A standardization of experimental conditions will facilitate comparison of the relative benefits of treatments evaluated by different investigators in different laboratories. The goal is to improve the speed and efficiency of translation from the laboratory to patients in the clinic.

Early phase versus later phase experiments

The progression of pathology in the mdx mouse is influenced by growth [11] and may be divided into three main phases: the pre-weaning phase (0–3 weeks of age) which is strongly influenced by growth and corresponds roughly to the first 6 months of human patients (see Box 1 in [4]); the post-weaning phase, with an acute onset of pathology around 3 weeks, followed at about 8 weeks by the adult phase with a reduced low level of chronic damage that persists throughout life. Experiments in the pre-weaning phase aim to evaluate the ability of a drug to reduce/delay the first round of necrosis. These studies on very young mice, which may not practically be available from central breeders, are frequently performed on pups of the same litter, without gender distinctions, from in-house colonies. These should be replenished every 20 generations to avoid gene drifting (see also http://jaxmice.jax.org/genetichealth/drift.html) and care should be taken to optimize litter size and to avoid any unnecessary stress (see below). Also, the number of pups per cage is known to affect post-natal growth due to issues of feeding [12]; this relates to the strength of stimulus for milk production (may be a problem with small litters) and sufficient availability of milk supply (an issue with large litters). Ideally therefore, numbers of pups in litters can be adjusted neonatally by moving pups between mothers to increase or decrease the number and provide a standardised number in each litter (e.g. seven pups/mother).

The guidelines proposed in the following chapters of this manuscript are relevant to standardization of experiments in the post-weaning and adult phase.

Methodological standards

There are a number of methodological issues that require standardization. These include animal care, several aspects of experimental design (including randomization, outcome measure selection, blinding and consideration of sample size and power) as well as strategies for reporting study results.

1. Animal care

A number of aspects of animal care may contribute to the variability observed in mice from different laboratories (see data collection in [2]). A standardized approach to animal care is expected to improve the reproducibility of the mouse dystrophic phenotype between experiments and laboratories. Grounds et al. [4] provided a detailed analysis of the impact of animal care on animal physiology; key measures to avoid such confounders are summarized as follows:

- Purchase mice from central distributors. This also potentially allows for a larger number of age-matched, same gender animals to be available within a reasonable timeframe, shorter than would normally be possible when relying on in-house animal breeding facilities.
- Report food regimen. A comparison of food composition from several food suppliers, including the standard diet used by Jackson Labs for mdx mice, showed great differences in the content of vitamins, minerals and crude fat (data not shown). Changing the source of the protein content from soy to casein, or from meat to fish, has dramatic consequences on some biological events and may affect the impact of a drug trial [4]. It is therefore recommended that within the same laboratory, the supplier of mouse diet be carefully maintained constant throughout different drug trials, thus avoiding changes in diet. It is advisable to report the food supplier in published data and manuscripts, and eventually, food composition should be considered when interpreting the results. Cage bedding should also be considered as some bedding (e.g. wheat kernels) contains high levels of vitamin E (anti-oxidant) and can...
have effects on mdx pathology [13]. Shell grit bedding or dried corn cob with nothing added is recommended.

- Reduce animal stress. Frequent human intervention is stressful to mice and changes in personnel who handle mice are a source of variation in breeding and behavior. Such changes should be avoided. Ideally, persons handling with mice should receive specific education on this topic. A recent review describes the use of environmental enrichment as a strategy to reduce anxiety [14,15].

- Consider circadian rhythm. Circadian rhythm is known to influence many physiological processes in mice [16,17]. Moreover, inflammation is known to play a role in DMD and mdx pathology [18–20] and is influenced by cortisol, the secretion of which is regulated by circadian rhythm. Finally, several studies in conditions such as inflammation, high blood pressure, and cancer, show a differential effect of drugs depending on the timing of daily administration [21–23]. Therefore, to minimize any variation due to the circadian rhythm, care should be taken to handle, treat and sample animals at the same time each day when performing experiments.

- Choose gender appropriately. It is preferable to use male animals for experiments, especially for pharmacological studies, given the potential impact of hormonal factors on drug metabolism. For non-pharmacological studies (e.g. molecular replacement of dystrophin), however, it may be reasonable to use female animals given the finding of consistent pathology in homozygous female mdx mice. At the very least, gender should be clearly stated and animals of only one gender used in a given experiment.

2. Experimental design

2.1. Outcome measure selection

Factors to consider in the selection of outcome measures for pre-clinical therapeutic studies include (a) reproducibility, (b) objectivity, (c) blinded assessment, (d) relevance to disease biology (e.g. muscle histology), and (e) similarity to human clinical trials endpoints (e.g. ability to walk and muscle strength testing). Depending on the presumed mechanism of action and the intended target of the experimental agent, additional outcome measures (e.g. to assess cardiac function) may be appropriate. Although a primary outcome measure should be designated for power calculations, it is appropriate to evaluate several outcome measures, in part because pre-clinical therapeutic studies should be hypothesis generating. Therefore, a core set of outcome measures, to be considered as a minimum, is proposed that should be performed in all pre-clinical therapeutic studies. In vivo non-invasive physiological readouts are assessed longitudinally during the treatment protocol, while terminal readouts are assessed at the end of the experiment. In addition, optional additional tests that may add value to specific therapeutic approaches are recommended. However care should be taken not to overload mice with tests in order to avoid too much stress. Standardized protocols for the assessment of most of the recommended parameters have already been produced by specialized working groups of experts and are reported in brackets (pdfs available on www.treat-nmd.eu/sopdmd [10]).

2.1.1. Core set of parameters in mdx pre-clinical efficacy studies

- Body weight of each individual animal should be measured at the beginning and end of every experiment, as well as at regular intervals during long-term studies.

- At least two of the following three functional in vivo readouts should also be tested: mouse strength (tests: whole body tension (DMD_M.2.2.006) or grip strength (DMD_M.2.2.001)), capacity for exercise (tests: treadmill or wheel running (DMD_M.2.1.003), spontaneous activity (test: open field digiscan (DMD_M.2.1.002)).

- Terminal tests include histological analyses of H&E stained muscle sections (e.g. % necrosis (DMD_M.1.2.007), % centronucleated (DMD_M.1.2.001) or unaffacted myofibres, myofibre size (DMD_M.1.2.001), fibrosis in exercised or older mdx mice, blood parameters (e.g. creatine kinase), muscle force/function (in vitro (DMD_M.1.2.002) or in situ isometric force (DMD_M.2.2.005)) and, when the therapy is supposed to induce a change in their expression, dystrophin/utrophin detection (western blot or immunofluorescence).

2.1.2. Optional set of parameters in mdx pre-clinical efficacy studies

- Physiological/functional tests in vivo include anatomical imaging (MRI), echocardiography (DMD_M.2.2.003), respiratory function tests (DMD_M.2.2.002), coordination tests (wire test, 4-limb grid test), other in vivo muscle force assessments (e.g. using a dynamometer), cognitive tests.

- Terminal tests include electrophysiology (ion channel function (DMD_M.1.2.005) and calcium homeostasis of skeletal muscle, sarcolemma integrity (Evans blue dye or Procion orange dye stains of muscle tissue or blood levels of lactate dehydrogenase and pyruvate kinase), fat content of tissue sections (oil red O stain), pro-inflammatory cytokines and transcription factors (ELISA, western blot, q-PCR, immunohistochemistry) and heart analyses such as left ventricular pressure/volume loops (DMD_M.2.2.004).

2.2. Muscle selection

One of the factors that contributes to discrepancies in results obtained from various laboratories may be the use of different muscles for a given outcome measure. These
discrepancies arise from several sources: (1) muscles are not equally affected by pathology (or exercise-induced damage), (2) muscle size may be an important consideration for the proper application of a method; for instance, perfusion in electrophysiological and physiological measurements is more efficient if applied to thin muscles, and (3) the fast/slow myofibre composition of muscle, muscle morphology and functional activity may play a role in the outcome of histological staining. The analysis of several different muscles from each mouse can improve the comparison of data between laboratories, although usually a specific muscle is used for a specific measurement (e.g. histology/immunocytochemistry on exercised quadriceps, RNA extracted from tibialis anterior for PCR, protein extracted from gastrocnemius for western blotting), in order to make the best use of all tissues from an individual mouse. The list in Table 1 is designed to guide scientists when deciding which muscle should be chosen for which assay and is meant to help reduce variation between, and within, laboratories.

2.3. Sample size and power

It is essential that animal experiments are designed with sufficient power to address the question at hand, and thereby to avoid the risk of both false positive and false negative results. This requires the investigator to (1) select the outcome measures that will be used to quantify the effect of a potential therapeutic agent (see recommendations in Section 3), (2) define the difference in outcome measure between treated and untreated mice that is thought to be clinically relevant, and (3) estimate the number of animals required to assess a drug-induced benefit in pathological aspects that are normally not very severe in adult mdx mice [25]. Several types of exercise with different impacts on mouse pathology were tested and are described in detail in [4]. If experiments are to be conducted in exercised conditions, we recommend the use of the same exercise regimen, including timing of the exercise sessions, to avoid inter-laboratory differences.

2.4. Randomization and blinding

Evidence from therapeutic studies in animal models (e.g. SOD1 mouse) of other human diseases (e.g. ALS) suggests that randomization is an essential element of experimental design if the effects of confounding are to be avoided [24]. The advantage of randomization over an approach such as stratification is that it controls for both known and unknown confounding factors. Known confounders in the mdx animal model include litter effects and animal weight and at a minimum, experimental groups should be stratified based on these characteristics. Optimally, within each stratum, animals should be randomly assigned to treatment groups and blinded assessment of outcome should be performed.

2.5. Exercise

A means to exacerbate muscle damage and pathology in mice is the use of exercise, a strategy that can be useful to assess a drug-induced benefit in pathological aspects that are equally affected by standard exercise regimes that induce muscle damage and contraction-induced pathological injury; but badly affected during onset of necrosis.

Table 1

Examples of mdx muscles widely used for morphological, biochemical and functional outcome measures. This table should help in choosing muscles for the planned endpoint assessments. Legend: ** = particularly appropriate; *= appropriate; X = not suitable.

<table>
<thead>
<tr>
<th>Muscle name</th>
<th>Type</th>
<th>Morphology</th>
<th>Biochemistry</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diaphragm (DIA)[^a]</td>
<td>Fast</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Extensor digitorum longus (EDL)[^b]</td>
<td>Fast</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Gastrocnemius (GAS)[^c]</td>
<td>Mixed</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Quadriceps (QUAD)[^d]</td>
<td>Mixed</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Biceps/triceps (BB/TB)[^e]</td>
<td>Mixed</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Tibialis anterior (TA)[^f]</td>
<td>Fast</td>
<td>[ ]**</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Soleus (SOL)[^g]</td>
<td>Slow</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

\[^a\] Severe phenotype [28], should be used as much as possible.
\[^b\] Small size: size makes a muscle more suitable for functional assays like physiological recordings (see Section 3.1).
\[^c\] Large size, heterogeneous fiber type composition, affected by exercise [4,29,30].
\[^d\] Large size, affected by exercise [4,29,30].
\[^e\] As back-up if GAS/QUAD not sufficient.
\[^f\] Large size, accessible [29].
\[^g\] Small size. As a purely slow muscle, soleus is more preserved from dystrophic pathology.
\[^*\] Only in in situ studies.
\[^**\] Less affected by standard exercise regimes that induce muscle damage and contraction-induced pathological injury; but badly affected during onset of necrosis.

variations; a detailed protocol for mild forced treadmill exercise can be found on www.treat-nmd.eu/sopdmd (DMD_M.2.1.001).

2.6. Age at treatment start and at terminal sampling

Based on mechanism of action, different drugs may be more or less effective depending on the age at which treatment is initiated and on the time period over which the drug is administered (i.e. treatment duration). Due to the ultimate translational aim of the pre-clinical experiments, it is important to consider the relationship between the age of the mdx mice and possible equivalence in DMD boys. A comparison of developmental stages in mice and humans is described in details elsewhere [4] and summarized in Table 2.

A crucial point when comparing results from different laboratories is the sampling age for terminal experiments (physiology, histology, and molecular analysis), so that the efficacy of diverse interventions can be rigorously evaluated. We recognize the need for variation in the choice of starting age and duration of the experiment, depending on the anticipated clinical impact of the experimental agent, and therefore, in the interest of consistency, we propose inclusion of particular sampling time points as a common denominator to increase comparability between laboratories. We suggest a standard, common sampling age of 12 weeks for medium term experiments (±2 weeks for endpoints whose evaluation cannot be accomplished in one day for all animals, such as physiological recordings) and of 6 months and 1 year for long term experiments.

3. Reporting results

Pre-clinical therapeutic studies aim to deliver information about how effective a new treatment may be when tested in patients. Therefore, when designing such studies and reviewing the results it is particularly important to look at three levels of benefits: (1) The difference in the outcome measures between the treated and untreated groups of mdx mice, taking into account any effect of the control/vehicle-treated group of mdx mice. Thus, in addition to demonstrating statistical significance, the extent of the benefit should be shown. This can be calculated as: Extent of benefit (%) = [treated mdx - untreated mdx] ÷ [wild type - untreated mdx] × 100. (2) The magnitude of the treatment effect in mdx mice relative to the difference between untreated mdx and wild type/normal C57Bl/10SnScmice, which has been referred to as the recovery score [26,27], should also be reported (see Section 3.1). (3) Ideally, some reference should also be made to the anticipated translational benefit to patients of each specific mouse outcome, which for most outcomes still needs to be formulated (see Section 3.2 and Box 1).

3.1. The recovery score

The recovery score is a tool that can be used to compare different therapies applied to mice, or results obtained by different laboratories with the same therapy. Firstly, the defect for a specific parameter caused by the absence of dystrophin must be quantified by comparing data obtained for wild type and mdx mice using a given assay. Then, the difference between untreated and treated mdx with the same assay is calculated. The recovery score may then be calculated as follows (see also DMD_M.1.1.001):

\[
\text{Recovery score} = \frac{|\text{treated } \text{mdx} - \text{untreated } \text{mdx}|}{|\text{wild type} - \text{untreated } \text{mdx}|} \times 100
\]

A score of 100% indicates that the parameter in treated mdx mice is equal to that of control wild type mice, and 0% indicates that no improvement was obtained relative to untreated mdx mice [27].

This score is scale-free and permits an easy comparison of treatment efficacy, independent of whether an increase or a decrease in measured values is expected. Moreover, it relates the benefit to the degree of severity of the given outcome measure. For instance, if muscle force in a specific muscle is severely affected by the absence of dystrophin, small differences between untreated and treated mdx groups that are statistically significant would still yield a low recovery score. Therefore this calculation represents a tool to evaluate the effective recovery achieved by the treatment tested. Although this implies the effort to include a wild type group of mice in any pre-clinical therapeutic study, we encourage the calculation of the recovery score in all studies where this effort is feasible.

3.2. Anticipated translational benefit

The anticipated translational benefit estimates the efficacy of treatment as measured in dystrophic mice (or dogs) in relation to a possible patient outcome. This will probably be different for each parameter being examined (see Box 1). The aim is to provide a basis for the justification of decisions to move an intervention (identified as ‘promising’ in pre-clinical trials) forward into clinical trials for DMD patients. Only where a sufficiently high translational benefit is demonstrated are clinical benefits likely to result. For example, if a specific parameter has a high recovery score in mice, but low translation efficacy in humans, this

<table>
<thead>
<tr>
<th>Time point</th>
<th>Mice</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 week</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Human</td>
<td>Newborn</td>
<td>~3 mo.</td>
</tr>
</tbody>
</table>

Box 1. Discussion of anticipated translational benefit, illustrated by dystrophin protein levels.

To date, the main attempt to evaluate the relative translational benefit from an animal species to human subjects has focused on the minimal levels of dystrophin protein required for functional stabilisation of dystrophic myofibres. Many factors need to be considered. This protective effect will depend not only on the amount of dystrophin protein within an individual myofibre, but also the extent of distribution within all myofibres, the size of the nuclear domain (how far dystrophin protein extends along the sarcolemma from the myonucleus where the mRNA is generated) and when during development the dystrophin must be produced. This situation was considered by Chamberlain (1997) who concluded, from analysis of mosaic transgenic mice and viral vector delivery with suboptimal doses into mdx mice, that >50% of myofibres need to express dystrophin, and that these must accumulate approximately 20% of wild-type levels of dystrophin for a significant correction of the muscle pathology in mice [31,32]. A low level of uniform dystrophin expression was more effective than high levels of dystrophin in a variable pattern. Studies in humans validate these mdx mice findings, since a similar level of dystrophin replacement (20–30% or more) and distribution in 60–70% of myofibres is significantly protective for human muscles, as indicated by studies of dystrophin levels in mildly affected patients and female DMD carriers [33–35]. These data show a high correlation between the benefits in mdx mice and humans with respect to dystrophin replacement indicating that a drug-induced improvement in dystrophin levels will have a high translational benefit. Indeed, equivalence in translational benefit is demonstrated by similar results with gentamicin upregulation of dystrophin in dystrophic mice [36] and humans [37]. This tight correlation may not be the case for other interventions that modify complex events downstream of the dystrophin deficiency, e.g. with drugs that target inflammation or oxidative stress to try and reduce the severity of the dystrophopathy.

Clearly access to human tissues to generate these data is more limited than for the mouse, although measurements using human blood samples are possible in some cases. The least invasive approach to estimate the anticipated translational benefits is the measure of some physiological functions. While these have not yet been explored, it would clearly be important to do so in order to understand what effect size is required in animal models in order to expect a clinically relevant effect in humans. For example, in the mdx mouse, a 15% increase in muscle strength is far from what is considered to be a high recovery score; however, if a therapy is able to produce the same magnitude of effect in humans, leading to maintained muscle function and hence quality of life for DMD boys, a clear benefit would be evident. What is critical to determine is whether a protective functional effect of 15% in the mdx mouse may actually have a far lesser benefit (say only 5%) in the larger muscles of DMD boys and therefore be of relatively little value (compared with the more comparable benefits between species for dystrophin). This anticipated translational benefit is the key issue to address for each parameter. Only where a sufficiently high translational efficacy is apparent are clinical benefits likely.

should be considered when evaluating the merit of this parameter as a useful pre-clinical outcome.

Conclusion

This article proposes that pre-clinical therapeutic studies in the mdx mouse should be conducted, similar to human clinical trials, in a more standardized and methodologically rigorous fashion in order to improve the potential for translation into therapeutic agents for patients with DMD. The Network of Excellence TREAT-NMD provides a platform for experts in the field and leading scientists to discuss and present their ideas and observations from pre-clinical research in mdx mice. The application of common experimental conditions and the use of SOPs in mdx mice should increase comparability of outcomes from different laboratories. This panel has sought agreement on several crucial issues that we believe are necessary if we are to accelerate the translation of potential new treatments from the bench to the clinic.

The primary recommendations for a more harmonized evaluation of pre-clinical therapeutic studies in the mdx mouse are:

- Consistency in animal care: careful attention to diet, care in reducing stress and human interventions to the minimum necessary and to do this at the same time each day.
- Careful design of the experimental schedule: choice of animal gender, adequate power, stratification and randomization of mice to control and treated groups, standard sampling times at 12 weeks (medium term) and 6 months/1 year (long term).
- Use of a standard set of core outcome measures to evaluate treatment response.
- Use of standardized protocols to assess the outcomes.
- Acceptance of common processes for the reporting of results; these should be based on strong data obtained with an appropriate number of samples for each specific
readout and should consider the recovery score (and ideally the translational benefit) in addition to the pure statistical significance.

Acknowledgements

Our thanks go to: Annemieke Aartsma-Rus, Thomas Meier, Stefanie Possekul, Markus Rüegg, Volker Straub for their valuable advice.

References