

# Nonclinical Studies that Support Viral Vector-Delivered Gene Therapies: An EFPIA Gene Therapy Working Group Perspective

Michael W. Bolt,<sup>1</sup> Laurence O. Whiteley,<sup>1</sup> Jessica L. Lynch,<sup>2</sup> Brian Lauritzen,<sup>3</sup> Antonio R. Fernández de Henestrosa,<sup>4</sup> Timothy MacLachlan,<sup>5</sup> Peter Ulrich,<sup>6</sup> Binu K. Philip,<sup>7</sup> Prathap Kumar Mahalingaiah,<sup>7</sup> Claudette L. Fuller,<sup>8</sup> and David R. Compton<sup>9</sup>

<sup>1</sup>Pfizer, Drug Safety Research and Development, Cambridge, MA 02139, USA; <sup>2</sup>Jansen Research and Development, LLC, Nonclinical Safety, Spring House, PA 19477, USA; <sup>3</sup>Novo Nordisk A/S, Global Discovery and Development Sciences, 2760 Maaloev, Denmark; <sup>4</sup>ESTEVE Pharmaceuticals, Department of Toxicology, 08028 Barcelona, Spain; <sup>5</sup>Novartis Institutes for Biomedical Research, Department of Preclinical Safety, Cambridge, MA 02139, USA; <sup>6</sup>Novartis Institutes for Biomedical Research, Safety Assessment Group, 4002 Basel, Switzerland; <sup>7</sup>Abbvie Preclinical Safety, North Chicago, IL 60064, USA; <sup>8</sup>Merck and Co., Safety Assessment and Laboratory Animal Resources, West Point, PA 19486, USA; <sup>9</sup>Sanofi US, Preclinical Safety Development Projects, Bridgewater, NJ 08807, USA

**Nonclinical development strategies for gene therapies are unique from other modalities. The European Federation of Pharmaceutical Industries and Associates (EFPIA) Gene Therapy Working Group surveyed EFPIA member and nonmember pharmaceutical and biotechnology companies about their current practices for designing and implementing nonclinical toxicology studies to support the development of viral vector-delivered *in vivo* gene therapies. Compiled responses from 17 companies indicated that these studies had some variability in species selection, study-design elements, biodistribution, immunogenicity or genomic insertion assessments, safety pharmacology, and regulatory interactions. Although there was some consistency in general practice, there were examples of extreme case-by-case differences. The responses and variability are discussed herein. Key development challenges were also identified. Results from this survey emphasize the importance for harmonization of regulatory guidelines for the development of gene-therapy products, while still allowing for case-by-case flexibility in nonclinical toxicology studies. However, the appropriate timing for a harmonized guidance, particularly with a platform that continues to rapidly evolve, remains in question.**

## INTRODUCTION

Gene therapy has recently emerged as a modality that the biopharmaceutical industry is incorporating into drug-development pipelines. The nonclinical development strategy for a viral vector-based gene therapy is unique and quite different from that of a small molecule (SM) or monoclonal antibody (mAb).<sup>1,2</sup> For example, nonclinical toxicology studies with SMs and mAbs typically involve repeat doses that are administered for as long as 9 or 6 months, respectively, but studies with a gene therapy involve a single dose, and animals are followed for a period of time that is determined on a case-by-case basis. Doses tested in nonclinical toxicology studies are often substantially higher (50×) than the predicted efficacious exposure (for SMs) or

10× the highest clinical exposure (for mAbs), but for viral vector-based gene therapies, doses are typically at or slightly above the clinically efficacious dose. How much higher than the clinically efficacious dose is determined on a case-by-case basis, often limited by transgene overexpression and influenced by multiple factors, such as tropism, transgene product potency, promoter strength, patient population, and other available therapies. Nonclinical studies with gene therapies typically include biodistribution endpoints, which are rarely included in SM or mAb studies.<sup>3</sup> Lastly, regulatory requirements for SMs and mAbs are well established and harmonized between various regions but are more flexible and evolving for gene therapies.

It is important to design nonclinical toxicity studies appropriately; otherwise, this could unnecessarily lead to issues with patient safety or prevent patients from receiving an optimal efficacious dose. However, sponsors are often challenged to minimize nonclinical study duration so that potentially life-saving medicines can be tested in patients sooner, as well as to consider opportunities to refine animal use. This is often accomplished by using a single species or single sex for a species (when appropriate) or by reducing the study duration, number of animals per group, dose, or number of dose groups (which may decrease the amount and the time to make the test article). Unfortunately, how much these parameters can be reduced for a viral vector-based gene therapy without impacting patient safety or regulatory acceptance is not clear.

Given that there are very few gene-therapy programs that have been approved for marketing and the unique development strategies,

---

Received 14 July 2020; accepted 27 August 2020;  
<https://doi.org/10.1016/j.omtm.2020.08.017>

**Correspondence:** Michael W. Bolt, PhD, Pfizer, Drug Safety Research and Development, 1 Portland St., Cambridge, MA 02139, USA.

**E-mail:** [michael.bolt@pfizer.com](mailto:michael.bolt@pfizer.com)



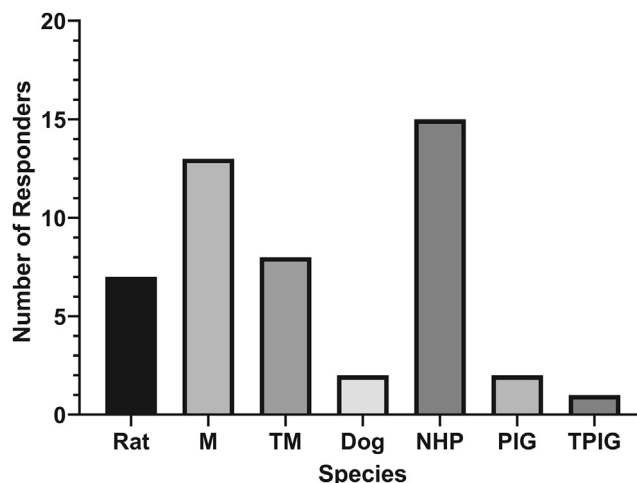
sponsors are less certain about what nonclinical studies would be acceptable to support clinical development of viral vector-based gene therapies. Although consultations with regulatory agencies are always a good option, this can slow down development, introduce differences in opinion/requirements between agencies, and potentially overwhelm agency reviewers. In addition, viral vectors have development challenges that are unique to this modality and continue to emerge with greater experience. With this background, the European Federation of Pharmaceutical Industries and Associates (EFPIA) Gene Therapy Working Group surveyed EFPIA member and nonmember companies that currently have gene-therapy programs in their research and development pipelines. The purpose of this survey was to gather information on current practices for nonclinical studies conducted with viral vector-delivered, *in vivo* gene-therapy products with the intent of determining best practices as well as identifying opportunities for harmonization. Results from this survey provide insight as to what studies biopharmaceutical/biotechnology companies are conducting, what endpoints they are evaluating, when and how they are evaluating them, and the key challenges they have been facing during the development process. These survey results may assist investigators when designing future nonclinical toxicology packages for viral vector-based gene therapies.

## RESULTS

### Company Information

Survey responses from 17 companies were provided. Of the 17 companies, 1/17 had a single gene therapy that was in the early stages of development, and so it could not respond to many of the questions. Responses from this company were included when appropriate. Of the 17 companies that responded, 8/17, 4/17, and 5/17 were considered large (greater than 10,000 employees), medium (1,000 to 10,000 employees), or small (less than 1,000 employees), respectively. Company-size categorization of small-, medium-, or large-sized companies was consistent with that used in other surveys.<sup>4</sup> There were 5/17 companies that considered themselves to be working for a biotechnology company (4/5 small and 1/5 middle sized), whereas 12/17 identified as working for a pharmaceutical company (1/12 small, 3/12 middle sized, and 8/12 large). Most (10/16) had 90% or 100% of their gene-therapy programs initiated in house (versus in-licensed), demonstrating that the survey responders had significant experience in gene therapy development. Only 2 companies had <50% of their gene therapy programs developed internally.

When asked, “How many gene therapy programs does your portfolio contain at each stage of development?” the response available to select in the survey was a number range (i.e., 1–2, 3–5, 6–10, 11–20, or >20 compounds) in order to help maintain anonymity of responses. All participating companies had assets (1–2 to 11–20) that were in the exploratory toxicity phase, and 7, 8, and 3 companies had 1–2 assets in the first-in-human (FIH)-enabling/good laboratory practice (GLP) toxicity phase, phase 1, or phase 2, respectively. In addition, 1 company had 1–2 assets in phase 3, and 1 company had 3–5 assets that were in the marketing phase. The therapeutic area for which the gene therapy assets were being developed included CNS disorders



**Figure 1. Responses to the Survey Question “Which Species Are Typically Used in Toxicity Studies? Select All That Apply”**

M, mouse; TM, transgenic mouse; NHP, nonhuman primate; PIG, pig or minipig; TPIG, transgenic pig or minipig. N = 16.

(12/17); liver-targeted diseases (9/17); lysosomal storage diseases (8/17); ophthalmology, blood disorders, or musculoskeletal diseases (5/17 each); or cardiac, oncology, respiratory, or anti-infective indications (1/17 or 2/17 each). When asked, “What percent of the gene therapy programs are for rare diseases?” 10/16 responded with a value  $\geq 75\%$ , 4/16 responded with a value of 50%, and 2/16 responded with 0%.

The most used vector to deliver the gene therapy was adeno-associated virus (AAV; 14/17), but a few companies listed lipid nanoparticles, naked plasmids, and/or nonviral DNA as their most used vector. In addition to AAVs, some companies were also using lentivirus or adenovirus to deliver their gene therapies. Most companies used cell-type-specific (14/17) and/or constitutive (13/17) promoters, with a minority using inducible promoters (3/17). The majority of companies considered promoter strength, as primarily defined by the level of expression of transgene, and to a lesser extent, infectivity and specificity, when choosing the promoter (13/15).

### Species Selection

Species typically used in toxicology studies are listed in [Figure 1](#). Selection of an animal species was based on permissiveness/susceptibility to infection with the viral vector, transduction profile, pharmacological response to the transgene, comparable physiology and anatomy to humans, immune tolerance to the gene therapy product, and feasibility of the delivery system/procedure. A single species was used for nonclinical toxicology studies by 7/16 companies, whereas 1/16 were not able to use a single species, and 8/16 had not tried to use a single species. The use of 1 species was justified because of the following: (1) the pharmacologic target organ was efficiently transduced; (2) there was already available information on vector safety; (3) it resulted in optimal human transgene expression, which

predicted a human response; and/or (4) the disease state was available in that species. In addition, some companies were limited to a single species because of pre-existing neutralizing antibodies (nAbs) in other species. Interestingly, 1 company responded that it never used 2 species for toxicity studies for a gene therapy program. In the absence of a pharmacologically active species, only 2/16 used surrogate genes/expression constructs, whereas 6/16 did not, and 8/16 companies replied not applicable (NA).

When questioned about animal age, 6/16 companies (evenly split between large and small companies) had conducted studies using sexually mature animals, involving rodents (3/6), nonhuman primates (NHPs) (2/6), or both rodent and NHPs (1/6). Sexually mature animals were included in the nonclinical safety program based on the following: the patient population (1/6), to allow for the inclusion of fertility endpoints (1/6), to evaluate semen clearance and assess germline transmission (2/6), to avoid regulatory misinterpretation of immaturity changes in peri-pubertal sex organs as target organs (1/6), or animal availability (1/6). When developing programs with patient populations less than 18 years of age, 7/16 companies used juvenile animals in their nonclinical safety program, 1/16 did not, and 8/16 responded NA. If it was not proposed by the sponsor, regulatory authorities requested the use of juvenile animals to support pediatric trials for some companies (3/16) but not all the time (3/16 companies were not requested; 10/16 responded NA).

Most companies (13/16) reported that prior to starting a study, they screen nonrodents for pre-existing antibodies (Abs) to vector capsid; primarily for nAbs (6/13), but some screened for both nAbs and binding Abs (bAbs; 4/13), whereas others solely screen for bAbs (3/13). In 11/13 companies, animals that had pre-existing nAbs would be excluded from study, and 2/13 responded NA because they only evaluated for bAbs. However, animals that had pre-existing bAbs were often excluded from study (6/13), whereas others were not (2/13); NA was the response for 5/13 companies. Companies that did not exclude animals with pre-existing Abs from the study did so because of the following: (1) Ab titers were below a cutoff value, (2) they were the animals with the lowest Ab titers, (3) bAbs aided uptake, and (4) to determine the relationship of pre-existing Abs to pharmacology/toxicology. Animals with titers that were included in studies were randomly assigned to groups or placed in the control group.

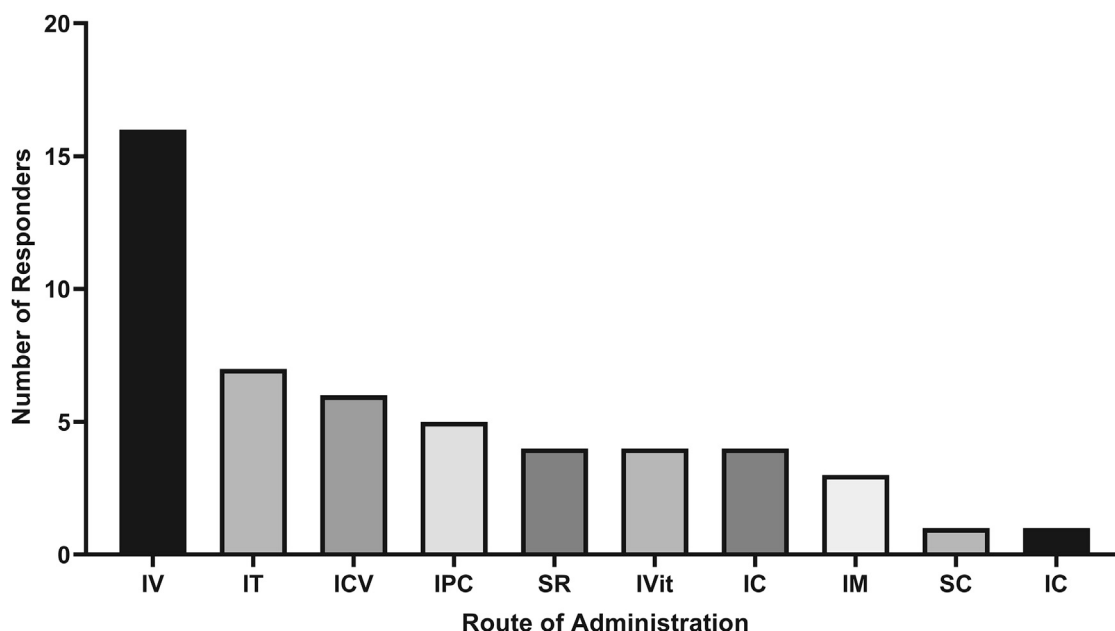
Immunosuppression was not a common practice, with 12/16 never immunosuppressing, 3/16 sometimes immunosuppressing, and 1/16 always immunosuppressing animals. Companies decided to immunosuppress because there was the following: (1) a potential for infusion reactions or transient immune responses but not anti-drug Ab responses or (2) a known acute/unexpected inflammatory response. One company stated that the use of NHPs was the only criteria for determining the need for immunosuppression. Immunosuppression was intermittent and for part of the study (3/4), or for the entire duration of the study (1/4). A few companies disclosed that they typically immunosuppressed animals with methylprednisolone/rituximab (10 mg/kg each) or prednisone/prednisolone (1 mg/kg orally or by intramuscular injection).

### Study Design Elements

Toxicity endpoints were included in efficacy studies always (7/17) or sometimes (8/17), whereas 1/17 responded NA, and 1/17 did not respond. By adding toxicity or biodistribution endpoints to efficacy studies, 5/16 and 5/16, respectively, were able to use that data as substitutes for conducting a stand-alone toxicity study in at least one case. In the case where the vector serotype and promoter had previously been evaluated (with a different transgene), a complete/full toxicology package with the new transgene (and same vector and promoter as before) was conducted all the time (7/16) or usually (3/16), with 5/16 responding NA. Interestingly, 1 company responded that it would never conduct a complete/full toxicology package in this situation. The routes of administration that companies were using are listed in [Figure 2](#).

The default duration for FIH-enabling toxicity studies with a gene therapy was 3 months (8/17 rodent, 6/17 nonrodent) or 6 months (5/17 rodent, 7/17 nonrodent), whereas 3/17 did not respond to the question, and 1/17 responded NA. However, it should be noted that an FIH-enabling study as short as 15 days in nonrodents was reported by at least 1 company. Interestingly, study durations were typically shorter for biotechnology companies when compared to pharmaceutical companies (6-month studies were default for 1/5 rodent and 1/7 nonrodent studies for biotechnology companies, versus 6-month studies were a default duration for 4/5 rodent and 6/7 nonrodent studies for pharmaceutical companies). Not counting for control groups, companies responded that FIH-enabling toxicity studies generally included 2 to 3 dose groups. Each group typically included 10 animals/sex/group (ranging from 5 to 20 animals/sex/group) for rodent studies and 3 animals/sex/group (ranging from 3 to 6 animals/sex/group) for nonrodent studies, generally with a comparable group size in cases where studies involved only one sex. Typically, a single dose was administered to the animals in toxicity studies (10/17), but some companies (3/17) had experience with administering more than 1 dose systemically (intravenous) or locally; 4/17 responded NA. Criteria used to select the high dose in a toxicity study are listed in [Figure 3](#).

Necropsy was conducted at a single time point for 7/17 companies (4/7, 2/7, and 1/7 at an intermediate, late, or unspecified time point, respectively) or at multiple time points for 7/17 companies (2/7 at an early and late time point and 5/7 at an early, intermediate, and late time point), whereas 3/17 companies did not respond. Early time points ranged between 3 days to 3 months (7/14), intermediate time points ranged between 1 and 3 months (9/14), and late time points ranged between 3 and 12 months (9/14 responders). Criteria used to select the timing for an interim necropsy included the following: to evaluate for acute toxicity, to measure tissue vector levels over time (including identification of time-to-peak transgene expression), to assess for anti-drug Abs, and/or to evaluate for biodistribution. By comparison, the timing of the terminal necropsy was selected based on the duration(s) needed to reach post-peak and stable transgene expression, to establish clearance of transgene and vector, to show toxicity and/or to show reversal/lack of



**Figure 2. Responses to “What Route of Exposure Do You Typically Use? Select All That Apply”**

IV, intravenous; IT, intrathecal; ICV, intracerebroventricular; IPC, intraparenchymal; SR, intraocular subretinal; IVit, intraocular intravitreal; IC, intracisternal; IM, intramuscular; SC, subcutaneous; IC, intracardiac. N = 17.

progression of effects seen at earlier time points based on feedback from regulatory authorities, or to evaluate for anti-drug Ab formation.

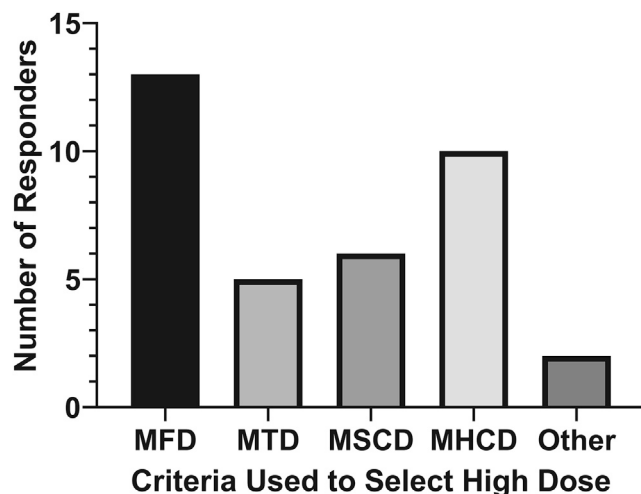
Companies responded that all (2/17), or most/some (11/17) parameters in FIH-enabling studies were conducted under GLP conditions. Parameters reported to be evaluated under non-GLP conditions included biodistribution, nonstandard clinical pathology biomarkers, immunohistochemical analyses, Ab assays (particularly those for pre-screening for nAbs), enzyme-linked immunosorbent spot (ELISpot) for cellular immune response, optical coherence tomography, or electroretinography. In addition, if toxicity were evaluated in more than 1 species, then a toxicity study in 1 of the 2 species may have been conducted under non-GLP conditions.

### Biodistribution

Biodistribution studies were conducted by 16/17 companies (8/16 large, 3/16 medium, and 5/16 small); 1/17 companies did not respond. Most companies (15/16) reported conducting biodistribution studies for AAV vectors. However, there were also cases of other viral vectors, such as adenovirus (2/16), lentivirus (1/16), or other nonviral vectors, such as lipid nanoparticles (2/16). Slightly more than one-half of the companies (9/16) noted doing exploratory biodistribution studies to support species selection for FIH-enabling toxicology studies, and this was generally an even split among small, medium, and large companies. Whereas 3/16 companies (1 large, 1 medium, and 1 small) stated they typically conduct a stand-alone biodistribution study, 9/16 companies only evaluate biodistribution as an endpoint in a toxicity study. The remaining 4/16 companies (all large) noted a

mix between doing stand-alone biodistribution studies and including biodistribution as an endpoint in a pharmacology/toxicology study, with 1 company pointing out that it will include biodistribution as an endpoint for all evaluations going forward. The companies conducting stand-alone studies responded that they usually use 2 dose groups.

A variety of species have been used for evaluating biodistribution, including, most often, mice (12/16) and monkeys (11/16), with rats (6/16), transgenic mice (5/16), dogs (2/16), and minipig (3/16) being used to a lesser extent. No company noted any use of transgenic rats or minipigs or rabbits for biodistribution studies. From the 15 companies that responded to have used rodents, biodistribution evaluations were typically conducted in 4–5 animals per sex/group/time point, with smaller companies trending toward the higher numbers. By comparison, typically 3 animals/sex/group/time point were used to evaluate for biodistribution by the 13 companies that responded to have used nonrodents for these analyses. However, 1–2 companies reported using one more or one less nonrodent per group. The timing of biodistribution evaluations ranged from day 1 to 365 days post-dose. Seven of 16 companies evaluated at an early, intermediate, and late time point, 1/16 evaluated at an early and intermediate time point, 3/16 evaluated at an early and late time point, 3/16 evaluated at a single time point, and 2/16 stated they collected at an early, intermediate, late, and only time point but did not specify which applied. An early time point ranged between 24 h and 3 months post-dose, with an average of 30 days and median of 14 days; an intermediate time point ranged between 7 days and 3 months; and a late time point ranged between 1 month and 12 months, with an average



**Figure 3. Responses to “In a FIH-Enabling Toxicity Study, What Criteria Do You Use to Select the High Dose? Select All That Apply”**

Multiples were typically 2–10× the highest clinical dose or up to 30× the clinical starting dose. MFD, maximum feasible dose; MTD, maximum tolerated dose; MSCD, multiple of starting clinical dose; MHCD, multiple of highest clinical dose; Other, as high as possible or as high as limited by product concentration. N = 14.

and median of 4 months. When evaluating for biodistribution, companies analyzed between 5 and 22 tissues (14/17; 3/17 did not respond), with 4/14 collecting 5–9 tissues, 4/14 collecting 10 tissues, 4/14 collecting 12–15 tissues, and 2/14 collecting  $\geq 20$  tissues. Criteria used to select tissues included regulatory guidance, the route of administration, the tropism of the vector, organs that enable assessments of shedding/germline transmission, or organs of concern based on known toxicity, transgene expression, efficacy, or the promoter used. One company responded that it collects “a full tissue list.” When evaluating for biodistribution, 14/16 are most often measuring DNA, and 10/16 are most often measuring mRNA in whole tissue. However, responders also evaluated for biodistribution by measuring for mRNA by *in situ* hybridization (2/16 often, 2/16 sometimes, and 5/16 rarely), the transgene protein (5/16 often, 4/16 sometimes, 2/16 rarely), or a surrogate marker (2/16 sometimes). The presence of vector in target tissues was most often measured by quantitative polymerase chain reaction (PCR) but also by droplet digital (dd)PCR, reverse transcriptase (RT)-PCR, and *in situ* hybridization (Table 1).

When using a capsid and promoter that have been previously tested (with a different transgene) and assuming the same route of exposure, 7/17 companies would always repeat biodistribution evaluations with the new transgene (and same capsid and promoter), 5/17 would sometimes repeat biodistribution evaluations, and 1/17 would never repeat biodistribution evaluations; 4/17 did not respond. Among the companies that responded that they sometimes or never repeat biodistribution studies in this situation, the rationale for leveraging the prior data included the belief that the capsid and route of administration drive distribution (and further, some would not repeat biodistribution if the promoter was changed). Consistent with the use of

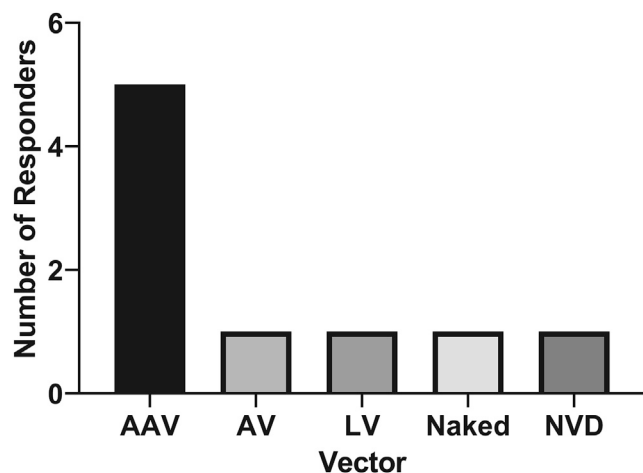
previously obtained data, 5/17 companies have, in at least 1 case, been able to cross-reference previously conducted studies with the same vector and promoter (but with a different transgene) to support not conducting a biodistribution study with the new transgene, 1/17 was not able to do so, 10/17 responded NA, and 1/17 did not respond. Lastly, 7/17 companies stated that they did not need to conduct additional biodistribution studies after FIH trials had initiated, whereas 9/17 responded NA, and 1/17 did not respond.

#### Genomic Insertion Assessments

There were 7/17 companies that have conducted insertional studies, 3/17 have not, 6/17 responded NA, and 1/17 did not respond. The majority (5/7) of insertional assessments were conducted with AAVs (Figure 4), which is consistent with the fact that AAV was the most commonly used vector by companies (14/17) responding to this survey. Insertional assessments were conducted based on a request from a regulatory agency (2/7), because companies routinely conducted insertional assessment studies for gene-therapy programs (2/7) or because they are not using an AAV (2/7 that used lentivirus); 1/7 did not state why it conducted them. One (1/2) company that was requested by a regulatory agency to conduct these studies also mentioned that this request was based on AAV presence in the gonads. Integration sites were evaluated/will be evaluated by linear amplification-mediated (LAM)-PCR (5/7 companies) or target-enrichment sequencing (TES) (1/7), and 1/7 did not disclose the method chosen. A response from one (1/7) company also suggested that it used a tiered approach evaluating target tissue samples for vector copy number by qPCR and integration site analysis by LAM-PCR and then subsequent additional evaluation to characterize integration profile and clonal dominance potential by deep sequencing. All companies that responded (3/3) stated that the integration assessment method was acceptable to regulatory authorities, although in one instance, the US Food and Drug Administration (FDA) requested details on method validation (for a new method that was based on the LAM-PCR technique with various modifications/enhancements). When asked if the generated insertional data were useful for risk assessment, 4 companies responded: 3/4 companies responded “yes,” whereas 1/4 responded “no.” The insertional data were not considered useful because they confirmed the low risk for integration already known in the literature. For the companies that responded “yes,” 1/3 found the insertional data useful because they confirmed low risk, 1/3 commented that the study was ongoing, and 1/3 did not respond.

#### Immunogenicity

Most companies (16/17) stated they evaluated immunogenicity for gene therapy programs; 1/17 did not respond. Vectors used when immunogenicity was assessed included AAV (13/16), adenovirus (2/16), lentivirus (1/16), naked DNA/plasmid (1/16), or nonviral DNA (1/16). Immunogenicity was usually assessed as an endpoint on a toxicity study (14/16), whereas 1/16 stated it evaluated for immunogenicity in both stand-alone studies and as an endpoint in a toxicity study; 1/16 did not respond. Immunogenicity evaluations were conducted under non-GLP conditions (7/16) versus GLP-validated assessments (6/16), whereas 1/16 responded NA, and 2/16 did not



**Figure 4. Responses to "For What Vector Types Have You Evaluated for Insertional Mutagenesis/Integration/Insertion? Select All That Apply"**  
AAV, adeno-associated virus; AV, adenovirus; LV, lentivirus; Naked, naked/plasma; NVD, nonviral DNA. N = 7.

respond. Non-GLP immunogenicity evaluations were used with AAV vectors (7/7), whereas GLP evaluations were used with AAV (5/6), lentivirus (1/6), adenovirus (1/6), and nonviral DNA (1/6) vectors. The type of immunogenicity assays used included Ab detection (15/16), cell-mediated response (11/16), and CD8<sup>+</sup> expression (1/16), and 1/16 did not respond. The transgene product (13/16) and vector capsid (11/16) were the antigens that most respondents said they assayed for, at least sometimes (2/16 did not respond). Immunogenicity resulted in an adverse event always (1/16), sometimes (3/16), or never (3/16); 5/16 responded that adversity was not determined, 2/16 did not respond, and 2/16 selected other. They selected other because a decrease in transgene protein activity might have been observed or because they saw signs of inflammation, but it was not considered adverse. One company stated that it selected the "not determined" option because it was not clear if the effects observed were secondary to a cytotoxic T lymphocyte response. For the companies that responded always, sometimes, or other (6/16), the adverse immune response was Ab mediated (2/6), cell mediated (1/6), or not determined (3/6), and it affected transgene expression or biological activity for 3/6 companies. Findings consistent with cytokine release syndrome were not observed after dosing a gene therapy (12/16), and 4/16 responded NA.

### Safety Pharmacology

Safety pharmacology endpoints were routinely integrated into general toxicity studies by 12/17 companies, 4/17 did not include safety pharmacology endpoints; 1/17 did not respond. Safety pharmacology endpoints routinely evaluated involved cardiovascular (9/12), respiratory (5/12), central nervous system (8/12), and/or body temperature (1/12) effects; 2/12 did not respond. Responders included safety pharmacology endpoints based on the biology or route of the transgene (for example, CNS evaluations for therapies delivered to the brain). One company responded that it always conducts cardiovascular evaluations

**Table 1. How the Presence of a Vector in Target Tissues Is Assessed during Biodistribution Evaluations**

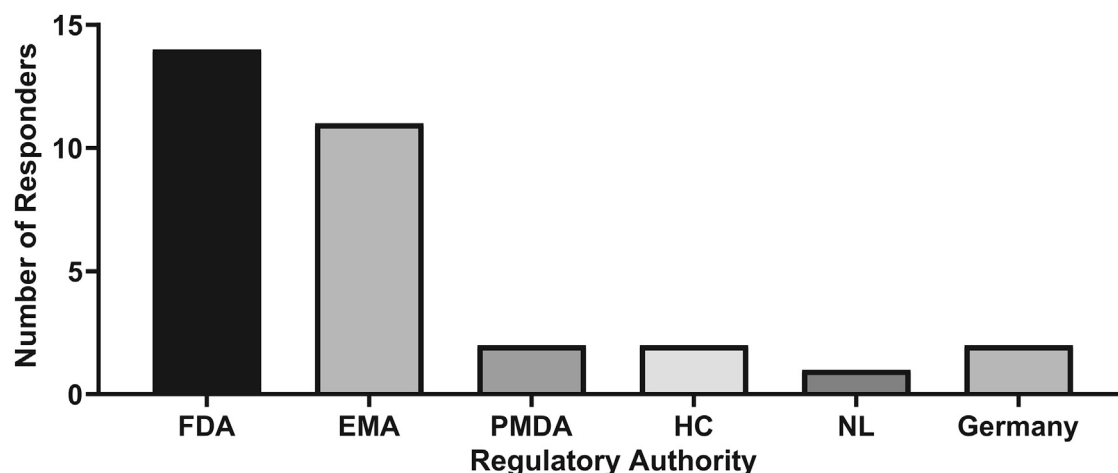
	Most Often	Sometimes	Rarely	Never	No Response
ISH	3	3	2	3	6
RT-PCR	5	2	1	3	6
ddPCR	5	1	1	5	5
qPCR	13	1	0	0	3

ISH, *in situ* hybridization; PCR, polymerase chain reaction; RT-PCR, reverse transcription PCR; ddPCR, droplet digital PCR; qPCR, quantitative PCR.

for a systemic study. Companies that do not routinely integrate safety pharmacology endpoints into general toxicity studies also do not conduct a stand-alone safety pharmacology study (4/4), because it did not make scientific sense, mice were used as toxicity species (making cardiovascular assessments difficult), or the FDA had specifically indicated that safety pharmacology studies were not needed for their program. These arguments were acceptable to regulatory authorities in most (3/4) cases.

### Regulatory Interactions

There were 14/17 companies that had regulatory interactions for their gene therapy products with the FDA, the European Medicines Agency (EMA), the Pharmaceuticals and Medical Devices Agency (PMDA) in Japan, Health Canada, the Medicines Evaluation Board in the Netherlands, and/or the Paul Ehrlich Institute (PEI) in Germany (see Figure 5). The majority of companies have had interactions with the FDA (14/17) or EMA (11/17). In addition, 9/17 companies have had interactions with both the FDA and EMA, and 2/17 companies have had interactions with the FDA, EMA, and PMDA. There were 3/17 companies with no interactions with regulatory agencies. There were 7/17 companies that sought regulatory feedback from the same agency for more than one program in the same therapeutic area, of which 5/7 had received consistent feedback from the same agency, and 2/7 had not. The differences in opinion between regulators were on the need for insertional assessments or feedback on study-design elements. Similarly, there were 10/17 companies that had received feedback from more than 1 agency for the same gene therapy program, of which 6/10 companies had received consistent feedback across the various agencies, and 4/10 companies had not. Differences in opinion on the number of species and interim sacrifices and/or the need for pediatric studies and the age of animals in those studies occurred among the FDA, EMA, PMDA, and/or PEI. There were 15/17 companies that think that now is the right time for a guidance on gene therapy from the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), whereas 2/17 did not. Other challenges that sponsors are facing with the development of a gene therapy are listed in Table 2, some of which include chemistry, manufacturing, and control (CMC) issues. For example, 1/17 companies had to conduct additional toxicity studies to show comparability between different lots, whereas 7/17 did not, 7/17 responded NA, and 2/17 did not respond.



**Figure 5. Responses to the Question “Which Regulatory Agencies Have You Interacted with for Your Gene Therapy Programs? Select All That Apply”**

FDA, US Food and Drug Administration; EMA, European Medicines Agency; PMDA, Pharmaceuticals and Medical Devices Agency in Japan; HC, Health Canada; NL, Medicines Evaluation Board in the Netherlands; Germany, the Paul Ehrlich Institute. N = 14.

## DISCUSSION

Survey responses were from 17 biotechnology or pharmaceutical companies that had <1,000 to >10,000 employees, with assets being developed for a wide range of indications. Although responses were limited to only 17 companies, it should be noted that at the time of the survey, these 17 companies collectively had >92 gene-therapy assets in various stages of development. Furthermore, given that gene therapies are a relatively novel platform, it is often difficult to get companies to participate in surveys, because it may reveal information that previously offered them a competitive advantage. Another limitation of this survey is that for any given question, the number of responses was often <17, and the number of responses varied between questions. This was often because the responder was unable (because the company did not have the relevant experience with their molecules) or unwilling to provide an answer. Regardless, companies responding to the survey (and the gene therapy assets they are developing) were considered a good representation of the biopharmaceutical industry as experts in gene therapy development.

Study design elements (such as species selection and study duration) are critical for evaluating the potential for toxicity for the test article. When setting these parameters, factors, such as pharmacological relevance, permissiveness of the species to the transgene, time to expected peak transgene expression, predicted target organs, frequency and dose-responsive nature of lesions, potential for immune response, identification of dose that causes no effect and/or no adverse effects, minimization of animal use, and regulatory requirements,<sup>5,6</sup> are all considered. Given the complex interplay among these various factors, it was not surprising that the survey results for study design elements demonstrated variability in responses. For example, although monkeys and mice were the most commonly used species, several other species were also used for toxicity studies with gene therapy products. This is likely because *in vivo* animal models need to be permissive for transduction with the

viral serotype being used and have anatomy that is representative of that in humans.<sup>7</sup> This was the case for species selection for Luxturna (AAV2-hRPE65v2), where nonclinical toxicity studies were conducted in dogs (a pharmacologically relevant species) as well as NHP (a more appropriate species for toxicology studies for anatomical and immunological reasons).<sup>8</sup> Furthermore, the selected species should allow for an evaluation of known potential risks associated with the vector. For example, dorsal root ganglia toxicity has been observed with AAV9, AAVh68, and/or AAV1 in NHPs and/or piglets,<sup>9–11</sup> so use of one of these species may be preferred when using these AAV serotypes. However, it should be noted that a detailed comparison of susceptibility to dorsal root ganglia toxicity in other relevant species used in toxicology studies (e.g., rat and dog) has not been reported, so species other than NHP may be appropriate. Finally, it should be noted that the use of two species in a nonclinical toxicology package is not always necessary, as confirmed by the 7/16 companies that responded that they used a single species. However, appropriate scientific justification for species selection should be provided in the regulatory submission documents. Study duration (typically 3 to 6 months) was also a parameter that had high case-by-case variability. However, with the assumption that peak transgene expression has occurred, and the potential for toxicity (and reversibility) has been established, one should consider, on a case-by-case basis, if toxicity studies longer than 3 months in duration are of value, especially since long-term follow-up (up to 15 years in some cases) is occurring in the clinical setting.

Biodistribution can be evaluated in stand-alone studies or as an endpoint in a toxicity study. Biodistribution analysis coupled to toxicology studies offers the advantage of allowing the ability to match biodistribution results with histopathology data in the case of potential findings. The assessment of both biodistribution and toxicology data from the same animal would also be in alignment with the 3R (replace, reuse, refine) principles. On the other hand, current

**Table 2. Key Challenges When Developing a Gene Therapy**

Study Elements
● Are long-term (i.e., 6-month studies) toxicity studies always necessary?
● What is the best method to select your doses for toxicity studies?
● What is the best way to conduct toxicity studies when you have immunogenicity?
● How many tissues need to be collected for biodistribution analyses?
● Are biodistribution assessments needed if a construct with the same capsid and promoter (but different transgene) has already been tested?
● Are integration assessments necessary for low/nonintegrating vectors if they are only confirming low risk already known in the literature?
Data Interpretation
● How do biodistribution/expression/adverse events in animals translate to humans?
● Lack of historical data
Variability of Production (i.e., CMC Characteristics) between In-House Methods, between Sites within a Company, between Company and Commercial Manufacturer, and between Lots by the Commercial Manufacturer
● What <i>in vitro</i> comparisons are sufficient to show comparability?
● How different does a product need to be to warrant additional toxicity testing?
● When necessary, what studies are needed to qualify differences?
Late-Stage Development
● What studies are needed to support women of childbearing potential?
Regulatory Considerations
● Lack of regulatory consensus between agencies
● Lack of clear guidance on how to do late-stage development (developmental and reproductive toxicology [DART] and carcinogenicity strategy)

guidance and publications do not specify particular time point samplings, only to analyze biodistribution at the expected time of gene-therapy product-peak expression and at several later time points to evaluate its clearance from tissues, which likely explains the high variability (early, intermediate, and/or late time points ranging from 1 to 365 days post-dose) observed in the survey responses for this parameter. However, it should be noted that the lack of rigid guidance around the timing of biodistribution evaluations provides flexibility in study design elements that are influenced by the route of administration, specificity of a gene therapy vector to a tissue, and dose levels. The selection of tissues to evaluate for biodistribution of the gene-therapy product is a key parameter for biodistribution assessments. Again, except for a list of nine key tissues outlined in FDA guidance and the International Pharmaceutical Regulators Programme (IPRP) publication (blood, injection site(s), gonads, brain, liver, kidneys, lung, heart, and spleen),<sup>12,13</sup> guidance (such as other tissues depending on product, route of administration, and tropism) is limited, which likely explains the high variability in the number of tissues collected by companies responding to this survey (5–22 tissues). In this survey, qPCR was the method used by most companies for determining presence of the transgene, likely because of its high level of sensitivity (a few copies of the recombinant nucleic acid can be detected) and specificity (amplifying the studied transgene). Although this method is acceptable to regulatory authorities,

one must be diligent to ensure that tissue samples used in the PCR amplification are free from cross-contamination during necropsy collection and sample processing. A potential opportunity raised by some companies developing viral vector-based gene therapies is the need to continue to assess for biodistribution of a test article in all previously tested tissues when the capsid and promoter are the same, and only the transgene has been changed. Perhaps future experience, accumulation of data, and advances in the field in gene therapy will favor limiting the number of tissues to be analyzed for biodistribution. If there was a concern about differential biodistribution, some companies suggested that a very limited distribution analysis could be included as part of the FIH-enabling toxicology study instead of repeating biodistribution evaluations in all tissues with a gene therapy with a previously tested capsid and promoter.

AAV-based vectors are considered to have relatively low risk for insertional mutagenesis because they predominantly remain in episomal form (nonintegrating) and rarely get integrated into host cell DNA to cause an adverse effect, such as hepatocellular carcinoma observed in neonatal mice administered an AAV2 vector expressing human B-glucuronidase gene from a B-actin promoter and cytomegalovirus (CMV) enhancer (AAV-GUSB).<sup>14,15</sup> However, results from this survey demonstrate that insertional assessments for AAV-based gene therapies are being conducted by some, but not all, companies, which suggests that there are some circumstances where a literature review (instead of conducting studies) is adequate. For companies that were conducting integration assessments, most were using LAM-PCR methodology, which is consistent with available guidance.<sup>16</sup> Interestingly, results from this survey indicate that there may be limited value in conducting insertional assessments for AAV-based gene therapies, because the studies are only confirming the known low risk for tumorigenicity, but this was based on a low number of responding companies (n = 4).

Immune responses to a gene therapy can have many different effects, such as preventing transgene expression (nAbs), increasing liver transduction (bAbs),<sup>17</sup> or causing toxicity via complement activation<sup>18</sup> or inflammation.<sup>19</sup> In this survey, most companies screened animals prior to dosing for pre-existing nAbs, but fewer companies screened for pre-existing bAbs. It should be noted that the survey did not ask what the titer cutoff was for including animals in the study. Given that the impact on bAbs on transgene expression is not fully understood, sponsors may wish to evaluate bAbs in future studies. The majority of companies were not immunosuppressing animals, which is likely a reflection of the desire not to have immunosuppressed subjects in clinical trials (unless absolutely necessary). However, should immunosuppression be necessary in nonclinical studies, the selection of which type of immunosuppressant should generally be considered in close collaboration with clinical colleagues so that the type and duration of immunosuppression are compatible with the proposed clinical regimen.

This survey identified that there were some differences in interpretation among regulatory authorities. As such, most of the companies



**Table 3. Concepts for Consideration If/When an ICH Guidance on Gene Therapy Is Prepared**

- Generally, a 3-month study may be sufficient to support registration, unless a risk-based approach (based on product properties such as peak expression, target organ concerns) indicates longer duration studies are necessary. Long-term follow-up would occur in the clinic.
- Literature-based evaluations on the risk for insertional mutagenesis should be acceptable, under some circumstances. It should be assumed that with AAVs, there is a low risk for integration, and the need to confirm the risk for insertional mutagenesis with product-specific insertional assessment should be case-by-case and not default. Although product attributes or patient population may define the need for further evaluations, health authorities and sponsors should align on criteria for when a literature assessment only would be acceptable.
- Previously determined biodistribution data in the literature or with other vectors of the same capsid serotype could be used to justify minimal or no additional biodistribution evaluations with constructs where a well-characterized vector has not changed (only the transgene has changed), and route of administration and dose levels are the same. Limited additional assessments for expression may need to be conducted when new promoters are used.
- Evaluations for shedding do not always need to occur in all matrixes (e.g., urine, feces, saliva, semen, etc.). Regulators and sponsors should align on appropriate matrixes and the need to gather these data, on a program-specific basis.
- Guidance on what nonclinical studies are necessary to support repeat dosing, dosing patients with pre-existing antibodies, dosing women of childbearing potential, or pediatric populations should be established. General principles of ICH S6<sup>20</sup> should be considered and exceptions noted where appropriate for gene and cell therapy.

participating in this survey responded that now would be the right time for an ICH guidance on gene therapy. An ICH guidance for the development of gene-therapy products may provide clarity and consistency for sponsors on issues, such as those listed in Table 2. Should such a guidance be written, results from this survey have led to the creation of potential recommendations for consideration (see Table 3). Whereas an ICH guidance is typically considered of value, some companies thought that because the field is evolving so rapidly, it may be more appropriate to wait until more is known about developing gene therapy products before creating an ICH guideline. This may prevent the need to revise the document multiple times. Other companies caution that based on the complexity of the platform, a unifying guidance may not be appropriate at this time because it may be too restrictive. This is particularly true when considering requirements for studies to support life-threatening indications with unmet medical need and how they compare to those for gene therapy products for which there is already a non-gene therapy treatment available for patients. Regardless, this working group recommends that any harmonized guidance be forward thinking so that it would be applicable to next-generation approaches.

## MATERIALS AND METHODS

The EFPIA Gene Therapy Working Group created 111 questions designed to gather information about the following: (1) the company responding and its extent of experience with using gene therapies, (2) general information on nonclinical study design elements and species selection, (3) biodistribution endpoints, (4) genomic insertional assessments, (5) immunosuppression and immunogenicity, (6) safety

pharmacology and reproductive toxicology studies, and (7) regulatory interactions.

For many of the questions, the responders were asked to select from a list of predetermined responses, but the option for the response of “other (please specify)” was usually included in these situations. In addition, for some questions, more than one response could be selected, if appropriate. Furthermore, some questions also gave the option for the selection of NA, but the reason for selecting NA was not asked for or suggested. However, sometimes a company chose not to respond to a question.

The number of questions asked was dependent on the response, so not all companies responded to all 111 questions (for example, if the responder answered “no” to a question about having conducted reproductive toxicity studies, then the responder was not asked the follow-up questions about reproductive toxicity studies). Since the number of companies responding varied from question to question, data were calculated, interpreted, and presented as a fraction of the number of companies responding to a specific question. Survey questions generally pertained to vector-delivered *in vivo* gene-therapy programs, although responses from a small number of companies that used lipid nanoparticles, naked plasmids, and/or nonviral DNA were included. The survey did not apply to gene editing, oligonucleotide therapeutics, modified oncolytic viruses, or *ex vivo* genetically modified human cells.

The survey was distributed electronically by a third party (the IQ DruSafe secretariat) to the survey champion at 34 different companies. The survey champion had to work actively for a pharmacology/biotechnology company (as opposed to a consulting company or contract research organization), which worked on at least 1 gene-therapy program. Companies had 3 weeks to respond to the survey, after which, the survey was closed. Electronic responses were provided back to the IQ DruSafe secretariat, who provided the compiled responses to the EFPIA Gene Therapy Working Group. The secretariat ensured that responses were only provided by the point of contact for each company, that only one set of responses was submitted per company, and was the only person to know the identity of the responder to prevent correlating a response to an individual company. This arrangement also allowed the EFPIA Working Group to request the secretariat to ask the responder any follow-up questions. Although the EFPIA Gene Therapy Working Group did not know the identity of the responders, it was provided with coded individual responses (e.g., Company A, Company B, etc.) so that answers from an individual company could be followed throughout the survey. The secretariat was the only person to know which company was associated with the coded individual response.

The survey was executed using Survey Gizmo, tables were generated using Microsoft Word, and figures were generated using GraphPad Prism software.

The EFPIA Gene Therapy Working Group subsequently evaluated the responses to identify industry trends and challenges

associated with developing viral vector-delivered *in vivo* gene therapy products.

#### AUTHOR CONTRIBUTIONS

All authors were involved in creating the survey questions, analyzing the survey responses, and authoring the journal article.

#### CONFLICTS OF INTEREST

The opinions expressed by the authors of this article do not necessarily reflect the opinions of the companies for which they currently work. M.W.B. and L.O.W. are paid employees of Pfizer. J.L. is a paid employee of Jansen Research. B.L. is a paid employee of Novo Nordisk. A.R.F.d.H. is a paid employee of ESTEVE Pharmaceuticals. T.M.L. and P.U. are paid employees of Novartis. B.P. and P.K.M. are paid employees of Abbvie. C.L.F. is a paid employee of Merck. D.R.C. is a paid employee of Sanofi. As such, these authors have stock interests in the companies for which they work.

#### ACKNOWLEDGMENTS

All animal studies were conducted in accordance the Institutional Animal Care and Use Committee guidelines. The authors would like to thank the EFPIA Preclinical Development Expert Group and other members of the EFPIA Gene Therapy Working Group (especially Giovanna Rizzetto, Philippe Detilleux, Lutz Mueller, Rajni Fagg, Roberto Nitch, Alison Wolfreys, Nancy Claude, Mick Fellows, Joshua Gamse, Manuela Braun, Cristina Rocha, Andreas Rothfus, Markus Stephan-Gueldner, and Gianni Dal Negro) for their guidance, as well as IQ DruSafe (especially Anne Marie Stanley and Maja Leah Marshall) for their assistance with conducting the survey.

#### REFERENCES

- Elsallab, M., Bravery, C.A., Kurtz, A., and Abou-El-Enin, M. (2020). Mitigating Deficiencies in Evidence during Regulatory Assessments of Advanced Therapies: A Comparative Study with Other Biologicals. *Mol. Ther. Methods Clin. Dev.* 18, 269–279.
- Vestergaard, H.T., D'Apote, L., Schneider, C.K., and Herberts, C. (2013). The evolution of nonclinical regulatory science: advanced therapy medicinal products as a paradigm. *Mol. Ther.* 21, 1644–1648.
- Silva Lima, B., and Videira, M.A. (2018). Toxicology and Biodistribution: The Clinical Value of Animal Biodistribution Studies. *Mol. Ther. Methods Clin. Dev.* 8, 183–197.
- Freifeld, L. (2019). 2019 Training Industry Report. *Training* 56, 18–31.
- U.S. Department of Health and Human Services, Food and Drug Administration, Center for Biologics Evaluation and Research. Guidance for Industry: Preclinical Assessment of Investigational Cellular and Gene Therapy Products. November 2013. <https://www.fda.gov/media/87564/download>.
- European Medicines Agency, Science Medicines Health, Committee for Advanced Therapies (CAT). Guideline on the quality, non-clinical and clinical aspects of gene therapy medicinal products. March 22, 2018, EMA/CAT/80183/2014. [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-quality-non-clinical-clinical-aspects-gene-therapy-medicinal-products\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-quality-non-clinical-clinical-aspects-gene-therapy-medicinal-products_en.pdf).
- MacLachlan, T. (2013). Turning the corner with viral-based gene therapy-development of the rogue biopharmaceutical. In *Nonclinical Development of Novel Biologicals, Biosimilars, Vaccines and Specialty Biologicals*, L.M. Plitnich and D.J. Herzyk, eds. (Elsevier), pp. 35–64.
- US Food and Drug Administration. FDA Advisory Committee Briefing Document, Spark Therapeutics, Inc., Luxturna (voretigene neparvovec). October 12, 2017. <https://www.fda.gov/media/108385/download>.
- Hordeaux, J., Hinderer, C., Goode, T., Katz, N., Buza, E.L., Bell, P., Calcedo, R., Richman, L.K., and Wilson, J.M. (2018). Toxicology Study of Intra-Cisterna Magna Adeno-Associated Virus 9 Expressing Human Alpha-L-Iduronidase in Rhesus Macaques. *Mol. Ther. Methods Clin. Dev.* 10, 79–88.
- Hinderer, C., Katz, N., Buza, E.L., Dyer, C., Goode, T., Bell, P., Richman, L.K., and Wilson, J.M. (2018). Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose Intravenous Administration of an Adeno-Associated Virus Vector Expressing Human SMN. *Hum. Gene Ther.* 29, 285–298.
- United States Securities and Exchange Commission. Form S-1 Registration Statement Under the Securities Act of 1933. February 3, 2020. <https://www.sec.gov/Archives/edgar/data/1787297/000104746920000652/a2240560zs-1.htm>.
- International Pharmaceutical Regulators Programme. Expectations for Biodistribution (BD) Assessments for Gene Therapy (GT) Products. April 12, 2018. [https://admin.iprp.global/sites/default/files/2018-09/IPRP\\_GTWG\\_ReflectionPaper\\_BD\\_Final\\_2018\\_0713.pdf](https://admin.iprp.global/sites/default/files/2018-09/IPRP_GTWG_ReflectionPaper_BD_Final_2018_0713.pdf).
- U.S. Department of Health and Human Services, Food and Drug Administration, Center for Biologics Evaluation and Research. Long Term Follow-Up After Administration of Human Gene Therapy Products, Guidance for Industry. January 2020. <https://www.fda.gov/media/113768/download>.
- Donsante, A., Vogler, C., Muzyczka, N., Crawford, J.M., Barker, J., Flotte, T., Campbell-Thompson, M., Daly, T., and Sands, M.S. (2001). Observed incidence of tumorigenesis in long-term rodent studies of rAAV vectors. *Gene Ther.* 8, 1343–1346.
- Donsante, A., Miller, D.G., Li, Y., Vogler, C., Brunt, E.M., Russell, D.W., and Sands, M.S. (2007). AAV vector integration sites in mouse hepatocellular carcinoma. *Science* 317, 477.
- European Medicines Agency, Science Medicines Health, Committee for Advanced Therapies (CAT). Reflection paper on management of clinical risks deriving from insertional mutagenesis. April 19, 2013, EMA/CAT/190186/2012. [https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-management-clinical-risks-deriving-insertional-mutagenesis\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-management-clinical-risks-deriving-insertional-mutagenesis_en.pdf).
- Fitzpatrick, Z., Leborgne, C., Barbon, E., Masat, E., Ronzitti, G., van Wittenberghe, L., Vignaud, A., Collaud, F., Charles, S., Simon Sola, M., et al. (2018). Influence of Pre-existing Anti-capsid Neutralizing and Binding Antibodies on AAV Vector Transduction. *Mol. Ther. Methods Clin. Dev.* 9, 119–129.
- Solid Biosciences Inc. Solid Biosciences Provides SGT-001 Program Update. November 12, 2019. <https://www.solidbio.com/about/media/press-releases/solid-biosciences-provides-sgt-001-program-update>.
- Wang, D., Tai, P.W.L., and Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* 18, 358–378.
- ICH. Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals S6(R1). [https://database.ich.org/sites/default/files/S6\\_R1\\_Guideline\\_0.pdf](https://database.ich.org/sites/default/files/S6_R1_Guideline_0.pdf).