Gene Therapy Researchers' Assessments Of Risks And Perceptions Of Risk Acceptability In Clinical Trials

Claire T. Deakin[1,2], Ian E. Alexander[1,3], Cliff A. Hooker[2,4] and Ian H. Kerridge[2,5]

[1]Gene Therapy Research Unit, Children's Medical Research Institute and The Children's Hospital at Westmead, New South Wales, Australia; [2]Centre for Values, Ethics and the Law in Medicine, School of Public Health, Sydney Medical School, The University of Sydney, New South Wales, Australia; [3]Discipline of Paediatrics and Child Health, Sydney Medical School, The University of Sydney, New South Wales, Australia; [4]Faculty of Education and Arts, University of Newcastle, New South Wales, Australia; [5]Department of Hematology, Royal North Shore Hospital, New South Wales, Australia.

Although recent clinical trials have demonstrated the increasing promise of gene therapy, they have also illustrated the difficulties of assessing risks, given the inherent uncertainty of trial outcomes. An international survey was conducted to investigate gene therapy researchers' perceptions and assessments of risks in clinical trials. Data from respondents (*n* = 156) demonstrated researchers' perceptions of clinical context and the strength of preclinical evidence strongly influenced risk assessments and judgments of acceptable risk levels. Professional experience in clinical care, and particularly care of children, predicted favorable attitudes toward nonanimal preclinical models and trial initiation when sub-optimal treatments were available. The potential for adverse events to impact negatively on the gene therapy field and on public trust were relevant considerations when planning a trial. Decisions about clinical trials appear to be influenced not only by the clinical context and preclinical evidence, but also subjective factors reflecting the experience of researchers, value-judgments about risk and benefit, and attitudes toward preclinical models, uncertainty, adverse events, and the perceived needs of patients. It is clear that risk assessment in clinical research involves moral and scientific judgment. Identifying moral assumptions and qualitative assessments underpinning the design and conduct of research may facilitate future decision-making in clinical trials.

INTRODUCTION

Although the promise of gene therapy has long been recognized, it is only in recent years that this promise has begun to be realized. Therapeutic efficacy following gene therapy has now been reported for several inherited diseases, including X-linked severe combined immunodeficiency, adenosine deaminase deficiency, chronic granulomatous disease, Leber's congenital amaurosis, adrenoleukodystrophy, β-thalassemia, and hemophilia B; as well as leukemia and HIV.[1–18] The lag in the transition from promising early laboratory results to clinical benefit underscores the challenges and uncertainties faced in the clinical translation of complex therapeutic products. Such challenges continue to pose difficulties for gene therapy research and are likely to confront other emerging research fields, including research in stem cell therapy and nanomedicines.

One of the most significant areas of translational uncertainty concerns the extent to which therapeutic efficacy and theoretically predicted risks evaluated in preclinical models will be observed in human research subjects, and the possibility that unknown side-effects may occur. The development of leukemia in 5 of 20 infants treated in the French and British gene therapy trials for X-linked severe combined immunodeficiency particularly demonstrates the difficulties with predicting risk in clinical research, when the only available knowledge about safety and efficacy is generated in preclinical models with limited predictive capacity.[19–21] Although vector-mediated insertional mutagenesis had been predicted as a possible risk, the likelihood of this resulting in leukemia was considered remote on the basis of preclinical data obtained from cell culture and small animal models.[22–25]

Although preclinical studies provide quantitative read-outs of safety and efficacy in disease models, the interpretation of data during the planning and conduct of clinical trials involves extrapolation, and may depend upon the attitudes of researchers to risk, benefit, and the validity of the preclinical models used. For example, attitudes toward risk may become more influential when there is greater uncertainty about extrapolating from preclinical safety and efficacy testing. Little is known, however, about how researchers view risks in clinical research, and about how the attitudes of researchers toward risks and perceptions of acceptable risk levels can influence their assessments of risks during the design and conduct of clinical trials.

We have conducted a survey to investigate gene therapy researchers' assessments of risks and benefits and their perceptions of acceptable levels of risks in phase I clinical trials. The factors selected for investigation in this study as potential contributors to decision-making about clinical trials of gene therapy include the nature of the clinical context for decision-making and

Correspondence: Ian H. Kerridge, Centre for Values, Ethics and the Law in Medicine, The University of Sydney, Medical Foundation Building K25, Sydney, NSW 2006, Australia. E-mail: ian.kerridge@sydney.edu.au

the attitudes of researchers toward uncertainty, adverse events (AEs) and the validity and utility of preclinical models, as revealed by researchers' interpretations of different levels of preclinical evidence. Demographic factors, such as age, gender, professional background as clinician or scientist, and cultural background, were also investigated.
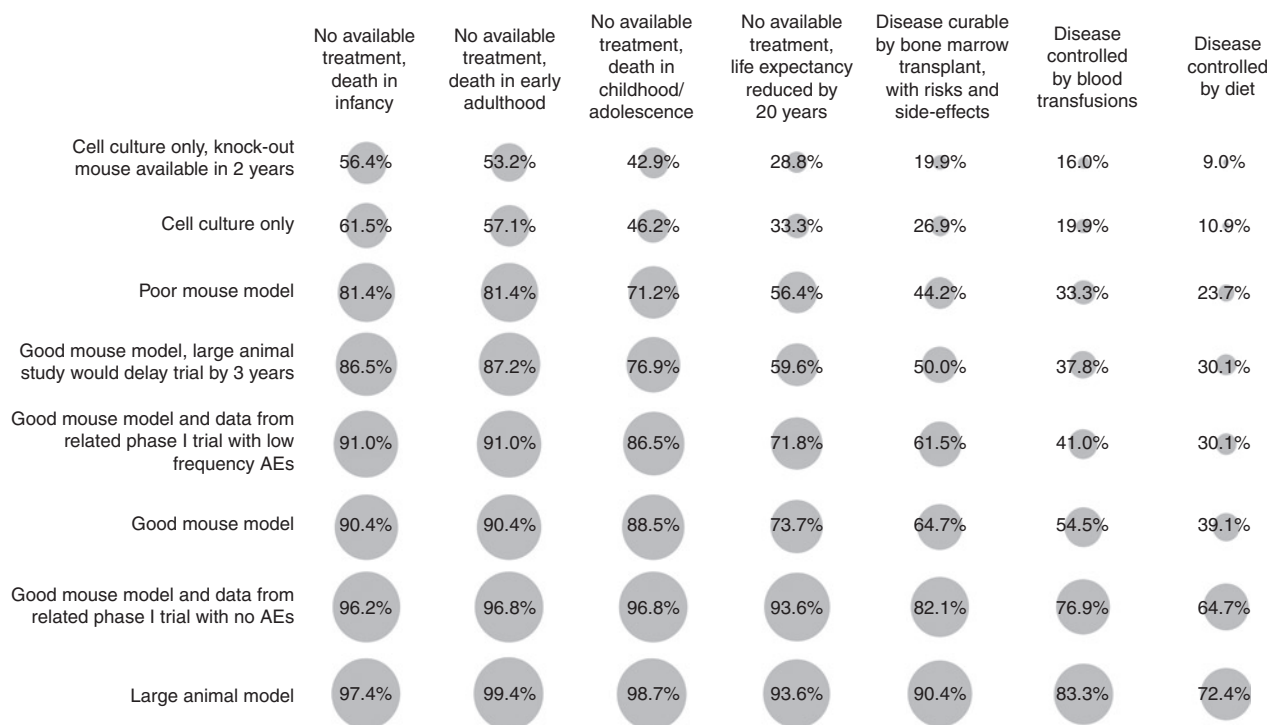
RESULTS

Demographics of respondents

A total of 156 complete survey responses were received from the mail-out to 1,733 members of gene therapy societies and 287 principal investigators of phase I gene therapy trials who received direct invitations to participate. This yields a response rate in the range of 7.7–9.0%, depending upon the extent of cross-membership between these groups, which is unknown because data from respondents was deidentified. To illustrate the representativeness of this sample of 156 responses with respect to the general population of 2,020 researchers who were invited to participate, it is possible to calculate the margin of error associated with estimating a proportion of 50% in the general population, which is the most conservative proportion to estimate.[26] The margin of error associated with a sample of 156 respondents is 7.6% with a 95% confidence interval. To achieve a margin of error of 5.0% with a 95% confidence interval, 323 respondents would be required. This suggests that the accuracy of proportions estimated using this sample population would not be substantially increased if the sample size were increased by more than twofold.

All major demographic groups of interest were represented, including gender, age, experience in a clinical trial, country in which research is conducted, years of research experience, professional role, and involvement in the clinical care of children (Table 1). These demographic groups were used for the subsequent cross-tabs analyses. Seventy-five per cent of respondents reported involvement in a clinical trial, almost 60% of respondents conducted their research in Northern America and almost 70% of respondents identified themselves as scientists.

Table 1 Demographic details of survey respondents

Demographic group	Proportion and number of responses for demographic categories		
Gender	38.5% female ($n = 60$)	61.5% male ($n = 96$)	
Age category	31.4% 40 years and under ($n = 49$)	68.6% over 40 years ($n = 107$)	
Experience in a clinical trial	75% yes ($n = 117$)	25% no ($n = 39$)	
Regional location	57.7% Northern America ($n = 90$)	19.9% Europe ($n = 31$)	22.4% Australia, Asia, and other ($n = 35$)
Professional role	69.9% scientists ($n = 109$)	4.5% clinicians ($n = 7$)	25.6% clinician-scientists ($n = 40$)
Involvement in the clinical care of children	16.0% yes ($n = 25$)	69.2% no ($n = 108$)	14.7% not applicable ($n = 23$)
Years of experience in research field	38.5% 10 years or less ($n = 60$)	37.8% 11–20 years ($n = 59$)	23.7% over 20 years ($n = 37$)

	No available treatment, death in infancy	No available treatment, death in early adulthood	No available treatment, death in childhood/ adolescence	No available treatment, life expectancy reduced by 20 years	Disease curable by bone marrow transplant, with risks and side-effects	Disease controlled by blood transfusions	Disease controlled by diet
Cell culture only, knock-out mouse available in 2 years	56.4%	53.2%	42.9%	28.8%	19.9%	16.0%	9.0%
Cell culture only	61.5%	57.1%	46.2%	33.3%	26.9%	19.9%	10.9%
Poor mouse model	81.4%	81.4%	71.2%	56.4%	44.2%	33.3%	23.7%
Good mouse model, large animal study would delay trial by 3 years	86.5%	87.2%	76.9%	59.6%	50.0%	37.8%	30.1%
Good mouse model and data from related phase I trial with low frequency AEs	91.0%	91.0%	86.5%	71.8%	61.5%	41.0%	30.1%
Good mouse model	90.4%	90.4%	88.5%	73.7%	64.7%	54.5%	39.1%
Good mouse model and data from related phase I trial with no AEs	96.2%	96.8%	96.8%	93.6%	82.1%	76.9%	64.7%
Large animal model	97.4%	99.4%	98.7%	93.6%	90.4%	83.3%	72.4%

Figure 1 Percentage of respondents who agreed it would be appropriate to recruit subjects to a phase I clinical trial for each of the hypothetical clinical scenarios based on data generated in each of the preclinical models. The radii of the circle behind each of the percentage values represents the relative proportions of respondents who agreed it would be appropriate to recruit subjects to a phase I trial. AEs, adverse events.
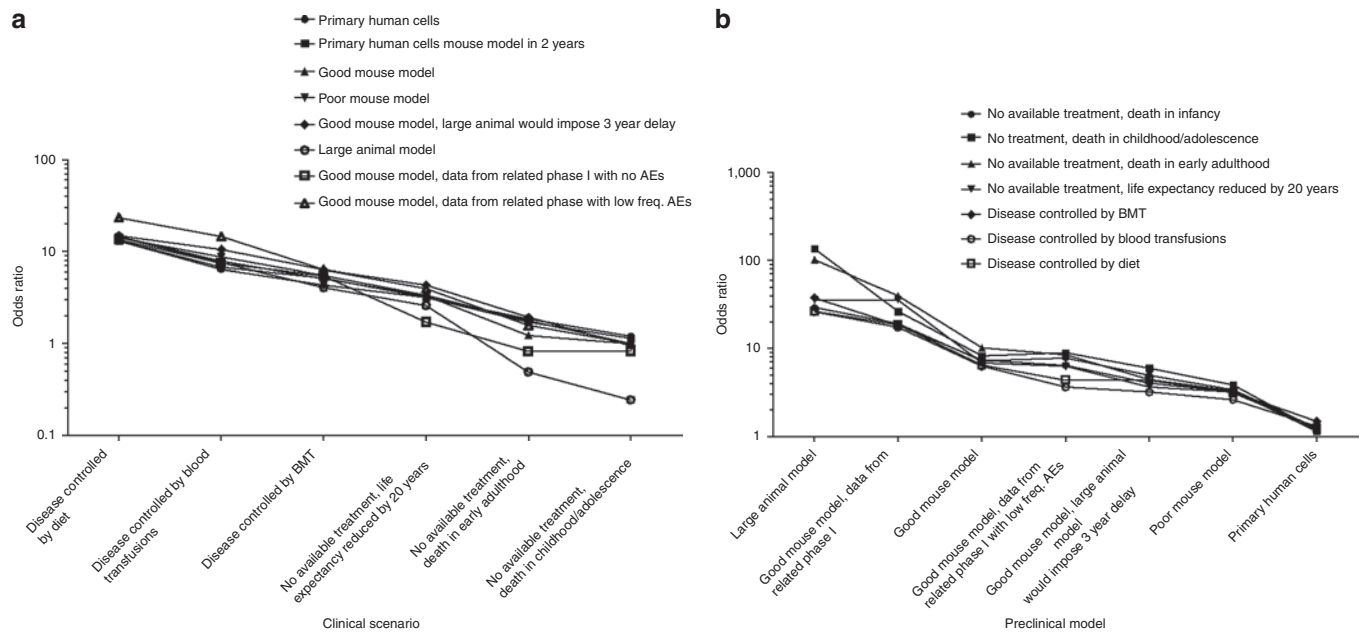
Figure 2 Consistent patterns of odds ratios when separate GEE models were fitted for each of the preclinical models and clinical scenarios. (**a**) Odds ratios representing the likelihood of disagreeing with the initiation of a clinical trial when the hypothetical clinical scenario was changed from the scenario involving an untreatable disease and death in infancy, which invoked the greatest support. Separate GEE models were fitted for each of the types of preclinical model. (**b**) Odds ratios representing the likelihood of agreeing with the initiation of a clinical trial when the type of preclinical model was changed from a cell culture-based model when a mouse model could be generated in 2 years. Similarly, separate GEE models were fitted for each of the clinical scenarios. AEs, adverse events; BMT, bone marrow transplant; freq., frequency; GEE, generalized estimating equations.

Influence of perceptions of clinical context and attitudes toward preclinical evidence on assessments of risks in a hypothetical clinical trial

Overall, the proportion of respondents who agreed that it would be appropriate to recruit subjects to a phase I clinical trial depended upon both the hypothetical clinical scenario and the strength of the preclinical evidence (**Figure 1**). The least amount of support for a trial was observed with the combination of the scenario where disease could be controlled by diet and the cell culture model when a knock-out mouse could be generated within 2 years. The greatest amount of support was observed with the scenario involving an untreatable disease with death in infancy and a large animal model. For the scenarios where no treatments were available, support for initiating a clinical trial generally increased with disease severity, more rapid disease progression and younger patient age. Support was lower for the scenarios where treatments with varying limitations were available.

Analysis of population-averaged effects across the responses to the panel of questions using generalized estimating equations (GEEs) demonstrated levels of support for initiating a clinical trial were highly dependent on both the clinical scenario ($P <$ 0.001) and the preclinical model ($P <$ 0.001). There was also a significant interaction between the clinical scenario and preclinical model variables ($P <$ 0.001). For each preclinical model, the effects of changing the clinical scenario were examined relative to the scenario which invoked the greatest support, where disease was untreatable and resulted in death in infancy (**Figure 2a** and **Supplementary Table S1** online). The likelihood of disagreement with initiating a trial progressively decreased as treatments became

unavailable, life expectancy was reduced and the disease became more severe and progressive. Changing the clinical scenario to any where treatments were available had a highly significant effect, regardless of the type of preclinical model (**Supplementary Table S1** online). Changing the scenario to either of the two scenarios where no treatments were available and life expectancy extended into adulthood had a significant effect for most of the preclinical models. Overall, the effects of changing the clinical scenario from one which was untreatable and caused death in infancy were the strongest for the types of preclinical models which involved small or large animals.

Similarly, separate GEE models were fitted for each of the clinical scenarios to examine the effects of changing the preclinical model, relative to the model that generated the least support for a trial, which was the cell culture model where a knockout mouse could be generated within 2 years (**Figure 2b** and **Supplementary Table S2** online). The likelihood of agreeing with the initiation of a trial progressively decreased as the strength of preclinical evidence decreased, from the large animal model through to the cell culture model (**Figure 2b**). For all of the clinical scenarios, the effects of changing the type of preclinical model to a small or large animal model were highly significant (**Supplementary Table S2** online). In general, the effect of changing the type of preclinical model was strongest for the clinical scenarios where treatments were available and quality of life was considered acceptable. Taken as a whole, the GEE analyses confirmed the overall pattern that support for a clinical trial increased with stronger preclinical models, greater disease severity and reduced life expectancy and treatment availability for the condition under study.
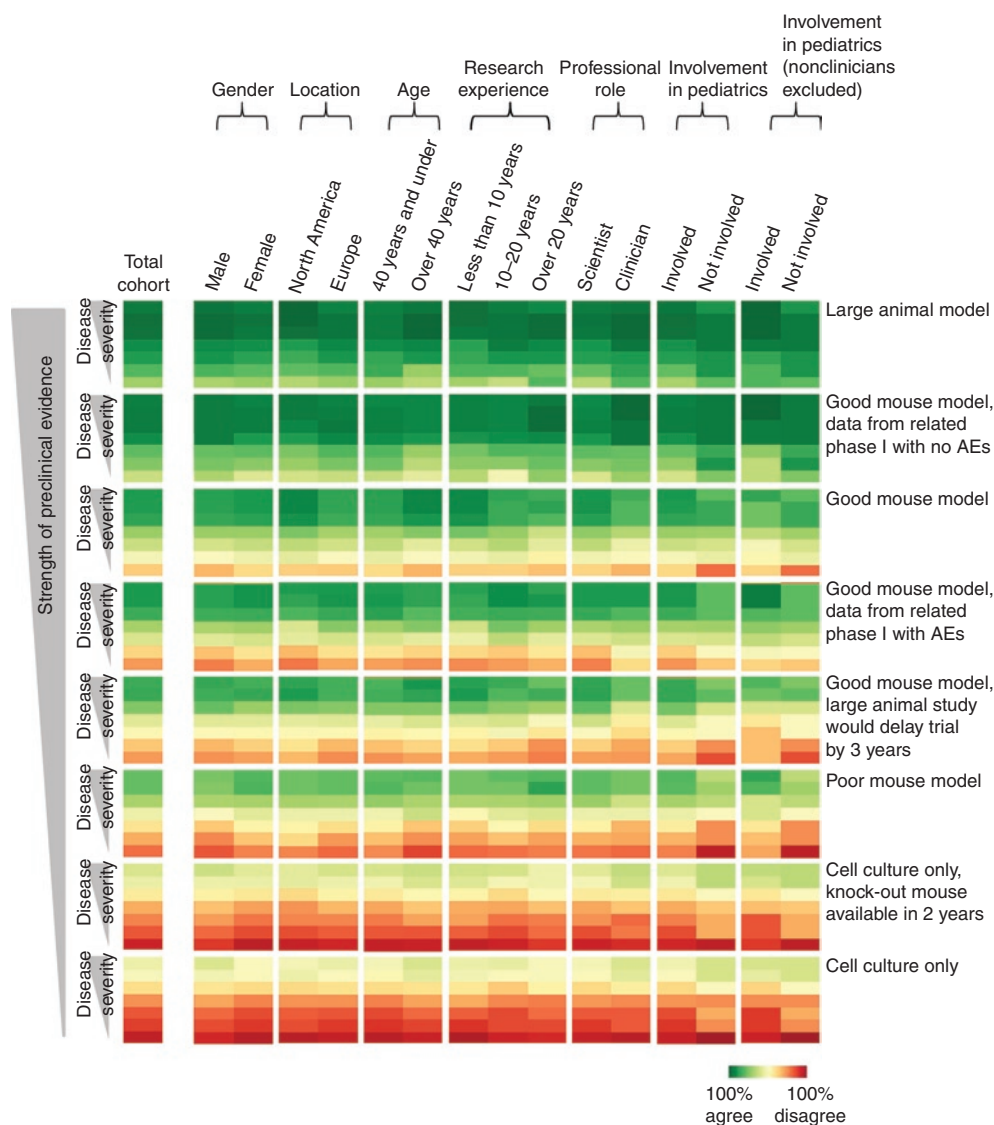
Figure 3 Heat-map representing the proportions of respondents in each demographic group who supported initiation of a clinical trial, for each of the clinical scenarios and types of preclinical models. Dark green represents 100% agreement with initiating a trial, while dark red represents 100% disagreement (0% agreement). The columns are separated by white spaces according to the demographic groups of interest, with the total cohort represented in the left-hand-most column. Each of the eight groups of seven rows separated by a white space represents one of the eight types of preclinical evidence, ordered according to decreasing strength from top to bottom. These eight rows contain groups of seven rows, which represent each of the seven different clinical scenarios. The seven clinical scenarios are ordered according to decreasing disease severity from top to bottom: (i) no available treatment, death in infancy; (ii) no available treatment, death in childhood or adolescence; (iii) no available treatment, death in early adulthood; (iv) no available treatment, life expectancy reduced by 20 years; (v) disease curable by bone marrow transplantation, with risks and side-effects; (vi) disease can be controlled by blood transfusions; (vii) disease can be controlled by diet. AEs, adverse events.

Differential risk assessment by demographic groups when making decisions about initiating a hypothetical clinical trial

To explore the potential influence of demographic factors, responses were aggregated according to gender, age, geographic location, professional role, involvement in pediatric care, and years of research experience. A heat-map was employed as a visualization tool for the purpose of guiding the analytical process. The proportions of respondents in each demographic group who supported a hypothetical clinical trial in each of the clinical scenarios and for each of the preclinical models were represented using a heat-map (**Figure 3**). The pattern of colors displayed for the total

cohort of respondents in the first column of the heat-map indicated that higher support for a trial was correlated with increased disease severity and strength of preclinical evidence. When the respondents were aggregated according to demographic factors, the heat-map patterns for the different demographic groups were similar to the pattern for the total cohort, indicating overall that levels of support were similar. Similar color patterns were also observed between demographic categories. In general, demographic factors appeared to exert little effect when the preclinical evidence was strong and when the disease scenarios were incurable and associated with higher severity, lower quality of life and shorter life expectancy. Gender, age, regional location, and years

Table 2 Effects of professional experience on decision-making in the least severe clinical scenarios when preclinical evidence involved cell culture-based models or there was a related phase I study with low frequency adverse events

Professional experience	Type of evidence	Clinical scenario	χ²P value	Odds ratio[a]	Odds ratio P value
Professional role (Clinicians versus scientists)	Primary human cells	Disease controlled by blood transfusions	0.017	2.73	0.015
		Disease controlled by diet	0.010	3.94	0.010
	Primary human cells, mouse model within 2 years	Disease controlled by blood transfusions	0.009	3.09	0.012
		Disease controlled by diet	0.031*	3.52	0.028
	Good mouse model with data from related phase I trial, with low frequency adverse events	Disease controlled by blood transfusions	0.006	2.63	0.007
		Disease controlled by diet	0.003	2.96	0.003
	Changed from primary human cells, mouse model within 2 years to good mouse model	No available treatment, death in infancy	0.023	0.40	0.025
		No available treatment, death in childhood/adolescence	0.005	0.33	0.007
		No available treatment, death in adulthood	0.014	0.41	0.016
		Disease curable by bone marrow transplant, with risks and side-effects	0.019	0.43	0.021
Involved in clinical care of children (Yes versus no or not applicable)	Primary human cells	Disease controlled by blood transfusions	<0.001	5.44	<0.001
		Disease controlled by diet	0.008	4.71	0.008
	Primary human cells, mouse model within 2 years	Disease controlled by blood transfusions	<0.001	6.57	<0.001
		Disease controlled by diet	0.012	4.86	0.008
	Good mouse model with data from related phase I trial, with low frequency adverse events	Disease controlled by blood transfusions	0.011	3.07	0.013
		Disease controlled by diet	0.009	3.09	0.012
Involved in clinical care of children (Yes versus no; nonclinicians excluded)	Primary human cells	Disease controlled by blood transfusions	0.019	4.62	0.024
	Primary human cells, mouse model within 2 years	Disease controlled by blood transfusions	0.015	5.50	0.021

[a]Odds ratios are expressed as the likelihood that support for a hypothetical trial for the given preclinical data and clinical scenario was associated with the first of the demographic subgroups specified in the "Professional experience" column, rather than the second of the demographic groups.
*Indicates P value calculated using Fisher's exact test.

Table 3 Percentage of respondents who agreed with statements about attitudes toward research involving human subjects

Statement	Percentage of whole cohort agreeing (%)
Acceptability of proceeding to clinical trial without extensive toxicological data, for severe disease when treatment does not exist	30.8
Absence of adverse events in small animal models do not enable meaningful predictions about long-term safety in humans	81.4
It is more important to have a greater prospect of benefit in trials involving children than in trials involving adults	64.7
For a progressive disease, it is preferable to recruit patients with advanced disease than patients with an early disease stage	68.6
In the context of severe disease when no treatment is available, it may be acceptable to proceed to trial when preclinical measurements of efficacy fall short of statistical significance (i.e., $P \geq 0.05$)	72.4
The potential for adverse events to have a negative effect on public support and trust is a relevant consideration when planning a trial	79.7
The potential for adverse events to have a negative effect on the field of research is a relevant consideration when planning a trial	79.1
Whether personal decision-making about clinical trials has been affected by the report of an adverse event	57.1

of research experience did not appear to influence decision-making outcomes. However, in some situations where the preclinical evidence was weaker and the disease scenarios were less severe, it appeared that there were some differences in levels of support according to professional background of the respondents as a clinician or scientist and involvement in pediatric care. Therefore, the effect of professional experience on decision-making was examined for the questions involving the two least severe disease scenarios and the two cell-culture preclinical models (**Table 2**).

When the effects of professional experience were examined, clinicians and clinician-scientists were more likely than scientists to support a phase I trial on the basis of data generated in cell culture models when disease could be treated by either blood transfusion or diet. Respondents involved in the clinical care of children were similarly more likely to support a phase I trial when the preclinical evidence was weaker and the disease was less severe. These associations persisted when nonclinicians were excluded from analysis for the cell-culture model when disease could be treated by blood transfusion, but not when disease could be treated by diet. On the other hand, scientists and respondents not involved in pediatric care were more likely to decide against initiating a trial when low frequency adverse events occurred in a related phase I trial. When

Table 4 Descriptions of preclinical evidence provided to respondents for decision-making about the initiation of a clinical trial

Abbreviated description of preclinical evidence	Description of preclinical evidence as provided to respondents in the survey instrument
Cell culture only	A statistically powered study demonstrating therapeutic efficacy in relevant primary human cells in culture AND evidence of safety in a validated cell culture assay
Cell culture only, knock-out mouse available in 2 years	A statistically powered study demonstrating therapeutic efficacy in relevant primary human cells in culture AND evidence of safety in a validated cell culture assay, if a knock-out mouse model is expected to be generated within 2 years
Good mouse model	A statistically powered study demonstrating safety and therapeutic efficacy, in correcting a mouse phenotype with no evidence of toxicity or adverse events in long-term follow up, using a mouse model where the phenotype is the same as the human disease (no large animal model available)
Poor mouse model	A statistically powered study demonstrating safety and therapeutic efficacy, in correcting a mouse phenotype with no evidence of toxicity or adverse events in long-term follow up, using a mouse model that involves the same gene but where the phenotype differs from the human disease (no large animal model available)
Good mouse model, large animal study would delay trial by 3 years	A statistically powered study demonstrating safety and therapeutic efficacy, in correcting a mouse phenotype with no adverse events, using a mouse model where the phenotype is the same as the human disease and when a large animal model is available but a large animal study would delay trial initiation by 3 years
Large animal model	A statistically powered study demonstrating safety and therapeutic efficacy in a large animal model, in correcting a disease phenotype that is the same as the human disease and with no adverse events in long-term follow up
Good mouse model and data from related phase I trial with no adverse events	Convincing preclinical safety and efficacy data in a mouse model and data from a phase I trial targeting a different disease of the same target tissue, which used the same technological intervention and demonstrated therapeutic efficacy and no adverse events
Good mouse model and data from related phase I trial with low frequency adverse events	Convincing preclinical safety and efficacy data in a mouse model and data from a phase I trial targeting a different disease of the same target tissue, which used the same technological intervention and demonstrated therapeutic efficacy and a low frequency of serious and life-threatening adverse events (e.g., <10%)

Table 5 Clinical scenarios provided to respondents for decision-making about the initiation of a clinical trial

Scenario abbreviations	Description of clinical scenario as provided to respondents in the survey instrument
Scenario 1 No available treatment, death in infancy	No available treatment, presents at birth or soon after, death in infancy
Scenario 2 No available treatment, death in childhood/adolescence	No available treatment, rapidly progressive, poor quality of life, death in childhood/adolescence
Scenario 3 No available treatment, death in adulthood	No available treatment, slowly progressive, poor quality of life during final 10 years, death in early adulthood
Scenario 4 No available treatment, life expectancy reduced by 20 years	No available treatment, quality of life acceptable until end-stage, life expectancy reduced by 20 years
Scenario 5 Disease curable by bone marrow transplant, with risks and side-effects	Disease can be cured by allogeneic bone marrow transplantation, but prior myeloablation is required and transplant may be associated with acute and chronic graft-versus-host disease and other serious and secondary effects e.g., malignancy. Survival rate is 65% at 10 years. Quality of life in survivors is reduced but generally regarded as acceptable
Scenario 6 Disease controlled by blood transfusions	Disease can be controlled by regular blood transfusions, which over time may lead to iron overload and cardiac failure. Life expectancy and quality of life are reduced, although quality of life is considered acceptable in most cases
Scenario 7 Disease controlled by diet	Disease can be controlled by diet, but compliance is poor and low grade neurocognitive damage is likely. Life expectancy is normal in most patients and quality of life is generally acceptable

the preclinical model was changed from a cell culture model to a good mouse model, scientists were more likely than clinicians and clinician-scientists to change their decision from disagree to agree in three of the scenarios where treatments were unavailable and the scenario where disease could be treated by bone marrow transplantation.

Involvement in pediatric care had a marked impact upon decision-making when included as a variable in the GEE models fitted for the cell culture model, the cell culture model when a mouse model could be available in 2 years, the good mouse model when there was a related phase I study with no AEs, and the good mouse model when there was a related phase I study with a low frequency of AEs (odds ratios of 7.42, 8.54, 4.06, and 5.13, respectively; P

values of 0.006, 0.003, 0.044, and 0.024, respectively). Similarly, professional experience in pediatric care affected decision-making outcomes when analyzed as a variable in the GEE models fitted for the scenarios where disease could be controlled by blood transfusions or by diet (odds ratios of 9.55 and 6.50, respectively; P values of 0.002 and 0.011, respectively). When nonclinicians were excluded from the GEE analysis, the effect of involvement in pediatric care remained for the GEE models fitted for the cell culture model, the cell culture model when a mouse model could be available in 2 years, and also the scenario where disease could be cured by bone marrow transplantation (odds ratios of 2.75, 2.87, and 2.52, respectively; P values of 0.049, 0.041, and 0.034, respectively).

Attitudes toward research involving human subjects

Responses to questions about attitudes toward research involving human subjects correlated with the responses to decisions about initiating a hypothetical clinical trial regarding the types of clinical scenarios and preclinical model more likely to lead to support for a trial. Respondents emphasized the importance of benefit in pediatric trials and preclinical safety testing, despite recognizing the limitations of small animal models. The importance of benefit in trials involving children was expressed by almost two-thirds of respondents (**Table 3**). Over two-thirds of respondents preferred recruitment of patients with advanced disease as compared with the patients with early stage disease.

Although over 80% of respondents considered the absence of AEs in small animal models did not facilitate meaningful predictions about long-term safety outcomes in humans, ~70% thought a clinical trial should not proceed without extensive toxicological data for a severe and incurable disease (**Table 3**). On the other hand, over 70% of respondents considered it may be acceptable to proceed to clinical trial in the context of a severe and incurable disease, when preclinical measurements of therapeutic efficacy fell short of clinical significance in a statistically powered animal study.

When asked about their own decision-making about a clinical trial, respondents considered the potential negative impact of AEs as a relevant factor, and indicated their personal decision-making had been affected by the report of an AE in their own field. Nearly 80% of respondents identified the potential for AEs to have a negative effect on public support and trust or on their own field of research as relevant considerations when planning a clinical trial (**Table 3**). Almost 60% of respondents indicated their own decision-making had been affected by the report of an AE in a trial in their field.

DISCUSSION

This study investigated gene therapy researchers' perceptions of risk, risk assessment, the validity and utility of preclinical models and AEs, as well as the potential influence of these perceptions on decision-making about gene therapy trials. Our data obtained from a subset of gene therapy researchers reveals that the characteristics of the target disease and existing therapies, as well as the availability and practicalities of animal models, are highly influential when decisions are made about the design and conduct of clinical trials. Decisions about clinical trials are influenced by a number of subjective factors, such as length and kind of experience, value-judgments regarding risk and benefit, attitudes toward the strength and validity of preclinical evidence, and perceptions of the need for a novel treatment. This represents the first empirical study of how gene therapy researchers assess risks in clinical trials, the factors which influence decision-making in clinical research, and how gene therapy researchers balance these factors. Previous empirical studies of decision-making about clinical trials have evaluated the experience of research participants, with a particular focus on consent and motivations for research participation.[27–32]

There are a number of potential limitations to this study which may arise from the relatively small sample size of 156 respondents, the heterogeneity of respondents, and limited number of qualitative questions included. The low response rate (7.7–9.0%) is consistent with the literature regarding online surveys, which describes response rates ranging from 1% to more than 60% depending upon

the characteristics of the target population, the survey length, mode and strategy for delivery, and salience of the survey topic.[33,34] In addition, our response rate is likely to reflect both the large number of complex questions in the survey, necessary for obtaining meaningful data, and also the likely time-poor nature of the target research population, which renders any dataset of this nature difficult to obtain. The consequence of this low response rate is that these findings are exploratory in nature, have limited potential for generalizability, and need confirmation in larger studies.

Related to the low response rate is the possibility of an ascertainment bias in our data, which may arise if the characteristics of survey respondents differ significantly from those of the general population of gene therapy researchers. Despite these limitations to generalizability, the findings and analyses reported here nevertheless have internal validity with respect to the sub-population of gene therapy researchers who responded to the survey. Furthermore, it is also possible that the data may be influenced by a positive ascertainment bias. Seventy-five per cent of our respondents reported having previous involvement in a clinical trial, which is likely to be a higher proportion than in the general population of gene therapy researchers. Although the low response rate necessitates careful and qualified interpretation of the data, the 156 responses obtained in this study reflect opinions in the field and are useful for identifying factors which may influence decision-making and which can be followed up in further studies.

Although the cohort of respondents were heterogeneous with respect to professional role, years of research experience, age, and cultural background, this heterogeneity also enabled identification of demographic factors which may predict attitudes and perceptions, such as professional experience. It should also be noted that few of the "clinicians" in this analysis identified themselves solely as clinicians, with the majority indicating a dual role as scientists and clinicians. Although this probably reflects the targeting of the survey to a community of researchers, further research and larger numbers of responses is required to further explore differences between individuals who identify solely as either scientists or clinicians.

Details of the clinical context for a trial, including the particular characteristics of the disease and the availability of alternative or effective therapy, were key factors in decision-making about clinical research. This was evident from both the responses of participants to hypothetical clinical trials and to the questions regarding their attitudes to research. When the needs of the patients are perceived as greater and there is increased certainty of a negative outcome under the existing clinical course, then researchers were more likely to support a clinical trial, even when the preclinical evidence was relatively weak. By contrast, when patients generally experienced an acceptable quality of life and sub-optimal treatments were available, then respondents were more divided about the appropriateness of initiating a phase I trial. In these situations, differences emerged between demographic groups, and particularly between scientists and clinicians.

What is clear from our data is that consideration of the particular clinical context and the potential for benefit shaped judgments of acceptable risk during the risk-benefit calculus for phase I trials, even though the primary goal of such trials is to evaluate safety, and not therapeutic efficacy. Respondents to this survey preferred to recruit patients with advanced disease as subjects in

trials involving adults, suggesting exposure to uncertain risks was considered more acceptable when research participants had "less to lose", even when the state of disease progression would reduce the potential for benefit. By contrast, in the pediatric setting, respondents generally considered the risk-benefit ratio should be skewed in favor of benefit. At the same time, when the target disease was severe and incurable, it was considered acceptable to proceed to a clinical trial even when measurements of therapeutic efficacy fell short of statistical significance in a statistically powered preclinical study, such that the prospect of benefit was unlikely. In other words, if the clinical context was dire, then it was more acceptable for the risk-benefit ratio to be skewed in favor of risk.

Gene therapy researchers' support for a hypothetical clinical trial was also strongly influenced by their attitudes toward the strength and utility of evidence generated in different preclinical models. Even though respondents placed great store in the use of preclinical models, they were clearly mindful of the limited predictive capacity and relatedness to human disease of some preclinical models, particularly cell culture-based models and small animal models. Attitudes toward the preclinical models perceived as weaker, particularly primary human cells, were also a key point of difference between scientists and clinicians, and pediatricians in particular.

Conversely, support for initiating a clinical trial was greatest when a large animal was used as a preclinical model across all clinical scenarios, with support being almost universal when treatments were nonexistent. Although it is self-evident that large animal models would be a preferred standard of evidence for preclinical studies, in reality such models are not readily available, are costly to maintain and impose protracted timelines. As a consequence, it is inevitable that researchers must consider the best way to communicate to prospective research participants the remaining uncertainties about safety and efficacy following preclinical studies in small animals. These may include how the limitations of models may affect attempts to extrapolate stable phenotype correction in mice (or other demonstrations of efficacy) to long-term potential benefit in humans, the potential for differences in immune responses between species, and the impact of short lifespans of small animals on assessments of long-term safety.

The possibility that scientists and clinicians may assess risks in clinical research differently highlights the influence of personal experience of disease upon decision-making; in this instance, gained through a professional role in clinical care and the care of children in particular. This finding is consistent with earlier research comparing perceptions of risk held by experts and laypersons.[35,36] It seems likely that those who work in healthcare may have a greater "everyday" experience of the burden of severe and incurable disease and the impact of such diseases on quality of life. This understanding may inform researchers' evaluations of acceptable levels of risk in research and decisions about when it is appropriate to initiate a trial, such as whether to delay a trial to develop a better preclinical model or an improved vector. Conversely, scientists may have greater experience in interpreting research evidence. Involvement of both scientists and clinicians, therefore, is necessary for maximally informed decisions. Discussion and decision-making in ethics committees may also be more rigorous when committee members have a heterogeneous range of personal experiences of disease.

This study also provides preliminary empirical support for the notion that SAEs can delay research, not from an ethical or regulatory perspective, but as a result of researchers' own decision-making about clinical trials. The gene therapy researchers in our study were clearly concerned not only about the short-term negative consequences of AEs for research participants and their families, but also about the longer-term negative consequences of AEs for public support and trust and for the gene therapy field as a whole. The majority of respondents took these latter, longer-term consequences into account when planning a clinical trial. Moreover, most researchers who had been involved in a clinical trial stated that their own decision-making had been influenced by the report of an AE in their field. These findings suggest there can be unrecognized variables which affect decision-making and it would be interesting to investigate whether other fields, which have experienced SAEs but without the same degree of media attention, would be similarly influenced.

This study has identified a complex matrix of ethical, social, experiential and evidence-related factors which may influence decisions about clinical trials, and which may be differentially evaluated by different individuals. Although the results of our study are intriguing, further research is required to confirm the demographic predictors identified. It is unclear how explicitly and transparently these factors are considered and how they are weighed during decision-making. Judgments about risk, benefit, uncertainty and the adequacy of evidence are inescapably value-laden. This suggests that what is needed is a more sophisticated discussion about how values influence the decision-making and the risk assessments of not only researchers, but also regulators, ethics committees and research participants, and how decision-making within and between these groups may differ. It also suggests that optimal decision-making requires involvement of a mixture of people with diverse experience. A list of factors which have been identified as potential influences on decision-making and specific questions that need to be asked of any clinical study may assist the evaluation of such factors and improve transparency and rigor in decision-making about clinical trials. Such questions could be used to guide researchers' evaluation of the preclinical evidence and clinical context, and their design of studies, selection of participants and communication of risks to potential participants.[37] This research represents a first step toward identifying some of the factors which require consideration during decision-making.

Gene therapy researchers' assessments of risks in clinical trials are influenced strongly by their perceptions of the clinical context, the characteristics and needs of the patient cohort, and the strength and validity of preclinical evidence. Other nonscientific factors, including personal experience of disease and concerns regarding the potential negative effects of AEs, may also influence the design and conduct of clinical trials in gene therapy. This exploratory study provides important insights into how researchers make decisions about clinical trials and makes it clear that trial design is as much a moral issue as a scientific one. Recognition that not only the decisions of researchers, but also those of research participants, ethics committees and regulatory authorities, requires the exercise of moral judgment may facilitate more sophisticated thinking about the design of first-in-human trials and the means by which such trials can be effectively communicated to potential participants.

MATERIALS AND METHODS

Study design. A quantitative internet-based survey was developed for investigating researchers' assessments of risks and perceptions of acceptable risk levels in phase I clinical trials, using a combination of short vignette questions, questions about the attitude of researchers to specific ethical issues about risks in clinical research, and questions about researchers' personal experience of decision-making about risk assessment (**Supplementary Appendix** online). The vignette questions required respondents to make hypothetical decisions about whether it would be appropriate to recruit participants to a phase I clinical trial. Respondents were presented with eight different descriptions of preclinical data (**Table 4**) in the context of seven different clinical scenarios (**Table 5**), and thus made a total of 56 hypothetical decisions. This approach required respondents to assess the risks in each hypothetical vignette and also revealed respondents' perceptions about acceptable levels of the potential risks in each of the hypothetical scenarios. The descriptions of preclinical data in this section were designed to reveal the attitudes of respondents toward uncertainty, AEs and the validity and utility of preclinical models. The scenarios were constructed to elucidate respondents' perceptions of the need for a novel treatment, and incorporated a range of disease severities, life expectancies, patient ages, qualities of life, availabilities and limitations of existing treatments, and uncertainties of outcome if the existing clinical course was followed. The questions about specific ethical issues about risks were designed to show the attitudes of researchers toward the needs of the patients, the strength of preclinical evidence, and the acceptability of research involving children. The questions about researchers' personal experience of decision-making asked about the influence of potential and actual AEs. These questions about researchers' attitudes and experience were designed to give further insights into how researchers assess risks and perceive acceptable risk levels.

An internet-based approach for the survey was chosen to minimize the time required to complete the survey and to facilitate representation of international gene therapy researchers, even though the trade-off was likely to be a low response rate. The survey was distributed on the authors' behalf by the American Society for Gene and Cell Therapy (ASGCT), the European Society for Gene and Cell Therapy (ESGCT), the Australasian Gene Therapy Society (AGTS), the British Society for Gene and Cell Therapy (BSGCT), and the French Society of Cell and Gene Therapy (SFTCG) to their respective mailing lists. In addition to this mail-out, the survey was sent directly to principal investigators of phase I trials of gene therapy, whose names were identified using the clinical trial registration website ClinicalTrials.gov and whose contact details could be accessed publicly. This enabled the most directly relevant data to be obtained from researchers who had personally been involved in clinical trial decision-making. Research was conducted with the approval of the Human Research Ethics Committee of the Children's Hospital at Westmead, Sydney, NSW, Australia.

Construction and validation of the survey instrument. Survey questions were developed to investigate researchers' perceptions and assessments of risks in clinical trials (**Supplementary Appendix** online). Survey questions were all multiple choice, with the exception of two optional open-ended text questions (data not shown). Part A of the survey instrument collected the demographic details of respondents. Part B of the survey contained the 56 hypothetical questions about whether it would be appropriate to recruit participants to a phase I trial, and respondents were given the options of "strongly agree", "agree", "disagree", and "strongly disagree". The hypothetical clinical scenarios were designed to represent examples of real diseases which are treated in phase I trials of gene therapy. Naming of specific diseases and usage of detailed medical jargon were avoided to ensure accessibility for researchers without medical training. The numbers of scenarios and descriptions of preclinical data in Part B of the survey reflected a balance between covering a representative range of variables contributing toward the nature of the clinical scenario, and minimizing the amount of time required to complete the survey. Parts C and D of the survey instrument examined the attitudes of researchers to ethical issues about risks in clinical research and AEs, and included questions about researchers' personal experience of decision-making about the conduct of clinical trials.

The internet-based survey instrument was constructed using the open source online survey application LimeSurvey, hosted by the LimeService platform (LimeSurvey.com, Hamburg, Germany). The survey instrument was validated for readability and content validity by piloting on a representative sample of researchers involved in clinical research.

Data analysis and interpretation. Statistical analysis of the survey data was conducted using the Statistical Package for the Social Sciences software (SPSS; Version 20, IBM, Armonk, NY) and heat-maps were generated using R.[38] Related responses were combined, such as "agree" and "strongly agree", "disagree" and "strongly disagree", "very relevant" and "somewhat relevant", and "completely irrelevant" and "somewhat irrelevant". Trends across the whole survey cohort were analyzed using GEEs, a method for modeling binary outcomes and which allows for intra-individual correlations.[39–41] GEE analysis enabled examination of patterns in each individual respondent's willingness to support a clinical trial, averaged across the whole survey cohort. The clinical scenario, the type of preclinical model and professional experience in the clinical care of children were examined as variables in the GEE model. The effects of changing the clinical scenario and preclinical model were also examined individually, by fitting separate GEE models for each preclinical model and for each clinical scenario, respectively.

Cross-tabs analyses were performed to analyze the association of binary outcomes with different demographic groups, using Pearson χ^2 tests, Fisher's exact tests and Mantel–Haenszel odds ratios. In all analyses, a P value of <0.05 was deemed statistically significant.

Respondents were aggregated on the basis of gender (female or male), age (40 years and under or over 40 years), regional location (North American or Europe), years of research experience (10 years and under, 11–20 years or more than 20 years), professional role (scientist or clinician/ clinician-scientist), and involvement in the clinical care of children (yes or no/not applicable; with and without the exclusion of nonclinicians). Responses from Australia and Asia were excluded from the analysis based on regional location, to enable investigation of cultural differences in risk perceptions in the regions with the greatest volume of clinical trial activity, North America and Europe.

SUPPLEMENTARY MATERIAL
Table S1. Analysis of the effects of varying clinical scenarios for GEE models fitted individually for each of the type of preclinical evidence.
Table S2. Analysis of the effects of varying preclinical evidence for GEE models fitted individually for each clinical scenario.
Appendix - Survey instrument.

REFERENCES
1. Cavazzana-Calvo, M, Hacein-Bey, S, de Saint Basile, G, Gross, F, Yvon, E, Nusbaum, P et al. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**: 669–672.
2. Hacein-Bey-Abina, S, Hauer, J, Lim, A, Picard, C, Wang, GP, Berry, CC et al. (2010). Efficacy of gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med* **363**: 355–364.
3. Gaspar, HB, Parsley, KL, Howe, S, King, D, Gilmour, KC, Sinclair, J et al. (2004). Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* **364**: 2181–2187.
4. Gaspar, HB, Cooray, S, Gilmour, KC, Parsley, KL, Adams, S, Howe, SJ et al. (2011). Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. *Sci Transl Med* **3**: 97ra79.

5. Aiuti, A, Slavin, S, Aker, M, Ficara, F, Deola, S, Mortellaro, A *et al.* (2002). Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science* **296**: 2410–2413.
6. Aiuti, A, Cattaneo, F, Galimberti, S, Benninghoff, U, Cassani, B, Callegaro, L *et al.* (2009). Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N Engl J Med* **360**: 447–458.
7. Cavazzana-Calvo, M, Hacein-Bey, S, Basile, G, Gross, F, Yvon, E, Nusbaum, P *et al.* (2000). Gene Therapy of Human Severe Combined Immunodeficiency (SCID)-X1 Disease. *Science* **288**: 669–672.
8. Ott, MG, Schmidt, M, Schwarzwaelder, K, Stein, S, Siler, U, Koehl, U *et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med* **12**: 401–409.
9. Bainbridge, JW, Smith, AJ, Barker, SS, Robbie, S, Henderson, R, Balaggan, K *et al.* (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* **358**: 2231–2239.
10. Maguire, AM, Simonelli, F, Pierce, EA, Pugh, EN Jr, Mingozzi, F, Bennicelli, J *et al.* (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med* **358**: 2240–2248.
11. Cartier, N, Hacein-Bey-Abina, S, Bartholomae, CC, Bougnères, P, Schmidt, M, Kalle, CV *et al.* (2012). Lentiviral hematopoietic cell gene therapy for X-linked adrenoleukodystrophy. *Meth Enzymol* **507**: 187–198.
12. Cavazzana-Calvo, M, Payen, E, Negre, O, Wang, G, Hehir, K, Fusil, F *et al.* (2010). Transfusion independence and HMGA2 activation after gene therapy of human ß-thalassaemia. *Nature* **467**: 318–322.
13. Nathwani, AC, Tuddenham, EG, Rangarajan, S, Rosales, C, McIntosh, J, Linch, DC *et al.* (2011). Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med* **365**: 2357–2365.
14. Kang, EM, Choi, U, Theobald, N, Linton, G, Long Priel, DA, Kuhns, D *et al.* (2010). Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. *Blood* **115**: 783–791.
15. Porter, DL, Levine, BL, Kalos, M, Bagg, A and June, CH (2011). Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N Engl J Med* **365**: 725–733.
16. Kalos, M, Levine, BL, Porter, DL, Katz, S, Grupp, SA, Bagg, A *et al.* (2011). T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. *Sci Transl Med* **3**: 95ra73.
17. Brentjens, RJ, Rivière, I, Park, JH, Davila, ML, Wang, X, Stefanski, J *et al.* (2011). Safety and persistence of adoptively transferred autologous CD19-targeted T cells in patients with relapsed or chemotherapy refractory B-cell leukemias. *Blood* **118**: 4817–4828.
18. Mitsuyasu, RT, Merigan, TC, Carr, A, Zack, JA, Winters, MA, Workman, C *et al.* (2009). Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nat Med* **15**: 285–292.
19. Hacein-Bey-Abina, S, Von Kalle, C, Schmidt, M, McCormack, MP, Wulffraat, N, Leboulch, P *et al.* (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**: 415–419.
20. Hacein-Bey-Abina, S, Garrigue, A, Wang, GP, Soulier, J, Lim, A, Morillon, E *et al.* (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest* **118**: 3132–3142.
21. Howe, SJ, Mansour, MR, Schwarzwaelder, K, Bartholomae, C, Hubank, M, Kempski, H *et al.* (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest* **118**: 3143–3150.
22. Stocking, C, Bergholz, U, Friel, J, Klingler, K, Wagener, T, Starke, C *et al.* (1993). Distinct classes of factor-independent mutants can be isolated after retroviral mutagenesis of a human myeloid stem cell line. *Growth Factors* **8**: 197–209.
23. Hacein-Bey, H, Cavazzana-Calvo, M, Le Deist, F, Dautry-Varsat, A, Hivroz, C, Rivière, I *et al.* (1996). gamma-c gene transfer into SCID X1 patients' B-cell lines restores normal high-affinity interleukin-2 receptor expression and function. *Blood* **87**: 3108–3116.
24. Hacein-Bey, S, Basile, GD, Lemerle, J, Fischer, A and Cavazzana-Calvo, M (1998). gammac gene transfer in the presence of stem cell factor, FLT-3L, interleukin-7 (IL-7), IL-1, and IL-15 cytokines restores T-cell differentiation from gammac(-) X-linked severe combined immunodeficiency hematopoietic progenitor cells in murine fetal thymic organ cultures. *Blood* **92**: 4090–4097.
25. Hacein-Bey, S, Gross, F, Nusbaum, P, Hue, C, Hamel, Y, Fischer, A *et al.* (2001). Optimization of retroviral gene transfer protocol to maintain the lymphoid potential of progenitor cells. *Hum Gene Ther* **12**: 291–301.
26. National Statistics Service. (2012). Sample size calculator <http://www.nss.gov.au/nss/home.nsf/pages/Sample+Size+Calculator+Description?OpenDocument>.
27. Catt, S, Langridge, C, Fallowfield, L, Talbot, DC and Jenkins, V (2011). Reasons given by patients for participating, or not, in Phase 1 cancer trials. *Eur J Cancer* **47**: 1490–1497.
28. Jenkins, V, Solis-Trapala, I, Langridge, C, Catt, S, Talbot, DC and Fallowfield, LJ (2011). What oncologists believe they said and what patients believe they heard: an analysis of phase I trial discussions. *J Clin Oncol* **29**: 61–68.
29. Miller, SM, Hudson, SV, Egleston, BL, Manne, S, Buzaglo, JS, Devarajan, K *et al.* (2012). The relationships among knowledge, self-efficacy, preparedness, decisional conflict, and decisions to participate in a cancer clinical trial. *Psychooncology* (doi:10.1002/pon.3043).
30. Joffe, S, Cook, EF, Cleary, PD, Clark, JW and Weeks, JC (2001). Quality of informed consent in cancer clinical trials: a cross-sectional survey. *Lancet* **358**: 1772–1777.
31. Joffe, S, Cook, EF, Cleary, PD, Clark, JW and Weeks, JC (2001). Quality of informed consent: a new measure of understanding among research subjects. *J Natl Cancer Inst* **93**: 139–147.
32. Truong, TH, Weeks, JC, Cook, EF and Joffe, S (2011). Outcomes of informed consent among parents of children in cancer clinical trials. *Pediatr Blood Cancer* **57**: 998–1004.
33. Cook, C, Heath, F, and Thompson RL. (2000). A meta-analysis of response rates in web- or internet-based surveys. *Educ Psychol Meas* **60**: 821–836.
34. Shih, TH, Fan, X. (2009). Comparing response rates from web and mail surveys: a meta-analysis. *Educ Res Rev* **4**: 26–40.
35. Slovic, P (1987). Perception of risk. *Science* **236**: 280–285.
36. Fischhoff, B, Lichtenstein, S, Slovic, P, Derby, SL and Keeney, RL. (1981). *Acceptable Risk*. Cambridge University Press: Cambridge MA.
37. Deakin, CT, Alexander, IE and Kerridge, I (2010). The ethics of gene therapy: balancing the risks. *Curr Opin Mol Ther* **12**: 578–585.
38. R Development Core Team. (2012). R: A language and environment for statistical computing. <http://www.r-project.org/>.
39. Zeger, SL, Liang, KY and Albert, PS (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics* **44**: 1049–1060.
40. (2008). *Handbook of Longitudinal Research: Design, Measurement, and Analysis*. Elsevier/Academic Press: Burlington, MA.
41. Rabe-Hesketh, S and Skrondal, A. (2008). *Multilevel and Longitudinal Modeling Using Stata*. Stata Press: College Station, TX.