Meeting FDA Guidance recommendations for replication-competent virus and insertional oncogenesis testing

Kenneth Cornetta,[1,2] Tsai-Yu Lin,[1,2] Danilo Pellin,[3] and Donald B. Kohn[4]

[1]Gene Therapy Testing Laboratory, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA; [2]National Gene Vector Biorepository, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA; [3]Gene Therapy Program, Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Harvard Medical School, Boston, MA, USA; [4]Departments of Microbiology, Immunology and Molecular Genetics, and Pediatrics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Integrating vectors are associated with alterations in cellular function related to disruption of normal gene function. This has been associated with clonal expansion of cells and, in some instances, cancer. These events have been associated with replication-defective vectors and suggest that the inadvertent exposure to a replication-competent virus arising during vector manufacture would significantly increase the risk of treatment-related adverse events. These risks have led regulatory agencies to require specific monitoring for replication-competent viruses, both prior to and after treatment of patients with gene therapy products. Monitoring the risk of cell expansion and malignancy is also required. In this review, we discuss the rational potential approaches and challenges to meeting the US FDA expectations listed in current guidance documents.

INTRODUCTION

Gammaretroviral, lentiviral, and other integrating vectors have been associated with adverse events in animals and humans. Vector integration can disrupt gene expression, or the regulatory regions within the vector can gene alter expression. Clonal cell expansion and leukemias have been documented in clinical trials using gammaretroviral and lentiviral vectors. The risk of an adverse event is related, in part, to the number of vector integrations within a transduced cell population. Therefore, inadvertent exposure to a replication-competent virus arising during vector manufacture would significantly increase the risk of dysregulated cell growth. The US FDA has developed specific guidance documents that list testing and monitoring expectations for investigators conducting gene therapy trials. In this review, we discuss the current monitoring recommendations and the scientific evidence on which these guidelines are based.

RETROVIRUSES AND MALIGNANCIES IN GENE THERAPY

The first approved clinical uses of gene transfer utilized vectors based on gammaretroviruses, including the murine leukemia viruses (MLVs).[1,2] MLVs do not carry oncogenes, but cause malignancy in mice as a consequences of virus integration (insertional oncogenesis or IO). Replacing the gammaretroviral genome with a transgene of interest generates a replication-defective virus (vector) that is capable of high-efficiency gene transfer (Figure 1A). The vector must be packaged into a virion, which is accomplished by independently expressing the viral genes (Figure 1B). The membrane-bound vector particle contains two copies of the vector RNA, and the viral gene regions (*gag* and *pol*) supply the capsid and matrix structural protein, an envelope glycoprotein to facilitate infection of a target cell, a protease to cleave precursor proteins, and reverse transcriptase and integrase to facilitate integration of the vector DNA into the target-cell genome. While the lentivirus genome is more complex, and there are design differences, the approach to generating vectors is similar.

In early clinical trials, treatment-related malignancy was a recognized risk of retroviral vectors, but the risk was believed to be low. This assumption was based on murine studies, which found that multiple integration events in the same cell were required for oncogenesis,[3,4] while retroviral vectors typically resulted in a low number of integrations per cell. A greater risk was believed to be inadvertent exposure of patients to a replication-competent retrovirus (RCR). Recombination was noted in early vector packaging systems, whereby the vector and viral genes recombined, restoring replication competence.[5] These RCRs had properties similar to those of the wild-type virus, including the ability to cause lymphoma in mice.[6] Most concerning, RCRs resulted in lymphoma in immunosuppressed non-human primates.[7] These findings have been the basis for the thorough RCR (and later replication-competent lentivirus or RCL) screening requirements issued by the US FDA and other regulatory bodies.

The assumption that replication competence was required for IO proved incorrect, as early studies using MLV-based retroviral vectors led to secondary leukemias in a subset of patients.[8–11] In most of these cases, the strong enhancer within the long-terminal repeats (LTRs) of
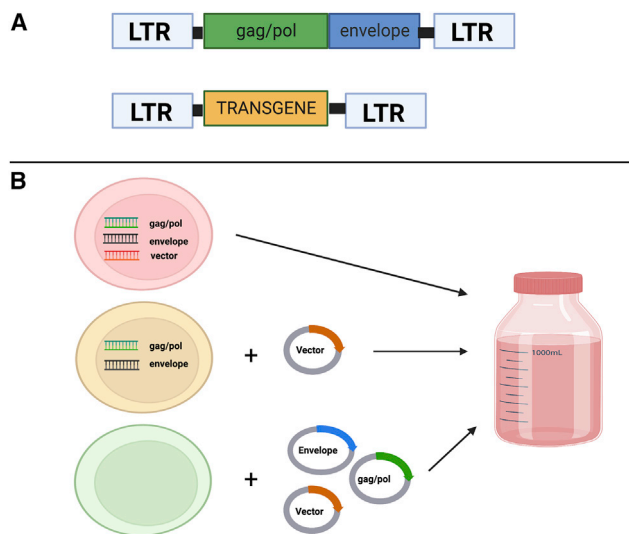
Figure 1. Vector design and manufacture methods

(A) Top: a schematic of a gammaretroviruses shows the gag/pol gene region that encodes the structural proteins and enzymes required for virion formation, reverse transcription, and integration. These membrane-bound viruses also contain a viral envelope glycoprotein, which facilitates entry into the target cell. Bottom: almost all the gag/pol region and envelope sequence can be replaced by the transgene of interest, making the vector replication defective. The long-terminal repeats (LTR) are retained, as they contain the viral promoter and enhancer that drive transgene expression and are also required for integration of the vector into the target-cell DNA. (B) Generation of retroviral particles requires expression of vector and packaging DNA. Packaging cell lines are created by transfecting viral genes into immortalized cell lines and selecting clones that integrate and stably express the DNA. The line may also contain stable integration of the transgene (gene of interest, GOI), or the transgene DNA can be introduced by transient transfection. An increasingly common method of viral vector production is to use transient transfection of all required sequences to manufacture vector particles.

gammaretroviruses, and its interaction with specific primitive cell lineages, appeared to play an important role in cancer development. For example, MLV-based retroviral vectors used in treating X-SCID resulted in T cell leukemias from insertions near *LMO-2*, a gene important in the development of *de novo* human T cell acute lymphocytic leukemia.[8,12] These severe adverse events led to a major change in vector design, and the FDA has addressed monitoring for malignancy within specific guidance documents.[13,14] In this review, we will discuss the current FDA recommendations related to RCR and IO that affect vector and cell product testing as well as patient monitoring.

REGULATORY GUIDANCE FOR RCR TESTING

Acknowledging that vector integrations are associated with severe adverse events, regulators have developed specific guidance for RCR testing. In the United States, the first RCR-specific guidance was published in November of 2006[15] and was replaced in 2020 by *Testing of Retroviral Vector-Based Human Gene Therapy Products for Replication Competent Retrovirus During Product Manufacture and Patient Follow-up: Guidance for Industry.*[13] Figure 2 illustrates the approach
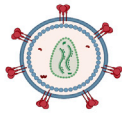
Figure 2. Replication-competent retrovirus (RCR) testing requirements

The current recommendations by the US FDA for testing vector products and *ex vivo*-transduced cells and for monitoring patients. EOP cells, end-of-production cells; *, whichever is less.

to testing. While many in the gene therapy field use "retrovirus" to refer to gammaretroviruses, the FDA guidance uses the broader definition. Therefore, lentivirus, foamy virus, and other viruses belonging to the *Retroviridae* family are subject to the guidance. This guidance is intended to supplement additional recommendations provided in the two other gene therapy documents, *Long Term Follow-Up After Administration of Human Gene Therapy Products: Guidance for Industry*[13] and *Chemistry, Manufacturing, and Control (CMC) Information for Human Gene Therapy Investigational New Drug Applications (INDs): Guidance for Industry.*[16]

CHALLENGES TO RCR DETECTION

To generate an RCR, a retroviral vector must incorporate the necessary genes (structural proteins, reverse transcriptase, and integrase) to restore replication competence. The most likely source will be the genes used in manufacturing vector particles. Given that state-of-the-art vector packaging systems have been designed to minimize homology between vector and viral packaging genes, RCRs will contain multiple sites of recombination and the sites are predicted to have novel sequences. Furthermore, incorporation of endogenous retroviral sequences into an RCR has been reported.[17] Co-infecting viruses could also be co-opted. For these reasons, designing molecular
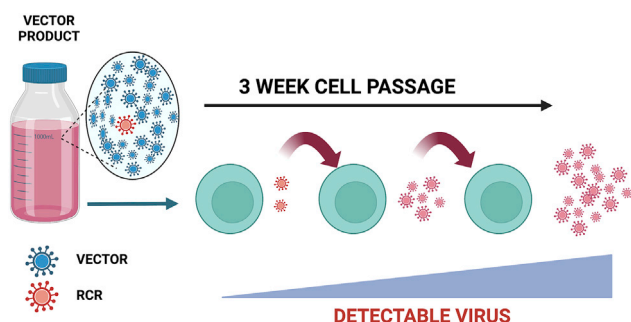
Figure 3. Biologic assays for replication-competent retrovirus detection
The diagram represents an analysis of vector supernatant, which typically contains replication-defective vector particles in excess of 10^9 particles per milliliter. The challenge is to detect an RCR present at very low numbers within the test material. This is facilitated by exposing vector products to permissive cells, which can amplify the RCR to high levels that are easily detectable, while the replication-incompetent vector particles dissipate as the cells expand.

screening assays is challenging, and biologic assays have been the gold standard for detecting a true RCR.

Detecting RCRs in vector products

Detecting RCRs in vector products presents a few unique challenges. First, vector and viral particles contain viral structural proteins, reverse transcriptase, and integrase, so protein analysis is not helpful in RCR detection. While the viral genes required for vector manufacturing are expressed independent of the vector genome, cellular, packaging plasmid, and viral DNAs are released into the supernatant and can also be incorporated into vector particles. This leads to false-positive results in molecular assays such as PCR. DNase treatment of vector products decreases but does not eliminate viral DNA. Therefore, biologic assays are used for testing vector products (Figure 3). This type of assay requires a cell line that has high susceptibility and can amplify the virus to high titer (amplification cells). The assay needs to be tailored for the parent virus (e.g., MLV, HIV-1) and vector pseudotype (envelope). The FDA has recommended a minimum of five cell passages (approximately 3 weeks) of culture to identify slow-growing viruses. Table 1 provides a list of detection methods used in RCR assays.

The second challenge to RCR testing is the volume of material that must be analyzed. The FDA continues to allow testing of 5% of supernatant but also allows an alternative, stating "we now recommend that sufficient supernatant be tested to ensure a 95% probability of detection of RCR if present at a concentration of 1 RCR/dose equivalent."[14] The guidance also recommends that 1% (up to 10^8) of the end-of-production (EOP) cells be tested. For unconcentrated vector products, the volume to be studied can be significant. Concentrated products may have smaller volumes, but pose a greater chance of receptor interference, whereby the replication-defective vector particles overwhelm cell receptors, preventing an RCR from entering the amplification cells. When testing for RCR, a high ratio of amplification cells to the test article (supernatant or EOP cells)

and diluting the concentrated vector to a defined level are important for maximizing the sensitivity of RCR detection. Concentrated vector can also prove toxic to amplification cells, and appropriate dilutions for RCR testing must be determined experimentally. Therefore, investigators seeking to utilize a new retrovirus or pseudotype should begin developing an appropriate assay early in vector development.

Ex vivo cell products

Significant changes have been made in the 2020 FDA RCR guidance compared with the prior guidance. Now, all *ex vivo* products must be tested regardless of the culture period prior to infusion. A biologic assay can be utilized, and the new guidance now permits the use of sensitive molecular assays such as PCR "particularly, when time constraints are present." While not specifically referenced in the guidance, the Agency has recognized the potential for false-positive PCR results due to the carryover of packaging DNA (plasmid or integrated viral DNA sequences) in the vector supernatant. The amount of packaging DNA will vary among the manufacturing methods and can vary from lot to lot of the final cell product. When RCR is absent, the amount of packaging DNA that will be carried over to *ex vivo* transduced cells is highest at the time of transfection and will decrease as the cells expand in culture. It is assumed that an RCR will increase during cell culture, and regulators have accepted a decrease in packaging DNA overtime as evidence of an RCR-negative cell product. We recommend testing cell products early and late in the *ex vivo* expansion, which has reliably documented a decrease in packaging DNA. If a single sampling is planned, careful preclinical studies must document consistently negative PCR products. For protocols with short culture periods where multiple sampling is not possible, continued culture of a subset of cells after the cell product is harvested may be required. This is feasible for products that will be frozen; for fresh products to be administered directly after culture, early discussion with the FDA should be initiated to come to consensus on meeting RCR testing expectations. Investigators should also note that the FDA guidance does leave open the possibility of discontinuing RCL testing after "accumulated manufacturing and clinical experience that demonstrates that your transduced cell product is consistently RCR-negative." The number of RCL-negative products is not specified and will need to be negotiated with the Agency. A goal of 10 negative lots may be a starting point for the discussion.

Patient monitoring

The 2020 RCR guidance has decreased the requirements for RCR monitoring, and an algorithm is illustrated in Figure 2. The preferred methods are serologic- or molecular (PCR)-based assays. A potential pitfall of the qPCR methods is a false-negative result due to mutation or replacement of the target sequence in the RCR. Challenges with the serology method include lack of antibody development in immune-compromised patients or, conversely, antibody development to vector proteins infused during the initial treatment.

Table 1. Summary of assays for replication-competent retrovirus and lentivirus

Authors	Year	Summary	Reference
Fischinger et al.	1968	S+/L− cell culture assay: only replication-competent MLV infection of S+/L− cells can support MSV to change cell morphology and cause focus formation	PMID: 4298946
Klement et al.	1969	XC cell rapid assay: MLV infection and replication in XC cell line results in rapid development of cytopathic effect (CPE)	PMID: 4186808
Bassin et al.	1970	S+/L− cell culture assay: only replication-competent MLV infection of S+/L− cells can support MSV to change cell morphology and cause focus formation	PMID: 4319233
Bassin et al.	1971	XC cell rapid assay: MLV infection and replication in XC cell line results in rapid development of CPE	PMID: 4925356
Bassin et al.	1982	extended S+/L− cell culture assay	PMID: 6959413
Kimpton et al.	1992	uses CD4-LTR/β-gal reporter HeLa cells to detect Tat-containing RCL	PMID: 1548759
Cornetta et al.[6]	1993	RCR detection using PCR for vectors propagating in 3T3 cells	PMID: 8280796
Chang et al.	1995	developed an immunohistochemical staining for HIV-1 p24 in indicator cells for RCL that lacks Tat accessory protein	PMID: 7645208
Forestell et al.	1996	marker assay using LacZ-expressing retroviral vector in *Mus dunni* cells	PMID: 8844623
Martineau et al.	1997	PCR and an ELISA to detect RCR by targeting anti-MLV-specific immunoglobulin in patients' samples	PMID: 9215740
Kim et al.	1998	developed an immunohistochemical staining for HIV-1 p24 in indicator cells for RCL that lacks Tat accessory protein	PMID: 9571629
Long et al.	1998	RCR detection using PCR for MLV in patients' peripheral blood leukocytes	PMID: 9625255
Chang et al.	1999	RCL detection using immunostaining of capsid and p24 ELISA	PMID: 10505094
Hugin et al.	1999	immunochemical detection assay for RCR-infected foci on cell monolayers	PMID: 10092145
Long et al.	1999	PCR assay to detect RCR in autopsy and biopsy specimens from patients	PMID: 10210141
Nemunaitis et al.	1999	PCR assay for detecting RCR in patients	PMID: 10365660
Metharom et al.	2000	marker rescue assay in combination with a reverse transcriptase assay	PMID: 10894263
Chen et al.	2001	noted false-positive results from carryover of packaging DNA in vector products	PMID: 11259201
Reeves et al.	2002	comparison of S+/L− and marker rescue assays for detecting RCR	PMID: 12396630
Escarpe et al.	2003	biologic assay for detecting RCL using C8166 cells and p24 ELISA	PMID: 12907156
Sastry et al.	2003	biologic assay for RCL detection using C8166 cells, PCR, and p24 ELISA	PMID: 14599817
Segall et al.	2003	detection of RCL using marker rescue assay and p24 ELISA	PMID: 12842435
Uchida et al.	2004	real-time RT-PCR to detect RCR from permissive cells infected by RCR samples; PEI beads were used to enrich RCR in the culture supernatants, where	PMID: 15536044

(Continued on next page)

Table 1. Continued

Authors	Year	Summary	Reference
		RCR RNA was extracted before qRT-PCR quantification	
Sastry et al.	2005	product-enhanced reverse transcriptase (PERT) assay for RCR and RCL	PMID: 16218784
Miskin et al.	2006	PERT assay for RCL detection of equine infectious anemia virus RCL	PMID: 16208418
Bauer et al.	2008	*in vivo* immune-deficient mouse model to assess RCR and RCL	PMID: 18461052
Cornetta et al.	2011	RCL assay screen of clinical lentiviral vector products using PCR and p24 ELISA	PMID: 21179010
Farley et al.	2012	qRT-PCR assay for RCL detection of equine infectious anemia virus	PMID: 23121195
Aloia et al.	2013	reporter inGluc-MLV-DERSE assay for detecting RCR	PMID: 22402321
Farley et al.	2015	RCL assay for dendritic cell-targeting lentiviral vectors using DC-SIGNR-expressing 293 cells and p24 ELISA and PERT	PMID: 26029728
Corre et al.	2016	"RCL-pooling assay" for detecting RCL in vector batches using sequential pooling	PMID: 26886834
Cornetta et al.[18]	2018	extended S+/L− cell culture assay to detect RCR	PMID: 30211249
Skrdlant et al.	2018	qPCR assay for RCL by targeting VSV-G	PMID: 29034262
Gludish et al.	2020	constructed TZM-GFP cells for analysis of rare and early HIV-1 infection	PMID: 33199722
Stuelke et al.	2020	ultrasensitive p24 digital ELISA viral outgrowth assay for measuring RCL	PMID: 32849659

Assays developed for the detection of replication-competent retrovirus and lentivirus are listed in chronological order. Assays for latent lentivirus are not included.

RCR testing is also discussed in the FDA *Guidance for Long-Term Follow-up*, which recommends 15-year monitoring for individuals treated with integrating vectors. This includes both clinical trial participants and those treated with a licensed product. If a patient develops evidence of malignancy or a clinical syndrome where retroviral infection is a consideration (e.g., fever; headache; muscle aches and joint pain; rash; persistent sore throat and painful mouth sores; swollen lymph glands, mainly on the neck; diarrhea; or unexplained weight loss), the RCR testing should be performed to exclude this as an etiologic agent. It is possible that a shorter monitoring period might be considered by the Agency for a non-integrating retroviral vector if lack of long-term persistence can be documented.

RESULTS OF RCR TESTING

When manufacturing vector products, initial clinical trials using gammaretroviral vectors utilized packaging cell lines that were prone to RCR development. Improvements in vector design and packaging systems have greatly decreased RCR development. One example is the PG13 cell line, a murine packaging cell line that generates vector particles expressing the gibbon ape leukemia virus (GALV) envelope.[19] The Indiana University Vector Production Facility has certified 29 vector products using this line, all of which passed RCR screening (D. Bischof, personal communication). While the FDA has stated a preference for packaging cell lines,[14] given their potential to provide consistent batch-to-batch final vector products, the field has been moving toward transient vector production methods for gammaretroviral vectors. Transient methods decrease the cost and time of manufacturing by eliminating the need for clone isolation and generation of a master cell bank. Eliminating these steps also decreases the chance of recombination and, to date, RCRs have not been reported. For lentiviral vectors, an RCL is highly unlikely, given the vector design, packaging system design, and frequent use of transient transfection methods.[20] RCLs remain a theoretical concern, as none have been reported to date.

The experience with screening *ex vivo*-transduced cell products supports the safety of current manufacturing methods in generating RCR-free vector. The NHLBI-funded National Gene Vector Biorepository (NGVB) screened 282 gammaretroviral vector-transduced cell products administered in 14 clinical trials; all screened negative for RCR.[18] The NGVB also reported negative RCL screening of 460 lentiviral vector-transduced cell products from 375 subjects enrolled in 26 clinical trials.[21] June and colleagues reported the absence of RCL in 17 vector lots used in manufacturing 375 T cell products given to 308 patients.[22] The NGVB also provides post-trial monitoring by qPCR for RCR/RCL exposure, and 4,543 samples from over 70 clinical trials have been negative. At this writing, there has been no report of RCR or RCL exposure in patients participating in human clinical gene therapy trials.
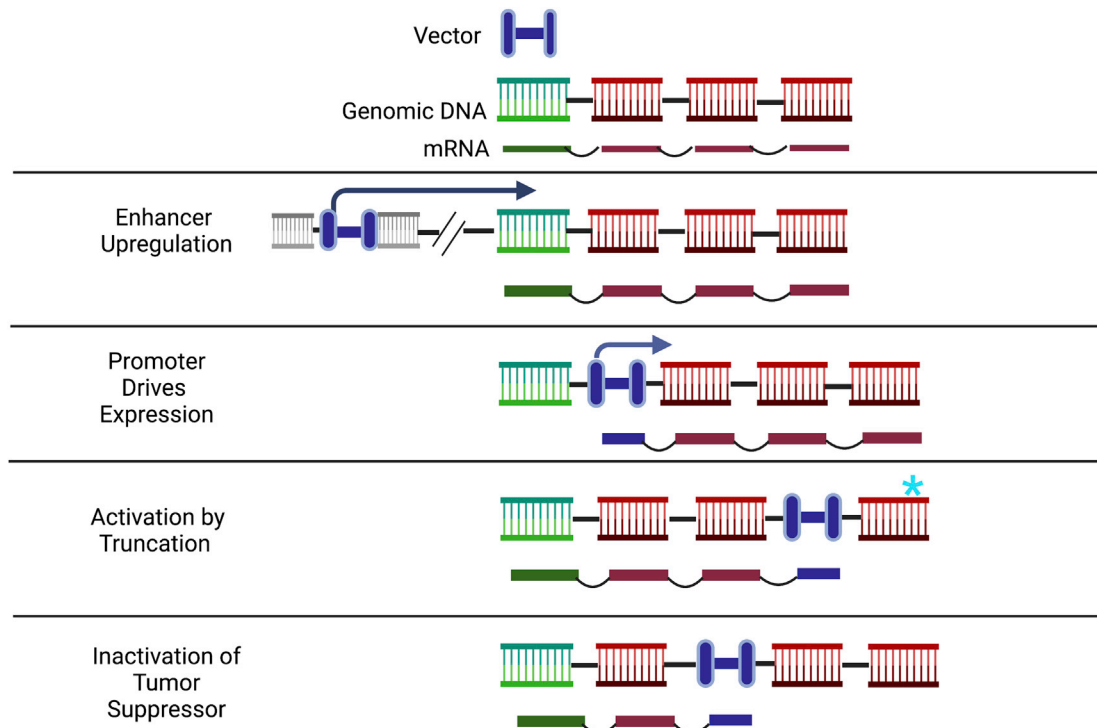
Figure 4. Mechanisms for insertional alteration of gene expression

The most common mechanism for insertional oncogenesis has been oncogene promoter upregulation by the enhancer element within gammaretroviral vectors. While vector promoters driving oncogene transcription have been reported in murine cancers, this has not been noted in human cancers arising after vector insertion. Activation by truncation has been shown after insertion led to the truncation of mRNA, removing a let-7 site that normally regulated gene expression (noted by asterisk). This led to transient clonal expansion without cancer formation. Integration and disruption of gene expression has also inactivated a tumor-suppressor gene, which led to clonal expansion without malignancy in a patient with an inherited hypomorphic mutation in the non-transduced allele. Green, promotor region; red, exon.

INSERTIONAL ONCOGENESIS

Retroviruses may cause oncogenic transformation by "hijacking" cellular proto-oncogenes (e.g., acutely transforming sarcoma viruses) or expressing proteins that alter intracellular signaling, cell-cycle regulation, and/or DNA repair (e.g., TAX protein in human T cell leukemia viruses that cause human adult T cell leukemia).[23] In contrast, the gammaretroviruses, such as MLV, cause leukemia by the process of IO. MLV tends to integrate near the 5′ end of transcriptionally active genes, where the potent enhancer elements they contain in their LTR may transactivate transcription of the adjacent host cell genes (Figure 4). Dysregulated expression of one or more cellular proto-oncogenes can initiate a process that leads to clonal expansion, cellular transformation, and development of hematological malignancies. Integrating vectors may also integrate into tumor-suppressor genes and disrupt their function or interfere with normal processing of cellular gene transcripts, altering their activity (see below). In some cases, transactivation of a single proto-oncogene may be sufficient to initiate this cascade, but multiple integrants activating different signaling pathways may cooperate to cause transformation.[24] The potential for multiple integrants in the same cell to increase transformation risks underlies the regulatory mandate to use transduction conditions that minimize high vector copies per cell, although the optimal limits are not well defined.

To date, all of the clinical trials reporting secondary malignancies have utilized gammaretrovirus-based vectors, with the exception of one trial for the treatment of cerebral adrenoleukodystrophy that used a lentiviral vector (Table 2).[25,26] A common factor in reported cases is the use of the MLV promoter/enhancer to drive gene expression, including the lentiviral vector in question. Incorporating the MLV promoter/enhancer into lentiviral vectors has been shown to increase *in vitro* immortalization of murine progenitor cells[27] and was associated with the development of malignancy in a non-human primate.[28] Together, these findings support the MLV LTR with its potent enhancer elements as a major factor in IO. In fact, a clinical trial of gene therapy for X-SCID safely used a gammaretroviral vector in which the LTR enhancer elements were self-inactivated ("SIN" vector); no clonal expansion or clinical leukoproliferation has occurred in more than 8 years of follow-up.[29]

Unlike gammaretroviruses, HIV-1 has a significantly lower incidence of transformation.[30] HIV-1-associated malignancies are mostly due to its immunosuppressive activities (e.g., non-Hodgkin's lymphoma from outgrowth of EBV-transformed B cells, HHV-8-transformed endothelial cells in Kaposi's sarcoma, or cervical cancer associated with human papilloma virus). The HIV-1 accessory proteins that

Table 2. Reported events of clonal expansion or frank leukoproliferation in clinical gene therapy trials using hematopoietic stem cells

Condition	Vector	Cases/total	Gene(s)	Mechanism	Consequence	Reference
SCID-X1	RV: MLV-*IL2RG*	6/20	*LMO2*	LTR *trans*-activation	T ALL	PMID: 18688285
			CDKN2A			PMID: 18688286
XCGD	RV: SFFV-*CYBB*	3/3	*MECOM*	LTR *trans*-activation	MDS/AML	PMID: 20098431
WAS	RV: CMMP-*WASP*	7/10	*LMO2*	LTR *trans*-activation	ALL/AML	PMID: 24622513
			MECOM			
β-Thalassemia	LV: β-globin LCR-*HBB* (cHS4 insulator in LTR)	1/1	*HMGA2*	HMGA2 transcript truncation from splicing into insulator	transient clonal expansion	PMID: 20844535
ADA SCID	RV: GIADAl (MLV-*ADA*)	1/≥22	*LMO2*	LTR *trans*-activation	T ALL	https://ir.orchard-tx.com/news-releases/news-release-details/orchard-statement-strimvelisr-gammaretroviral-vector-based-gene
X-ALD	LV: lenti D (MND LTR-*ABCD1*)	3/67	*MECOM*	LTR *trans*-activation	MDS/AML	https://investor.bluebirdbio.com/news-releases/news-release-details/bluebird-bio-provides-update-severe-genetic-disease-programs
SCID-X1	LV: Cl20-i4-EF1α-hγcOPT (cHS4 insulator in LTR)	8/8	*HMGA2*	HMGA2 transcript truncation from splicing into insulator	transient clonal expansion	PMID: 35764638

mediate immunosuppression are removed from lentiviral vectors, eliminating the putative oncogenic potential of HIV-1 in current lentiviral vector systems. Moreover, lentiviral vectors generally lack strong enhancers and utilize the SIN LTR, further improving their safety profile.[20]

Vector insertions that alter cell growth do not always result in malignancy. Fraietta et al. reported transient expansion of a T cell clone transduced with a lentiviral vector expressing a chimeric antigen receptor (CAR-T).[31] The patient had a germline hypomorphic mutation of the methylcystosine dioxygenase TET2 allele, and the vector insertion disabled the remaining wild-type allele of this tumor-suppressor gene. Cavazzana-Calvo et al. reported a clonal expansion of erythroid progenitors in a patient treated with a lentiviral vector for thalassemia.[32] In this case, a vector integrant within an intron of the *HMGA2* gene led to splicing of the *HMGA2* transcript to a cryptic splice site in an insulator element in the LTR. There was transcription termination from the 5′ LTR polyadenylation signal and elimination of a 3′ let-7 binding site normally contained in the 3′ untranslated region of the *HMGA2* transcript, preventing normal microRNA regulation of the *HMGA2* transcript. The clone eventually dissipated and did not become transformed. A recent report observed a similar phenomenon in patients with X-linked SCID treated with a lentiviral vector also containing an insulator in the LTRs, which caused premature termination of *HMGA2* transcripts and transient clonal expansion.[33] Recent adverse events in a clinical trial for sickle cell anemia point out multiple factors important to consider in IO risk, including the underlying disease, the preparative regimen, and the vector insertion.[34] In this study, two individuals developed acute myeloid leukemia/myelodysplasia (MDS) but only one patient had vector in the leukemic cells.[35] Investigations into the mechanisms involved in these events are ongoing. Altered cell growth does not appear unique to retroviral or lentiviral vectors. Clonal cell expansion has been documented in ca-

nines treated for hemophilia with a factor VIII-expressing adeno-associated virus (AAV) vector.[36] Moreover, certain regulatory elements in AAV vectors have the ability to cause liver tumors in mice, although the relevance to human gene therapy is being evaluated.

It is important to point out that malignancies have been reported in only one cell target, specifically, the hematopoietic stem and progenitor populations. In contrast, gammaretroviral vectors have been used extensively in the modification of peripheral blood T cells (e.g., CAR-T cells) without reports of secondary cancers. The risk to other cell populations, and what proto-oncogenes would be at risk for IO in different cell types, is currently unknown. Understanding this risk will be important, particularly for *in vivo* administration of integrating vectors.

Insertional oncogenesis assays for preclinical vector assessments

Given the different risks inherent to different vector configurations and cellular targets, preclinical genotoxicity studies are an essential component of clinical gene therapy development. Unfortunately, a reliable assay that predicts IO risks in humans has yet to be developed. *In vivo* assays have used transplantation of transduced murine or human cells into mice to assess the biosafety of vectors in the hematopoietic system. At the completion of in-life observations, typically 4–6 months, the mice are necropsied, with organs examined by gross appearance and histopathology, and hematopoietic cells from blood, marrow, spleen, and thymus are examined by flow cytometry to assess lineage differentiation, with vector copy number and vector integration site analyses performed to assess patterns of integration sites and possible clonality. Marrow may be serially transplanted to secondary recipients, which may increase the sensitivity for detecting expanding clones, although serial transplants may be inefficient using clinically relevant human bone marrow and peripheral blood stem cells. It is essential to be able to discriminate malignancies that arise
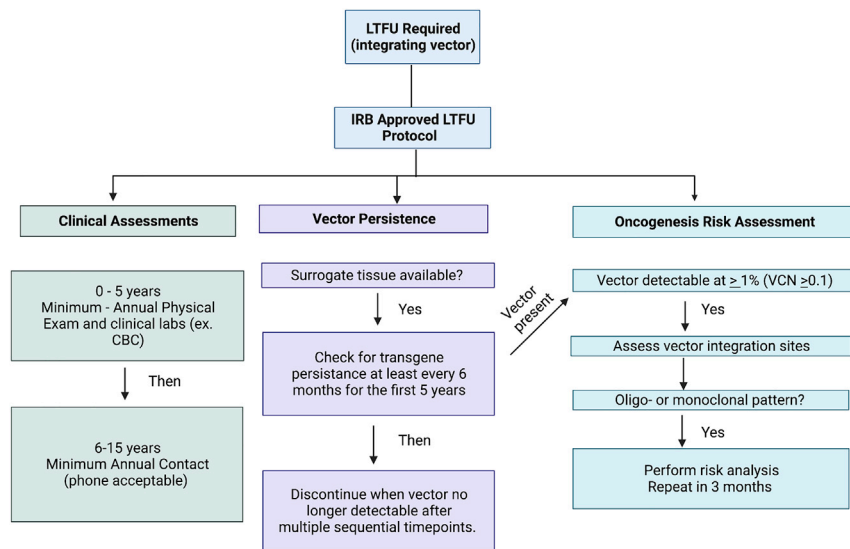
Figure 5. Long-term follow-up (LTFU) monitoring plan
An example algorithm for monitoring patients for insertional oncogenesis.

transformation.[40] This surrogate assay for geno-toxicity assessment (SAGA) is a robust and sensitive assay and reliably predicted the mutagenic risk for previously used vectors that had caused leukoproliferative severe adverse events in clinical trials. The role the assay will play in predicting IO risk with new vector systems and new cell targets will require further study.

Clinical monitoring for IO

A suggested schema for monitoring IO is shown in Figure 5 and is based on a 2013 FDA guidance.[37] As noted above, methods for assessing risk vary, and the schema will need to be reviewed by the governing regulatory authorities. When hematopoietic cells are the target, monitoring generally focuses on detecting leukemia or MDS. For 15 years, patients should be periodically questioned about recurrent fever, frequent infections, bleeding, and swollen lymph nodes and examined for evidence of hepatosplenomegaly or lymphadenopathy. Complete blood cell counts should be assessed for hyperleukocytosis, the presence of abnormal white blood cells, or new onset of anemia or thrombocytopenia. If abnormal blood cell counts (e.g., WBC count >20,000/μL or presence of blast cells) or clinical signs and symptoms suggest the potential presence of leukoproliferation or MDS (as above), a bone marrow analysis should be done, including histopathology, flow cytometry, karyotype, next-generation sequencing of a panel of genes related to clonal hematopoiesis and leukemia/MDS, and vector integration site analysis. Additional studies may be performed if abnormalities are found, using FISH, CGH, RNA-seq, or other relevant molecular analyses.

Standard clinical criteria for diagnosing leukemia and the specific subtype may be used; patients should be referred to the appropriate hematology/oncology physicians for treatment. Identification of a clonal expansion (>20% clonal frequency) or diagnosis of leukemia or MDA needs to be reported to the appropriate regulatory authorities.

At least in treatments using hematopoietic stem cells, the only cell type observed to undergo vector-related transformation, vector integration site analysis should be performed from blood and/or bone marrow using shear extension PCR[41] or a similar robust method on a routine schedule (e.g., 6, 12, 18, and 24 months after gene therapy) or upon abnormal findings as described above. Presence at any time of a clonal vector integrant at >10%–30% frequency (e.g., 10%–30% of all productive sequence reads or of unique sequences when using shearing methods) should trigger further investigations to determine if it represents a progressive clonal expansion. Complete blood cell

from murine donor gene-marked cells from the relatively frequent host cell malignancies that may arise, especially when recipients undergo total body radiation conditioning. One standard approach is to transduce and transplant marrow that contains a marker to distinguish the donor cells from the recipients, e.g., male into female, measuring X chromosome sequences, or between CD45.2/CD45.1 congenic mice. Alternative approaches include: (1) the use of tumor-prone mice (e.g., the cdkn2a$^{-/-}$ mouse), which may increase sensitivity but is complicated by high background rates of transformation and (2) xenografting of transduced human cells into an immune-deficient mouse host to assess effects on differentiation and possible clonal expansions. In general, these assays are not very sensitive and often fail to detect clonality or malignancy, even when using gammaretroviral vectors as positive controls. While the FDA has issued a guidance that outlines an approach to preclinical safety studies,[37] the document lacks specific recommendations on how these studies are designed or the numbers of mice to be studied.

In an attempt to increase sensitivity, Ute Modlich and colleagues in Hannover, Germany, developed an *in vitro* insertional mutagenesis (IVIM) assay to quantify the transforming potential of vectors in hematopoietic stem and progenitor cell targets.[27,38] This assay uses murine bone marrow enriched for stem/progenitor cells by depleting the cells expressing markers of lineage differentiation (lin$^-$ cells). This assay provides quantitative information on the relative transforming activity of different vector constructs but is largely limited to potential for myeloid cell transformation. Nevertheless, it does provide a relatively straightforward assay that can provide a good assessment of one aspect of vector safety. Zhou et al. developed a similar *in vitro* transformation assay to evaluate T lineage transformation using murine thymocytes.[39] More recently, Hannover investigators have developed a more advanced cell-culture-based transformation assay that analyzes the gene expression signature of the transduced cells to identify those they characterized using machine learning as representing

Review

count, physical examination, and integration site testing should be repeated and a bone marrow analysis done as detailed above.

There is no specific requirement for interval monitoring of integration site during long-term follow-up in the absence of clinical indication. If there were a prominent clone present at the last study-mandated test (e.g., 1%–10% of total adjacent to a known gene of risk at 24 months end of study visit), it may be prudent to follow it annually until it is known to be stable or decreasing.

Bioinformatic assessments of IO

Integration site analysis (ISA) is an essential tool in meeting the FDA recommendations for IO monitoring. The assay can be viewed as the combination of three main steps requiring highly interdisciplinary teams: a lab protocol for amplifying vector-genome junctions found in a population of transduced cells,[42–48] sequencing of a library of DNA amplicons, and bioinformatics analysis.[49–53]

In the past two decades, protocols to perform ISA have been continually updated by incorporating novel techniques and strategies to improve their reliability and efficiency. The genomic DNA isolated from a clinical sample is fragmented using a cocktail of restriction enzymes or, as recently proposed, by sonication to provide more unbiased coverage of the host genome.[42–44] A combination of adapters and linkers are then ligated to the DNA fragments generated, and the genomic region flanking integration sites are selectively amplified through PCR using primers specifically designed to target the provirus LTR and the linker sequences. Next-generation sequencing (NGS) had a transformative impact on the field of ISA. In addition to increasing the number of distinct integration sites retrieved, by introducing oligonucleotide barcodes in the amplicon design, NGS allows for identifying DNA fragments derived from different biological samples or technical replicates within a single run. Indexing strategies using oligos containing random sequences (unique molecular identifiers or UMIs) can also be leveraged to aid the quantitation of integration-site frequency in a population of cells.[50,51,54,55] Over time, Illumina platforms have become the NGS technology employed by most ISA protocols, triggering the need for novel and more sophisticated bioinformatic pipelines capable of managing short paired-end sequencing readouts made of tens of millions of reads. The operations performed by most bioinformatic pipelines can be divided into three main tasks: preprocessing of reads, mapping to the reference genome, and identification of integration site coordinate and clone abundance estimation.

Significant progress has been made in improving ISA accuracy over the years and currently, it provides the most comprehensive screening method to detect IO. However, reproducibility issues, high sensitivity to contamination, and PCR biases are still relevant and, in clinical settings, independent tests, such as qPCR, usually verify the contribution of IS clones potentially associated with IO events. While a complete survey of the PCR protocols and bioinformatics analysis tools used in the ISA field is beyond this review's scope, the reader is directed to the references mentioned below.

Interpreting ISA data

A primary goal of ISA is to investigate the presence of a persistent monoclonality and clonal expansion, defined as the increase over multiple time points of a clone harboring a particular vector integration site. Indeed, the longitudinal analysis of ISA results derived from multiple patients and cell types allows for a more comprehensive characterization of the *in vivo* clonal dynamics of the gene therapy product, as demonstrated in several publications.[56–61]

In gene therapy, a clone is defined as a group of cells deriving from a common and unique transduced cell and therefore sharing the same integration site. Clone size or abundance is the relative frequency of the cells belonging to a particular clone in the total number of vector-containing cells found in a sample and is estimated based on ISA readouts.[50–52,54,55] The definition of clonality differs across the different fields of biology, and in the context of gene therapy, it involves both clonal richness, the number of distinct clones present in the system, and evenness, how homogeneous the clones are in size.

If vector integrations were to have limited or no impact on gene expression and the cell product infusion results in a successful engraftment, we would expect ISA to reveal a highly polyclonal population of cells. This scenario consists in a large number of clones having similar sizes. Many factors can have an impact on clonal dynamics. In addition to alteration in individual clone growth caused by gene dysregulation due to vector insertion, factors such as the underlying disease, patient's genetics, prior or concurrent therapies, cell dose, and preexisting clonal patterns can lead to oligoclonal or even monoclonal configurations. Our understanding of the determinants and risks associated with these scenarios is still unclear, and although they may or may not represent the early stages of malignancy, if persistent, they need to be reported to regulatory bodies and the patient needs to be closely monitored.

Currently, there is no consensus on a specific numerical index or threshold for the definition of oligo- and monoclonality based on ISA.[62,63] A meaningful evaluation of the clonality must consider several factors, such as the dynamic component of clonality evolution, follow-up time point, and patient's clinical history and course. Similarly, there are no formal criteria for discriminating between normal clonal growth trajectories and expansion. According to the FDA guidelines, clone expansion consists of detecting a clone with an increasing frequency in two or more consecutive ISAs. However, the relevance of such a situation dramatically depends on the overall clonality, time point, cell source, importance of the cell type to the disease, and results of other treatment endpoints (e.g., vector copy number, target protein levels). ISA is a fundamental tool for assessing the outcome of gene therapy applications and monitoring their safety. Still, the interpretation of ISA results needs to be critically reviewed based on clinical parameters and validated by independent assays when possible. As our understanding of *in vivo* clonal dynamics deepens, guidelines should be updated and incorporate the novel knowledge available. The definition of context-aware rules that consider patient and treatment information and clinical course for

recognizing oligo- and monoclonality and clonal expansion would benefit the field of gene therapy.

The above recommendations are focused on hematopoietic cell targets, since retroviral vectors have been the most common vector systems utilized in these cells and adverse events have been observed. As new target cells are engineered (e.g., mesenchymal stem cells, induced pluripotent stem cells, and others), the risk of IO will need to be assessed. Moreover, what type of malignancy might be observed is unknown and will need to be considered in developing long-term monitoring plans.

OTHER CONSIDERATIONS

Unlike FDA regulations, which are codified in the *Code of Federal Regulations* (CFR) and are legally binding, FDA guidance documents state that they represent the Agency's "current thinking," which leaves open the possibility of alternatives. The FDA encourages development and discussion of an RCR and IO monitoring plan early in clinical trial planning (at the pre-IND stage). Changes are possible to monitoring requirements, but they should be submitted to the FDA for review and concurrence prior to implementation. Negative RCR testing results can be submitted with the IND annual report or in a *Development Safety Update Report*, but positive RCR testing results for any treated patient must be reported immediately in an IND safety report (21 CFR 312.32). While the risk of IO must be addressed in clinical trials using integrating vectors, robust assessment methods are lacking. Therefore, early engagement with regulatory agencies to define the type of safety studies suitable for the intended clinical application is prudent. Reporting IO results should include information about the method used, the sensitivity, and the frequency of prominent or expanding clones (>10%–30% of recovered integration sites). These findings merit reporting to FDA for discussions of follow-up monitoring plan.

CONCLUSIONS

The known adverse events with retroviral vectors have warranted careful monitoring of patients receiving this class of vectors. We now have decades of experience with these vectors, and to date, there has been no evidence of patient exposure to RCR, indicating that manufacturing methods have greatly decreased the risk of RCR development. Early use of gammaretroviral vectors led to leukemia in a subset of patients, but experience with lentiviral vectors and SIN retroviral vectors demonstrate that IO risk can be greatly decreased when the MLV enhancer is removed from vector constructs. While the FDA guidance documents have opened the door for reduced RCR/RCL monitoring, new vector systems and cell targets will likely require continued extensive monitoring for RCR and IO events. Ideally, data from adverse event investigations will be reported in scientific journals, and the NIH should consider establishing a database to assist research efforts in preventing future events. Finally, these recommendations are those of the authors based on their interpretation of FDA guidance documents and publicly available information. They are meant to assist in developing a monitoring plan and initiating discussion with regulatory agencies. Regulators may require

modifications to a proposed plan given the perceived risk of the vector type, vector transgene, target cell, method of manufacture, and patient population.

AUTHOR CONTRIBUTIONS

All authors contributed to the conceptualization, writing, review, and editing of this article.

DECLARATION OF INTERESTS

Indiana University has licensed technology to Charles River Laboratories and Genezen Labs of which K.C. and T-Y.L. receive a portion of the fees. D.B.K. is a paid scientific advisory board member for Allogene Therapeutics, ImmunoVec, Pluto Therapeutics, MyoGene Bio, and Innoskel and an *ad hoc* consultant for Cimeio Therapeutics, TransformaTx, and bluebird bio.

REFERENCES

1. Blaese, R.M., Culver, K.W., Miller, A.D., Carter, C.S., Fleisher, T., Clerici, M., Shearer, G., Chang, L., Chiang, Y., Tolstoshev, P., et al. (1995). T lymphocyte-directed gene therapy for ADA-SCID:initial trial results after 4 years. Science *270*, 475–480.

2. Rosenberg, S.A., Aebersold, P., Cornetta, K., Kasid, A., Morgan, R.A., Moen, R., Karson, E.M., Lotze, M.T., Yang, J.C., Topalian, S.L., et al. (1990). Gene transfer into humans-immunotherapy of patients with advanced melanoma, using tumor infiltrating lymphocytes modified by retroviral gene transduction. N. Engl. J. Med. *323*, 570–578.

3. Cuypers, H.T., Selten, G., Quint, W., Zijlstra, M., Maandag, E.R., Boelens, W., van Wezenbeek, P., Melief, C., and Berns, A. (1984). Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. Cell *37*, 141–150.

4. van Lohuizen, M., Verbeek, S., Krimpenfort, P., Domen, J., Saris, C., Radaszkiewicz, T., and Berns, A. (1989). Predisposition to lymphomagenesis in pim-1 transgenic mice: cooperation with c-myc and N-myc in murine leukemia virus-induced tumors. Cell *56*, 673–682.

5. Muenchau, D.D., Freeman, S.M., Cornetta, K., Zwiebel, J.A., and Anderson, W.F. (1990). Analysis of retroviral packaging lines for generationof replication-competent virus. Virology *176*, 262–265.

6. Cornetta, K., Nguyen, N., Morgan, R.A., Muenchau, D.D., Hartley, J.W., Blaese, R.M., and Anderson, W.F. (1993). Infection of human cells with murine amphotropic replication-competent retroviruses. Hum. Gene Ther. *4*, 579–588.

7. Donahue, R.E., Kessler, S.W., Bodine, D., McDonagh, K., Dunbar, C., Goodman, S., Agricola, B., Byrne, E., Raffeld, M., Moen, R., et al. (1992). Helper virus induction T cell lymphoma in nonhuman primates after retroviral mediated gene transfer. J. Exp. Med. *176*, 1125–1135.

8. Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science *302*, 415–419.

9. Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D., 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.

10. Braun, C.J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., Modlich, U., Beier, R., Göhring, G., Steinemann, D., et al. (2014). Gene therapy for Wiskott-Aldrich syndrome–long-term efficacy and genotoxicity. Sci. Transl. Med. 6, 227ra33.

11. Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kühlcke, K., Schilz, A., Kunkel, H., 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.

12. Schwarzwaelder, K., Howe, S.J., Schmidt, M., Brugman, M.H., Deichmann, A., Glimm, H., Schmidt, S., Prinz, C., Wissler, M., King, D.J.S., et al. (2007). Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution in vivo. J. Clin. Invest. 117, 2241–2249.

13. United States Food and Drug Administration (2020). Long Term Follow-Up after Administration of Human Gene Therapy Products: Guidance for Industry (U.S. Food and Drug Administration).

14. U.S Food and Drug Administration (2020). Testing of Retroviral Vector-Based Human Gene Therapy Products for Replication Competent Retrovirus during Product Manufacture and Patient Follow-Up (Guidance for Industry).

15. U.S Food and Drug Administration (2006). Guidance for Industry- Supplemental Guidance on Testing for Replication Competent Retrovirus in Retroviral Vector Based Gene Therapy Products and during Follow-Up of Patients in Clinical Trials Using Retroviral Vectors.

16. U.S Food and Drug Administration (2020). Chemistry, Manufacturing, and Control (CMC) Information for Human Gene Therapy Investigational New Drug Application (INDs) (Guidance for Industry).

17. Chong, H., Starkey, W., and Vile, R.G. (1998). A replication-competent retrovirus arising from a split-function packaging cell line was generated by recombination events between the vector, one of the packaging constructs, and endogenous retroviral sequences. J. Virol. 72, 2663–2670.

18. Cornetta, K., Duffy, L., Feldman, S.A., Mackall, C.L., Davila, M.L., Curran, K.J., Junghans, R.P., Tang, J.Y., Kochenderfer, J.N., O'Cearbhaill, R., et al. (2018). Screening clinical cell products for replication competent retrovirus: the national gene vector biorepository experience. Mol. Ther. Methods Clin. Dev. 10, 371–378.

19. Miller, A.D., Garcia, J.V., Von Suhr, N., Lynch, C.M., Wilson, C., and Eiden, M.V. (1991). Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. J. Virol. 65, 2220–2224.

20. Schambach, A., Zychlinski, D., Ehrnstroem, B., and Baum, C. (2013). Biosafety features of lentiviral vectors. Hum. Gene Ther. 24, 132–142.

21. Cornetta, K., Duffy, L., Turtle, C.J., Jensen, M., Forman, S., Binder-Scholl, G., Fry, T., Chew, A., Maloney, D.G., and June, C.H. (2018). Absence of replication-competent lentivirus in the clinic: analysis of infused T cell products. Mol. Ther. 26, 280–288.

22. Marcucci, K.T., Jadlowsky, J.K., Hwang, W.T., Suhoski-Davis, M., Gonzalez, V.E., Kulikovskaya, I., Gupta, M., Lacey, S.F., Plesa, G., Chew, A., et al. (2018). Retroviral and lentiviral safety analysis of gene-modified T cell products and infused HIV and oncology patients. Mol. Ther. 26, 269–279.

23. Boxus, M., and Willems, L. (2009). Mechanisms of HTLV-1 persistence and transformation. Br. J. Cancer 101, 1497–1501.

24. Modlich, U., Kustikova, O.S., Schmidt, M., Rudolph, C., Meyer, J., Li, Z., Kamino, K., von Neuhoff, N., Schlegelberger, B., Kuehlcke, K., et al. (2005). Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. Blood 105, 4235–4246.

25. Eichler, F., Duncan, C., Musolino, P.L., Orchard, P.J., De Oliveira, S., Thrasher, A.J., Armant, M., Dansereau, C., Lund, T.C., Miller, W.P., et al. (2017). Hematopoietic stem-cell gene therapy for cerebral adrenoleukodystrophy. N. Engl. J. Med. 377, 1630–1638.

26. U.S Food and Drug Administration (2022). Cellular, Tissue, and Gene Therapies Advisory Committee June 9, 2022- June 10, 2022 Meeting Briefing Document-FDA- 125717. https://www.fda.gov/media/159132/download. (Accessed 22 November 2022).

27. Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M.H., Schambach, A., Charrier, S., Galy, A., Thrasher, A.J., et al. (2009). Insertional transformation of hematopietic cells by self-inactivating lentiviral and gammaretroviral vectors. Mol. Ther. 17, 1919–1928.

28. Espinoza, D.A., Fan, X., Yang, D., Cordes, S.F., Truitt, L.L., Calvo, K.R., Yabe, I.M., Demirci, S., Hope, K.J., Hong, S.G., et al. (2019). Aberrant clonal hematopoiesis following lentiviral vector transduction of HSPCs in a Rhesus macaque. Mol. Ther. 27, 1074–1086.

29. Hacein-Bey-Abina, S., Pai, S.Y., Gaspar, H.B., Armant, M., Berry, C.C., Blanche, S., Bleesing, J., Blondeau, J., de Boer, H., Buckland, K.F., et al. (2014). A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. N. Engl. J. Med. 371, 1407–1417.

30. Mellors, J.W., Guo, S., Naqvi, A., Brandt, L.D., Su, L., Sun, Z., Joseph, K.W., Demirov, D., Halvas, E.K., Butcher, D., et al. (2021). Insertional activation of STAT3 and LCK by HIV-1 proviruses in T cell lymphomas. Sci. Adv. 7, eabi8795.

31. Fraietta, J.A., Nobles, C.L., Sammons, M.A., Lundh, S., Carty, S.A., Reich, T.J., Cogdill, A.P., Morrissette, J.J.D., DeNizio, J.E., Reddy, S., et al. (2018). Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. Nature 558, 307–312.

32. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., et al. (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassemia. Nature 467, 318–322.

33. De Ravin, S.S., Liu, S., Sweeney, C.L., Brault, J., Whiting-Theobald, N., Ma, M., Liu, T., Choi, U., Lee, J., O'Brien, S.A., et al. (2022). Lentivector cryptic splicing mediates increase in CD34+ clones expressing truncated HMGA2 in human X-linked severe combined immunodeficiency. Nat. Commun. 13, 3710.

34. Liggett, L.A., Cato, L.D., Weinstock, J.S., Zhang, Y., Nouraie, S.M., Gladwin, M.T., Garrett, M.E., Ashley-Koch, A., Telen, M.J., Custer, B., et al. (2022). Clonal hematopoiesis in sickle cell disease. J. Clin. Invest. 132, e156060.

35. Jones, R.J., and DeBaun, M.R. (2021). Leukemia after gene therapy for sickle cell disease: insertional mutagenesis, busulfan, both, or neither. Blood 138, 942–947.

36. Nguyen, G.N., Everett, J.K., Kafle, S., Roche, A.M., Raymond, H.E., Leiby, J., Wood, C., Assenmacher, C.A., Merricks, E.P., Long, C.T., et al. (2021). A long-term study of AAV gene therapy in dogs with hemophilia A identifies clonal expansions of transduced liver cells. Nat. Biotechnol. 39, 47–55.

37. U.S Food and Drug Administration (2013). Guidance for Industry Preclinical Assessment of Investigational Cellular and Gene Therapy Products. https://www.fda.gov/media/87564/download. (Accessed 22 October 2022).

38. Modlich, U., Bohne, J., Schmidt, M., von Kalle, C., Knöss, S., Schambach, A., and Baum, C. (2006). Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. Blood 108, 2545–2553.

39. Zhou, S., Fatima, S., Ma, Z., Wang, Y.D., Lu, T., Janke, L.J., Du, Y., and Sorrentino, B.P. (2016). Evaluating the safety of retroviral vectors based on insertional oncogene activation and blocked differentiation in cultured thymocytes. Mol. Ther. 24, 1090–1099.

40. Schwarzer, A., Talbot, S.R., Selich, A., Morgan, M., Schott, J.W., Dittrich-Breiholz, O., Bastone, A.L., Weigel, B., Ha, T.C., Dziadek, V., et al. (2021). Predicting genotoxicity of viral vectors for stem cell gene therapy using gene expression-based machine learning. Mol. Ther. 29, 3383–3397.

41. Cohn, L.B., Silva, I.T., Oliveira, T.Y., Rosales, R.A., Parrish, E.H., Learn, G.H., Hahn, B.H., Czartoski, J.L., McElrath, M.J., Lehmann, C., et al. (2015). HIV-1 integration landscape during latent and active infection. Cell 160, 420–432.

42. Paruzynski, A., Arens, A., Gabriel, R., Bartholomae, C.C., Scholz, S., Wang, W., Wolf, S., Glimm, H., Schmidt, M., and von Kalle, C. (2010). Genome-wide high throughput integrome analysis by nrLAM-PCR and next-generation sequencing. Nat. Protoc. 5, 1379–1395.

43. Gabriel, R., Eckenberg, R., Paruzynski, A., Bartholomae, C.C., Nowrouzi, A., Arens, A., Howe, S.J., Recchia, A., Cattoglio, C., Wang, W., et al. (2009). Comprehensive genomic access to vector integration in clinical gene therapy. Nat. Med. 15, 1431–1436.

44. Schmidt, M., Schwarzwaelder, K., Bartholomae, C., Zaoui, K., Ball, C., Pilz, I., Braun, S., Glimm, H., and von Kalle, C. (2007). High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). Nat. Methods 4, 1051–1057.

45. Sherman, E., Nobles, C., Berry, C.C., Six, E., Wu, Y., Dryga, A., Malani, N., Male, F., Reddy, S., Bailey, A., et al. (2017). INSPIIRED: a pipeline for quantitative analysis of sites of new DNA integration in cellular genomes. Mol. Ther. Methods Clin. Dev. 4, 39–49.

46. Cesana, D., Calabria, A., Rudilosso, L., Gallina, P., Benedicenti, F., Spinozzi, G., Schiroli, G., Magnani, A., Acquati, S., Fumagalli, F., et al. (2021). Retrieval of vector integration sites from cell-free DNA. Nat. Med. 27, 1458–1470.

47. Rae, D.T., Collins, C.P., Hocum, J.D., Browning, D.L., and Trobridge, G.D. (2015). Modified genomic sequencing PCR using the MiSeq platform to identify retroviral integration sites. Hum. Gene Ther. Methods 26, 221–227.

48. Benedicenti, F., Calabria, A., Cesana, D., Albertini, A., Tenderini, E., Spinozzi, G., Neduva, V., Richard, A., Brugman, M.H., Dow, D.J., et al. (2019). Sonication Linker Mediated-PCR (SLiM-PCR), an efficient method for quantitative retrieval of vector integration sites. Hum. Gene Ther. 30, A214–A215.

49. Calabria, A., Leo, S., Benedicenti, F., Cesana, D., Spinozzi, G., Orsini, M., Merella, S., Stupka, E., Zanetti, G., and Montini, E. (2014). VISPA: a computational pipeline for the identification and analysis of genomic vector integration sites. Genome Med. 6, 67.

50. Berry, C.C., Nobles, C., Six, E., Wu, Y., Malani, N., Sherman, E., Dryga, A., Everett, J.K., Male, F., Bailey, A., et al. (2017). INSPIIRED: quantification and visualization tools for analyzing integration site distributions. Mol. Ther. Methods Clin. Dev. 4, 17–26.

51. Firouzi, S., López, Y., Suzuki, Y., Nakai, K., Sugano, S., Yamochi, T., and Watanabe, T. (2014). Development and validation of a new high-throughput method to investigate the clonality of HTLV-1-infected cells based on provirus integration sites. Genome Med. 6, 46.

52. Afzal, S., Fronza, R., and Schmidt, M. (2020). VSeq-toolkit: comprehensive computational analysis of viral vectors in gene therapy. Mol. Ther. Methods Clin. Dev. 17, 752–757.

53. Spinozzi, G., Calabria, A., Brasca, S., Beretta, S., Merelli, I., Milanesi, L., and Montini, E. (2017). VISPA2: a scalable pipeline for high-throughput identification and annotation of vector integration sites. BMC Bioinf. 18, 520.

54. Robin, J.D., Ludlow, A.T., LaRanger, R., Wright, W.E., and Shay, J.W. (2016). Comparison of DNA quantification methods for next generation sequencing. Sci. Rep. 6, 24067.

55. Rosewick, N., Hahaut, V., Durkin, K., Artesi, M., Karpe, S., Wayet, J., Griebel, P., Arsic, N., Marçais, A., Hermine, O., et al. (2020). An improved sequencing-based bioinformatics pipeline to track the distribution and clonal architecture of proviral integration sites. Front. Microbiol. 11, 587306.

56. Biasco, L., Pellin, D., Scala, S., Dionisio, F., Basso-Ricci, L., Leonardelli, L., Scaramuzza, S., Baricordi, C., Ferrua, F., Cicalese, M.P., et al. (2016). In vivo tracking of human hematopoiesis reveals patterns of clonal dynamics during early and steady-state reconstitution phases. Cell Stem Cell 19, 107–119.

57. Scala, S., Basso-Ricci, L., Dionisio, F., Pellin, D., Giannelli, S., Salerio, F.A., Leonardelli, L., Cicalese, M.P., Ferrua, F., Aiuti, A., et al. (2018). Dynamics of genetically engineered hematopoietic stem and progenitor cells after autologous transplantation in humans. Nat. Med. 24, 1683–1690.

58. Wang, G.P., Berry, C.C., Malani, N., Leboulch, P., Fischer, A., Hacein-Bey-Abina, S., Cavazzana-Calvo, M., and Bushman, F.D. (2010). Dynamics of gene-modified progenitor cells analyzed by tracking retroviral integration sites in a human SCID-X1 gene therapy trial. Blood 115, 4356–4366.

59. Biasco, L., Scala, S., Basso Ricci, L., Dionisio, F., Baricordi, C., Calabria, A., Giannelli, S., Cieri, N., Barzaghi, F., Pajno, R., et al. (2015). In vivo tracking of T cells in humans unveils decade-long survival and activity of genetically modified T memory stem cells. Sci. Transl. Med. 7, 273ra13.

60. Six, E., Guilloux, A., Denis, A., Lecoules, A., Magnani, A., Vilette, R., Male, F., Cagnard, N., Delville, M., Magrin, E., et al. (2020). Clonal tracking in gene therapy patients reveals a diversity of human hematopoietic differentiation programs. Blood 135, 1219–1231.

61. Pellin, D., Biasco, L., Aiuti, A., Di Serio, M.C., and Wit, E.C. (2019). Penalized inference of the hematopoietic cell differentiation network via high-dimensional clonal tracking. Appl. Netw. Sci. 4, 115.

62. Bystrykh, L.V., and Belderbos, M.E. (2022). Measures of clonal hematopoiesis: are we missing something? Front. Med. 9, 836141.

63. Del Core, L., Cesana, D., Gallina, P., Secanechia, Y.N.S., Rudilosso, L., Montini, E., Wit, E.C., Calabria, A., and Grzegorczyk, M.A. (2022). Normalization of clonal diversity in gene therapy studies using shape constrained splines. Sci. Rep. 12, 3836.