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 inadver- tent exposure to a replication-competent virus arising during vector manufacture would significantly increase the risk of treatment-related adverse events. These risks have led regulato- ry 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 expan- sion 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 inte- gration can disrupt gene expression, or the regulatory regions within the vector can gene alter expression. Clonal cell expansion and leuke- mias 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-compe- tent 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 moni- toring 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 ex- pressing the viral genes (Figure 1B). The membrane-bound vector particle contains two copies of the vector RNA, and the viral gene re- gions (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 re- sulted 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 is- sued 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 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 re- sulted in T cell leukemias from insertions near LMO-2, a gene impor- tant 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 pub- lished 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 to testing. While many in the gene therapy field use "retrovirus" to refer to gammaretroviruses, the FDA guidance uses the broader defi- nition. 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 In- dustry[13] and Chemistry, Manufacturing, and Control (CMC) Informa- tion 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 neces- sary 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 ho- mology 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 retro- viral sequences into an RCR has been reported.[17] Co-infecting viruses could also be co-opted. For these reasons, designing molecular 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 su- pernatant 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 suscep- tibility 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 recommen- ded 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 super- natant 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. Concen- trated 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. Concen- trated vector can also prove toxic to amplification cells, and appro- priate dilutions for RCR testing must be determined experimentally. Therefore, investigators seeking to utilize a new retrovirus or pseudo- type should begin developing an appropriate assay early in vector development.

Ex vivo cell productsSignificant 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 bio- logic 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-pos- itive 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 docu- mented 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. 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 su- pernatant 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 suscep- tibility 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 recommen- ded 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 super- natant 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. Concen- trated 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. Concen- trated vector can also prove toxic to amplification cells, and appro- priate dilutions for RCR testing must be determined experimentally. Therefore, investigators seeking to utilize a new retrovirus or pseudo- type 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 bio- logic 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-pos- itive 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 docu- mented 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. 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 recombi- nation 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 sup- ports the safety of current manufacturing methods in generating RCR-free vector. The NHLBI-funded National Gene Vector Bio- repository (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 clin- ical 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.

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 regu- lation, 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 gam- maretroviruses, such as MLV, cause leukemia by the process of IO. MLV tends to integrate near the 5^0 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 func- tion 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 coop- erate to cause transformation.[24] The potential for multiple integrants in the same cell to increase transformation risks underlies the regulato- ry 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 expres- sion, including the lentiviral vector in question. Incorporating the MLV promoter/enhancer into lentiviral vectors has been shown to in- crease in vitro immortalization of murine progenitor cells[27] and was associated with the development of malignancy in a non-human pri- mate.[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" vec- tor); 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 Vector insertions that alter cell growth do not always result in malig- nancy. Fraietta et al. reported transient expansion of a T cell clone transduced with a lentiviral vector expressing a chimeric antigen re- ceptor (CAR-T).[31] The patient had a germline hypomorphic mutation of the methylcystosine dioxygenase TET2 allele, and the vector inser- tion 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 thalas- semia.[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 5^0 LTR polyadenylation signal and elimination of a 3^0 let-7 binding site normally contained in the 3^0 untranslated re- gion of the HMGA2 transcript, preventing normal microRNA regula- tion of the HMGA2 transcript. The clone eventually dissipated and did not become transformed. A recent report observed a similar phenom- enon 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 under- lying disease, the preparative regimen, and the vector insertion.[34] In this study, two individuals developed acute myeloid leukemia/myelo- dysplasia (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 canines treated for hemophilia with a factor VIII-expressing adeno-asso- ciated 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 pop- ulations. In contrast, gammaretroviral vectors have been used exten- sively 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, partic- ularly for in vivo administration of integrating vectors.

Insertional oncogenesis assays for preclinical vector assessmentsGiven 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 hu- man cells into mice to assess the biosafety of vectors in the hemato- poietic 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 integra- tion site analyses performed to assess patterns of integration sites and possible clonality. Marrow may be serially transplanted to sec- ondary recipients, which may increase the sensitivity for detecting ex- panding 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 from murine donor gene-marked cells from the relatively frequent host cell malignancies that may arise, especially when recipients un- dergo total body radiation conditioning. One standard approach is to transduce and transplant marrow that contains a marker to distin-

guish 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 tu- mor-prone mice (e.g., the cdkn2a$^{-/-}$ mouse), which may increase sensitivity but is complicated by high background rates of transfor- mation and (2) xenografting of transduced human cells into an im- mune-deficient mouse host to assess effects on differentiation and possible clonal expansions. In general, these assays are not very sen- sitive and often fail to detect clonality or malignancy, even when using gammaretroviral vectors as positive controls. While the FDA has is- sued 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 he- matopoietic stem and progenitor cell targets.[27,38] This assay uses mu- rine 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 rela- tively straightforward assay that can provide a good assessment of one aspect of vector safety. Zhou et al. developed a similar in vitro trans- formation 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 ana- lyzes the gene expression signature of the transduced cells to identify those they characterized using machine learning as representing transformation.[40] This surrogate assay for geno- toxicity assessment (SAGA) is a robust and sen- sitive assay and reliably predicted the mutagenic risk for previously used vectors that had caused leukoproliferative severe adverse events in clin- ical trials. The role the assay will play in predict- ing 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 guid- ance.[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 lymph- adenopathy. Complete blood cell counts should be assessed for hyper- leukocytosis, 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/mL or presence of blast cells) or clinical signs and symptoms suggest the potential presence of leukoprolifera- tion 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 inte- gration 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 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 integra- tion 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 contin- ually 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 en- zymes or, as recently proposed, by sonication to provide more unbi- ased 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 provi- rus 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 intro- ducing oligonucleotide barcodes in the amplicon design, NGS allows for identifying DNA fragments derived from different biological sam- ples or technical replicates within a single run. Indexing strategies us- ing 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 set- tings, 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 integra- tion 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 vec- tor-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 in- volves 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 pop- ulation of cells. This scenario consists in a large number of clones hav- ing similar sizes. Many factors can have an impact on clonal dy- namics. In addition to alteration in individual clone growth caused by gene dysregulation due to vector insertion, factors such as the un- derlying 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. Simi- larly, 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 dis- ease, and results of other treatment endpoints (e.g., vector copy num- ber, 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 tar- gets, since retroviral vectors have been the most common vector sys- tems 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 develop- ment and discussion of an RCR and IO monitoring plan early in clin- ical 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 informa- tion 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 devel- opment. 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 sci- entific journals, and the NIH should consider establishing a database to assist research efforts in preventing future events. Finally, these rec- ommendations 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.

ACKNOWLEDGMENTS

This project has been funded in part with federal funds from the Na- tional Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract no. 75N92019D00018 (K.C., T-Y.L., D.P.) and the contents are the views of the authors alone. Figures were created with BioRender.

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 Labora- tories 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 Allo- gene Therapeutics, ImmunoVec, Pluto Therapeutics, MyoGene Bio, and Innoskel and an ad hoc consultant for Cimeio Therapeutics, TransformaTx, and bluebird bio.