

Proposal for ISCT position statement on novel assays for the quality control and potency assessment of adoptive cellular immunotherapies

, Ben Weil¹ Patrick J. Hanley² & Mark Lowdell^{2,3}

¹Centre for Cell Gene & Tissue Therapeutics, Royal Free Hospital, London, UK

²Center for Cancer and Immunology Research, Center for Cancer and Blood Disorders, Children's Research Institute, Children's National Medical Center and The George Washington University, Washington DC, USA

³Dept of Haematology, Cancer Institute, University College London, London, UK

Key words: Quality control; quality assurance; in-process QC; release assays

Abstract:

Translation of cell and gene therapies from pre-clinical experiments to clinical trials and final drug licensing brings requires the development, verification and even validation of the assays essential for the definition of the drug product. The technical and scientific challenges in doing this are far greater than they seem at first and are compounded by a lack of approved standards for assays used to support (c)GMP manufacture. This paper highlights some of those challenges and proposes solutions based on the experience of our colleagues using similar assay platforms in regulated pathology laboratories.

Scope: this paper will examine the challenges of in-process QC, release assays and in-vivo monitoring of cellular immunotherapies manufactured ex-vivo as medicines. It will encompass T cells, NK cells and mesenchymal stromal cells (MSC) from autologous and allogeneic sources both native and gene modified. The immunotherapy may be anti-cancer, anti-infection or anti-inflammatory. It sets the scene for the following series of reviews papers on analysis of specific cellular therapies in clinical trial and development towards licensed medicines.

A good rule for QC assays is "just because it can be measured, doesn't mean that it should be measured"; make sure that a QC assay provides a "need to know" result and not just a

“nice to know”. “Nice to know” assays should be for information only and the results should be kept outside of the manufacturing records.

It is also essential to avoid the McNamara Fallacy¹. This was originally applied to business practice but is equally relevant to QC assays:

The McNamara Fallacy –

- 1. measure whatever can be easily measured**
- 2. disregard that which can't be easily measured**
- 3. presume that what can't be measured easily really isn't important**
- 4. say that what can't be easily measured really doesn't exist.**

The challenge:

Adoptive immunotherapies have mostly arisen from the related field of hematopoietic stem cell transplantation; either as anti-leukaemia effector cells aiming to replicate or enhance the graft-versus-tumour effect (GvT) or as anti-viral effector cells to reconstitute the immune system after allogeneic HSCT. More recently this has expanded into the production of MSC and regulatory T cells (Treg) for the treatment of refractory graft-versus-host disease in the same group of patients. What these have in common is a relative ease of small scale manufacture and a high degree of scientific complexity; unique characteristics in the area of pharmaceutical medicines and ones which drive their development within academic units. The strengths of this approach are the diversity of products and applications and the speed with which they can get to early phase clinical trial. From a conventional drug development point of view these are all weaknesses at the same time and are compounded by the enormous challenge of defining these products to an appropriate standard to move beyond early phase trials and to a marketing authorisation as a licensed medicine – something which is often completely outside the mind-set and experience of the clinician/scientist who conceived the idea initially.

The pathway from proof-of-concept trials to an application for marketing approval provides plenty of opportunity for development of the product specification but the nature of the starting materials for these adoptive immunotherapy products means that a product

specification is almost certainly never going to reach the standards applied to conventional pharmaceuticals; even biologics such as recombinant molecules and monoclonal antibodies. Fortunately medicines regulators are fully aware of this challenge and acknowledge it but, nonetheless, it is beholden on us as cell therapy drug developers to create ever better assays to define our products and to work out ways to validate these to an appropriate standard. Many of these assays continue to come from research labs with little development as quality control assays as understood by pharmacists and regulatory agencies.

The solution: One way to approach this dilemma is to apply the same standards to cell therapy QC assays as are applied to diagnostic pathology assays since most share the same platforms and many problems have already been solved. Moreover, it seems logical to apply the same standard to the therapeutic product as to the diagnostic tests which identified the patient as eligible for the therapy. This means that a list of target acceptance criteria for assays can be drawn up on the basis of the technical platform used; irrespective of the specific assay.

A good starting point is to apply the Cotlove rule recommended by the 1976 College of American Pathologists Conference on Analytical Goals in Clinical Chemistry, namely, that analytic variance (intra-assay CV) should be less than one fourth of the appropriate biological variance (patient cohort CV).

For this it is helpful to define some commonly used terms in diagnostic pathology.

DEFINITIONS (CLIA, CLSI, CUMITECH 31A)

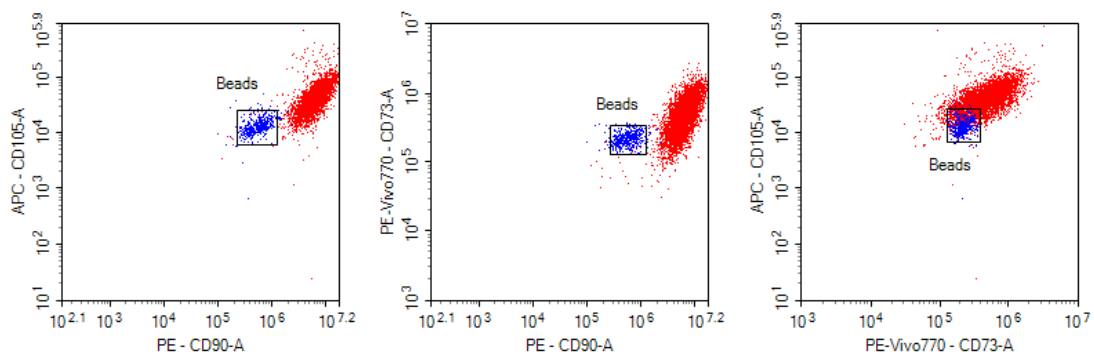
- Verification – “...one-time process performed to determine or to confirm a test’s expected performance prior to implementation in the clinical laboratory...“Does the test work?””
- Validation – “...ongoing process of monitoring a test, procedure, or method to ensure that it continuously performs as expected ...“Does the test still work?””
- Assay sensitivity – “how well does the assay correctly predict a clinical outcome?”
- Assay specificity – “how well does the assay incorrectly predict a clinical outcome?”
- Inter-assay cv – what is the cv (sd/mean) from the same product tested repeatedly – i.e. multiple samples taken from a single batch of product and prepared and analysed identically

- Intra-assay cv - what is the cv (sd/mean) from the same test sample analysed repeatedly – i.e. a single prepared sample from the example above analysed repeatedly.

Probably the most widely used assay platform in this field is flow cytometry, the complexity of which is widely under-estimated. Criteria for an acceptable flow cytometry based assay could look like this:

1. Standardised fluorescence beads must be run prior to every assay for each fluorochrome used in the assay and the bead population must fall into pre-set analysis regions which are operator independent
2. For each target antigen there must be at least three different hybridomas available which are conjugated with the fluorochrome of choice and which bind to the same epitope
3. The choice of fluorochrome used for each antibody must be optimised with respect to the signal:noise ratio of the antigen
4. The “all bar one” method should be used to set up the detectors and fluorescence compensation matrices of the flow cytometer
5. Antibody binding beads can be run with each experiment to confirm the stability of the analyser, antibody and fluorochrome such that the bead population must fall into pre-set analysis regions which are operator independent –

Human MSC labelled with anti-CD90 and anti-CD105 in the presence of antibody binding beads (ThermoFisher) as a control for instrument performance and adequate cell labelling



6. The analysis template must be locked (if possible) or the analyses must be printed and retained in the batch record for QC approval. Changes in the analysis regions to account for normal biological variation must require QC approval
7. Qualified assays should have an intra-assay cv of no greater than 5% and an inter-assay cv of no more than 10%.
8. When analysing rare events (fewer than 5% of the total population) at least 100 “positive” events must be acquired (irrespective of the total number of cells analysed) since this will give a cv of 10% according to Poisson statistics.
9. Isotype control samples should only be used to provide the “negative” control where that can be justified; an “all bar one” method should be used in preference
10. Positive controls for critical antigens should be devised wherever possible
11. Reference cell samples (fixed or cryopreserved) should be used as quality assurance controls where possible
12. All flow cytometers used for QC should be on permanent maintenance contracts and be part of an internal QC programme and external QA programme for CD4 and or CD34 enumeration
13. The assay must be shown to be independent of the flow cytometer used and be able to provide comparable results (e.g. within the +/- 2SD of the results obtained from the original flow cytometer) on at least one other instrument of a different type.
14. All assays should be Verified and Validated according to the definitions above wherever possible

Conventional haematology blood analysers are another widely used tool in this field but often without the awareness that the technology used to identify and count cells differs widely between instruments. Moreover these machines are optimised to count and calculate blood cell differentials in whole blood samples which are preserved with EDTA. The morphology of EDTA preserved lymphocytes and monocytes in whole blood compared to T and monocytes cells from the same donor in isotonic cell culture medium is different so the cell count and the differentials are unlikely to be the same. The breadth of the field in terms of devices for providing cell counts seems to increase annually but criteria can be set for acceptability:

1. Intra-assay cv of <3%

2. Inter-assay cv of <8%
3. Operator independent analysis
4. Linearity of enumeration in serial dilutions over the range of cell densities which are likely in the product during manufacture and at release

These two examples are simple and other analytical platforms will set greater challenges for standardisation but it is only by setting standards for our analytical tools that we can start to advise regulatory agencies on the appropriateness of individual QC and release assays for products filing for marketing authorisation. This is where international pathology accreditation expertise can be used for the benefit of the field.

More challenging are the functional assays used for potency assessment. As cell therapy drug developers we are often unaware that the concept of potency assays for release of a medicine is relatively new. Functional assays for drugs arose out of the development of biologics in the 1990s rather than more conventional chemically synthesised pharmaceuticals. No one would expect a potency assay to be performed on a batch of cyclophosphamide or even of aspirin.

Regulatory agencies are very aware of the unique challenges posed by potency assays and considerable flexibility already exists in current guidance. Nonetheless it is important to be aware of the limitations of potency assays to establish acceptable ranges for results; even if these give very wide margins of acceptance due to the innate variability of the material and the assay. It is expected that a potency assay will be evaluated during the clinical development of the product so that the range of acceptable results can be determined from clinical data; i.e. correlation to determine predictability of the potency assay. The GMP standards for the European Union, EUDRALEX, state, "Validated analytical methods are not necessarily required during product- and process-development activities or when used in characterization studies" and "Analytical methods should be scientifically sound (e.g., specific, sensitive, and accurate) and provide results that are reliable".

It is assumed that the results will also be meaningful but therein lies another challenge for cell and tissue medicines. Can we measure a function in vitro which is relevant to the in vivo effect and, moreover, is that result likely to predict the in vivo benefit or adverse event?

These two questions are often conflated and lead to the design of unreasonably complicated assays. A potency assay should be designed to determine that the cells have

the capacity to perform a relevant function; e.g. interferon secretion in response to relevant peptide antigens or in vitro proliferation in response to antigen or mitogen stimulation.

Assays such as this demonstrate that the cells have “potency”; they do not tell you whether the dose of cells being delivered will have the relevant clinical effect which is what we often try to achieve. Identifying a level of functional activity (we deliberately don’t use the term “potency” here) which predicts clinical efficacy is unreasonable since the factors determining clinical efficacy are innumerable and no other type of pharmaceutical product is held to such a standard. Indeed this is even more unreasonable for adoptive cell therapies where the product is usually planned to proliferate in vivo so even the cell dose given is not the effective dose and the correlation of a numerical indicator of functional potency pre-expansion with post-expansion efficacy is highly unlikely.

An excellent, example has been reported from the CAR-T program at University of Pennsylvania². An adult patient who had remained refractory to treatment after an initial dose of autologous CD19-CAR-T for treatment of chronic lymphocytic leukemia received a second dose 70 days later. There was no detectable in vivo expansion of the CAR-T at one month post second infusion and there was no clinical effect on his tumour burden.

However, at around day +54 CAR-T cells appeared in his peripheral blood, expanded and resulted in complete remission. Subsequent analysis of the TCR of the CAR-T showed that >90% of these were derived from a single T cell clone; in other words just one cell from the CAR-T product expanded in vivo and delivered the clinical effect.

Nonetheless, potency assays are here to stay for cellular therapies and the selection of scientifically relevant analytes is critical as is the determination of acceptance criteria which are likely to be specific to the type of product and the clinical application as discussed in the subsequent articles.

Having established a secure basis of our assays we can proceed to discuss the use of the results which they provide and advise regulatory agencies on interpretation of these results.

How do we establish the identity of an adoptive cell therapy?

The nature of this challenge is plainly predicated on the type of cell but a useful exemplar is a typical T cell immunotherapy. These can be autologous or allogeneic and each has different options for establishing identity. Autologous T cells may be genetically modified

such as CAR-T where a known novel gene is inserted which leads to a known novel protein expression. Here identity can be established by a validated immunophenotyping assay and, using CAR-19 CTL as an example, it could be valuable to create a fluorochrome conjugated rCD19 to enumerate the CAR-T and confirm target antigen binding in a single assay. As CAR-T cell therapies become more refined such as engineering the CD4:CD8 ratio of the final product or even the selection of functional subsets of T cells such as those with effector memory phenotype in preference to central memory the ability to reliably quantitate the specific cell dose will require more careful design of antibody and fluorochrome panels with concomitant validation.

Given the capacity of cell therapies to proliferate within the patient, it may only be important to deliver a threshold dose which is effective as determined from clinical trials. As allogeneic CAR-T are developed a more important safety consideration may be to avoid administering an excessive dose of cells which have the ability to mount an allogeneic response.

Suggested minimal information regarding identity:

Can we agree on a table here of assays which would be appropriate during development (e.g. Phenotype, DNA fingerprint, HLA typing etc) and those which would remain relevant at phase III and for the marketed product?

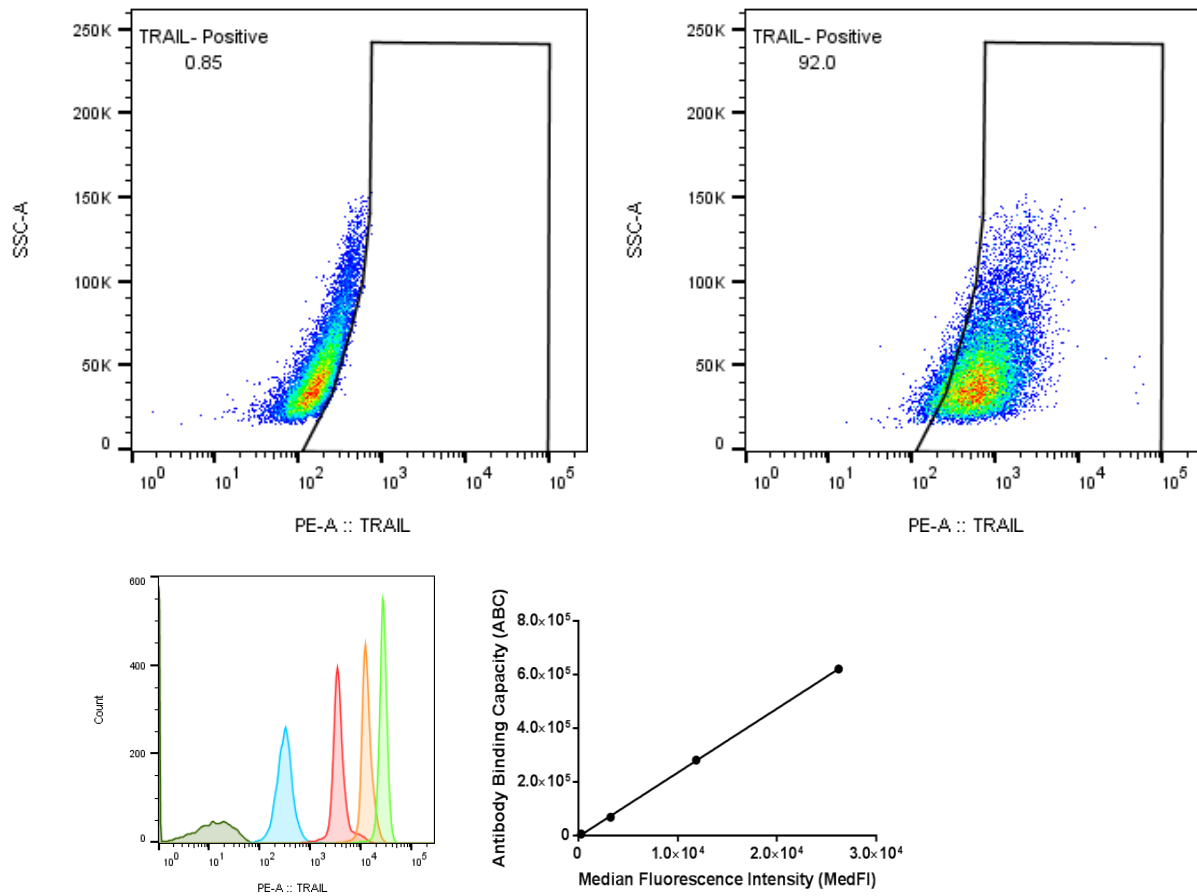
There are two basic criteria of identity assays:

1. To confirm that the product contains the right cell type(s)
2. To confirm that the product has come from the correct donor

Most identity assays to determine the right cell type rely upon phenotyping by flow cytometry because it is fast, reliable and semi-quantitative or even quantitative. However, it is important to consider what parameters are informative. When seeking to prove identity of a product the often used “percent positive” is relatively weak as an acceptance criterion. Semi-quantitative measurement of marker density may be a more valuable parameter to measure, especially if it can be related to potential potency such as the density of expression of an engineered T cell receptor or CAR or high-affinity FC-gamma receptor.

During the pre-clinical development of the MSC-TRAIL product at UCL, we designed a flow cytometric assay to identify the proportion of MSC successfully transfected with recombinant TRAIL and to determine the number of TRAIL molecules per cell using antibody binding capacity beads and an APC-conjugated anti-TRAIL antibody (figure 1). This had the advantage of being able to describe the drug product in terms of “dose of TRAIL” rather than cell number and we have been able to show a dose response in vitro with respect to potency. (figure 2)

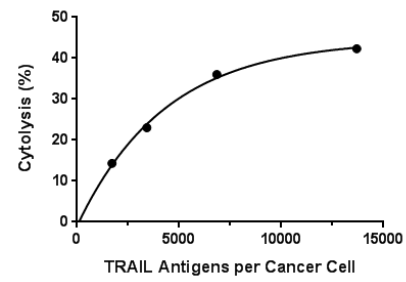
Figure 1: Lentiviral transduction of TRAIL induces TRAIL expression which quantifiable by antibody binding beads and is stable beyond passage 1



Passage #	#TRAIL antigens per MSCTRAIL
P1	20,986
P2	14,225
P3	14,456
P4	13,700

Figure 2: The ratio of TRAIL transduced MSC: target cells was adjusted to deliver a total dose of TRAIL molecules ranging from 1,713 to 13,700 to each tumour cell.

MSCTRIL Dose	# TRAIL Antigens per Cancer Cell	Percentage Cytolysis (%)
1,250	1,713	14.3
2,500	3,425	23.0
5,000	6,850	36.0
10,000	13,700	42.3



Identity assays to confirm the correct donor or correct parent master cell line of a working cell bank can be provided by HLA-typing (low resolution is adequate) or by DNA fingerprinting by STR which are widely available assays with established controls.

Robust identity assays tick two important regulatory boxes; first they support the potential potency of the product since the wrong cells are unlikely to have the pharmacologic effect. Second they address safety since only the cells tested in the early phase trials are known to be safe. Other cell populations may have a worse safety profile such as a higher propensity to transform. This second point is somewhat moot since, in contrast to other drugs, most cell therapies are derived from normal parent cells which are pre-programmed to die. Since the evolution of multicellular organisms there has been a requirement to eradicate dead and dying cells and the human reticuloendothelial system (RES) manages this with spectacular success. In most adoptive cell therapies the contaminating impurities are cells with little or no risk of adverse effect and which will die and be cleared by the RES. Reliable assays to track cell fate *in vivo* are lacking in the clinical setting and are insufficiently sensitive in current animal models. However these will be a terrifically valuable in supporting the safety profile of adoptive cell therapy, both autologous and allogeneic.

Are there identity and potency criteria which can be agreed upon now and thus begin to determine appropriate assays which can be harmonised and qualified with novel QA reagents?

Now that the field has matured to the stage of multiple licensed products on the market identity criteria for many cell therapies are becoming established and are discussed in detail in the following reviews. Examples of common products are given below

1. Autologous CAR-T
 - a. CD3+ve / CD56-ve >90%
 - b. Expression of the CAR in >50% of cells
 - c. Density of CAR expression within the range shown to be effective in previous *in vitro* target cell lysis assaysOR
 - d. Density of CAR expression within the range shown to be effective in previous clinical trials of the same CAR construct

2. Allogeneic CAR-T
 - a. All of the above plus
 - b. Density of donor MHC class I expression below the threshold shown to provoke allogeneic T cell response in vitro and predict protection from rejection in vivo
3. Directed HLA-matched allogeneic anti-viral T cells
 - a. CD3+ve / CD56-ve >90%; <2% of antigen presenting cells (DCs or monocytes)
 - b. Total absolute T cell dose below the threshold for a conventional donor lymphocyte infusion (TC-T) in the clinical setting in which the product will be used
 - c. Presence of detectable T cells secreting IFN-G and or IL-2 (or another cytokine shown to be predictable of outcome in early phase trials) by ELISPOT or i/c flow cytometry or Luminex in response to the viral peptides used in the manufacture of the product. In the case of multi-virus specific T cell products the composition of the product with respect to activity to all viruses should be documented to prevent the pre-emptive use of a combination product which is non-reactive to the reactivating virus. In this case the risk:benefit ratio with respect to anti-viral response versus GvHD risk is plainly adverse.
4. Allogeneic MSC
 - a. Phenotype according to ISCT criteria³
 - b. Immunosuppressive function in a CD3/CD28 mitogen stimulation of third-party normal donor T cells by flow cytometric CFSE dilution assay⁴
or
 - c. Tri-lineage chondrogenic, adipogenic and osteogenic differentiation by immunocytochemistry⁵

Are there assays which can be agreed upon now and thus begin to be harmonised and qualified with external QC reagents?

Despite the broad range of cell therapies in development there are a number of assays which are common to many and will appear often in the subsequent review articles in this series. As an initial aim perhaps the community needs to adopt the same approach as the US

Center for Disease Control did over 20 years ago in its first recommendations for standardisation of CD4 enumeration (MMWR 1992)⁶ and which have been updated regularly as technologies have advanced. A good starting point would be:

1. Flow cytometric measurement of surface expression of CD3, CD14, CD16, CD19, CD28, CD34, CD45RA, CD45RO, CD33, CD56, CD73, CD90, CD105, CD117 are all common so now may be the time to establish recommended fluorochromes for each antigen in a no-wash labelling protocol with internal counting beads or volumetric counting flow cytometer akin to the ISHAGE CD34 assay protocol
2. Flow cytometric assay of cell membrane integrity as measurement of viability using 7-AAD, Propidium Iodide, ToPro-3 iodide, Dapi/Hoescht etc with an agreed gating strategy to avoid the common fault of pre-gating on “live” cells by morphology and thus under-estimating the proportion of dead cells and debris in the drug product.

To enable this sort of standardisation will require new QC reagents and, ideally, new external QA schemes for production QC laboratories to sign up to; perhaps led by agencies such as the FDA in the USA, the MHRA National Institute of Biological Standards in the UK and the Paul Ehrlich Institute in Germany. In France the ANSM already runs a national QA scheme for CD34 enumeration of hematopoietic cell transplant products so there is a precedent.

Can we establish the risk of an adoptive cell therapy with release assays?

The risks of cell therapies are multiple but are principally:

- Infection through contamination
- Tumorigenicity
- Off target Autoreactivity
- Off target Alloreactivity

Infection risk of patient-specific cell therapies is readily controlled by donor screening, GMP compliance and microbial testing of products prior to release and assays for all of these are well established and routine. With respect to off-the-shelf products from master cell banks the testing is more extensive and has been defined by FDA and EUDRALEX (refs). These

standards require screening for adventitious viruses which may derive from raw materials used in the manufacture in addition to the donor screening testing.

Tumorigenicity is more controversial. Historically many regulatory agencies have required analysis of chromosomal instability of cell products which have been cultured “extensively” – without any definition of “extensively”. Typically this has been applied to MSC products from large cell banks where the cells in the final product may have undergone over 15 population doublings and is performed by visual analysis of G-bands by qualified cytogeneticists. Although still commonly used this assay is of dubious value since it typically involves analysis of 50 and maximally 200 mitotic cells from a dose of many millions. The incidence of a malignant tumour cell is likely to be far lower than 5% of the population and thus it is statistically a Poisson distribution. The coefficient of variance (cv) of a Poisson distribution is determined by the absolute number of positive events counted. If 5% of the cells in a sample of 200 mitotic cells are tumor cells that will equate to 10 positive events. The cv of a Poisson distribution is calculated as

$$\frac{100}{\sqrt{n}} \quad \text{where "n" = the number of positive events counted}$$

In this example we counted 10 events so the cv = 31.6 which exceeds the acceptable intra-assay cv of 3% for a diagnostic test by tenfold. To achieve a cv of 3% requires counting of 1100 abnormal mitotic cells which, even at 5% incidence requires screening over 22,000 mitoses! Even then, fewer than 0.02% of a typical dose of 2×10^6 cells per kilo for an 80kg adult will have been analysed so the sampling error is extreme.

In the case of allogeneic master cell banks the screening for tumorigenicity is also extensive and well described in the standards referenced above. Briefly these include in vivo testing and DNA stability and identity tests by STR profiling.

Off target autoreactivity is not a typical QC test required for autologous cell therapies and the clinical risk is determined from phase I and II clinical trials. It is challenging to conceive a relevant assay which could be informative as a release criterion.

It is tempting to consider the development of alloreactivity assays to control the risk of GvHD by allogeneic T cells. There is an extremely long history of putative assays to predict

GvHD potential of allogeneic HSCT and none has achieved widespread adoption due to the lack of published data on sensitivity and specificity of the assays in predicting clinical outcome. A recent review of prediction algorithms for GvHD after allogeneic hematopoietic cell transplantation did not even mention in vitro tests of alloreactivity as part of patient:donor risk stratification⁷. It is unlikely that any assay of alloreactivity can be used as a release criterion, not least because of the multifactorial risk factors associated with the patient clinical status, age, sex and conditioning regimen.

In the setting of adoptive immunotherapy of allogeneic anti-viral T cells post HSCT it is common to dose on total donor T cells rather than on the absolute number of virus-specific T cells in the product. In this way the patient can be protected from GvHD by limiting the total allogeneic T cell dose. The reliability of the assay to determine the absolute number of T cells in the product is essential.

Conclusions:

Cell therapies are the most exciting new medicines in development and they present new and complex challenges in production, definition, delivery and administration. They will continue to question established dogmas in drug development, not least in the regulatory aspects. In contrast to all previous medicines, cell therapies are mostly relatively easy to manufacture since single batches usually treat single patients and the clinical impacts are often profound with clear demarcation of “success” and “failure”; CAR-T cells being the archetypal example although the long history of successful use of TIL in melanoma is equally impressive. It is this latter aspect which presents the greatest opportunity and challenges to drug developers since small numbers of patients are often all that are needed to obtain data for submission for marketing approval. As an example, Kymriah™ achieved marketing approval in the US after 106 patients had been treated in a multi-centre clinical trial. This rapid pace to regulatory approval and licensing provides little time to optimise in process QC and release assays so this needs to be considered before the first trial is opened. With cell therapy drug development still largely within academic units, the need for optimised and harmonised assays requires multidisciplinary collaboration and leadership to ensure that the quality of these medicines is consistent.

References:

1. O'Mahony S. Medicine and the McNamara fallacy. *J. Royal Coll Physicians Edinb* 2017; 47: 281–7.
2. Fraietta JA, Nobles CL, Sammons MA et al. Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. *Nature* 2018;558:307-12.
3. Dominici M, LeBlanc K, Slaper-Cortenbach I, Marini F, Krause D, Deans R, Keating A, Prockop Dj, Horwitz E. Minimal criteria for defining multipotent mesenchymal stromal cells. The ISCT position statement. *Cytotherapy* 2006;8:315-17.
4. Ribiero A, Ritter T, Griffin M, Ceredig R. Development of a flow cytometry based potency assays for measuring the in vitro immunomodulatory properties of mesenchymal stromal cells. *Immunology Letters* 2016;177:38-46.
5. Deskins DL, Bastakloty D, Saraswati S, Shinar A, Holt GE, Young PP. Human mesenchymal stromal cells: Identifying assays to predict potency for therapeutic selection. *Stem Cells Trans Med* 2013 2:151-158.
6. CDC. Guidelines for the performance of CD4+ T-cell determinations in persons with human immunodeficiency virus infection. *MMWR* 1992;41(No. RR-8):1-17.
7. Lee C, Haneuse S, Wang H-L, Rose S, Spellman SR, Vernis M, Hsu KC, Fleischhauer K, Lee SJ, Abdi R. Prediction of absolute risk of acute graft-versus-host disease follow hematopoietic cell transplantation. **PLoS ONE 13(1)**: e0190610.
<https://doi.org/10.1371/journal.pone.0190610>