Sarcouncil Journal of Internal Medicine and Public Health

ISSN(Online): 2945-3674

Volume- 05| Issue- 03| 2026

Research Article **Received:** 01-05-2026 | **Accepted:** 26-05-2026 | **Published:** 13-06-2026

Data-Driven Labeling Architectures in Cell and Gene Therapy Manufacturing

Prathyusha Guttikonda

Independent Researcher, USA

Abstract: Labeling in cell and gene therapy (CGT) manufacturing occupies a unique operational position that has no direct parallel in conventional pharmaceutical production. Because CGT products are biologically variable, frequently patient-specific, and characterized by critical quality attributes that are determined only after manufacturing is complete, the sequential logic that governs traditional labeling workflows—wherein product characteristics are known before a label is applied—cannot be assumed. Regulatory frameworks governing advanced therapy medicinal products (ATMPs), including the applicable provisions of 21 CFR Part 1271 and EudraLex Volume 4 Annex 2, establish functional requirements for product identification and traceability without prescribing the structural architecture through which those requirements must be fulfilled. This regulatory flexibility has, in practice, allowed legacy practices such as two-stage labeling to persist in environments where digital manufacturing infrastructure has materially changed the operational conditions that originally justified them. The biological basis of late-stage data generation in CGT—rooted in the variability of living cellular starting materials, the complexity of gene transfer processes, and the post-manufacturing nature of release testing—makes deferred label completion an intrinsic feature of these manufacturing workflows rather than a procedural anomaly. Alternatives, including just-in-time labeling and minimal-label architectures grounded in digital traceability, offer operationally sound and regulatorily consistent responses to this challenge. A transition toward data-centric labeling models, in which the physical label functions as an identity anchor and validated digital systems carry the authoritative attribute record, represents the most coherent path forward for scalable, safe, and compliant CGT manufacturing.

Keywords: Advanced Therapy Medicinal Products, Cell And Gene Therapy Manufacturing, Labeling Traceability, Data-Centric Quality Systems, Digital Manufacturing Integration.

INTRODUCTION

The manufacture of advanced therapy medicinal products (ATMPs), and specifically cell and gene therapy (CGT) products, operates under a set of biological and procedural constraints that have no meaningful precedent in conventional pharmaceutical manufacturing. Small-molecule drugs and even most biologics are produced through processes where critical quality attributes (CQAs) are substantially predictable before a batch reaches its final container. The manufacturing process is designed to produce a known output, and quality control testing serves largely to confirm that the output falls within pre-established specifications. Labeling, in this context, is a downstream formality—an exercise in applying known data to a physical container at the appropriate point in the workflow [Rathore, A. S., & Winkle, H. 2009].

Cell and gene therapy manufacturing disrupts this logic at multiple levels. These products are frequently patient-specific, meaning that each manufacturing run begins from a unique biological starting material—typically autologous cells harvested from a single patient—whose characteristics cannot be standardized in advance [Kouris, N. A. *et al*., 2011]. The process transforms that material through a series of biological manipulations that introduce inherent variability at each stage. Potency, viability, fill volume, dose concentration, and expiry are not predetermined specifications that a batch is manufactured to meet; they are empirical outcomes that are measured after manufacturing is complete [Ferrer, L. *et al*., 2016]. A label cannot be finalized before these measurements exist, and these measurements do not exist until after the product has been manufactured, often tested under time-sensitive conditions, and characterized through a battery of release assays.

This creates a structural timing problem that sits at the intersection of regulatory compliance, operational safety, and quality system design. The problem is not adequately addressed by the regulatory frameworks currently governing ATMPs, which emphasize traceability and identification requirements without prescribing how manufacturers should manage interim labeling states [Martin, R. M. 2025]. It is also not resolved by the legacy practice of two-stage labeling, which was designed for environments lacking the digital infrastructure now available in modern biomanufacturing facilities [Szkodny, A. C., & Lee, K. H. 2022]. What is needed is a reconceptualization of what a label is required to do—and a clear-eyed assessment of what digital systems can do instead.

***Corresponding Author:** Prathyusha Guttikonda

DOI- https://doi.org/10.5281/zenodo.20682403

This article examines the structural conditions that produce labeling challenges in CGT manufacturing, analyzes the inadequacy of legacy practices in the current manufacturing environment, and proposes a transition toward data-centric labeling architectures grounded in digital traceability. Throughout, the analysis remains anchored to the regulatory frameworks and peer-reviewed evidence relevant to ATMPs, without reference to specific commercial entities.

The Biological Basis of Late-Stage Data Generation in CGT

Understanding why labeling presents a distinctive challenge in CGT manufacturing requires first understanding the biological characteristics of these products and how those characteristics shape the manufacturing process. Unlike conventional pharmaceuticals, which are synthesized from chemical precursors with well-characterized reactivity, CGT products are manufactured from living cells that vary between donors, respond differently to culture conditions, and undergo biological changes throughout the manufacturing process itself [Kouris, N. A. *et al*., 2011].

For autologous cell therapies, the starting material is harvested from the patient who will ultimately receive the product. This means that each manufacturing campaign begins from a cell population whose baseline characteristics—viability, expansion capacity, and phenotypic markers—are unknown until characterization assays are completed after collection [Ferrer, L. *et al*., 2016]. The manufacturing process must then be executed on this unique starting material, and the attributes of the final product—including potency, purity, identity, and dose—are determined by how that specific cell population responds to culture conditions, transduction or transfection, and downstream processing steps. These are not attributes that can be assigned in advance; they are measured outcomes.

Gene therapy products introduce additional complexity. Viral vector manufacturing requires characterization of vector genome titer, transduction efficiency, and the ratio of functional to non-functional particles, all of which are determined through post-manufacturing analytical testing [Logan, A. C. *et al*., 2002]. These measurements directly determine whether a product meets release specifications and what dose information will appear on the final label. Until these assays are complete, the label cannot carry accurate dose or potency information.

This late-stage data generation is not a manufacturing deficiency—it reflects the biological nature of the products themselves. It has been extensively documented in the ATMP literature as a defining feature of the manufacturing landscape that distinguishes CGT from other pharmaceutical sectors [Harrison, R. P. *et al*., 2018]. Quality systems designed for ATMPs must therefore accommodate a workflow in which labeling completion is necessarily deferred, and where interim product states require management through mechanisms other than finalized physical labeling.

The consequences of failing to accommodate this reality are operationally significant. Placeholder values on labels introduce transcription error risk at the point of correction. Manual interventions to update labels after attribute confirmation create documentation gaps and increase the probability of label mix-ups, particularly in facilities running multiple patient-specific campaigns simultaneously [Bravery, C. A. *et al*., 2013]. These risks are amplified in decentralized manufacturing environments, where products may be processed at sites remote from the clinical administration point and must move through complex logistics chains before reaching the patient.

Table 1: Key Biological Drivers of Late-Stage Attribute Determination in CGT Manufacturing [Ferrer, L. *et al*., 2016; Harrison, R. P. *et al*., 2018]

Attribute	Basis for Late Determination
Potency	Confirmed through post-manufacturing functional assays; cannot be predicted from process parameters alone
Viability	Measured at the end of manufacturing, influenced by culture conditions and donor cell characteristics
Dose / Fill Volume	Determined by final cell count or vector titer after downstream processing
Purity	Assessed via release testing following completion of all manufacturing steps
Vector Genome Titer	Quantified post-production through analytical characterization of viral vector batches

Publisher: SARC Publisher

Regulatory Frameworks: Functional Requirements Without Structural Prescriptions

The regulatory frameworks governing ATMPs in major jurisdictions share a common emphasis on product identification, chain-of-identity, chain-of-custody, and traceability. In the United States, the relevant framework under 21 CFR Part 1271 establishes requirements for labeling human cells, tissues, and cellular and tissue-based products with information necessary to identify the product and trace it through the distribution chain [Martin, R. M. 2025]. In the European Union, EudraLex Volume 4 Annex 2 sets out good manufacturing practice (GMP) requirements for biological medicinal products, including ATMPs, with particular emphasis on traceability and the management of product identity throughout manufacture [EudraLex-Volume, E. U. 2003].

Both frameworks are notably silent on the structural question of how labeling should be organized across the manufacturing timeline. They require that labels be accurate and that products be traceable—but they do not specify when during the manufacturing process a label must be complete, what information must be physically present on a label versus accessible through a connected digital system, or how interim product states should be managed before final attribute data is available. This regulatory flexibility is intentional; it reflects a recognition that ATMP manufacturing is diverse and that prescriptive structural requirements would constrain the operational innovation needed as the field matures [Marks, P., & Gottlieb, S. 2018].

However, this flexibility has had an unintended consequence. In the absence of clear regulatory guidance on interim labeling practices, manufacturers have defaulted to legacy structural models—primarily two-stage labeling—that predate the digital manufacturing environments now common in CGT facilities. These models were developed for environments where paper-based records, manual data entry, and limited system integration made physical interim labels the only practical mechanism for communicating product status across manufacturing stages [Szkodny, A. C., & Lee, K. H. 2022]. Their persistence reflects institutional inertia rather than regulatory necessity.

The functional requirements of the existing frameworks are, in fact, consistent with a range of labeling architectures beyond two-stage labeling. A label that carries a unique product identifier and links to a validated digital record containing complete attribute data satisfies the traceability requirements of both 21 CFR Part 1271 and EudraLex Volume 4 Annex 2 [4; 9]. The regulatory question is not whether the label is complete at all times—it is whether the product is accurately identifiable and traceable at all times. These are related but distinct requirements, and conflating them has produced labeling practices that impose unnecessary operational complexity without delivering proportionate compliance benefit.

Table 2: Regulatory Framework Requirements Relevant to ATMP Labeling [Martin, R. M. 2025; EudraLex-Volume, E. U. 2003]

Regulatory Instrument	Jurisdiction	Core Labeling Requirement
21 CFR Part 1271	United States	Product identification and traceability through distribution
EudraLex Volume 4 Annex 2	European Union	GMP traceability, chain-of-identity, product identification
21 CFR Part 11	United States	Electronic records and signatures' validity

The Limitations of Two-Stage Labeling in Modern CGT Facilities

Two-stage labeling, in its conventional form, involves the application of an initial label at an early point in manufacturing—typically identifying the product as quarantined, in-process, or conditionally designated for a specific patient—followed by the application of a completed label once all release attributes have been confirmed. The initial label may carry partial information; the final label carries the complete attribute set required for release and administration.

This approach was operationally rational in facilities where manufacturing execution systems and laboratory information management systems either did not exist or were not integrated with labeling workflows. In such environments, the physical label was the primary mechanism for communicating product status, and a two-stage approach allowed that communication to occur

Publisher: SARC Publisher

even before all product data was available [Szkodny, A. C., & Lee, K. H. 2022]. The limitations of the approach—including the risk of transcription errors during label updates, the complexity of managing multiple label states for concurrent campaigns, and the difficulty of maintaining label version control—were accepted as unavoidable costs of operating in a data-sparse environment.

Modern biomanufacturing facilities are not data-sparse. Integrated manufacturing execution systems provide real-time batch tracking, electronic status management, automated data capture across process stages, and dynamic linkage between process data and laboratory results [Charaniya, S. *et al*., 2010]. Laboratory information management systems manage release testing workflows, record analytical results, and can trigger downstream process events—including label generation—upon result confirmation [Randers-Eichhorn, L. *et al*., 1996]. In this environment, the physical interim label is no longer the only mechanism, or even the most reliable mechanism, for communicating product status through the facility. Digital systems can manage interim product states with greater accuracy, auditability, and configurability than physical labels can.

The persistence of two-stage labeling in this environment reflects a gap between available technology and adopted practice. It also reflects the validation burden associated with changing established quality management system procedures in regulated environments—a burden that is real but should be weighed against the operational risks that legacy labeling practices introduce. Marks and Gottlieb have argued that regulatory frameworks for advanced therapies are intended to support innovation rather than entrench legacy practice and that manufacturers operating in a flexible regulatory environment bear a corresponding responsibility to adopt practices commensurate with their technical capabilities [Marks, P., & Gottlieb, S. 2018].

The risks associated with two-stage labeling in multi-patient CGT facilities are not hypothetical. The simultaneous management of multiple patient-specific campaigns creates conditions in which label mix-ups—applying a corrected or finalized label to the wrong container—can have direct patient safety consequences. The severity of these consequences, given the patient-specific nature of most CGT products, elevates label accuracy to a patient safety issue rather than merely a quality documentation issue [Bravery, C. A. *et al*., 2013].

Table 3: Operational Limitations of Two-Stage Labeling in CGT Environments [Szkodny, A. C., & Lee, K. H. 2022; Bravery, C. A. *et al*., 2013]

Limitation	Operational Consequence
Manual label corrections after attribute confirmation	Elevated transcription error risk; documentation gaps
Multi-campaign concurrent management	Increased label mix-up probability in patient-specific workflows
Physical label as sole status carrier	No real-time synchronization with manufacturing execution data
Version control complexity	Risk of superseded label versions remaining in circulation

Alternative Architectures: Just-in-Time Labeling and Minimal-Label Models

Two operational alternatives to conventional two-stage labeling have emerged from the practical experience of CGT manufacturers and from the broader literature on digital transformation in biomanufacturing. These approaches differ in their technical requirements and operational implications, but both are grounded in a common principle: the physical label should carry only the information that is available and confirmed at the time of labeling, while additional attributes are managed through validated digital systems until they can be incorporated into the label or accessed via digital linkage.

The first alternative is just-in-time labeling. In this approach, final label printing is deferred until all required release attributes have been confirmed through quality control testing. The label is generated automatically by the manufacturing execution or laboratory information management system upon confirmation that all release criteria have been met, and it carries the complete attribute set from the moment it is applied [Charaniya, S. *et al*., 2010]. This eliminates the need for interim status labels entirely by collapsing the labeling

Publisher: SARC Publisher

event to the point of data completeness. It requires close integration between quality control data systems and label generation platforms, and it places operational demands on storage and logistics workflows that must accommodate products in an unlabeled or minimally identified state during the testing interval.

The second alternative employs a minimal-label architecture with external data anchoring. Under this model, a label is applied at an early point in manufacturing, carrying only the information that is confirmed and stable at that time—typically a unique product identifier, a lot or batch number, and basic storage condition requirements. All variable or late-determined attributes are maintained in a connected digital system and linked to the product via the unique identifier [Ferrer, L. *et al*., 2016]. The label functions as an identity anchor rather than a complete data carrier. Chain-of-identity and chain-of-custody are maintained through the digital record, which is updated in real time as attribute data becomes available. The physical label does not require updating because it was never intended to carry the complete attribute set.

Both architectures carry trade-offs that must be evaluated in the context of specific manufacturing environments. Just-in-time labeling reduces the risk of interim label errors but requires robust system integration and may introduce time pressure in release workflows, particularly for products with short shelf lives [Rader, C. 2020]. Minimal-label architectures reduce the complexity of interim label management but require sustained digital system availability, validated data linkage, and procedural frameworks that define the authoritative status of digital records relative to physical labels. Neither approach is universally superior; the appropriate choice depends on facility capabilities, product characteristics, and the specific operational risks being managed.

What both approaches share is a reconceptualization of the label's function. Rather than treating the physical label as the primary repository of product attribute data, they treat it as one element in an information architecture in which digital systems carry the authoritative record. This reconceptualization is consistent with the functional requirements of existing regulatory frameworks and with the direction of digital transformation across biomanufacturing more broadly [Lipsitz, Y. Y. *et al*., 2016].

Table 4: Comparative Characteristics of Alternative Labeling Architectures [Charaniya, S. *et al*., 2010; Rader, C. 2020]

Characteristic	Just-in-Time Labeling	Minimal-Label with Digital Anchor
Label applied	At the point of full data confirmation	At the early manufacturing stage
Information on the label	Complete attribute set at time of application	Unique identifier, lot number, storage conditions
System integration requirement	High — triggered by release data confirmation	Moderate — requires validated digital linkage
Interim product management	Unlabeled or minimally identified during testing	Labeled with an identity anchor throughout
Primary risk	Time pressure in short-shelf-life workflows	Dependence on digital system availability

Toward a Data-Centric Labeling Standard for ATMPs: Gaps and Recommendations

The transition from legacy labeling practices to data-centric architectures in CGT manufacturing is technically feasible and regulatorily consistent, but it is not yet supported by a coherent framework of guidance, standards, or validated implementation models. Several gaps must be addressed before data-centric labeling can be adopted at scale across the ATMP manufacturing sector.

The first gap is regulatory clarity. While existing frameworks permit data-centric labeling approaches, they do not explicitly address them. Manufacturers seeking to implement minimal-label architectures or just-in-time labeling systems must currently construct regulatory justifications from first principles, interpreting silence in the guidance documents as permissive rather than prohibitive [Martin, R. M. 2025; EudraLex-Volume, E. U. 2003]. Explicit regulatory guidance on the use of digital systems as authoritative data sources for ATMP labeling—addressing validation requirements, audit trail expectations, and the conditions under which a physical label may carry reduced attribute information—would substantially reduce the regulatory uncertainty that currently deters adoption.

Publisher: SARC Publisher

The second gap is in validation frameworks for dynamic label generation. Just-in-time labeling requires that label generation be triggered automatically by system events—specifically, the confirmation of release attribute data—in a manner that is accurate, auditable, and resistant to system failure. Validation of these workflows under applicable GMP frameworks, including 21 CFR Part 11 in the United States and the equivalent EU requirements for electronic records and signatures, requires methodologies that are not yet standardized for ATMP applications [EudraLex-Volume, E. U. 2003]. Development of validated templates and reference architectures for dynamic label generation would accelerate implementation across the sector.

The third gap is in standards for decentralized manufacturing environments. As CGT manufacturing increasingly moves toward decentralized models—processing at or near the clinical administration site rather than at centralized production facilities—the chain-of-identity and chain-of-custody challenges associated with labeling are amplified [Harrison, R. P. *et al*., 2018]. Products move through more complex logistics chains, may be processed at sites with more limited digital infrastructure, and must maintain unambiguous traceability across greater operational distances. Labeling standards designed for centralized manufacturing environments may not translate directly to decentralized contexts, and purpose-built guidance for decentralized ATMP labeling is currently limited.

The fourth gap is integration between process analytical technology (PAT) and labeling workflows. Process analytical technologies offer the potential to accelerate the generation of the late-stage attribute data that currently creates the labeling window problem—if real-time potency and viability measurements can be confirmed during manufacturing rather than after completion, the temporal gap between product availability and label completion narrows [Rader, C. 2020]. Realizing this potential requires not only the development and validation of real-time analytical methods but also the integration of PAT outputs into manufacturing execution and label generation systems in a manner that satisfies release testing requirements.

Addressing these gaps will require coordinated effort across regulatory bodies, manufacturing technology developers, and the ATMP manufacturing community. The regulatory frameworks governing ATMPs have been designed with sufficient flexibility to support this transition [Marks, P., & Gottlieb, S. 2018]; what is needed now is the operational and technical infrastructure to implement it. The fundamental principle is straightforward: the label identifies, the digital system records, and the two together constitute a complete and accurate traceability framework for patient-specific CGT products.

Table 5: Implementation Gaps for Data-Centric Labeling in ATMP Manufacturing [Martin, R. M. 2025; EudraLex-Volume, E. U. 2003; Marks, P., & Gottlieb, S. 2018]

Gap Area	Description
Regulatory clarity	No explicit guidance on digital record authority relative to physical label completeness
Validation frameworks	Absence of standardized templates for dynamic, event-triggered label generation under GMP
Decentralized manufacturing	Labeling standards designed for centralized production may not apply to distributed CGT sites
PAT integration	Real-time analytical outputs not yet systematically linked to label generation workflows

CONCLUSION

Labeling in CGT manufacturing is not a peripheral quality documentation task—it is a patient safety function. The patient-specific nature of most CGT products means that a labeling error does not affect a product lot shared across many recipients; it affects a single patient for whom no substitute product exists. This elevates label accuracy, chain-of-identity, and traceability from quality system priorities to direct clinical risk factors, and it demands that labeling architectures be designed accordingly. The core tension driving labeling complexity in CGT is temporal: critical quality attributes that must appear on a label are generated after the manufacturing process is complete, through release testing workflows that cannot be accelerated beyond the biological time constants of the assays themselves. This gap between product availability and data availability is not a manufacturing deficiency—it is a property of the

Publisher: SARC Publisher

biology. Any labeling architecture that does not explicitly account for it will produce either inaccurate interim labels, operationally burdensome correction workflows, or both. Digital manufacturing infrastructure—integrated execution systems, laboratory information management platforms, and validated electronic record frameworks—has created the technical conditions under which data-centric labeling is not only feasible but preferable. A label that anchors product identity while a connected digital system carries the complete and current attribute record satisfies the functional intent of applicable regulatory frameworks without requiring that all data be physically present on a label before it exists. Realizing this model at scale requires regulatory clarity on digital record authority, validated templates for dynamic label generation, and purpose-built guidance for decentralized ATMP manufacturing environments.

REFERENCES

1. Rathore, A. S., & Winkle, H. "Quality by design for biopharmaceuticals." *Nature biotechnology* 27.1 (2009): 26-34.
2. Kouris, N. A., Squirrell, J. M., Jung, J. P., Pehlke, C. A., Hacker, T., Eliceiri, K. W., & Ogle, B. M. "A nondenatured, noncrosslinked collagen matrix to deliver stem cells to the heart." *Regenerative medicine* 6.5 (2011): 569-582.
3. Ferrer, L., Kimbrel, E. A., Lam, A., Falk, E. B., Zewe, C., Juopperi, T., & Hoffman, A. "Treatment of perianal fistulas with human embryonic stem cell-derived mesenchymal stem cells: a canine model of human fistulizing Crohn's disease." *Regenerative medicine* 11.1 (2016): 33-43.
4. Martin, R. M. "101 Topics for Clinical Microbiology Laboratory Leaders: Accreditation, Verification, Quality Systems, and More." *John Wiley & Sons*, (2025).
5. Szkodny, A. C., & Lee, K. H. "Biopharmaceutical manufacturing: Historical perspectives and future directions." *Annual Review of Chemical and Biomolecular Engineering* 13.1 (2022): 141-165.
6. Logan, A. C., Lutzko, C., & Kohn, D. B. "Advances in lentiviral vector design for gene-modification of hematopoietic stem cells." *Current opinion in biotechnology* 13.5 (2002): 429-436.
7. Harrison, R. P., Ruck, S., Rafiq, Q. A., & Medcalf, N. "Decentralised manufacturing of cell and gene therapy products: Learning from other healthcare sectors." *Biotechnology advances* 36.2 (2018): 345-357.
8. Bravery, C. A., Carmen, J., Fong, T., Oprea, W., Hoogendoorn, K. H., Woda, J., & Van't Hof, W. "Potency assay development for cellular therapy products: an ISCT review of the requirements and experiences in the industry." *Cytotherapy* 15.1 (2013): 9-19.
9. EudraLex-Volume, E. U. "Volume 4 Good manufacturing practice (GMP) Guidelines." *European Commission* (2003).
10. Marks, P., & Gottlieb, S. "Balancing safety and innovation for cell-based regenerative medicine." *New England Journal of Medicine* 378.10 (2018): 954-959.
11. Charaniya, S., Le, H., Rangwala, H., Mills, K., Johnson, K., Karypis, G., & Hu, W. S. "Mining manufacturing data for discovery of high productivity process characteristics." *Journal of biotechnology* 147.3-4 (2010): 186-197.
12. Randers-Eichhorn, L., Bartlett, R. A., Frey, D. D., & Rao, G. "Noninvasive oxygen measurements and mass transfer considerations in tissue culture flasks." *Biotechnology and bioengineering* 51.4 (1996): 466-478.
13. Rader, C. "Bispecific antibodies in cancer immunotherapy." *Current opinion in biotechnology* 65 (2020): 9-16.
14. Lipsitz, Y. Y., Timmins, N. E., & Zandstra, P. W. "Quality cell therapy manufacturing by design." *Nature biotechnology* 34.4 (2016): 393-400.

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:
Guttikonda, P. "Data-Driven Labeling Architectures in Cell and Gene Therapy Manufacturing." *Sarcouncil Journal of Internal Medicine and Public Health* 5.3 (2026): pp 57-63.

Publisher: SARC Publisher