International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Representing Pandemic Impacts in Clinical Trial Data and Outputs

Jyoti (Jo) Agarwal[1], Sushil Tripathi, PhD[2]

Abstract: ***The COVID-19 pandemic disrupted clinical trial conduct through missed or modified visits, site and travel restrictions, changes in laboratory procedures, protocol deviations, treatment interruptions, and participant-level infection or related events. These impacts can appear across multiple collection sources and are difficult to summarize consistently when teams identify them independently. This paper describes a standards-oriented implementation framework for tracing pandemic-related information from collected data through SDTM, ADaM, and selected tables, figures, and listings (TFLs). The approach begins with a source inventory and a controlled impact taxonomy. It then links relevant records across source domains, represents visit modifications and other impacts in the appropriate standardized structures, derives subject- and event-level analysis indicators, and produces outputs from validated analysis data. Free-text classification can help locate candidate records, but requires review against the source text and related records. A focused reconciliation process is essential when operational, adverse-event, and visit data do not agree. The framework is intended as a reusable implementation aid, not as a report of study outcomes or a replacement for current CDISC and regulatory guidance. It can help programming and data standards teams improve traceability, consistency, and reviewability while preserving the distinction between pandemic-related disruption and disease-specific clinical events.***

Keywords: COVID-19 Clinical Trial Disruptions, CDISC Data Standards, SDTM and ADaM Implementation, Clinical Data Traceability, Pandemic-Related Protocol Deviations

1. Introduction

The COVID-19 public health emergency changed how many clinical studies operated. Participants could miss scheduled visits, complete assessments remotely, receive laboratory testing locally, or experience delays in study treatment and follow-up. Guidance issued during the emergency emphasized documenting changes to trial conduct and their effects on participants and study data (FDA 2020). CDISC also published guidance and examples for representing pandemic-related information in clinical data standards and analysis structures (CDISC 2020; CDISC 2023).

For statistical programmers, the practical challenge was to convert scattered operational information into a traceable set of standardized records and analysis variables. Relevant information could be collected in visit records, protocol-deviation records, adverse-event forms, concomitant medication forms, laboratory records, disposition data, and free-text comments. A single participant could have several distinct impacts, and a single event could be described differently across systems. A reliable implementation therefore needs explicit rules for identifying candidate records, resolving discrepancies, carrying source context forward, and deriving output populations.

This paper presents a generalized framework for that work. It presents no participant-level data and makes no claim about the performance of a particular study implementation.

2. Scope and Design Principles

The framework addresses study-conduct impacts related to an epidemic or pandemic. It distinguishes those impacts from clinical events that indicate infection or disease. For example, a missed visit attributed to pandemic restrictions and an adverse event representing the infection itself may both be relevant, but they answer different questions and should retain separate meanings in the data.

Four design principles guide the implementation. First, preserve the source record and its origin so reviewers can trace each derived indicator. Second, use a common impact taxonomy across functional sources while retaining domain-specific details. Third, apply the current version of CDISC guidance and study metadata rather than treating an example mapping as universally required. Fourth, generate TFLs only after source-to-analysis reconciliation and documented quality review.

3. Source Inventory and Impact Taxonomy

The first step is to list every collection source that could capture pandemic-related information. The inventory should identify the source owner, record key, relevant dates, text or coded fields, and known limitations. Common source groups include visit records, protocol deviations, adverse events, treatment exposure or interruption records, disposition, laboratory assessments, concomitant medications, and general comments. A study may use only a subset of these sources.

A shared taxonomy makes the source inventory useful for programming and review. Candidate categories include visit missed, visit conducted remotely or by an alternate method, assessment not performed, local laboratory used, protocol deviation, treatment interruption, treatment discontinuation, study discontinuation, hospitalization, infection-related event, and death. Categories should be adapted to the protocol and collection design. Teams should also record whether the relationship is direct, indirect, or uncertain when the source allows that distinction.

Licensed Under Creative Commons Attribution CC BY

Table 1: Example source inventory and impact categories. The applicable domain mapping depends on the study design and current implementation guidance

Impact signal	Typical source	Standardization and analysis use
Missed or modified visit	Visit records and visit-related comments	Represent the visit outcome, contact mode, and reason using the applicable SDTM visit guidance; derive an analysis indicator when defined by the analysis plan.
Protocol deviation	Deviation records and operational descriptions	Retain deviation classification and pandemic relationship; distinguish important from non-important deviations when the study collects that distinction.
Infection-related event	Adverse-event and disease assessment records	Keep event-level clinical information distinct from general study-conduct impact indicators.
Local laboratory assessment	Laboratory records and source annotations	Preserve the result and identify local testing where collected and relevant to interpretation.
Treatment or study discontinuation	Exposure and disposition sources	Reconcile the recorded status and reason across source records before deriving subject-level indicators.
Other operational impact	Comments, concomitant medication, or other forms	Use as a candidate source; retain the source link and require review before classifying or deriving an analysis flag.

4. Source Review and Record Linkage

Source identification should combine structured fields with targeted review of free text. Keyword searches or natural-language processing can identify candidate comments and descriptions, but they should serve as screening methods. A classifier may assign a provisional category such as missed visit, modified visit, or unclassified. A reviewer should compare the category with the original text, the scheduled and actual visit information, and related source records before accepting it. This review is especially important when the text contains ambiguous phrases, multiple impacts, or no explicit causal link to the pandemic.

Record linkage rules should be written before programming. They should define the participant key, event or visit key, date tolerances, matching priority, treatment of duplicate comments, and handling of records with incomplete identifiers. If one source identifies a modified visit while another indicates that the visit was not performed, the team should determine whether these describe the same visit, separate contacts, or a data-entry inconsistency. Unresolved conflicts should remain visible in a review log rather than being silently resolved by a derived flag.

SDTM Representation

SDTM mapping should follow the applicable SDTM Implementation Guide and CDISC guidance for disrupted studies. For visit impacts, the implementation should represent whether a visit occurred and how contact took place, when these details are collected and supported by the standards. The reason for a missed or modified visit and the relationship to the pandemic should be represented in the location specified by the applicable guidance and the study's standards strategy. Where a non-standard variable or supplemental qualifier is necessary, its metadata and rationale should be documented.

Other impact records should stay in the domain that represents their clinical or operational meaning. Protocol deviations should remain recognizable as deviations; adverse events should remain event records; local laboratory results should remain laboratory observations; and disposition should continue to describe treatment or study completion. A broad pandemic indicator can support cross-domain review, but it should not erase the underlying event, reason, timing, or classification. The reviewer guide should explain any implementation choices that generate validation findings or differ from a standard structure.

ADaM Derivations and TFL Production

Analysis datasets can provide subject-level and event-level indicators that support consistent summaries. Subject-level indicators may capture whether a participant experienced a specified pandemic-related impact, while event-level indicators retain the detail needed for event summaries and listings. Composite indicators should have transparent component definitions. Their derivations should specify source variables, handling of missing or conflicting source values, date rules, and whether the indicator refers to a direct clinical event or a broader operational relationship.

TFL shells should be mapped to these definitions before production. Potential outputs include summaries of treatment and study discontinuation attributed to the pandemic, protocol-deviation listings, missed or modified visit listings, participant impact listings, and listings of infection-related adverse events. Each output should state its population, denominator, source analysis variables, and inclusion criteria. Outputs that contain no qualifying records should be handled according to the statistical analysis plan and reporting conventions; empty listings should not be populated with invented or placeholder records.

A critical review step compares each output back to its contributing analysis records and then to the source data. The review should confirm that denominators use the intended analysis set, percentages use the stated denominator, reasons and categories reconcile, and the same participant is not counted inconsistently across related outputs. When source records disagree, the output should follow a documented adjudication rule or the issue should be escalated for data clarification.

Quality Control and Governance

A practical quality plan includes checks at each handoff. At source review, confirm that candidate text and structured records are classified against the agreed taxonomy. At SDTM, verify that records retain source identity and that visit, event, deviation, laboratory, and disposition concepts remain in their intended domains. At ADaM, independently review indicator derivations, component logic, traceability, and population membership. At TFL production, compare counts and records with analysis data, inspect display labels and footnotes, and

Licensed Under Creative Commons Attribution CC BY

verify that the denominator and output filters match the approved specification.

Data management, clinical operations, safety, standards, and programming teams should agree on definitions for terms such as “pandemic-related,” “missed visit,” “modified visit,” and “discontinued due to pandemic.” The definition should identify who makes the attribution, which source takes precedence, and how uncertainty is represented. These decisions belong in controlled specifications or a reviewer guide so that later data refreshes do not change the interpretation unnoticed.

5. Lessons and Limitations

A cross-source inventory reduces the risk that a relevant impact remains isolated in an operational comment or domain-specific field. A shared taxonomy reduces inconsistent labels across datasets and outputs. A visible reconciliation process is especially valuable for status and reason variables because operational records can lag behind event or exposure data. These are practical design implications; the framework does not report quantified effectiveness findings.

This paper describes a framework rather than a validated production package. It does not provide executable code, dataset specifications, a complete CDISC mapping, or study results. COVID-19 data collection practices and regulatory expectations varied by study and changed over time. Teams should verify the current standards version, FDA guidance, protocol, statistical analysis plan, and sponsor data standards before implementation. This manuscript contains no participant-level data, sponsor names, study identifiers, internal systems, proprietary variable names, or study-specific results. Any use of historical implementation knowledge in a public submission should complete applicable employer and legal review.

6. Conclusion

A defensible pandemic-impact reporting workflow links the source record, standardized representation, analysis derivation, and TFL inclusion rule. The core work is to define a shared impact taxonomy, review automated text classification, reconcile sources that describe the same event, and preserve traceability through SDTM and ADaM. This framework offers statistical programming and data standards teams a practical starting point for organizing pandemic-related information while maintaining clear distinctions among clinical events, trial-conduct changes, and derived analysis indicators.

References

[1] CDISC. 2020. “Guidance for Ongoing Studies Disrupted by COVID-19 Pandemic.” COVID-19 Therapeutic Area. https://www.cdisc.org/standards/therapeutic-areas/covid-19.
[2] CDISC. 2023. “COVID-19 Therapeutic Area User Guide v2.0.” https://www.cdisc.org/standards/therapeutic-areas/covid-19/covid-19-therapeutic-area-user-guide-v2-0.
[3] U.S. Food and Drug Administration. 2020. “Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency: Guidance for Industry, Investigators, and Institutional Review Boards.” https://www.fda.gov/media/137496/download.

Licensed Under Creative Commons Attribution CC BY