D-pack - Tool to effectively compile m5 structure

Rajesh Kulal, Zifo, Chennai, Tamil Nadu, India

Priyadarsini Gopalakrishnan, Zifo, Chennai, Tamil Nadu, India

BENGALURU, INDIA PharmaSUG SDE APRIL 11, 2026

Abstract

The m5 (Module 5) in the eCTD is an essential framework to display clinical data and results for regulatory submissions. This consistent structure aids in easy navigation for reference and review. However, manual compilation of the folder is arduous, long-drawn and prone to missing key documents.

Once the necessary inputs are placed in the study folder, the D-pack tool automates the m5 folder creation and compilation, consisting of SDTM/ADaM/TLFs datasets/reports, programs, define.xml, reviewer's guidelines, arm, QC documents, protocol, SAP and annotated CRF in their valid standard format as per guidelines. In addition, the tool performs a few checks and throws a warning for missing files, a datetime stamp mismatch and modlog verification for the right Sponsor name and Study ID. This reduces the manual error and time required for putting the files together to a great extent. A detailed walkthrough of the tool's working principle and advantages will be shared at the conference.

Introduction

Preparing a clinical submission is a complex and resource-intensive effort, often involving hundreds of interdependent activities and files. With the strict placement rules of the eCTD, building and maintaining an accurate Module 5 folder structure can become overwhelming, increasing the risk of errors, omissions, and rework. This is where the D-Pack tool comes into play by streamlining m5 creation, enforcing structure and consistency, and allowing teams to stay focused on the science rather than the logistics of submission assembly.

Module 5 Structure

The m5 section of the eCTD provides a standardized structure for regulatory submissions, ensuring that all clinical study reports, datasets, and supporting documentation are organized in a consistent, globally recognized format. This harmonized organization plays a crucial role in enabling regulatory reviewers to efficiently locate, interpret, and verify clinical evidence.

A major advantage of the m5 framework is its emphasis on ease of navigation for reference and review as seen in Figure 1. By categorizing clinical information into well-defined subsections such as tabular study listings, clinical study reports, additional study related documents, and literature references, the structure supports faster and more accurate regulatory evaluation. This standardized accessibility is essential because incomplete or poorly organized m5 submissions are a major cause of review delays and Refuse to File outcomes.

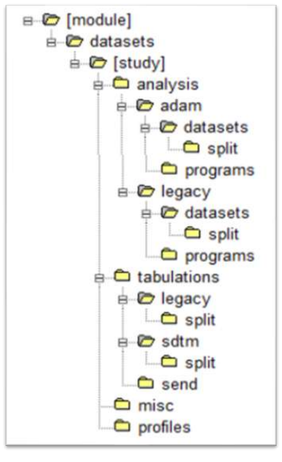

Figure 1. Folders in the m5 Structure

The m5 module follow strict formatting, and organizational rules. These requirements ensure consistency, and compliance across submissions. This promotes transparency and clarity of clinical trial data. By consolidating all clinical evidence it enhances regulatory understanding and supports well-informed decision making. The consistent format also improves clarity for downstream users such as auditors, data reviewers, and health authorities who depend on transparent documentation to assess the safety and efficacy of investigational products.

Manual Limitations

However, assembling this module manually introduces several operational challenges:

1. Time - Manual compilation consumes enormous effort

Constructing an m5 folder by hand means organizing hundreds of files into a very strict hierarchical structure. This process is long, tedious, and highly resource-intensive, particularly when repeated for multiple studies in a program which can lead to delays or additional rework, amplifying the overall time burden.

2. Standards Compliance – Often compromised in manual workflows

Manual assembly increases the likelihood of inconsistencies or missing regulatory requirements such as: Absent mandatory subfolders, Incorrect or inconsistent naming conventions, Missing CSRs, or datasets, Incorrect dataset formatting.
Such deviations are repeatedly cited as root causes of regulatory queries and submission deficiencies.

3. Human Errors – High risk when managing large volumes of data

With a vast quantity of files and strict placement rules, manual processes create opportunities for:

- Misplacing files
- Omitting essential documents
- Outdated files

D-Pack

To overcome the above mentioned errors, D-Pack Tool, developed via R and implemented using a 8 step process as described in Figure 3, integrates a robust layer of automated validation checks.

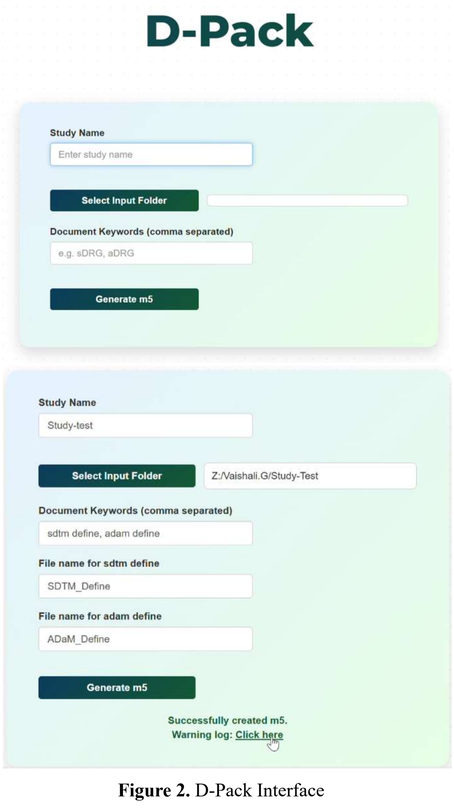

Figure 2. D-Pack Interface

These checks ensure that every file placed in the common input folder is placed into the m5 structure and is complete, consistent, and compliant before the final package is generated.

1. Modlog Verification

D-Pack automatically checks Study ID and Sponsor Name consistency across all files, ensuring uniformity and preventing errors that commonly occur when documents are sourced from multiple contributors.

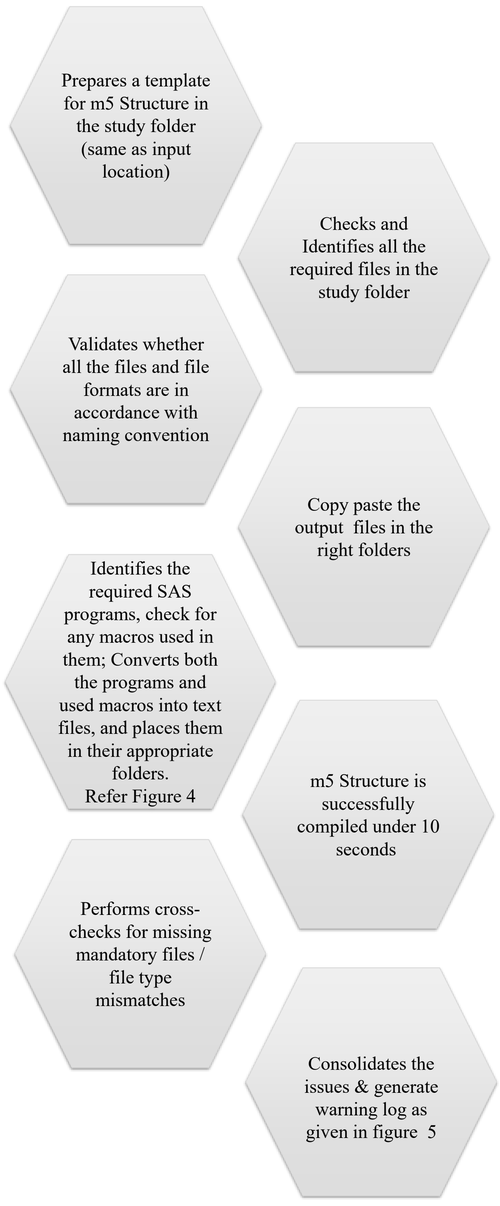

Figure 3. Workflow of D-Pack

2. Date–Time Stamp Validation

The tool validates the chronological order of outputs and validation documents, ensuring that timestamps follow expected sequencing. This preserves version integrity and supports regulatory traceability.

3. Missing Files Detection

D-Pack cross-checks the study folder against mandatory m5 requirements and instantly flags any missing components such as SDTM/ADaM datasets, TLFs, programs, define.xml, reviewer's guides, annotated CRF, ensuring nothing essential is overlooked.

4. Program–XPT Pair Validation

The tool verifies that each dataset has a corresponding program in the correct folder and highlights any missing pairs. This maintains clear traceability between data generation programs and their resulting outputs. All warnings, alerts, and validation notes generated from these checks are automatically consolidated into a single Excel report, giving teams a streamlined and actionable summary before the submission package is finalized.

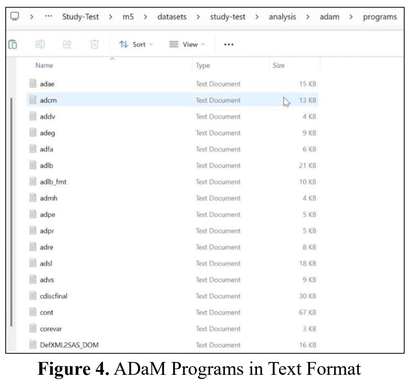

Figure 4. ADaM Programs in Text Format

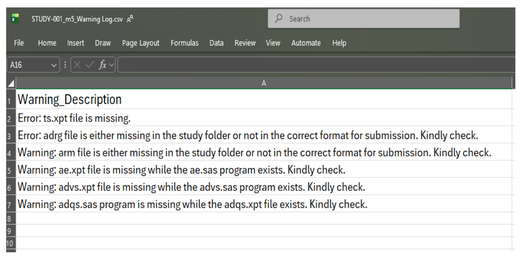

Figure 5. Consolidated Warning Log

Conclusions

Through its automated compilation, validation intelligence, and structured approach to organizing clinical documentation, D-Pack significantly modernizes and streamlines the creation of the m5 package. This reduces the manual effort, ensures standards compliance, and minimizes the human errors, thus making the m5 structure creation faster and more reliable for eCTD submissions. Its integrated checks ranging from modlog verification to program–XPT pairing ensure that every submission is both complete and compliant.

D-Pack will continue to be enhanced and future improvements include but are not limited to:

1. Split Datasets Support

D-Pack will incorporate the capability to automatically detect, manage, and organize split datasets, ensuring that large SDTM and ADaM domains exceeding file-size limits are handled smoothly and placed in the appropriate subfolders without manual intervention.

2. PMDA Submission Compatibility

The tool will be extended to support the specific structural and technical requirements of PMDA submissions, enabling seamless export options tailored to regional regulatory formats and ensuring compliance with Japan's eCTD lifecycle standards.

3. Legacy Studies Integration

D-Pack will introduce enhanced logic to accommodate legacy studies, allowing older formats, partial standards, and non-standard folder structures to be intelligently mapped into the correct m5 layout. This will help modernize historical packages, reduce rework, and support consolidated submissions that span multiple study generations.

These advancements will further strengthen D-Pack as a comprehensive submission-readiness solution bringing together automation, compliance, and future-proof design to support clinical documentation teams and regulatory stakeholders with greater efficiency and confidence.

References

- ICH – CTD structure & Module 5 requirements (ich.org)
- FDA eCTD Technical Conformance Guide, Module 5 (U.S. Food and Drug Administration)
- Electronic Common Technical Document Specification (ICH M2 EWG)

Contact Information

Your comments and questions are valued and encouraged. Contact the author at:

Name: Rajesh Kulal
Company: Zifo
E-mail: Rajesh.K@zifornd.com
Web: LinkedIn