PharmaSUG 2019 – Paper DS-49

Streamlining the Metadata Management Process Using SAS® Life Science Analytics Framework

Alex Ford, SAS Institute

ABSTRACT

It was not long into my clinical programming career before I discovered that CDISC is truly an acronym for "Can Do It Somewhat Correctly". Each run of a validation report uncovered new warnings or errors followed by tracking down the source of those issues to log and report for a define.xml. The latest release of the SAS® Life Science Analytics Framework (LSAF) provides a centralized framework where standards can be imported and live alongside a study and its data, managed by a graphical user interface. By associating a data standard, controlled terminology, and dictionaries with a study, team leads have the data and information necessary to produce a define.xml at the click of a button. Join us as we explore the metadata management features available in LSAF 5.1 which enable programmers of all levels to manage data standards correctly the first time, saving studies both time and money.

INTRODUCTION

LSAF is an end to end clinical platform that was specifically designed for life science customers and partners to help streamline the statistical progress of clinical trials. This is accomplished by housing all study related artifacts, providing a native code editor, establishing individual workspaces, mitigating inherent risks through audit trail and versioning, and in this latest release - providing a metadata management framework. This paper will take a deep dive into the metadata aspect and discuss how it assists clinical research teams to adhere to CDISC standards and enable Define.xml production at the click of a button.

The metadata management framework takes CDISC formats, controlled terminology, and external dictionaries and associates them with individual studies. This enables LSAF to run checks against the standards to report findings for reconciliation prior to define creation. LSAF also generates value level metadata and analysis results metadata which are combined with supporting documents in the define. From an admin perspective, metrics are provided giving team leads a high-level overview of associations between studies and standards for impact analyses. In the next sections we will walk through the look and feel of the framework and highlight how it makes your life easier as a programming team lead.

CLINICAL METADATA MANAGEMENT

Within LSAF, there is a specific section dedicated to Clinical Management which is broken down into the subsections Data Standards, Controlled Terminology, and External Dictionaries. These three subsections house the entirety of the metadata management materials which are associated with studies and used in the creation of a define package. The Data Standards consist of CDISC standards, while Controlled Terminology and External Dictionaries are flexible enough for manual definitions. Each entry is classified by type, has information concerning the standard on which they are based, and are versioned. The published state of a standard indicates whether that standard is in development, production, or retired. Only standards in the production state can be associated with new studies and only individuals with proper write access can develop new standards.

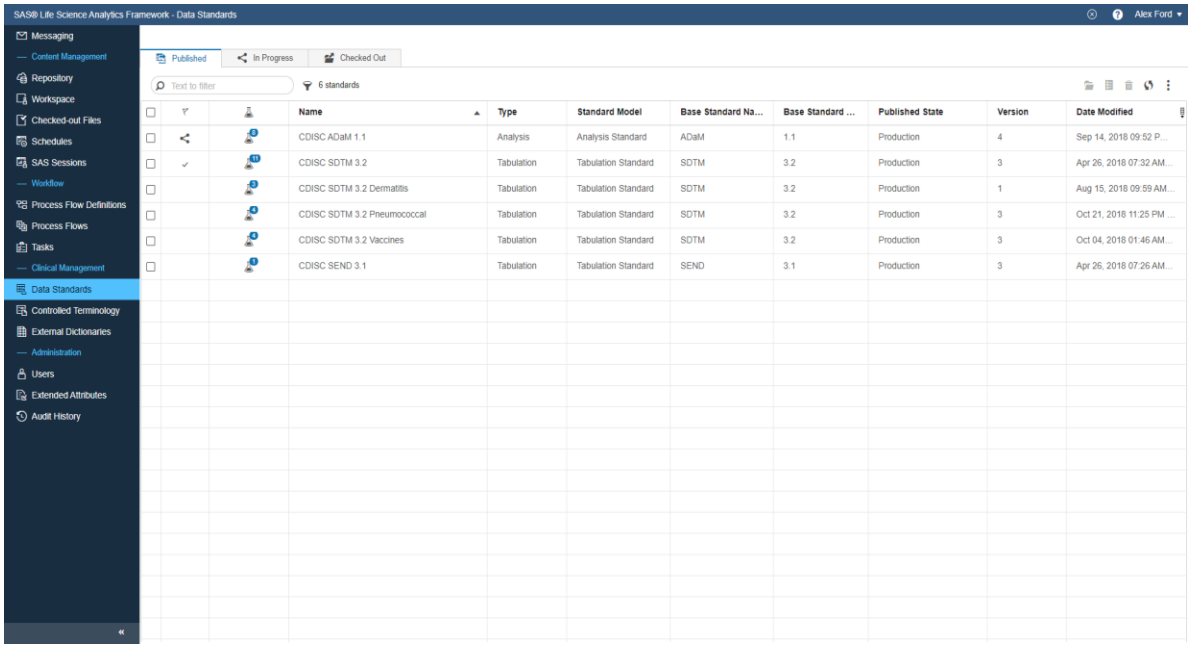

Figure 1. Data Standards

After picking a standard from the list in Figure 1 and associating it with a new study, study teams would begin by filling in the pertinent study information as it would be displayed in a CSR submitted to the FDA for review.

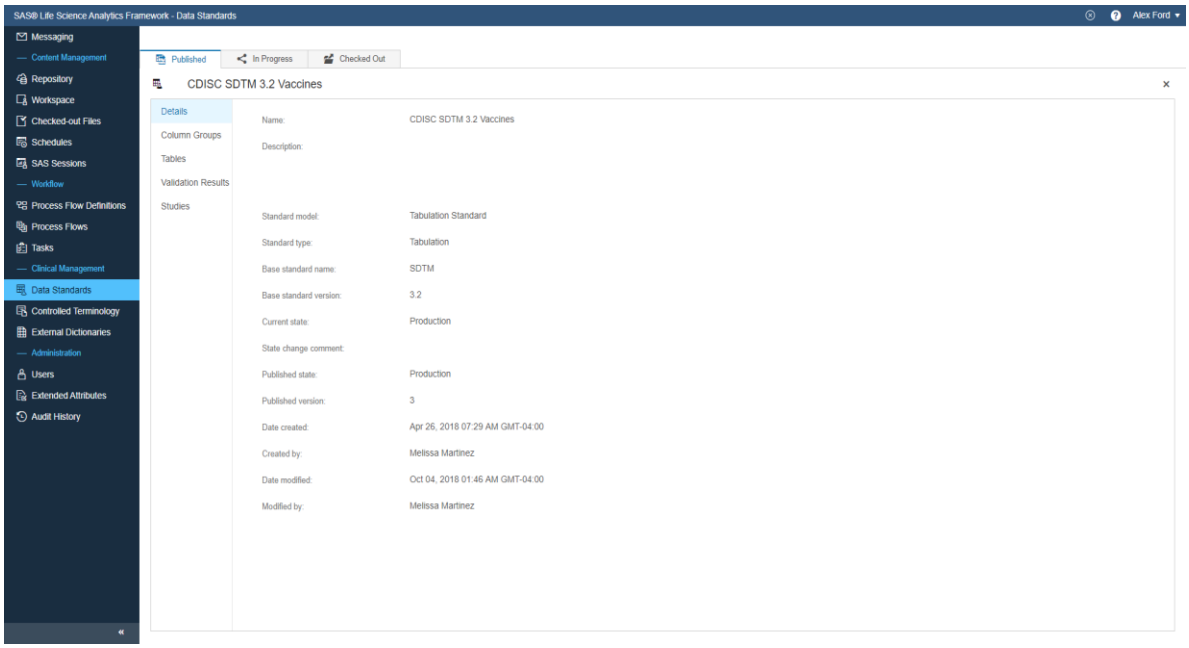

Figure 2. Study Details

Within each data standard, column groups are defined. The column groups house the variable names, labels, order, requirement, and computational methods. For example, here is the Associated Persons group:

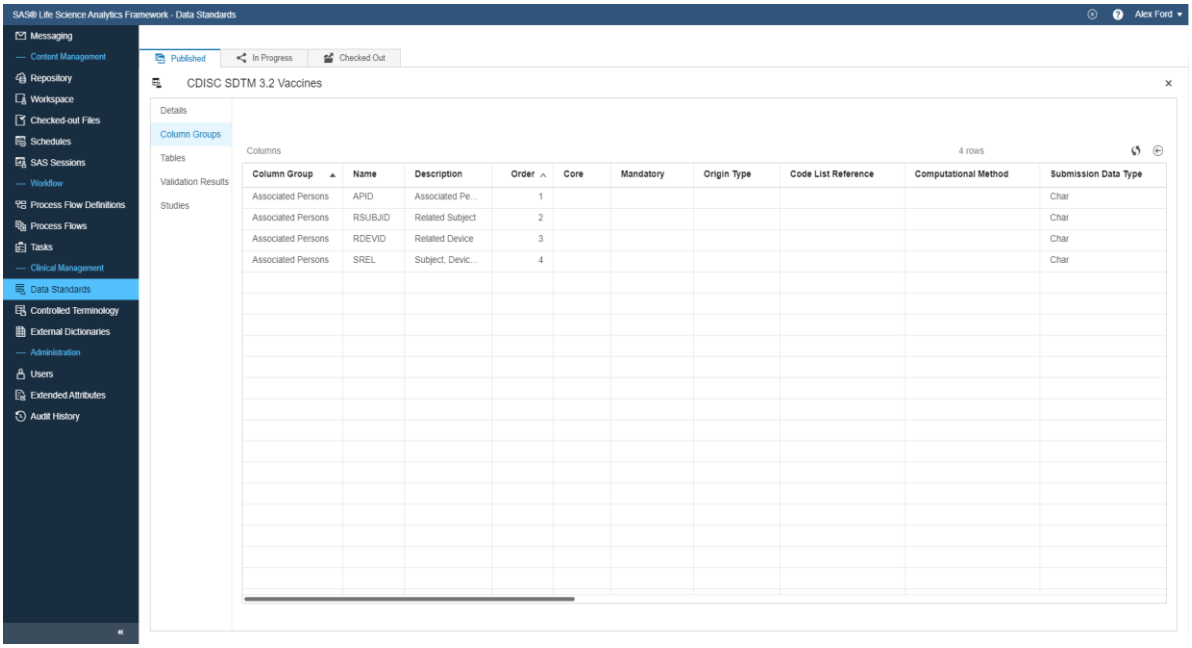

Figure 3. Column Groups

The Tables tab contains the name of each of the tables that is required as part of a particular data standard as well as its description, structure, order, and purpose.

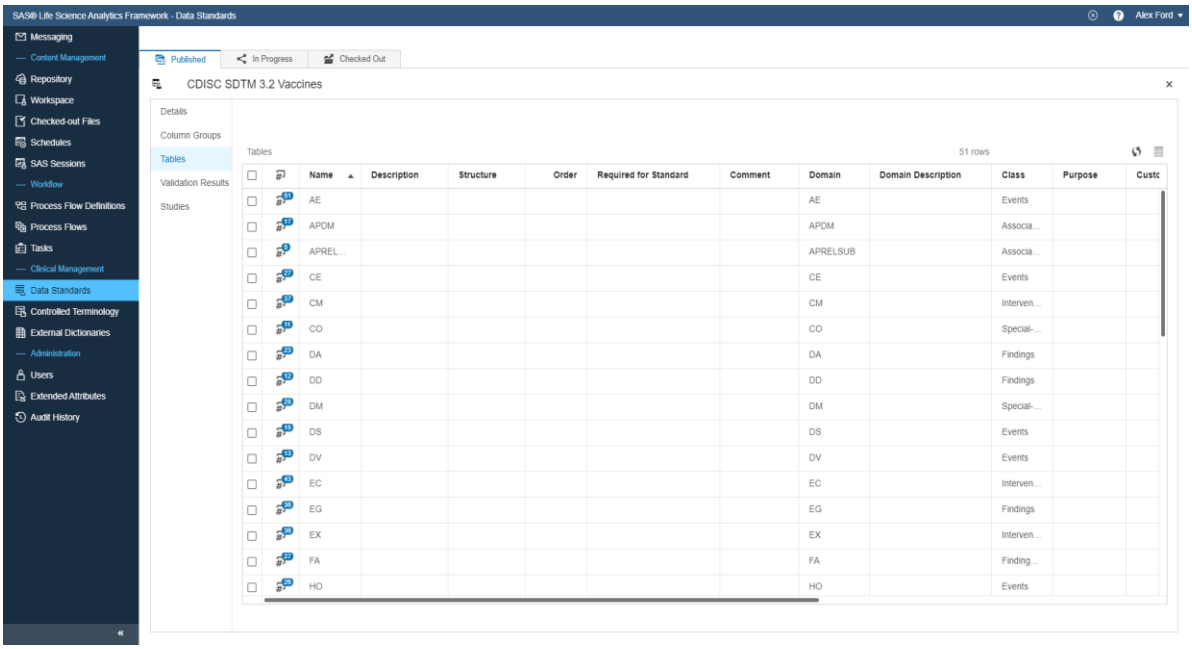

Figure 4. Tables

LSAF is able to run validation checks using the data standards that have been associated. The results of these checks are compliance reports which are exported to the study folder.

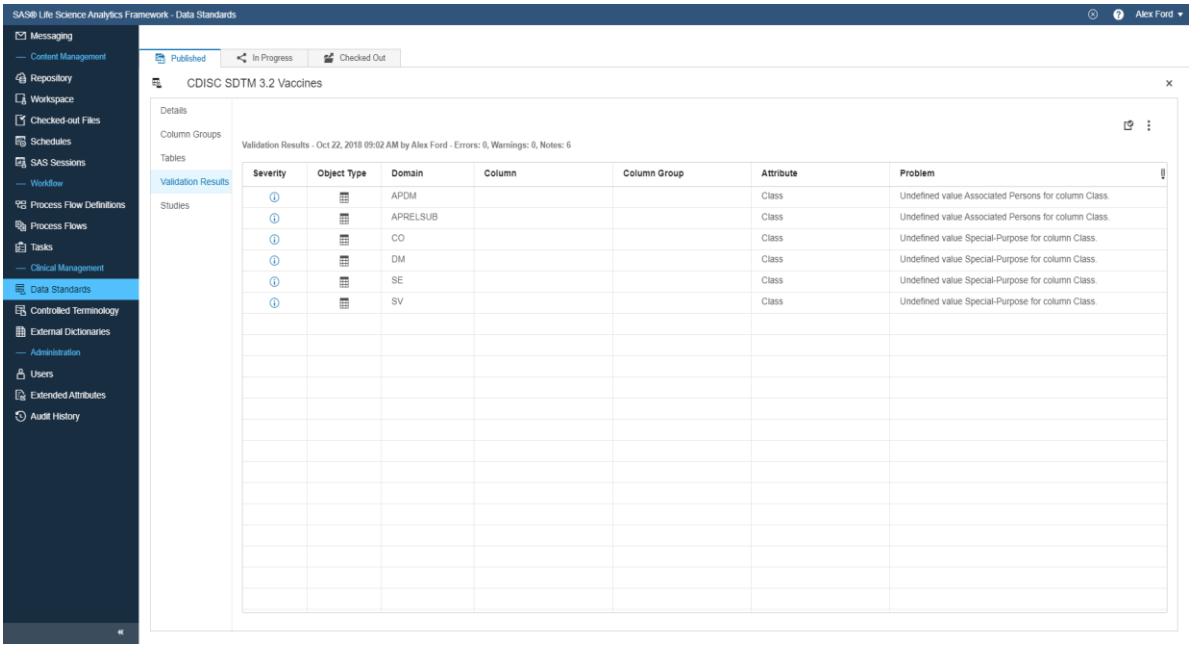

Figure 5. Validation Results

Under the studies tab there is additional information that is populated for the creation of the define package. All tables, the associated CT, and the relevant external dictionaries are defined at this level.

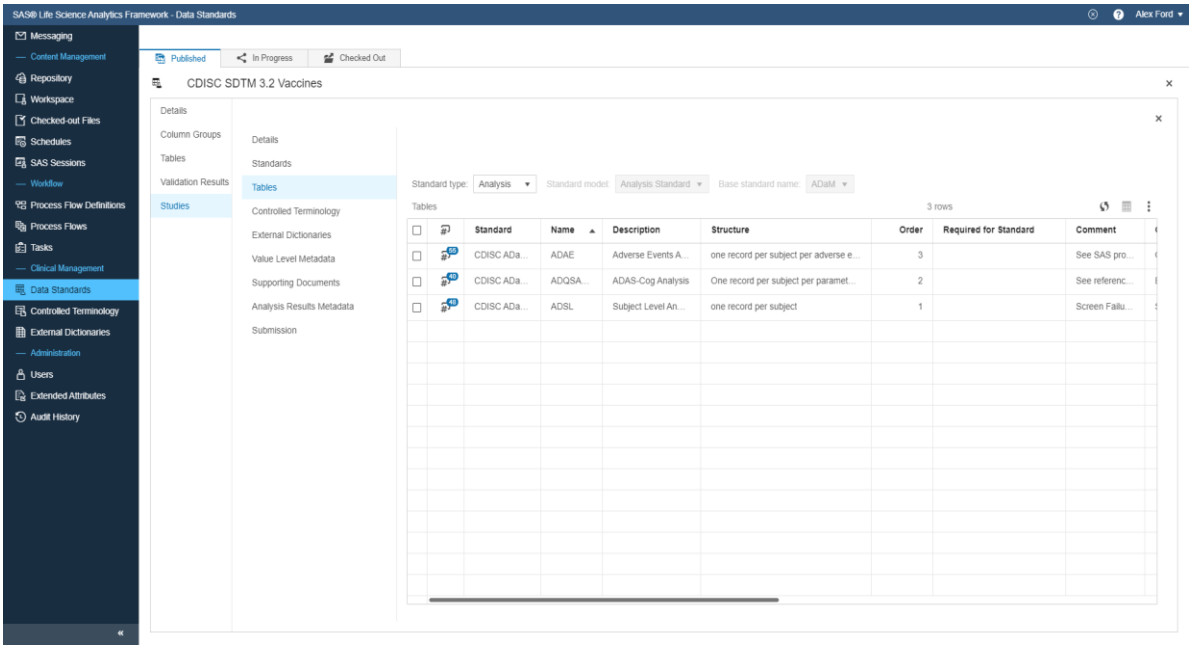

Figure 6. Study Definitions

The study specs, programs, and outputs are housed within LSAF, so the framework can derive Value Level Metadata and Analysis Results Metadata within the study level.

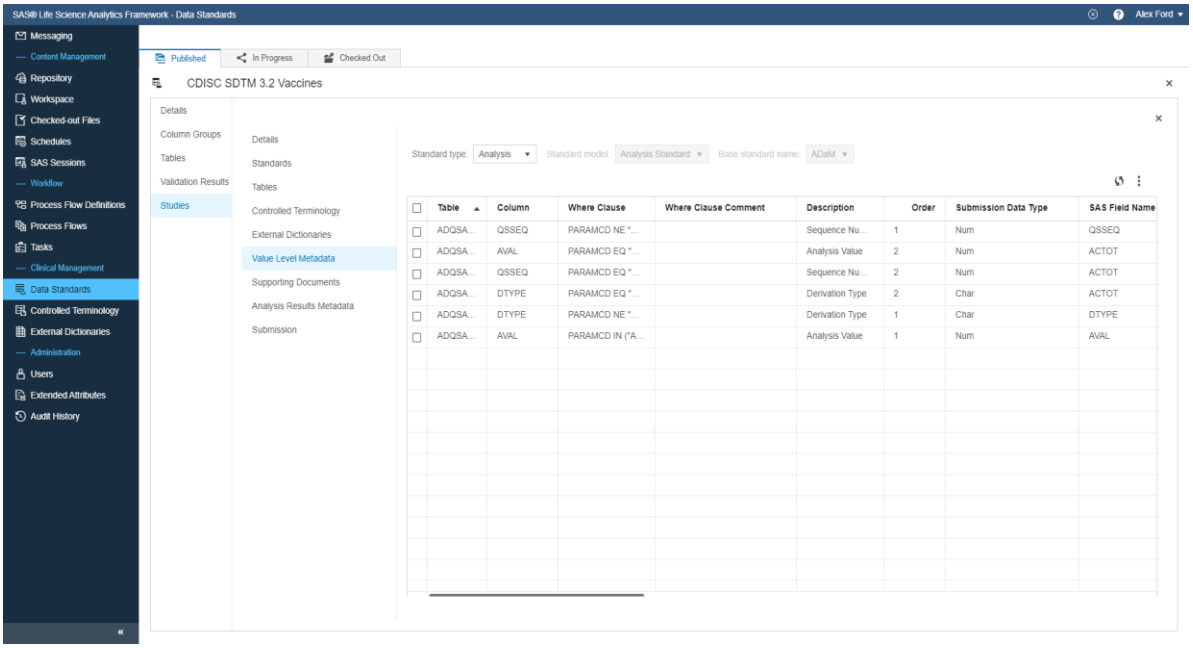

Figure 7. Value Level Metadata

Lastly, all relevant supporting documents are packaged at the study level. This could include the protocol, CRF, study specifications, SAP, computational algorithms, etc.

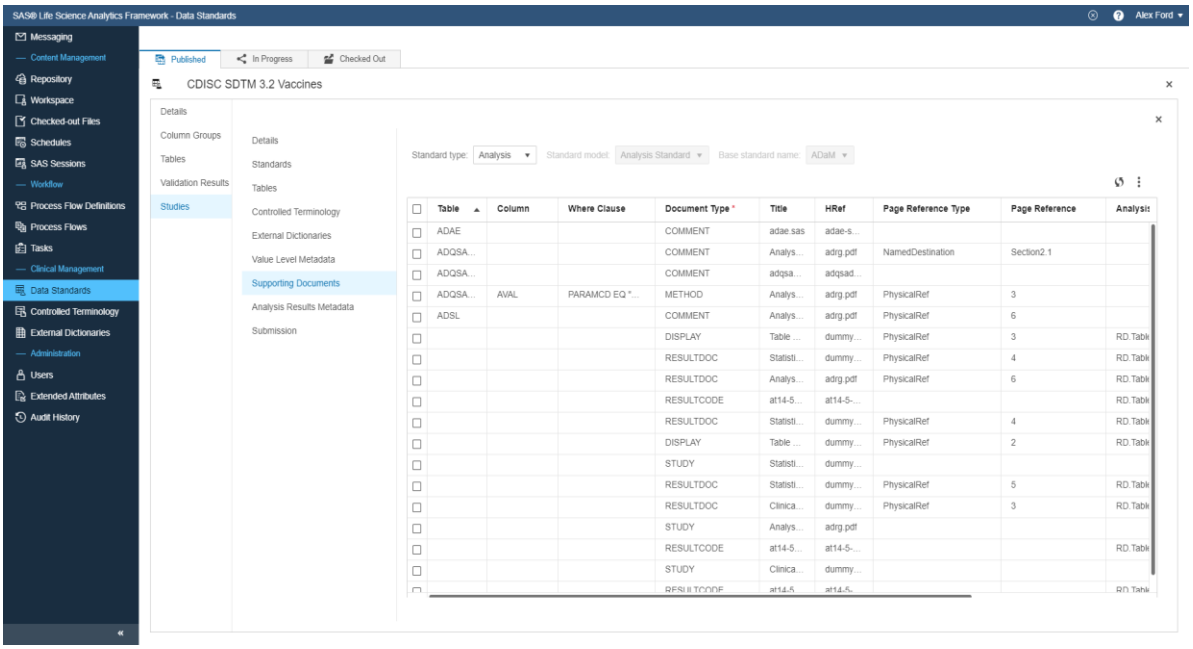

Figure 8. Supporting Documents

Once all study level assets have been defined and statistical development of tables, listings, and figures is complete, LSAF has a GUI interface that allows for creation of the define.xml. This GUI bundles together the data standard, VLM, ARM, CTs, External Dictionaries, and all manually input study level information and generates a Define.xml package.

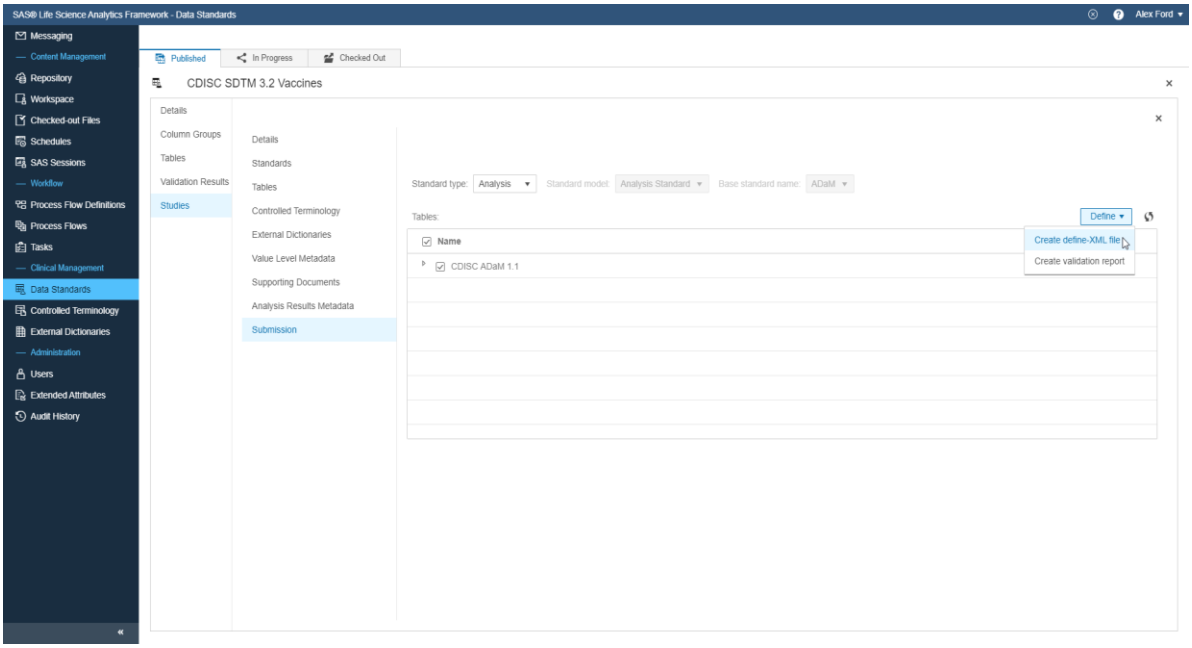

Figure 9. Define Creation

CONCLUSION

LSAF is intended to be an end to end clinical platform that assists in governing and streamlining the submission process. LSAF V5.1 has extensive metadata management capabilities which are designed to assist programming teams in meeting CDISC data standards prior to submission. The metadata management portion of the framework is capable of housing and versioning standards as they are established by a standards governor. Those standards are then associated with studies and used for compliance checks and define.xml generation. Traditionally, define creation has been a manual and disjointed process. LSAF empowers statistical teams to contain their entire development and QC process within a single unified system to save time and to enable users to complete CDISC compliant datasets correctly the first time.

REFERENCES

https://www.sas.com/en_us/software/life-science-analytics-framework.html

CONTACT INFORMATION

Any questions or comments are encouraged. Contact the authors at:

Alex Ford

SAS Institute

100 SAS Campus Drive

Cary, NC, 27513

Email: Alex.Ford@sas.com