



XML bulk upload facility

XML, XSD files and validation

Introduction

Many NJR users have databases containing information about some or all of their hip and knee joint replacement operations. For those surgical units performing a large number of procedures, this information may be considerable and internal links may exist to patient management, financial, and other systems. It is important that the NJR can make direct use of the relevant information in such internal databases automatically, rather than requiring the data to be manually re-entered via the NJR online web interface. This helps ensure data accuracy, timeliness of updates and prevents the need for double manual data entry.

The NJR is developing a data upload capability which will allow a large number of operation records to be uploaded to the NJR from users' own systems with minimal user intervention. Key to the operational success of this process is the use of agreed government standards. The NJR uses the e-GIF (electronic Government Interoperability Framework), which underpins all of government e-business information exchange. e-GIF mandates the use of XML (Extensible Mark up Language) as the standard for exchange of data between databases, because of its status as an international standard, its global use for e-business and the corresponding wide implementation of this standard within software products.

Example files

All example files related to this document are available from the NJR website.

XSD and sample XML instance files

The NJR Bulk Upload facility reads XML documents, which describe operations performed at the user's hospital. These documents are defined, described and validated by three XSD schemas. The central message schema defines the structure of the submission. This references two architectural schemas. These define the NJR code lists and some structured data types.

The NJR code list architectural schema is in turn described by an XML instance document. This associates a text description with each value using the technique. Here is an example:

'PatientGenderCodeList' describes patient genders. In the code list instance document, the description 'Male' is associated with the stored value of 1.

An example XML instance document can be found in bulkupload.zip on the NJR website.

The sample document represents the file that will need to be created and which can then be uploaded and processed via the Bulk Upload facility. This example contains 5 surgery records and each surgery record contains a number of procedures.

This document can be validated against the XSD schema known as 'message schema' (within bulkupload.zip)

Please note that the version number has been excluded from the filename. This has been done to avoid re-linking the file every time the file is amended. The schema itself holds the version number of this document.

The message schema refers to two architectural schemas. The first specifies the possible code list values.

The code list values are explained by the code list descriptions XML instance document.

The second architectural schema specifies the data types used to ensure sensible values are submitted for elements such as GMC code.

All of the schemas are subject to change although none are currently planned.

Validation of bulk upload XML files

Summary

Providing the submission instance document can be successfully validated against the message schema, it will then be subjected to a second tier of programmatic validation. This will perform some of the tests that are outside of the scope of the XSD for example is the ISO operation date valid and in the correct range? Providing the submission passes these tests, the contents will be saved into the NJR edit stack. The contents can then be checked, consolidated and submitted using the standard data entry system. Successfully uploaded records will appear within 24 hours.

If the submission instance document fails the XSD validation, the whole submission will be instantly rejected. The reason for the failure will be relayed.

Each submission can contain up to 100 operations. Each operation can have up to 4 procedures. The majority of the elements are optional. This means that if the user's system only collects certain parts of the data set, they can still take advantage of the bulk upload system. However, the more elements submitted via bulk upload, the less that will need to be manually entered into the data entry system.

Surgeons should be identified by providing their GMC code. If the identified surgeon is not recognised by the NJR or if the surgeon is not linked with the user's hospital, the surgery record will be rejected by the second tier of validation.

Components should be described using lot number and reference number. If any of these are not recognised by the NJR, the second tier of validation will quarantine them. The search tools in the data entry system can then be used to identify each quarantined component.

Validation process

A number of layers of validation will be applied to the data presented to the bulk upload. These are as follows:

Version validation:

- This validation is performed before the data is looked at.
- This checks that the version number supplied with the data matches that included with the latest code list download. This is very important as if both the NJR and the caller are not referring to the same code lists; there is an opportunity for confusion.
- If the version number does not match, the upload will be rejected with a defined error message.

XML upload size validation:

- This validation is performed before the data is looked at.
- If the size of the upload exceeds the defined maximum, the upload will be rejected with a defined error message.

Structural validation:

- This validation is performed before the data is looked at.
- This checks that the presented XML data conforms to the structure and relationships defined in the XSD schema.
- This checks that each value in the data that refers to a code list has an associated entry in the code list table. All code list values will be mandatory. Please note 'Not selected' will be a valid code list option.
- The XSD schema will reference these code lists dynamically.
- If the data does not conform, the upload will be rejected with a defined error message.
- Please note that structural errors will be reported back on a first error only basis. For example if there are two mandatory code list lookups out of range in an entity, only the first one will be reported back as the error.

Unique operation check:

- Each surgery record presented by the bulk upload must have a hospital unique code associated with it. This provides a means of referring to the record before it has an NJR Index number assigned. This number can be checked to prevent accidental repeat submissions. i.e. if the same number is seen twice, we reject the upload with an explanatory message.

Referential and integrity validation:

- This validation is performed as the first stage of examining the data.
- The formatting of all dates is checked to ensure that they are within a sensible range and use the unambiguous ISO formatting standard of <CCCCMMDD>.
- If any data item does not pass the test applied to it, the record will be rejected and details of each test fail will be returned.

MDS2 Dataset validation:

- This validation is performed as the third stage of examining the data.
- This checks that the MDS2 defined validation is conformed to. These tests include:
 - Ensuring mandatory fields are populated.
 - Ensuring fields that are conditionally mandatory are populated.
 - Ensuring fields that have complex inter-relationships with other fields have their constraints met. For example, a value for 'Complex Osteotomy' can only be set if 'Trochanteric Osteotomy' is 'Yes'.
 - Ensuring that no components are quarantined.
 - Ensuring that no component requests are outstanding (may not be relevant).
 - Ensuring that there are no component validation errors.

If any data item does not pass the test applied to it the details of each test fail will be returned. The record will be placed in the edit stack.

Please ensure that the instructions regarding the installation of the security certificate are followed to avoid any problems. The instructions can be found on the NJR Centre website.