This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 875299

Project acronym: UNICOM
Project full title: Up-scaling the global univocal identification of medicines in the context of Digital Single Market strategy
Call identifier: H2020-SC1-DTH-2019

WP4 IDMP implementation at National Drug Agencies

D4.7: SPAIN: Annual progress report on refactoring the national IT system, migration of national data and data interfaces to EMA´s SPOR

Version:	1.0
Status:	Final
Dissemination Level[1]:	PU
Due date of deliverable:	31.01.2021
Actual submission date:	31.01.2021
Work Package:	WP4: IDMP implementation at National Drug Agencies
Lead partner for this deliverable:	AEMPS
Partner(s) contributing:	
Deliverable type[2]:	R

Main author(s):

José Manuel Simarro	AEMPS
Ana López de la Rica	AEMPS

Resource consumption estimate:	Person months
	0,15

[1] Dissemination level: PU: Public; CO: Confidential, only for members of the consortium (including the Commission Services); EU-RES: Classified Information: RESTREINT UE (Commission Decision 2005/444/EC); EU-CON: Classified Information: CONFIDENTIEL UE (Commission Decision 2005/444/EC); EU-SEC Classified Information: SECRET UE (Commission Decision 2005/444/EC)

[2] Type of the deliverable: R: Document, report; DEM: Demonstrator, pilot, prototype; DEC: Websites, patent fillings, videos, etc.; OTHER; ETHICS: Ethics requirement; ORDP: Open Research Data Pilot

Revision history

Version	Date	Changes made	Author(s)
0.1	29.12.2020	First draft	Ana López de la Rica (AEMPS)
0.2	29.12.2020	Peer Review	José Manuel Simarro (AEMPS)
0.3	30.12.2020	Final draft for revision	Ana López de la Rica (AEMPS)
0.4	24.01.2021	Final with comments from NO sent to Empirica	Ana López de la Rica (AEMPS)
1.0	29.01.2021	Review and submission	EMP

Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Deliverable abstract

ISO IDMP implementation at Agencia Española del Medicamentos y Productos Sanitarios (AEMPS) by refactoring its RAEFAR database to ISO IDMP standard.

At this point, we have started mapping and cleansing of the medicinal product data in the current RAEFAR database and development of a new RAEFAR IT system.

Mapping and cleansing activities have been mainly focused on substances, organizations and referential lists.

We are also participating in a UAT run by EMA for testing the SPOR API V2 where two end points are being tested: creating a new medicinal product and consulting existing ones.

Regarding implementation of Clinical particulars, AEMPS was required from the NHS to include in its logical data model indications and adverse reactions and it is shared in this report.

Keywords: AEMPS, cleansing, mapping, logical data model clinical particulars

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of abbreviations

Abbreviation	Complete form
API	Application Program Interface
ATC	Anatomical Therapeutic Chemical
DADI	Digital Application Dataset Integration
eAF	Electronic Application Form
EMA	European Medicines Agency
EUTCT	European Union Telematics Controlled Terms
FHIR	Fast Healthcare Interoperability Resource
ID	Identification
IG	Implementation guideline
LOC ID	Location identity
MAAs	Marketing Authorisation Applications
MAH	Marketing Authorisation Holder
MedDRA	Medical Dictionary for Regulatory Activities
MPID	Medicinal Product Identifier
NCA	National Competent Authority
NHS	National Health System
OMS	Organization Management System
ORM	Object Relational Mapping
ORG ID	Organization Identity
PCID	Packaged Medicinal Product Identifier
PMS	Product Management Service
PLN	Natural Language Processing
PoC	Proof of Concept
QA	Quality Assurance
RMS	Referential Management System
SmPC	Summary of Product Characteristics
SMS	Substance Management System
UAT	User Acceptance Testing
WHOCC	World Health Organization Collaboration Centre

1 Executive summary

The objective of this report is to present the tasks performed at AEMPS in its Medicinal Products Information System during 2020 focused on achievement of compatibility with ISO IDMP standards, interoperability with SPOR system from EMA and progressive introduction of implemented changes in the National Health Service.

AEMPS as a trusted information source on medicinal products for the NHS is a major actor to unfold the implementation of ISO IDMP at national level.

This document explains how the gap analysis has been developed, how the data is being cleansed and SPOR synchronization and finally how the logical data model for clinical particulars is being implemented.

At the end some dependencies and relationships are briefly discussed as well as some risks and mitigations.

2 Transition strategy

AEMPS transition strategy to make its information system ISO IDMP compatible follows these principles:

- A genuine ISO IDMP implementation has not been considered as the estimated effort of rebuilding current systems is prohibitive from an effort/cost versus benefit point of view. For that reason, the objective has been adapting national systems to be ISO IDMP compatible making them SPOR interoperable.
- National control terms lists need to be matched with SPOR control terms lists. In the long run, a replacement might be considered but not in this current project.
- The major driver of the project is the capacity of consuming FHIR dataset from eAFs.

The introduction of ISO IDMP in our systems will be done performing the following tasks:

1. Gap analysis
2. Data cleansing and SPOR synchronization
3. Implementation of the identified changes
4. Automatic upload of eAF's datasets
5. Implementation of the logical data model for clinical particulars

During 2020 AEMPS has been focused in tasks 1,2 and 5 and in the following sections, all three will be extensively developed in order to explain milestones achieved.

2.1 Gap Analysis

AEMPS is developing a process of adapting the national database responsible for medicinal product management to make it ISO IDMP compliant considering the Chapter 2 of EU IG v.2 draft document and FHIR 4. This analysis has allowed us to identify which points in our database need to be modified and which ISO IDMP concepts should be included, if they are relevant and of interest to the National Health System. The following points have been detected which need to be modified:

- Manufactured Item

A medicinal product may contain, in the packaging, one or more manufactured items and one or more pharmaceutical products.

The full information on Manufactured Item as presented in the FHIR *ManufacturedItemDefinition* is:

- Identifier

- Manufactured Dose Form

- Unit of presentation

- Manufacturer

- Ingredient

All this information must be included in RAEFAR national database.

At the moment, only the information corresponding to Administrable Dose Form has been added, as it is showed in the following figures.

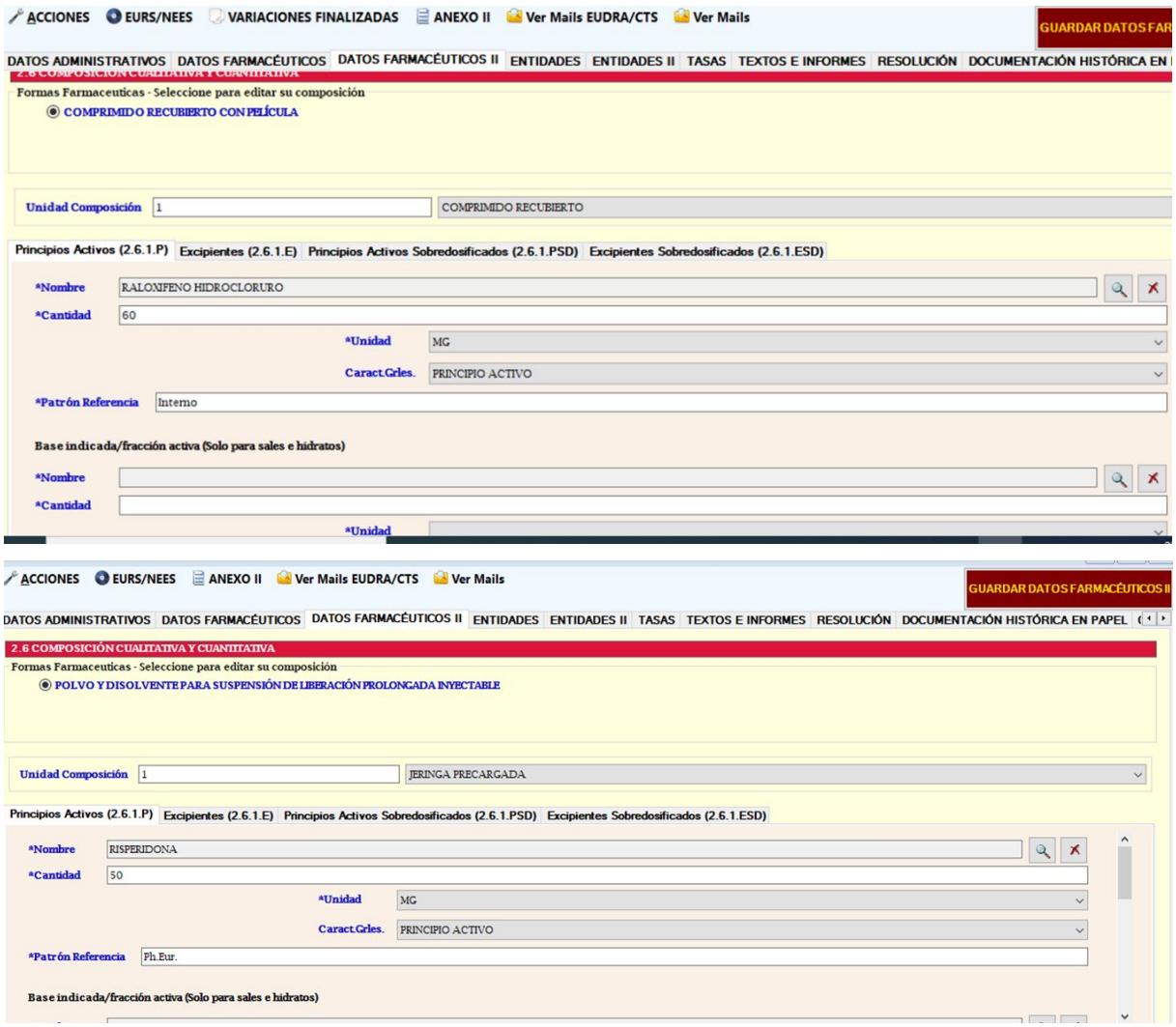

Figure 1: Administrable Dose Form – Active substances

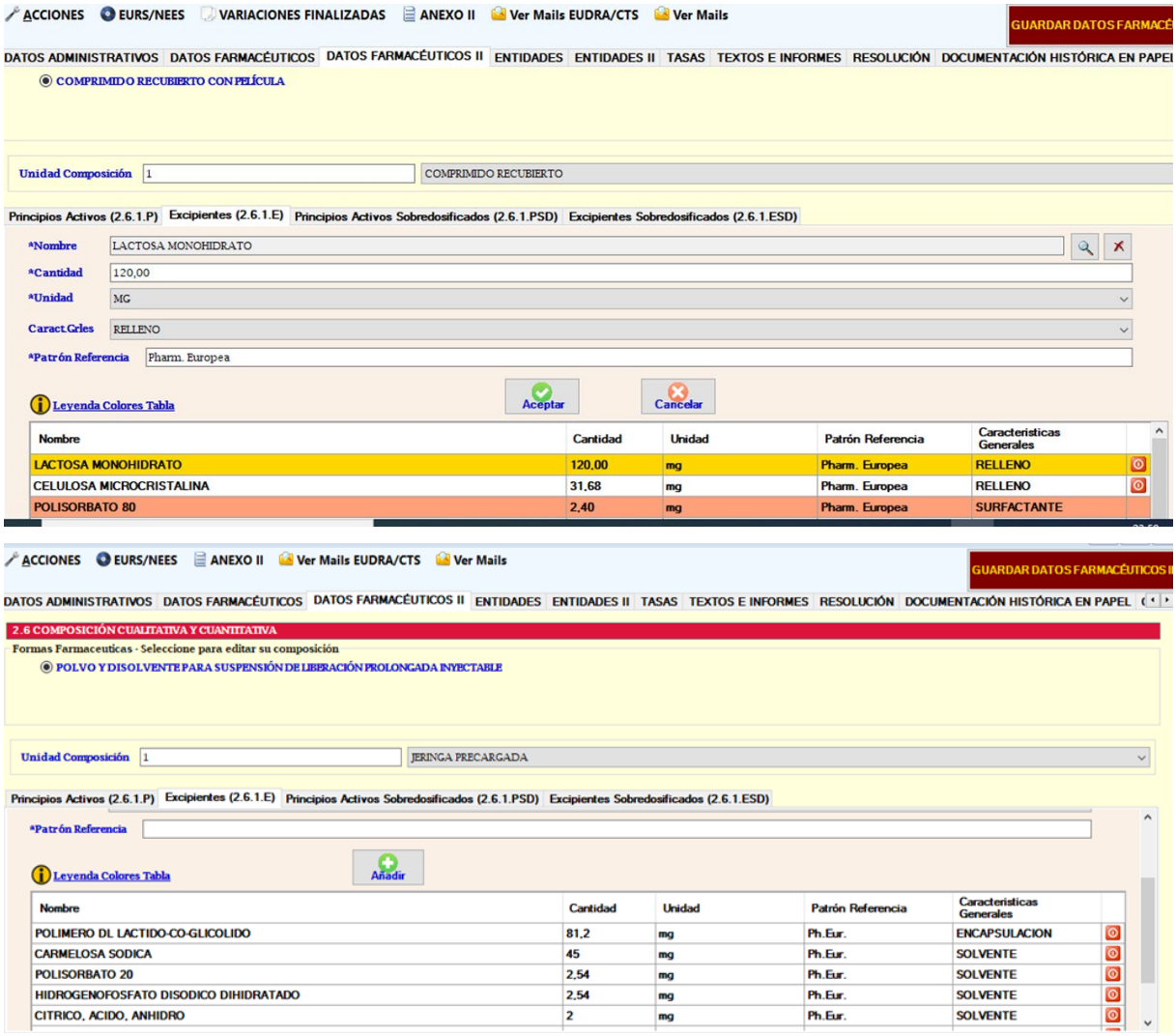

Figure 2: Administrable Dose Form information - Excipients

The aim, at this point, is to break down the information detailed in Chapter 2 of EU IG v.2 draft document for both Administrable Dose Form and Manufactured Item, to get the separate information from what each manufactured item contains.

- Presentation/Concentration Strength

The strength of the substance should be declared as a quantity of the substance contained in a Manufactured Item or Pharmaceutical Product. The strength must be provided based on a numerator and denominator value and unit, i. e. unit of presentation or unit of measure/concentration.

Currently, the strength of the pharmaceutical product or the one reflected in section 3.2.P.1 of Module 3 of the registration dossier is expressed. Consequently, the appropriate changes will be made to the system so that it can be indicated in both ways when possible, in order to be compatible with ISO IDMP and provide as much information as possible.

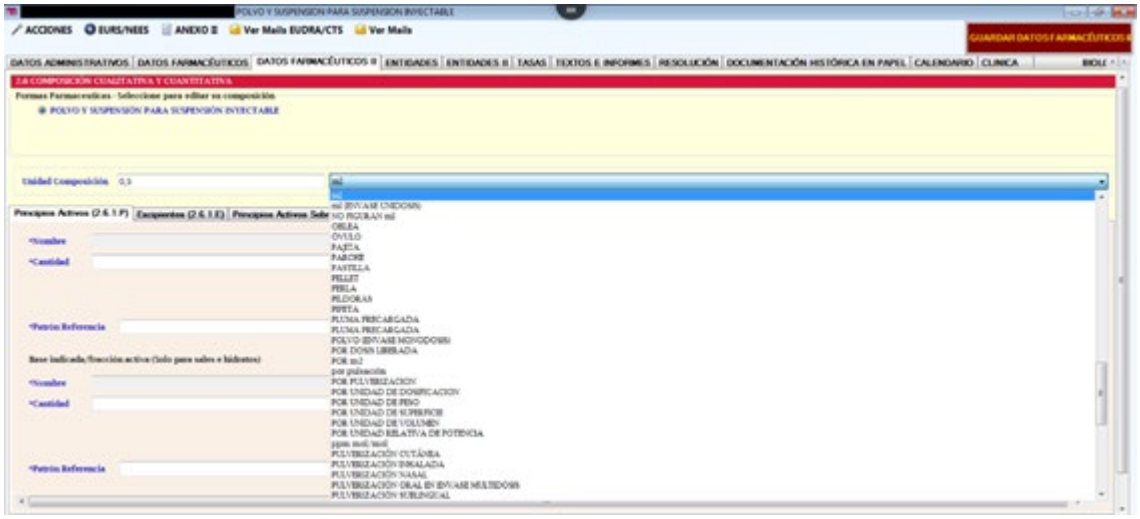

Figure 3: Strength (quantitative composition) – unit of presentation and unit of measure/concentration

- Packaging

The intermediate packaging is one of the concepts we are still working on; in fact, only one level of intermediate packaging is considered and the recursive relationship existing in ISO IDMP has to be implemented in a compatible way.

- Reference strength

Reference strength represents the strength (quantitative composition) of the active moiety of the active substance.

This is a term that it has been already included in our system, but there is a lack of developing in a more detailed way as explained in Chapter 2 IG document in order to appropriately address the Reference strength (Presentation) and Reference strength (Concentration).

The action to be done consists of reviewing the national modeling of the Reference Strength Concept to make it ISO IDMP compatible, as the concept is considered at national level but with less granularity in its recording.

Figure 4: Amlodipine/Amlodipine besilate- Reference strength

- Other changes

Composition expression, considering quantities in numerator and denominator, unit, and operator fields must be modified to be as similar as possible to ISO IDMP structure.

2.2 Data Cleansing and SPOR Synchronisation

AEMPS has been automatically processing electronic application forms (eAFs) submitted by applicants since April 2016. Thus, an integration of eAF's EUTCT codes has been rolling out since then in our national database.

We have been annotating our master data with the EUTCT code as they are included and available in the eAF, in order to be able to translate the EUTCT codes presented in the eAF to our internal IDs.

Current situation is shown in the following table:

AEMPS table name	SPOR system	SPOR list name	Entries total	Coded entries	Coded entries %	Duplicate entries %
MT08_CODS_ATC	RMS	ATC classification system - Human	6.639	6.639	100,00%	0,00%
MT13_PAISES	RMS	Countries	265	248	93,58%	4,20%
MT18_UNID_CANTI	RMS	Units of Measurement	304	164	53,95%	7,28%
MT202_TEJIDOS	RMS	Tissue	37	13	35,14%	0,00%
MT204_CODS_ATC	RMS	ATC classification system - Veterinary	7.220	7.220	100,00%	0,00%
MT24_TIPO_DISPOS	RMS	Packaging	111	92	82,88%	2,22%
MT25_TIPO_CONSER	RMS	Special Precaution for Storage	268	33	12,31%	17,86%
MT28_VARIACIONES_2010	RMS	Variation Classification (Human)	773	717	92,76%	1,42%
MT28_VARIACIONES_2010_VET	RMS	Variation Classification (Veterinary)	776	736	94,85%	0,96%
MT52_PR_ACTIVO	SMS	Active substances	18.432	12.032	65,28%	0,30%

MT54_FOR_FARM	RMS	Pharmaceutical Dose Form Combined Pharmaceutical Dose Form	680	516	75,88%	0,19%
MT55_VIA_ADMIN	RMS	Routes and Methods of Administration	101	94	93,07%	0,00%
VW_ESPECIES_DESTINO_EUTCT	RMS	Target species	354	334	94,35%	0,00%
VW_ESPECIES_DESTINO_LMR_EUTCT	RMS	Maximum Residue Limit Species	31	31	100,00%	0,00%
VW_EXCIPIENTES_EUTCT	SMS	Excipients	7.888	2.932	37,17%	5,44%

Table 1: RMS and SMS mapping synchronization

Some comments about the above table are:

- For the ATC codes, the WHOCC coding system is used instead of EUTCT ("Source Id" column of the RMS list)
- Units of measurement and unit of presentation are stored in the same AEMPS' table (MT18_UNID_CANTI)
- The variation classification is coded using RMS list's "short name" column, since EUTCT codes were not available in eAF, prior to version 1.24.0.0
- We store in only one table both simple and combined dose forms (MT54_FOR_FARM)

Currently, the synchronization process is one-way, reactive (not proactive) and is performed in a semi-automatic way.

The involved steps are:

- Incoming eAF (human or veterinary MAA or a variation) is automatically processed. When either a RMS term EUTCT, an organization (loc-id, org-id) ID or a substance EUTCT code cannot be translated, an alert is sent (via email) to the Standardization Area, who is in charge of, among other things, to administer AEMPS' master data. The eAF is then put on standby, provided that it cannot be loaded in AEMPS national database because it wouldn't have referential integrity with master data.
- Upon reception of an alert, operators look up the item (RMS term, OMS term or EUTCT term) in the list of not yet coded elements of the same type, using the rest of data describing it and included in the eAF. One of two possibilities occurs:
    - The element exists in the uncoded list. In this case, its translation info, either an EUTCT code or a (org-id, loc-id) pair is updated.

    - The element does not exist. In this case, a new record is added in the corresponding table and is populated with the information carried out in the eAF. Translation to Spanish, where appropriate, is performed.
- After updating master data, the eAF that triggered the update process is reprocessed. This time, the translation will succeed and it will be successfully loaded in the national database.
- We are considering the synchronization every 6 months in order to mitigate possible errors performing a revision.
- Regarding "Substances", we do not foresee important changes when migration takes place from EUTCT to SMS as per our current level of mapping to EUTCT and because SMS will remain including a reference to the EUTCT ID.

With "Organizations" a similar approach has been followed, mapping all the ORG-ID and LOC-ID received through the eAFs, nonetheless as "O" is not mandatory in the eAF the mapping process goes under willingness of the applicants. Making a data analysis of the Organizations information at national level.

AEMPS table name	SPOR subsystem	SPOR list name	Entries total	Coded entries	Coded entries %	Duplicate entries %
MT56_LABORAS	OMS	MAHs	13.234	4.363	30,13%	7,29%
MT56_LABORAS	OMS	Batch Releasers	3265	248	8,58%	4,20%
MT56_LABORAS	OMS	Active Substance Manufacturers	3304	364	10,35%	4,28%

Table 2: OMS mapping synchronization

During 2021 a communication action will be rolled out to request organizations involved in AEMPS regulatory activities to register at SPOR Portal and provide IDs assigned to AEMPS.

A solution to keep our national dictionary of organizations updated following to the changes recorded in "OMS" is under study to see if it is needed as this synchronization can be achieved via the eAFs data. A procedure to register MAHs, Batch Releasers and active ingredient Manufacturers that exist in our national database exist also in "O" but inconsistency data problem is also under discussion waiting for a coordinated action at European Network level.

2.3 Testing SPOR API V2

We are also participating in a UAT run by EMA for testing the SPOR API V2. Two end points are being tested: creating a new medicinal product and consulting existing ones. A high-level design of a client for this API has been started as well as a search for available FHIR code and libraries that potentially could be reused in our national project.

The high-level view of our client architecture is shown below, in the “create product” use case scenario.

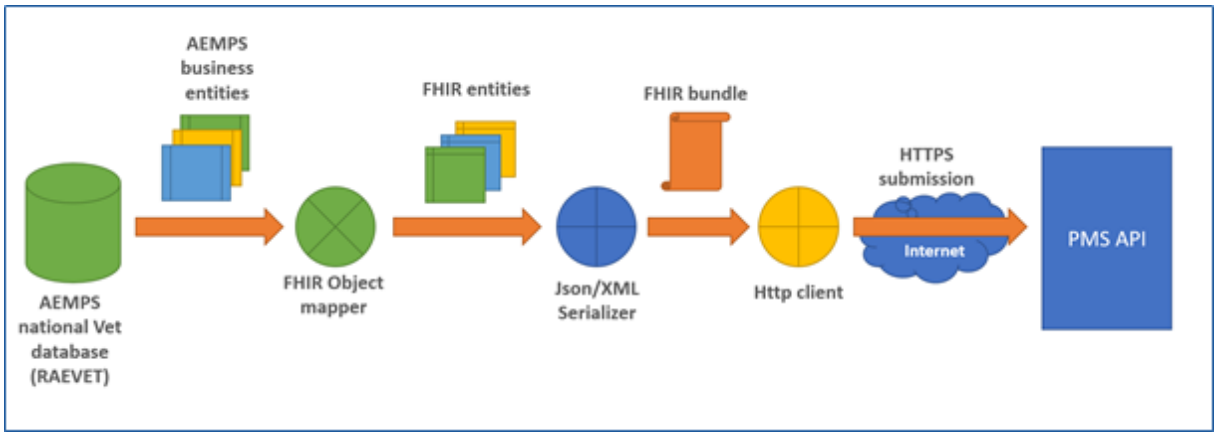

Figure 5: Client architecture for testing the API SPOR V2

Design guidelines we are implementing, or have planned to implement:

- Technology used will be Microsoft .NET framework, widely used at AEMPS.
- We use an ORM solution to extract data from AEMPS national Medicinal Product Database.
- FHIR entities are handled using a third-party open-source library (fhir-net-api).
- Mapping from AEMPS business objects to FHIR entities is performed using a mapping library (AutoMapper – open source).
- Once FHIR objects have been generated, a serialization operation is performed, to obtain a textual representation. Json or XML representation can be used indifferently since PMS API supports both (for being FHIR compatible).
- Serialization is performed using services implemented in the FHIR library mentioned above.
- Once the text document is generated, it is submitted to PMS API using the standard .NET System.Net.HttpClient API.

Current status of the ongoing tasks:

- We requested to use the SPOR PMS API in the testing environment. We have been granted access with unlimited quota usage of the API.
- We have been testing the API since June 2020. We usually meet EMA people for informal reporting of test results/found issues every week.
- We have been able to successfully interact with EMA PMS API in a manual way, using Postman as an API querying tool.
- **[Http submission]** We developed a small PoC application to make automated submission of data to the API, in form of a POST call to the API endpoint used to create a medicinal product with all its parts (FHIR resource of type bundle/transaction). Submitted content is based on examples given by EMA.
- **[FHIR serialization]** We coded a PoC application that, feeding FHIR entities instances of objects implemented in the fhir-net-api library with mock data, generates a textual representation of a FHIR resource (of type bundle/transaction) ready to be submitted to the API.

Our planned activities for the future are:

- An improvement to the FHIR serialization is needed, since we have detected some minor differences between the FHIR document we generate, and the ones provided by EMA. An exact match is needed, in order to successfully submit the data to the API.
- Finishing the definition of the mapping between data in our national medicinal products database and FHIR resources properties.
- Coding the FHIR object mapper, in line with the mapping defined in the above point.
- Querying Medicinal Product database to obtain the data needed to build the FHIR messages to be submitted.
- Integrating all the components in an end-to-end solution, candidate to go live after our internal QA.
- In parallel with all the above, going on with the UAT testing in collaboration with EMA.

3 Dependencies and relationships

A strong dependency exists with WP3 as the convergency between our national information systems and EMA SPOR system relies mainly on the interoperability with electronic application forms. The new FHIR datasets to be implemented at EMA DADI project for MAAs and Variations, will have to match with the ISO compatible design that is now undergoing.

WP1 for training and gaining knowledge on ISO IDMP and FHIR is also a trusted source of information to understand concepts like:

- Manufactured Item.
- Different pharmaceutical dose form usages.
- Presentation and Concentration strength concepts.
- Packaging modeling to make it ISO IDMP compatible.
- MPID and PCID definition.

3.1 Risks and Mitigations

Two major risks have been identified:

- The arise of changes in the IDMP Logical data model once an adaptation design is set and in place. These changes can be a consequence of the conclusions of WP1.

- Delays in the scheduled timelines of WP3 and non-availability of its deliverables. The forms will be developed mainly under the scope of the DADI project carried out by EMA.

A mitigation measure to be adopted is to create awareness of these risks among UNICOM partners, closely follow both WPs and warn on the impacts that decisions in these WPs may have over national projects.

4 Implementation of Clinical Particulars

AEMPS was required by the NHS to include in its logical data model indications and adverse reactions. The indications are relevant for some prescriptions due to reimbursement purposes and the adverse reactions are relevant to identify rare diseases and avoiding confusion in their detection.

The logical data model implemented for this purpose is shown in the below figure and it follows the IDMP logical model.

An activity of processing SmPCs using PLN techniques is scheduled for 2021 to populate this logical data model. Target of this activity is the identification and extraction of MedDRA terms from SmPC section 4.1 “Therapeutic indications” and section 4.8 “Undesirable effects”.

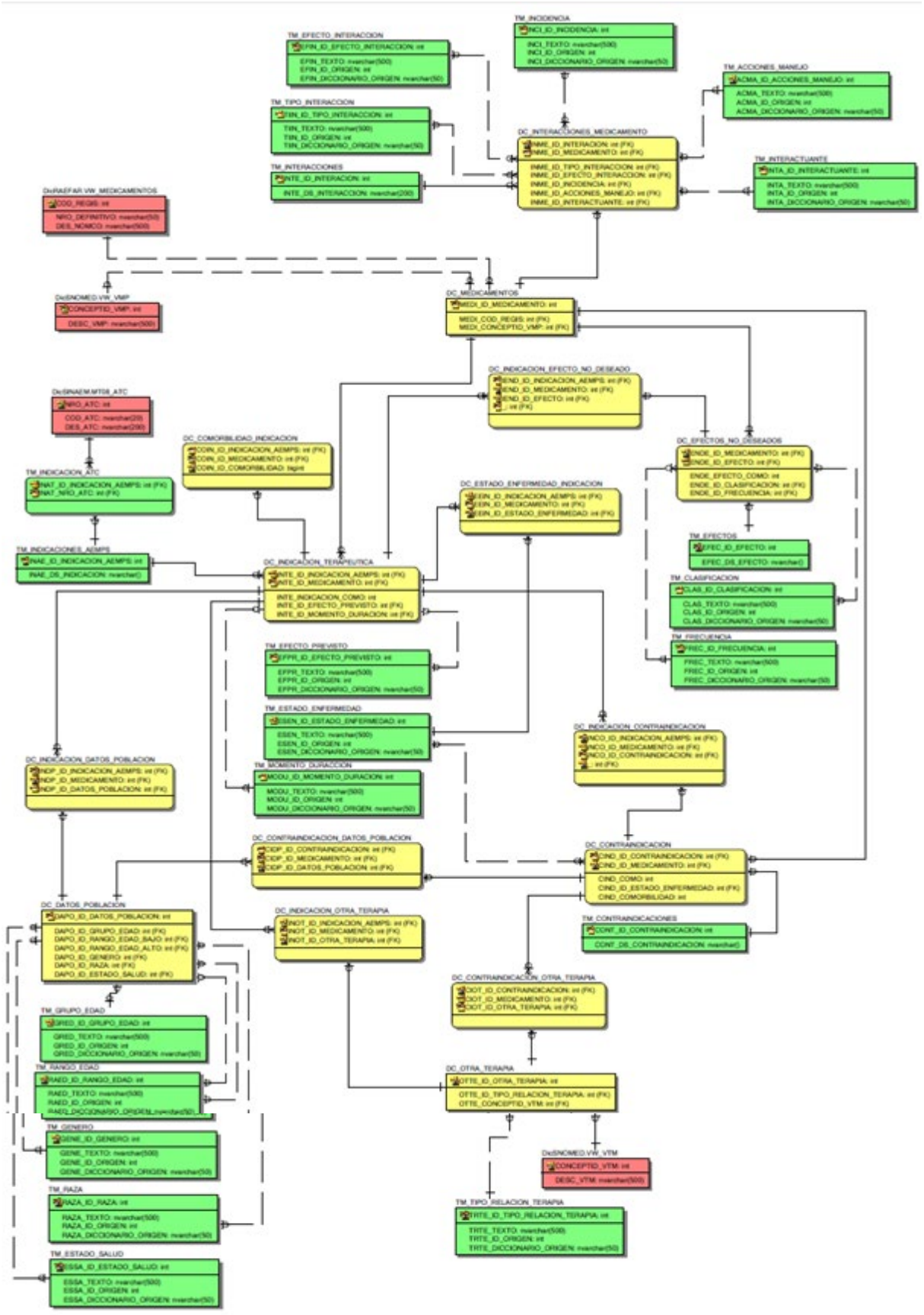

Figure 6: Clinical Particulars_logical data model