Deliverable

D3.5 Treatabolome database	
Version | Status	V1 | final
Work package	WP3
Lead beneficiary	INSERM-CRM (Gisèle Bonne)
Due date	31.12.2020 (M36)
Date of preparation	15.12.2020
Target Dissemination Level	Public
Author(s)	INSERM-CRM: Gisèle Bonne, Antonio Marques Atalaia CNAG-CRG: Carles Hernandez-Ferrer, Leslie Matalonga, Sergi Beltran
Reviewed by	Rachel Thompson (CHEO-RI), Hanns Lochmüller (CHEO-RI)
Approved by	Holm Graessner (EKUT)

The Solve-RD project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 779257.

Explanation according to GA Annex I:

The Treatabolome: flagging treatable genes and variants. The database will be connected to the RD-Connect genomic analysis platform and made accessible as part of the real-time analysis of the patients undergoing sequencing or exome analysis within Solve-RD as a proof of concept for the utility of the approach.

Partners involved: INSERM-CRM, UNEW, CHEO-RI, CNAG-CRG, all partner ERNs

Abstract:

To alert the treating geneticist or clinician about patients with treatable conditions at the time of reviewing NGS results, **INSERM-CRM** and **UNEW** initiated the development of a computer-readable and interoperable knowledgebase (termed the Treatabolome) that links treatable variants with the evidence for the treatment. The scheme adopted is to 1) produce systematic reviews – gathering published scientific evidence for each treatment of specific rare disease – led by experts of the participating ERNs and 2) transfer these systematic reviews into a an interoperable database newly developed by **CNAG** to enable the "flagging" of a potentially treatable variant. A pilot study was done by **UNEW** and **INSERM-CRM**, using the congenital myasthenic syndromes (Thompson et al 2018, PMID 30931400). In order to link treatments with disease identifiers, with **INSERM-Orphanet** they expanded the Orphanet nomenclature for CMS (Thompson et al 2018, PMID 30477555). A guideline/protocol paper to conduct these systematic reviews has been published (Atalaia et al 2020, PMID: 32787960). Seven other systematic reviews are ongoing by experts from ERN EURO-NMD and ERN-RND. A special issue of the *Journal of Neuromuscular Diseases* is planned by first quarter of 2021 to disseminate these systematic reviews and the concept of Treatabolome. The Treatabolome concept has been widely disseminates as well at different conferences, and in particular with a dedicated workshop held during the virtual ESHG conference in 2020. A '*Treatabolome task force*' gathering **INSERM-CRM, CHEO-RI, CNAG and JAX** has been created with the aim to optimize the work necessary for systematic reviews and their FAIR implementation in IT tools. This has allowed the CNAG team to develop a dedicated IT tool to make accessible any existing treatments at the level of the gene and/or variant levels. The Treatabolome DB is being connected to the RD-Connect Genome-Phenome Analysis Platform (GPAP) to be used as part of the real-time analysis of the patients undergoing sequencing or exome analysis within Solve-RD, providing thus a proof of concept for the utility of the approach.

Introduction:

Despite recent scientific advances, most rare genetic diseases do not currently have curative gene-based therapies available. However, in some cases, a confirmed genetic diagnosis immediately provides guidance on treatment, with drugs available that may significantly alter the disease course, improve functional ability and extend life expectancy. Nevertheless, many treatable patients remain undiagnosed or do not receive treatment even after genetic diagnosis. Next-generation sequencing is increasingly used as a first-line diagnostic tool for rare disease, and this includes patients with conditions that may be treatable. To alert the treating geneticist or clinician about patients with treatable conditions at the time of reviewing NGS results, we aimed to progressively develop a computer-readable and interoperable knowledgebase (termed the Treatabolome) in a way that links treatable variants with the evidence for the treatment.

Report:

In order to ensure the resulting database is reliable, this procedure requires the gathering of published scientific evidence for each treatment as the starting point. It is thus quite labor-intensive and requires input from disease experts. The task therefore foresees the generation of systematic reviews for specific disease areas led by experts from each of the participating ERNs in Solve-RD. The data resulting from the systematic reviews is captured in a form that allows it to be transformed by bioinformatics experts into a database that can be queried by analysis systems such as RD-Connect in order to enable the “flagging” of a potentially treatable variant.

In 2018, R Thompson (UNEW), H Lochmüller and G Bonne (INSERM-CRM) piloted this procedure on behalf of the neuromuscular ERN, EURO-NMD, using the congenital myasthenic syndromes (CMS) as an example. CMS are a group of clinically and genetically heterogeneous but frequently treatable neuromuscular conditions where the appropriate treatment varies depending on the underlying cause. They performed a systematic review of the evidence for pharmacological treatment of each CMS type, gathering evidence from 207 studies of over 1000 patients and stratifying by genetic defect. They assessed the strength and quality of the evidence and generated a dataset that provides the foundation for the planned computer-aided system. This was published as Thompson et al 2018 (PMID 30931400). In order to link treatments with disease identifiers, they also worked with INSERM-Orphanet to expand the Orphanet nomenclature for CMS (Thompson et al 2018, PMID 30477555).

As a next step, G Bonne and A Atalaia (INSERM-CRM) have expended the pilot concept for reuse by other disease areas. Within Solve-RD partners, interest in performing systematic reviews has been expressed by 7 researchers from the ERN Euro-NMD and ERN RND (for Charcot-Marie-Tooth diseases, channelopathies, mitochondrial disorders, metabolic myopathies, laminopathies, Parkinson disease and early-onset ataxias).

A ‘Treatabolome task force’ gathering INSERM-CRM, CHEO-RI, CNAG and JAX has been formed with the aim to optimize the work necessary for systematic reviews and their FAIR implementation in IT tools. First, to standardize the methodology for performing the reviews in a way that is easily transferrable to a database, we have followed and adapted the methodology of two pilot papers published for CMS and published a methodology paper, i.e., a guide to writing systematic reviews of rare disease treatments to generate FAIR-compliant datasets (Atalaia et al 2020, PMID: 32787960). Second, in order to valorize and disseminate the labor-intensive work of the systematic literature reviews, we have established link with the editors of the *Journal of Neuromuscular Disease* who have accepted to publish in a special issue devoted to the Treatabolome Concepts, the different treatment-focused systematic reviews prepared by RD experts. Third, to further publicize and disseminate the Treatabolome concept, and thus gathered additional systematic reviews and associated dataset related to other RD, it has been widely presented at different scientific meetings. A dedicated workshop has been organized at the virtual ESHG conference in June 2020 (ESHG2020, http://solve-rd.eu/the-treatabolome/). Its aim was to describe the concept and encourage community uptake and participation. Flyers, oral and poster presentations were presented not only at the Solve-RD Annual Meeting (5-6 March 2020, CNAG, Barcelona), but also at the virtual 10th European Conference on Rare Diseases and Orphan Products (14-15 May 2020) and the virtual 6th Congress of the European Academy of Neurology (23-25 May 2020).

In parallel of this intense work of preparation and collection of systematic reviews for various RD, the CNAG-CRG team have develop the IT tool for the Treatabolome database based on the supplementary tables available in the first pilot CMS review by Thompson et al. (PMID:30931400) and the guideline paper by Atalaia et al (MID: 32787960). The “Treatabolome task force” defined the vocabulary and the 3rd party databases needed to develop the tool (*treatabolome db*). A "disease" is defined as the combination of Orphanet Rare Disease Ontology (ORDO) term with, optionally, a single Online Mendelian Inheritance in Man (OMIM) term, and none to many Human Phenotype Ontology (HPO) terms. A "variant" is identified as the combination of a gene (as gene symbol and RefSeq id), a variant in cDNA

form (optionally a protein form can be provided), and its inheritance mode. A "treatment" is defined with an identifier coming from the Medical Subject Headings (MeSH), the Unified Medical Language System (UMLS), or Chemical Entities of Biological Interest (ChEBI); in combination with its clinical and biomarker effect. Finally, an entry into the *treatabolome db* will combine these three entities with the publication reporting them and two grades of evidence (GRADE and OCEBM).

In order to enable automated data intake from the SLRs and individual submissions, a manager tool and the first version of the web interface have been developed by the CNAG-CRG to query and use the data. This preliminary work was presented in a workshop at the 2020 European Society of Human Genetics meeting (ESHG2020, http://solve-rd.eu/the-treatabolome/, presentation 4: Treatabolome database).

For automatic data bulk-upload in the system, a template in MS Excel format has been created by the “Treatabolome task force” on the structure to collect the results from the SLRs. Future work includes the development of an automatic validator and uploader for the collected data through the template, as well as the update of the web query interface.

Conclusion:

Thanks to these actions, by the end of December 2020, 7 systematic reviews are almost ready for publication in the Treatabolome special issue of ***Journal of Neuromuscular Diseases*** to be published at the beginning of 2021: papers on Neuropathies (*accepted*, *accessible on line*), Channelopathies (*accepted*), Laminopathies (*under review*), Parkinson (*under review*), Metabolic Myopathies (under review), Ataxias (*in progress*), Mitochondrial Disorders (*in progress*) and Metabolic myopathies (*in progress*). Accordingly, as the IT tool and the template for transfer of dataset are ready, the production of datasets formatted in compliance with the Treatabolome Upload Tool produced by CNAG and their transfer a to CNAG (Barcelona) are in progress: Laminopathies (*dataset sent*), Channelopathies (*in progress*), Neuropathies (*in progress*), Parkinson (*in progress*), Metabolic Myopathies (*in progress*), Ataxias (*in progress*) and Mitochondrial Disorders (*in progress*). In order to engage people to write systematic reviews beyond the two participating ERNs (Euro-NMD and RND) so far, we will promote the model of the Treatabolome special issue in speciality journals related to other ERNs, first within Solve-RD partners, but also beyond to reach a maximum of rare diseases.

Finally, in order to facilitate in the near future, the labor-intense work of systematic review, the group led by P Robinson (JAX, USA) works on a closely related initiative that aims to develop a Medical Action Ontology to provide a structured vocabulary for medical procedures, interventions, therapies, and treatments for rare diseases and at the same time to use machine learning and literature mining approaches to extract treatments from the literature in an automated way. As this automated approach is highly complementary to the manual expert-led approach developed in this task, the “Treatabolome task force” aims at including bothuter-readable treatment annotations accessible at the point of diagnosis. Further sets of experts to see how the approaches can work towards the same end result of comp technical aspects of the Treatabolome database are also under discussion by the Treatabolome task force, including the development of systems to facilitate literature reviews and mechanisms to keep the content up to date.