The VEO Project and COVID-19

From a One Health perspective, global changes in global trends can act as drivers in emerging infectious diseases (EIDs).

VEO aims to answer the following key questions:

- How can data science and technology innovations improve infectious disease preparedness and response?
- Can we predict outbreaks? How can we use the novel tools to be better prepared? How can we bring outbreaks under control faster?
- Which are the barriers to the use of big data combined with citizen science? How do we resolve potential barriers?

What VEO hopes to achieve, is to take the data types that were brought together in COMPARE (sample, genomic, clinical, laboratory/microbiological and epidemiological data) and put it with other biodata and contextual data: novel lab and field phenotypic data, notifications, climate and environmental, citizen science, geospatial, macro-economic/demographic, and socio-cultural data.

VEO will use five scenarios and the key data drivers to develop the VEO platform. One of these scenarios is the Disease X scenario.

VEO started on 01 January 2020. The WHO announced COVID-19 outbreak as a pandemic on 11 March 2020. As the VEO Consortium brought together many of the Europe's best players in the area of EIDs, data capture and sharing, and zoonosis, many efforts within VEO could then be focused on COVID-19 and actions to be taken immediately as well as guidance in the case of future epidemics/pandemics.

Due to the COVID-19 pandemic, several WPs of VEO have had to explore how the resources of VEO should be focused to react to this Disease X scenario playing out before our very eyes. Certain work plans and time plans within VEO have been adjusted due to the outbreak. VEO is also supporting the EU COVID data initiative.

Below are the actions taken so far within the VEO Consortium to the COVID-19 pandemic.

WP 01, VEO Data Platform

WP 01 covers project objective 1: to develop and operate the core VEO data platform.

Partner 4, European Molecular Biology Laboratory (EMBL), Guy Cochrane, is leading this effort.

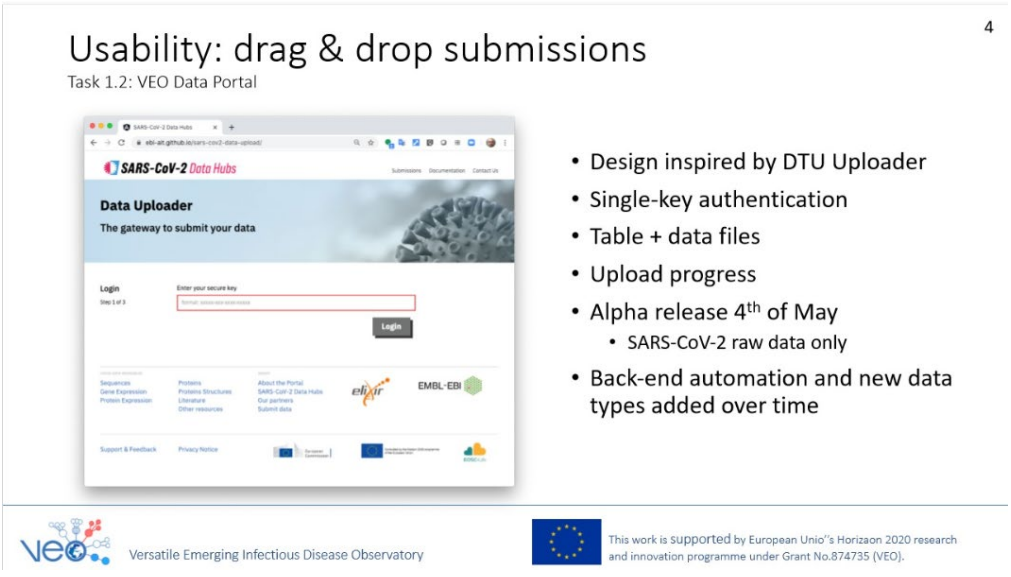
Due to the COVID-19 pandemic, work under WP 01 has been refocused to develop the SARS-CoV-2 Data Hubs; a new, specific WP (WP 15 COVID-19 Response) was initiated.

There is a need for a suite of analytical tools, storage, and a data sharing workspace to facilitate the sharing, analysis and reuse of raw and annotated SARS-CoV-2 genomic data. Sequences from around the world were being shared via other platforms, but none included any linked metadata or offered a way of sharing data and common analysis pipelines.

The European COVID-19 Data Platformwas started in March 2020 and launched 20 April 2020 and includes the SARS-CoV-2 Data Hubs and the COVID-19 Data Portal components.
https://www.eosc-life.eu/news/european-commission-launches-data-sharing-platform-for-covid-19/

Within the SARS-CoV-2 Data Hubs, work from WP 01 was accelerated to respond to the need of immediate international use. One of these elements includes adding a simplified data submission tool (drag & drop), installation of newly emerging analytical workflows, data visualization systems and substantial data mobilisation. In addition, national portals will be added to the COVID-19 data hub (Sweden is available now) to provide access to a number of national-level functions including support and links to national public health and clinical data systems.

Further improvements and additional tools will be added to the SARS-CoV-2 Data Hubs.

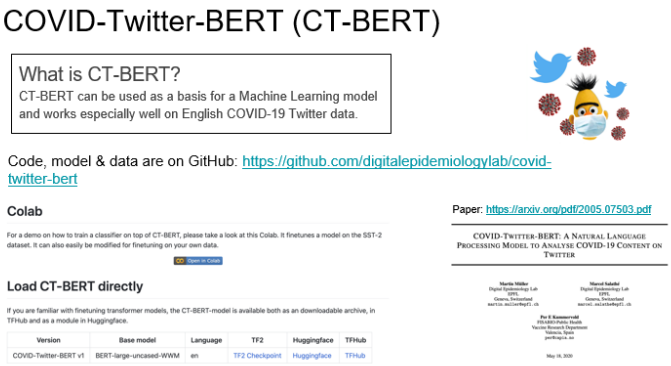


WP 02, VEO Analytical Platform: Advanced Datamining Tools

WP 02 covers project objective 2: To develop innovative cloud-based collaborative datamining tools, supporting data-intensive interdisciplinary collaborations of geographically distributed international teams on early warning, risk assessment and monitoring of EID, and integrate these into the VEO system.

Partner 13, Ecole Polytechnique Federale de Lausanne, led by Marcel Salathé, is leading efforts to use Twitter data as an additional type of data to indicate outbreaks. In addition, Marcel Salathé has been active in the digital contact tracing efforts related to COVID-19 as part of the DP-3T group.

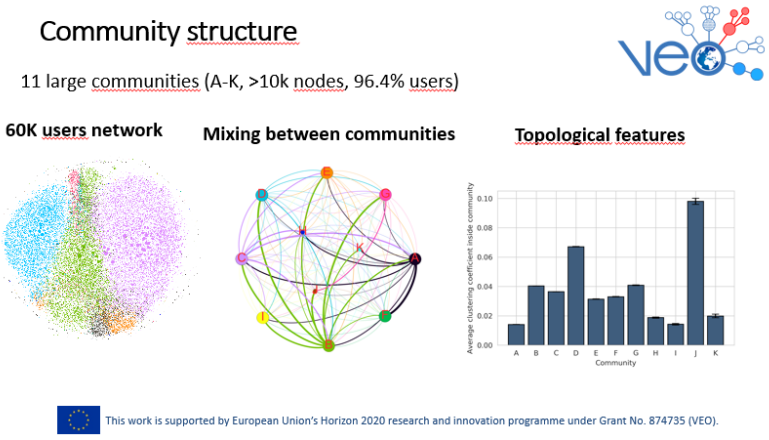
From January 2020, there have been 160 million tweets related to COVID. BERT (Bidirectional Encoder Representations from Transformers) is increasingly the preferred NLP technique for this type of data. COVID-Twitter-BERT (CT-BERT) works with English COVID-19 Twitter data. See article: https://arxiv.org/pdf/2005.07503.pdf . With the data the VEO Consortium members gathered and annotated, Twitter took notice and offered additional data (the full COVID stream) https://developer.twitter.com/en/docs/labs/covid19-stream/overview .



Partner 16, Alma Mater Studiorum - Universita di Bologna (UNIBO), led by Daniel Remondini, is undertaking a network analysis of social media data.

The aim of the analysis is to characterize Twitter network related to COVID-19 (English). They have the tweets, the users/nodes and weighted links (retweets).

Clear communities emerge, related to countries or to specific topics, with leading influencers. One of the aims is to characterize network evolution over time. How do communities, leading users and trend topics change, in relation to main COVID-19 events worldwide? Which picture is the social network structure providing about the pandemics concern?



Forecasting, nowcasting and tracking: Consortium members are called to serve

Partner 17, Università degli Studi di Padova (UNIPD), led by Luisa Barzon, is at the center of the European outbreak of COVID-19 in northern Italy.

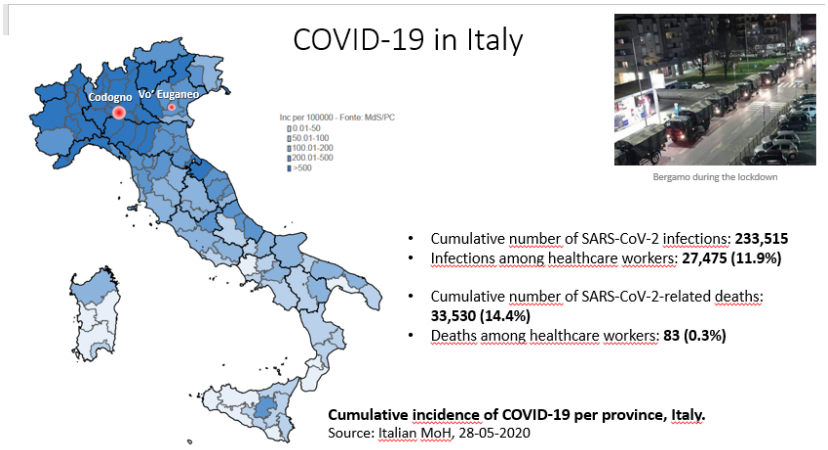
Dr. Luisa Barzon is the Co-leader of WP 05, Mosquito-Borne Diseases Use Case Scenario. WP 05 is aimed at Project objective 6 - To test and evaluate the VEO system in a mosquito-borne disease use case scenario and define key user requirements for VEO functionalities.

During the COVID-19 emergency, Dr. Barzon and her team were on the frontlines of the pandemic, and could gather and share firsthand knowledge of the spread, containment and related outcomes from the outbreak.

How did the COVID-19 pandemic develop in northern Italy? What was the epidemiology of COVID-19, what were the public health strategies?

Extensive case and contact investigation in the early epidemic phase provided relevant information on COVID-19 epidemiology and transmission dynamics that will be useful to WP 09, Disease X scenario.

At an initial glance, we can carry forward the following lessons: proactive surveillance and tracing, laboratory preparedness and capacity, and timely and extensive testing as crucial to prevent and control epidemics.



Partenr 11, Institut Pasteur (IP), led by Simon Cauchemez, IP, is exploring estimating the burden of COVID-19 in France by integrating multiple data types.

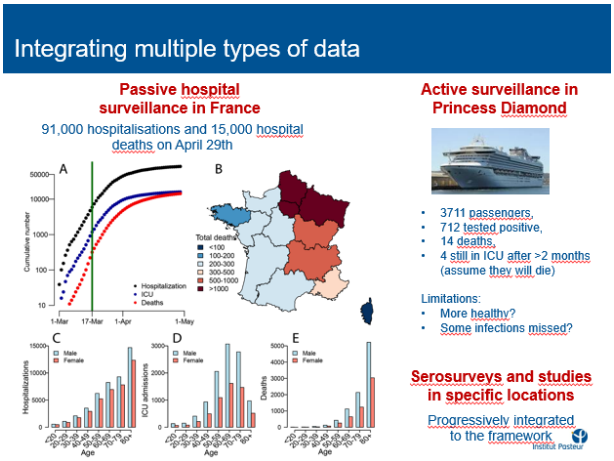
Simon Cauchemez is involved primarily with WP 04, Novel Host-Exposure Assessment Tools and Workflows.

The primary questions are:

- What has been the impact of the French lockdown on COVID-19 transmission?
- What is the proportion of the French population that has been infected by COVID-19 (national and regional)?
- Can we anticipate ICU admissions and bed occupancy to support planning?

Multiple types of data can contribute to answering these questions: passive hospital surveillance in France, active surveillance on the Princess Diamond cruise ship, serosurveys and specific location studies.

We need to explore limits of specific data and data types. The next steps include integrating serological data, what are the changes in transmission after lockdown, can we make short-term predictions based on this.



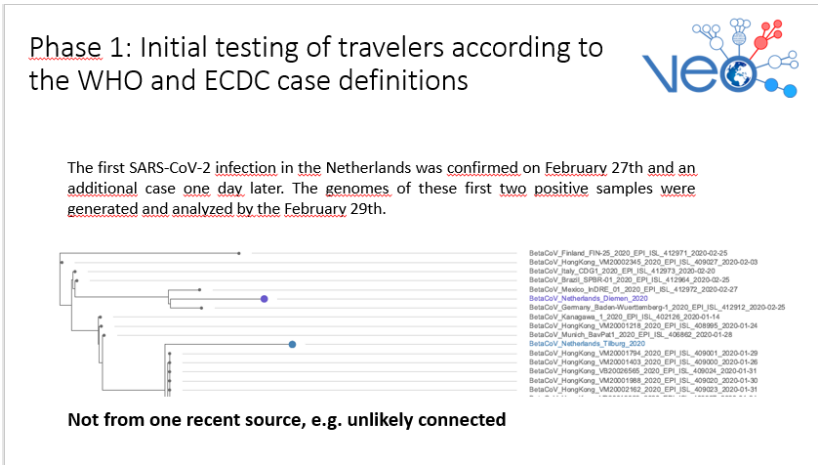
Partner 01, Erasmus Medical Center, (EMC) led by the work of Bas Oude Munnink, took up precision public health through real-time sequencing to track and support public health investigations.

Bas Oude Munnink primarily works in WP 09, Disease X Scenario. This WP aims at testing and evaluating the VEO system on its applicability in a disease x use case scenario - a fast spreading EID, with high case fatality by an unknown pathogen. This is also connected to WP6 concerning the fieldable sequencing, e.g. in the wildlife scenarios.

Bas Oude Munnink is part of the team at EMC who conducts the rapid SARS-CoV-2 whole genome sequencing for informed public health decision making in the Netherlands (https://doi.org/10.1101/2020.04.21.050633).

In the Netherlands, WGS was used during to facilitate informed public health decision making during three different diagnostic phases: Initial testing of travelers, screening high-risk contacts, and systematic sequencing during exponential growth phase.

There are additional outbreak investigations that are ongoing, for example: mink farms, hospitals, nursing homes, schools and slaughterhouses.

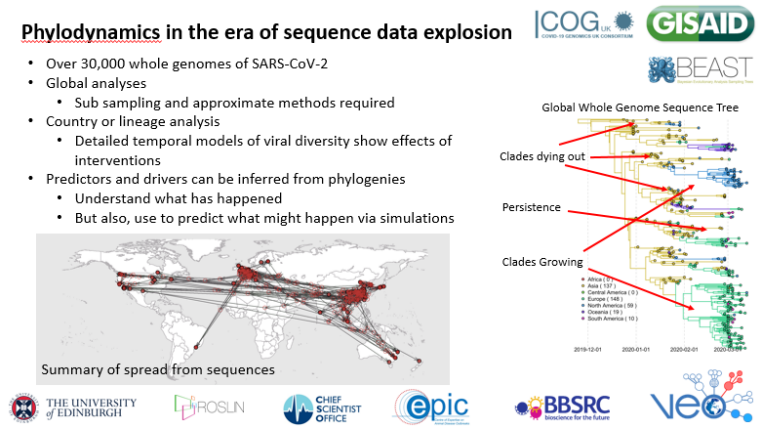


Partner 15, University of Edinburgh (UEDIN). Led by Sam Lycett, UEDIN, is looking at phylodynamics in the era of sequence data explosion.

Dr. Lycett is linked to WPs 06 (Zoonotic) and 09 (Disease X). WP 06 is the Zoonotic Wildlife Scenario, the main aim of this scenario is to substantially improve the prediction of incursion and spread of HPAIV, USUV, and WNV via wild birds in Europe, as well as to assess the risk in real time of the zoonotic and pathogenic potential of emerging virus strains.

Sequence data collected globally are shared via GISAID. This platform was originally set up for influenza, but now there are over 30,000 whole genomes of SARS-CoV-2. In the UK, COVID-19 Genomics Consortium has sequenced thousands of UK samples, and these data are organized into Global and UK lineages with the Pangolin lineage assigner. The team has developed timeline and trees for early introductions and spread of COVID-19 outside of Asia.

Large data sets of thousands of sequences cannot be processed using BEAST (the phylodynamic inference program of choice), so the data needs to be subsampled for using a suitable stratified method for global analysis. However, for a single country (with few introductions) or per single lineage, it is possible to generate time-resolved phylogenies and estimate viral effective population sizes over time, which can show the bottleneck or suppression effects from interventions. Predictors and drivers for outbreaks can also be inferred from phylogenies.

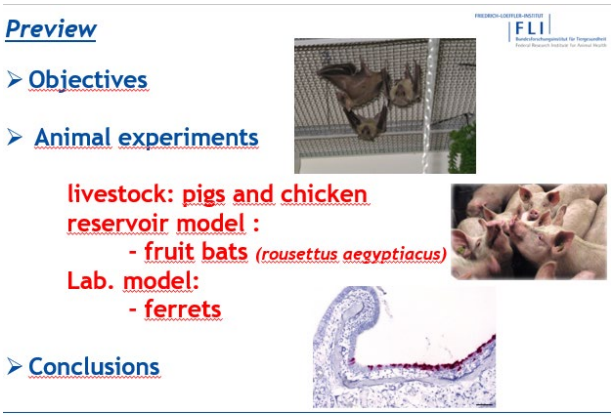


One Health: The human-animal interface

Partner 03, Friedrich Loeffler Institut (FLI), led by Martin Beer, FLI, is now looking into the host range of COVID-19 via fruit bats, swine, chickens and ferrets and further animal models will be studied. Martin Beer is the WP Leader for WP 06, Zoonotic Wildlife Scenario. WP 06 aims to improve the prediction of incursion and spread of HPAIV, USUV, and WNV via wild birds in Europe, as well as to assess the risk in real time of the zoonotic and pathogenic potential of emerging virus strains.

WP 06 will now include zoonotic SARS-CoV-2 scenarios; including animal reservoirs, zoonotic spillover (e.g., mink-human, cat-human), and fieldable sequencing. The work included the following questions:

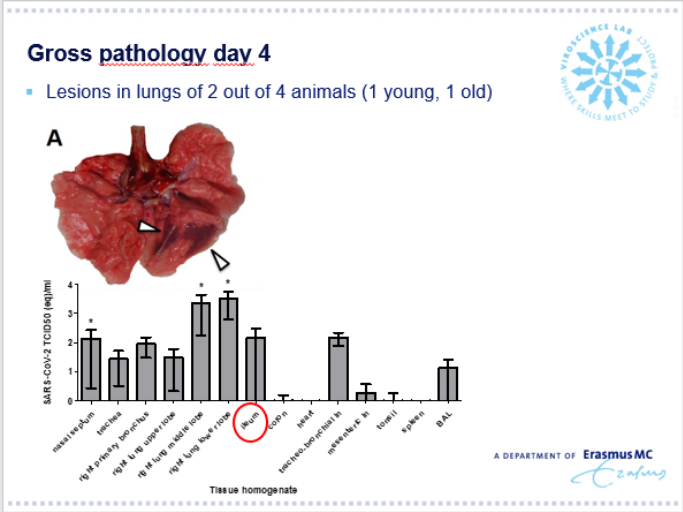
- Are important livestock animals (pigs, chicken) susceptible to SARS-CoV-2-infection?
- Infection of fruit bats as model for a SARS-CoV-2 reservoir host
- Laboratory animal model – are ferrets suitable? (infection, replication, transmission)



Partner 01, Erasmus Medical Center (EMC), with efforts from Bart Haagmans, is developing animal models of disease and transmission.

His work starts with studies on the COVID-19 interaction with the ACE-2 viral receptor. Given the high homology between human and non-human primate ACE-2, the pathogenesis of SARS-CoV-2 was studied in NHP and compared to results obtained previously with other coronaviruses such as MERS and SARS coronavirus. The team is also using the ferret model to study transmission of coronaviruses in more detail.

The next steps include integrating knowledge on pathogenesis and transmission of coronaviruses in humans, animals, and in vitro systems (e.g., organoids).

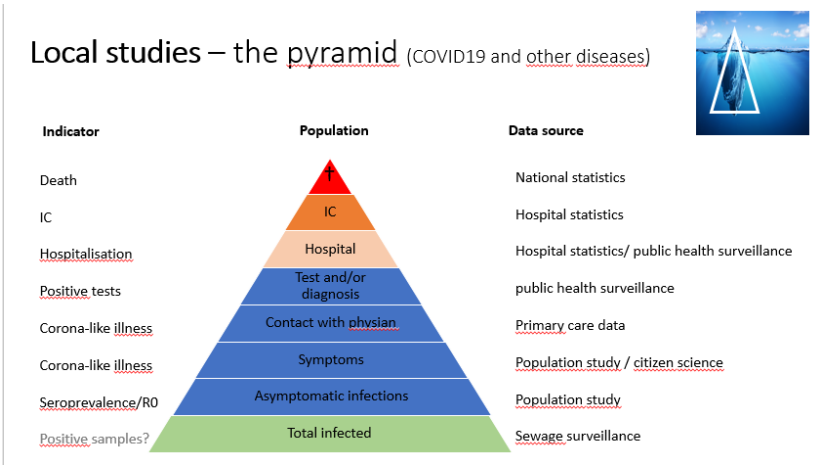


WP 08, Silent Epidemic Scenario

This WP is aimed at Project objective 6 - To test and evaluate the VEO system in a silent epidemics use case scenario and define key user requirements for VEO functionalities. The aim is to understand the full burden of 'silent epidemics'. This approach will be validated using two key challenges: i) emergence and circulation of AMR determinants; ii) emergence of circulation of widespread common pathogens.

Partner 02, Technical University of Denmark (DTU), led by Frank Aarestrup, is leading the effort to research the silent epidemic/endemic diseases. The studies are based on the pyramid surveillance, with classical isolate based surveillance at the top of the pyramid and the sewage based surveillance at the bottom of the pyramid.

Due to the COVID-19 outbreak, WP 08 has reviewed the pyramid surveillance work it will do and will now collect data related to COVID-19 from four cities across Europe, taking advantage of the major intervention/disruption due to lockdown. The VEO team is looking at common data from the four cities (clinical data, mobility data, sewage, etc.).



The above descriptions are the efforts from Consortium members related to the VEO project objectives, work plans, and deliverables. However, many members of the Consortium are continuing to support efforts to care for patients, respond to data requests and analyze raw and metadata related to the pandemic.

This work is supported by European Union's Horizon 2020 research and innovation programme under Grant No. 874735 (VEO).