

## Healthdata.be

Healthdata.be fait permet à tous les laboratoires cliniques belges de rapporter d'une manière standardisée – pour plusieurs projets - à une **plateforme centralisée pour la gestion et l'analyse de données**. Cette plateforme peut à son tour être utilisée par les chercheurs et par les équipes de surveillance. Pour optimiser les prestations et la flexibilité, healthdata.be va mettre la nouvelle architecture technique HD4DP2.0 à la disposition de tous les établissements de soins belges et en assurer également l'entretien. Une *Application Programming Interface (API)* sera développée pour l'échange de données cliniques, épidémiologiques et génomiques en format *Fast Healthcare Interoperability Resources (FHIR)* comme conçu par Health Level Seven International (HL7), une organisation qui se concentre spécifiquement sur la standardisation de données de santé. Etant donné la nature des données qui doivent être collectées, **les services d'une tierce partie habituée indépendante (eHealth)** seront utilisés tant pour la pseudonimisation que **pour le codage end-to-end** des données à caractère personnel. La traduction et l'harmonisation des codes et valeurs actuellement utilisés dans des systèmes terminologiques reconnus internationalement (p. ex. SNOMED CT, LOINC) sont prévues. Pour le projet HERA-BE-Incubator, healthdata.be assurera la collecte et le transfert de données dans le format demandé vers les différents utilisateurs finaux impliqués dans la surveillance et l'analyse d'épidémies. La figure ci-dessous donne un aperçu du *high level design* (Figure 2).

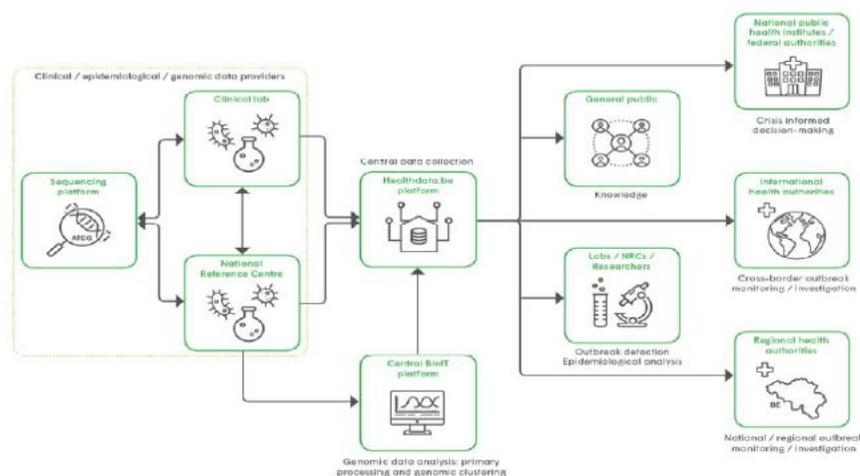


Figure 2. High level design de l'infrastructure nationale eHealth pour la surveillance génomique et épidémiologique des maladies infectieuses. Les données épidémiologiques et/ou génomiques, reçues au départ des centres nationaux de référence ou des laboratoires cliniques, seront collectées à l'aide d'un ensemble standardisé de variables sur la plateforme bio-IT centrale. Les données génomiques brutes seront immédiatement envoyées vers la plateforme bio-IT centrale où une analyse bio-informatique automatisée et harmonisée sera effectuée. Les documents FASTA et les indicateurs génomiques seront stockés sur la plateforme healthdata.be. Des données seront envoyées au départ de la plateforme centrale de données génomiques-épidémiologiques vers les différentes parties impliquées. Cette opération aura lieu à tout moment conformément aux besoins de la santé publique, aux mandats à disposition et aux règles relatives à la gestion de données.

## HERA-BE-INCUBATOR-WGS: VERS UNE NOUVELLE (H)ERE

### Infrastructure nationale pour la surveillance génomique-épidémiologique des maladies infectieuses

#### Introduction HERA-BE-Incubator

En réaction à la pandémie du COVID-19, la Commission européenne a créé l'Autorité européenne de préparation et de réaction en cas d'urgence sanitaire (HERA). Le but en est de renforcer l'union européenne de la santé en améliorant la préparation et la réaction à des foyers (pandémiques) nationaux et internationaux de maladies infectieuses dans le futur. Le projet HERA-BE-Incubator a été lancé par HERA et son implémentation en Belgique permet de développer une infrastructure nationale pour la collecte de données (SGE) 'Séquençage du génome entier' ainsi qu'une infrastructure durable dans laquelle les données génomiques peuvent être couplées aux données épidémiologiques et cliniques. Une telle infrastructure peut ainsi être utilisée plus largement pour anticiper les épidémies et gérer les risques qui y sont liés. Le projet vise le traitement transactionnel de volumes de données plus importants dans le cas d'une vaste épidémie ou d'une pandémie. **Pour l'instant, le projet HERA-BE-Incubator a été mis en place comme un projet *proof-of-principle* pour cinq Centres nationaux de référence (CNR) de microbiologie humaine et il fait appel à la plateforme healthdata.be. Le projet est coordonné par Sciensano et il ambitionne d'accueillir la participation d'autres laboratoires dans un futur proche.**

#### Centres nationaux de référence (CNR)

Les analyses génomiques sont de plus en plus souvent intégrées dans les activités des CNR mais aussi des laboratoires cliniques (laboratoires SGE non CNR). Ce dernier point s'est avéré clairement lors de la pandémie du COVID-19. Toutefois, les données sont actuellement réparties entre différents établissements et il se peut que cette tendance se renforce encore, à mesure que le séquençage pourra être implémenté plus facilement dans le futur par les laboratoires (cliniques). Le projet HERA-BE-Incubator prévoit, grâce à la centralisation des données génomiques des pathogènes et au développement de pipelines bio-informatiques automatisés et validés, de globaliser les données provenant de plusieurs sources. Grâce au traitement harmonisé, **les données pourront plus facilement être comparées entre elles et elles pourront être utilisées pour la détection de clusters**. De plus, la sensibilité pour l'identification d'une augmentation épidémiologique de certains variants pourra être renforcée et la surveillance de profils de résistance prévus de différents pathogènes pourra être étendue, par exemple. Le couplage des données épidémiologiques permettra d'éclaircir le contexte des épidémies et d'effectuer des analyses complémentaires.

*"Le projet renforcera la détection et l'analyse d'épidémies en centralisant la surveillance génomique et en la couplant aux données épidémiologiques et cliniques."*



Plus d'infos?  
Scannez le code QR



sciensano

Tableau 1: Opportunités pour les acteurs dans le domaine de la santé publique

### Qu'est-ce que ce projet peut signifier pour vous ?

Les données épidémiologiques et génomiques sont **couplées** et analysées ensemble.

Les données sont collectées **en temps réel**.

Détection **ponctuelle et sensible** des variants/épidémies.

Meilleures **gestion et analyse des épidémies** grâce à une disponibilité centralisée des données.

Support aux **activités de santé publique** et renforcement de la surveillance.

**Pipelines de bio-informatique** harmonisés et automatisés pour tous les laboratoires concernés.

Surveillance étendue de **profils de résistance antimicrobienne** prévus, à l'aide la génomique.

**Le téléchargement automatisé** de grands fichiers de données est possible.

Possibilité de **télécharger des fichiers FASTQ** vers ENA via le système.

*“Ancrer et renforcer l'infrastructure développée lors de la pandémie du COVID-19 et en faire un outil de préparation plus large pour une réaction aux foyers pandémiques”*

### Bio-informatique

Un environnement bio-informatique central sera mis en place avec les outils, les pipelines et les bases de données de bio-informatique nécessaires pour permettre un **traitement standardisé, harmonisé et traçable de toutes les données SGE pour les pathogènes**. De cette manière, un typage et une caractérisation génomiques entiers seront possibles, y compris une vaste gamme d'indicateurs (p. ex. cgMLST, prévisions de profils de résistance antimicrobienne (RAM), de gènes de virulence, de sérotypage, etc.). Un support pour tous les pathogènes pertinents sera graduellement intégré et les résultats seront dans ce cadre rapportés d'une manière claire et intuitive aux fournisseurs de données. Les indicateurs génomiques, obtenus pour les échantillons, seront introduits dans une base de données moléculaire nationale centralisée. Les données des échantillons eux-mêmes seront **réanalysées périodiquement à l'aide de banques de données de référence actualisées** afin de garantir que chaque fois que de nouvelles informations génétiques (p. ex. nouveaux gènes RAM, marqueurs cgMLST) sont disponibles, elles soient également reprises dans l'infrastructure génomique-épidémiologique nationale. De plus, des analyses de détection de cluster automatisées, basées sur une parenté génomique, seront implémentées afin de générer des avertissements automatiques pour de potentiels clusters. Tout ceci permettra des analyses de cluster plus approfondies.

*“L'approche centralisée et standardisée offre la possibilité d'une nouvelle ère, par un renforcement continu de la préparation de la Belgique à des situations de crise en santé publique”*

### Epidémiologie des maladies infectieuses et santé publique

Le réseau belge de 41 CNR assure l'expertise, le diagnostic et la surveillance d'une série de pathogènes importants pour la santé publique. Actuellement, la collection de données est centralisée, traitée et mise à la disposition des acteurs de la santé publique et de la population (collection de données CNR « classique »). Cela se fait toutefois indépendamment d'une **identification des patients**. Une transition vers un système dans lequel les données sont collectées sur la base de cette identification unique offrira de nouvelles opportunités, notamment le couplage de données. De plus, ce système permettra de mener des études rétrospectives et prospectives de cohorte, de cas-cas et de cas-control, qui à leur tour, pourront donner de nouvelles perspectives tant en ce qui concerne l'impact que l'épidémiologie des maladies infectieuses. Concrètement, les données seront collectées selon les **Définitions de collecte des données (DCD)**, qui comprendront non seulement les données CNR « classiques » mais aussi les **données génomiques, dont les données SGE**. Cette collection de données centralisée, facilitée par healthdata.be, sera disponible pour les différents acteurs de santé publique selon les besoins et les mandats. L'infrastructure nationale vise à renforcer chacune des trois phases de détection de signaux et de suivi à l'aide de meilleures données (Figure 1).

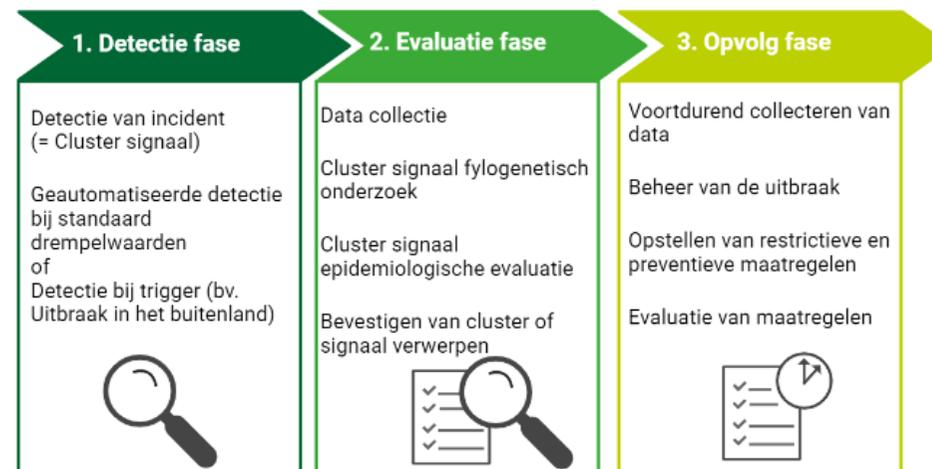


Figure 1. Représentation visuelle des trois phases de la détection de signaux et du suivi épidémiologiques, toutes trois importantes pour la préparation aux pandémies. Un renforcement régulier de toutes ces phases améliorera la préparation à une pandémie. Les trois phases comprennent : 1) détection d'épidémies potentielles, 2) évaluation d'épidémies potentielles et 3) suivi de l'épidémie, c'est-à-dire la gestion de l'épidémie. Les données génomiques centrales permettent une détection automatisée de clusters (phase 1) : sur la base de points communs dans le génome, des avertissements de clusters potentiels sont générés (= signal de cluster). Ces clusters potentiels exigent toutefois une évaluation individuelle (phase 2), y compris une analyse phylogénétique détaillée et une évaluation contextuelle et épidémiologique. Après confirmation du cluster ou du foyer épidémique, des mesures de prévention et de confinement peuvent être prises si nécessaire (phase 3) et un suivi peut être prévu. Notons toutefois que le projet HERA-BE-Incubator se concentre actuellement sur la création de l'infrastructure nécessaire et ne va pas implémenter en détail toutes les phases susmentionnées.