Proposal template Part B: technical description

STRENGTHENING ENVIRONMENTAL SURVEILLANCE TO ADVANCE PUBLIC HEALTH ACTION

List of participants in the ODIN consortium

Participant No. *	Participant organisation name, short name	Country
1. Rolf Lood (Coordinator)	Lund University, LU	Sweden
2. Adriana Krolicka	Norwegian Research Centre, NORCE	Norway
3. Tarja Pitkänen	The Finnish Institute for Health and Welfare, THL	Finland
4. Vito Baraka	National Institute for Medical Research, NIMR	Tanzania
5. Vivi Maketa	University of Kinshasa, UNIKIN	DR Congo
6. Marc Christian Tahita	Centre National de la Recherche Technologique et Scientifique, IRSS-DRCO	Burkina Faso
7. Lennart Martens	Flanders Institute for Biotechnology, VIB	Belgium
8. Trudie Lang	Oxford University, The Global Health Network, TGHN	United Kingdom
9. Bart Mesuere	Ghent University, UGENT	Belgium

1.   Excellence

1.1   Objectives and ambition

The ODIN project will strengthen genomics and bioinformatics capacity, as well as database management skills for generating, maintaining, and querying large data sets, in sub-Saharan countries. Further, the project will develop a genomic surveillance system relying on environmental monitoring of major communicable disease agents in community wastewaters and other environmental samples (wells, rivers, soil). In a timely, socially and ethically acceptable anonymous manner, environmental surveillance can detect outbreaks of poverty-related pathogens, waterborne diseases and antimicrobial resistance (AMR), as well as the means to convey this information to key stakeholders for efficient implementation of data-driven evidence for informed public health policies. By bringing together a multi-disciplinary team of leading experts and organisations within the fields of communicable disease epidemiology, microbiology, bioinformatics, water technology and environmental science, ODIN will provide an optimal research, development, capacity building, and implementation environment that will help set up a sustainable model for how genomics surveillance systems can be applied in sub-Saharan conditions and support safe drinking-water supply and sanitation. We will achieve our goals, through the following objectives, which are reflected in the project's work packages:

● Set up an environmental surveillance scheme for a set of chosen human pathogens and AMR determinants to be piloted and utilised in local sub-Saharan communities, supporting SDG3. (WP2)

● Develop a mobile wastewater and clean water surveillance system in order to efficiently and in a timely manner detect, gather data on, and report outbreaks in remote areas of the participating countries, contributing to rapid response. (WP3)

● Accelerate raw data processing and analysis through interactive, semi-automation, thus speeding up transfer of information about threat. (WP4)

● Set up a training program for improved capacity building that entails genome sequencing of environmental samples and subsequent bioinformatics processing of the produced genomic data containing spatial and temporal trends in the occurrence of human pathogens in sub-Saharan countries. (WP6)

● Develop standards to support practices for sharing genomic data, including beyond national borders, for public health measures. (WP7)

● Strengthen health systems by process(es) on how the generated data will be transferred to key stakeholders (clinics, governments, policy makers, etc) for efficient and timely actions upon outbreaks or detection of contaminated water. (WP4, WP6, WP7)

● Investigate, from an epidemiological and bacteriological angle, the role of different *Vibrio* species in causing cholera. (WP5)

● Reduce annual cases of disease, illness and death due to drinking water from contaminated sources by increasing awareness and providing protocols for efficient water and sanitation interventions, supporting SDG6 (WP5)

Figure 1. The Norse God Odin.

The project acronym is inspired by the Norse God Odin, who, according to mythology, sacrificed an eye to reach a higher level of wisdom. Through this sacrifice, and his ravens Hugin and Munin, he was able to see and know everything that was occurring, everywhere.

This project will go forward from, and beyond, a setting where much focus has been on individual pathogens (e.g. narrow scope) as well as individual measurements (e.g. patients), to build a system that is low-cost, robust, easy to maintain, but has the capacity, like Odin and his ravens, to detect "everything". The raw data will go through several adjustable analytical processes to create easily understandable reports for key stakeholders to take informed, and rapid, decisions.

Problem description

Water, and in particular clean water, is a resource we take for granted in many parts of the world. However, a vast majority of the global population do not have that luxury, and instead have to use water from sources that can be contaminated with infectious microorganisms. 80% of all infectious diseases worldwide are waterborne, accounting for 1.7 million human deaths annually, and 3.7% of the total DALYs (disability-adjusted life year) global burden of diseases **(1)**. More specifically, most cases are caused by faecal contamination of the water (e.g. pathogenic *E. coli*), or bacterial outbreaks like *Vibrio cholerae*, but can also be of viral nature (hepatitis A and E, polio). Such contaminations and outbreaks are therefore of high priority to surveil in order to reduce infections, and thus deaths, associated with poverty-related infections.

Cholera outbreaks are episodic in most parts of Africa, with several outbreaks per year. Other parts have endemics (e.g. Eastern Congo) of cholera – situations that are enhanced due to geopolitical circumstances. Cholera is mediated by the bacterium *Vibrio cholerae* carrying a cholera toxin (*ctx*) gene, disrupting the barrier function of the intestines leading to massive water loss and potential death if untreated. Existing vaccines are often only targeting certain serotypes of *V. cholerae*, are expensive, and do not generate lifelong immunity. It is therefore imperative that better surveillance systems, and implementations thereof, are put in place to enable timely detection, and to provoke efficient responses of these threats.

Genomics approaches to tackle infectious threats

The conventional approach to surveillance of infectious diseases faces several limitations including potential sampling bias and high cost, and these are especially pertinent in low- and middle-income countries with constrained resources. These limitations can be overcome by deployment of complementary environmental surveillance of sewage in septic tanks and wastewater treatment plants (WWTP). In low-resource settings, non-sewage systems are also important collectors of pathogens excreted by the local inhabitants. Wastewater-based epidemiology is an early warning tool that provides a novel surveillance strategy to identify known and unknown infectious pathogens, (re-)emerging infections, antimicrobial resistance and tracking of evolution/pathogenesis at population level in a cost-effective manner. Therefore, environmental surveillance can provide a useful complementary strategy to shed light on clusters and transmission of infections, as well as AMR, in resource-constrained settings and communities with fragile healthcare systems, and can contribute to improved regional, as well as global, health security. Indeed, using advances in next-generation (NGS) sequencing tools, it is now feasible to characterise microbiome and virome sequences at community level from environmental samples contaminated with human excreta to profile known and unknown pathogens, of epidemic or pandemic potential, and to chart ARG composition (resistome).

The methods used for the corresponding qualitative and quantitative analyses, which are still seen as `sophisticated`, have in fact become considerably less demanding thanks to the technological advances, and can thus assist the development of rapid and routine protocols for gene copy number quantification. Advances in NGS technologies have enabled rapid pathogen identification and control in Europe and North America, while in Africa, several platforms are being deployed. However, this capacity remains low and inadequate to the needs due to weak infrastructure, constrained resources and restricted capacity. Currently, NGS sequencing capacity in the region is limited and concentrated in only a few countries, mostly outside the public health institutes (PHLs). Therefore, strengthening the capacity of PHLs is required to harness the potential impact of sequencing in public health.

Recently, there has been increasing interest to use Oxford Nanopore Technologies (ONT) sequencers in several resource-limited settings due to its suitable features including portability, turn-around time, low capital investment cost, and low infrastructure requirement. Since its first introduction in 2014, ONT has grown rapidly to one of the most powerful sequencing technologies **(2)**. The broad range of applications of this technology includes genome assembly, full-length transcript detection, base modification detection, and especially urgent point-of-care testing that requires fast turnaround results as seen in the COVID-19 pandemic **(3,4)**. Nanopore sequencing has also been

used in AMR surveillance and detection in health care settings **(5)**. In Africa, ONT platforms have been used for real-time outbreak tracking during recent international epidemics, including Ebola, Zika, SARS-CoV-2, and the detection of bacterial infections and antimicrobial resistance. Despite the potential benefits, deployment of novel NGS tools in the African region remains limited due to several issues, including lack of sequencing infrastructure, data storage and shortage of computational skills.

Based on the success of using wastewater-based epidemiology during the COVID-19 pandemic worldwide, we here propose to implement complementary surveillance systems for multiple pathogens, adapting monitored pathogens to the context of the location. However, wastewaters only provide information on *what is* happening, and less information on *how* it has happened, which limits its impact with regards to addressing the problem. We will therefore complement this strategy with environmental monitoring using a developed mobile surveillance system with remote access, and subsequent interventions based on e.g. water and sanitation safety plans for the prevention and control of infection outbreaks.

Preliminary data

During a pilot test, tap water was collected from several locations at a cholera outbreak site, as well as from medical wards, restaurants, and village wells in Tanzania. None of the samples were positive for *Vibrio cholerae*, but all samples from the cholera outbreak site had high levels of *Vibrio metschnikovii*, a rare, and rather poorly characterised pathogen associated with human infections. In the few cases described, the bacterium tends to cause mainly intestinal symptoms **(6)** but had not yet been associated with cholera outbreaks before. For water sources outside of the cholera outbreak, all samples contained noticeable levels of faecal contamination (*e.g. E. coli*), while lacking in typical foodborne pathogens (*e.g. Helicobacter pylori*). Several resistance genes could also be detected in tap water in both hospitals as well as in regular water supplies (*e.g.* wells), mainly including *tetA, tetM,* and *bla*$_{CTXM-9}$. Importantly, *bla*$_{CTXM-1}$ genes were found to be highly enriched in bacteriophage populations, indicating that these can be mobilised and spread through transduction.

Innovative nature of Project ODIN

Project ODIN is ideally suited to the current call topic *Genomic epidemiology for surveillance and control of poverty-related and emerging/re-emerging infections in sub-Saharan Africa*. Our project is based on genomic surveillance of wastewater and environmental samples within sub-Saharan countries, which is a highly cost-efficient means to analyse prevalence and outbreaks of infectious agents and antimicrobial resistance (AMR). Furthermore, the project will develop easy-to-implement analytical tool packages, enabling different regions/countries to focus on selected pathogens from the same raw data (*e.g.* genomics data from wastewater), which thus adds a layer of flexibility and sustainability to the methods, as these can be adjusted based on different needs in different regions, at different time-points. We will not only focus on urban regions but will also add flexibility in the surveillance system by developing a unique mobile surveillance unit, targeting remote areas where outbreaks are common. Importantly, our project is not only focused on research and development/innovation, but includes two other, key pillars: capacity building, and implementation/dissemination, which will help ensure sustainability of the ODIN project. Currently, surveillance programs within the sub-Saharan countries are rather scattered, focus on single pathogens, and have little synergy or longevity, along with limited or no communication between research institutes and key stakeholders. This, combined with low incapacity in instrumentation, funding, and expertise in genomics and bioinformatics has hampered earlier attempts to strengthen capacity within these countries. Based on lessons thus learned we therefore strongly believe that for a project to be successful, it needs to:

1.   implement cost-efficient, static and mobile surveillance methods;
2.   develop flexible surveillance and intervention methods that can easily be adapted for changing needs;
3.   implement means to target regional surveillance needs through adapted analytical workflows;
4.   build local capacity to further develop, maintain, teach, and use the developed systems; and
5.   have strong local presence and involvement on all levels (public, research, stakeholders, government).

Project ODIN fulfils all of the above criteria and will thus serve to integrate a highly efficient surveillance system in sub-Saharan countries that will not only last during the proposed funding period, but that will be sustained, and can even expand, after the current funding scheme is ended. For this, we will have several high-level stakeholder meetings, including early in the process, to secure funding for the surveillance systems. All main goals are organised into work packages and tasks that can all be measured and verified, to make certain that they will be achieved in time. The composition of the consortium, described in more detail in section 3.2, is key for the feasibility of the project, allowing the objectives to be realistically achievable. Specifically, several PIs have first-hand experience in developing, setting up, and maintaining surveillance systems as well as developing portable, low-cost, mobile surveillance systems, bioinformatics pipelines for large-scale omics data, and human health risk assessment approaches for environmental exposure. Furthermore, we have a strong local presence (TZ, DRC, and BF) and the ability to implement the methods and competences through global and local networks (TGHN, TGHN-Africa, Africa

CDC, ALERRT, and WRC).

Ambition of Project ODIN

Our project goes well beyond any current surveillance ventures within the participating countries due to its ability to not only cover a large region thanks to its cost-efficient method (WBE), but also in that it is a *non-targeted sampling* (*e.g.* "all" data is collected through 16S amplicon sequencing and metagenomics), with a *targeted analytical process* added on top, allowing for individual regions to take advantage of the huge amount of information generated, yet filter it to fit their current needs - needs that will be different in different geographical regions, and at different time points. Moreover, the development of a mobile, low-cost instrument for timely assessment of contaminated water sources is an innovative approach that will facilitate a broad reach in both rural as well as urban regions. Our project is thus highly ambitious in terms of its reach (geographic, and in surveillance capacity) as well as its innovativeness, building upon novel methodologies that have recently been developed, and well-studied and validated, in Europe during the covid19-based WBE surveillance. The main objective of this project is therefore not to advance science per se, but to maximally leverage recent advances to improve health and save lives by building better surveillance systems that can inform government institutes in a timely manner on how to act to reduce the spread of infections.

Technology Readiness Level (TRL) of Project ODIN

The implementation of a static surveillance system will be built upon existing examples in Europe, and thus starts at a high TRL (TRL 7/8). What specifically needs to be developed is the analytical work package, where inspiration can be taken from other genomics analysis approaches. For the portable, low-cost surveillance module, prototypes have been developed earlier by NORCE (TRL 5/6), and will now be adjusted for the current needs. We therefore start at a relatively high TRL for some of the work, while other parts will be built from idea stage to application (e.g. *Vibrio* work). Importantly, this combination of established methods with added innovation is often a key for success, and will allow us to bring this technology platform to sub-Saharan Africa, implement the surveillance system, build capacity for genomic and bioinformatic surveillance, and improve quality of life for millions, through lower need of hospitalisation due to waterborne infections, and ultimately saving lives.

1.2 Methodology

Background of the environmental surveillance of human pathogens

Wastewater testing is a quick and reliable way to track the levels of human pathogens in a community over time, providing trend data that complements other surveillance data and that informs public health decision making. Wastewater-based epidemiology is an especially useful approach in low-resource settings where the availability of clinical testing as well as the willingness of individuals for being tested for communicable disease agents might be very low. In other words, wastewater surveillance does not depend on people's access to healthcare, people seeking healthcare when sick, or the availability of clinical testing. The ability of wastewater testing to reveal the occurrence of infected people that shed pathogens in their excreta (e.g. stool, urine, saliva, nasal excretions) in the sewerage network area even if they are not showing symptoms, has already been used for decades in the global efforts to eradicate poliovirus from the world[1]. In Finland, for example, the presence of poliovirus in wastewater samples has been tested with good national coverage since the 1960´s already **(7)**. While endemic in Afghanistan and Pakistan, poliovirus outbreaks have occurred in two European, four Eastern Mediterranean, and 27 African countries. The latter includes Burkina Faso and Democratic Republic of the Congo (DRC), classified as states infected with vaccine-derived poliovirus type 2 (cVDPV2), with two cases in 2021 in Burkina Faso, and with 28 **cases** in 2021, and 73 cases and one environmental sample positive for cVDPV2 so far in 2022 in DRC.

Since spring 2020, worldwide implementation of wastewater-based epidemiology (WBE) for tracking of SARS-CoV-2 at population level has yielded invaluable insight into the dynamics of the COVID-19 pandemic and has helped governments take appropriate actions to contain the spread of the virus. The European Commission has strongly encouraged Member States to put in place a national wastewater surveillance system targeted at data collection of SARS-CoV-2 and its variants (UE 2021/472[2]). In addition to Europe, such initiatives have also sprouted in North America and Australia. Now that this infrastructure exists for Covid-19, it can be relatively easily expanded to other viruses, bacteria, and fungi. For example, in the United States, the CDC is expanding the National Wastewater Surveillance System (NWSS[3]) by early 2023 from SARS-CoV-2 monitoring to tracking other disease-causing microbes such as the most concerning types of antibiotic resistance, influenza, norovirus, *Candida auris*, and, most recently, monkeypox[4]. Although wastewater testing has proved to be a reliable way to track levels of

---

[1] https://polioeradication.org/

[2] https://wastewater-observatory.jrc.ec.europa.eu/#/

[3] https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html

[4] https://www.inverse.com/science/wastewater-antibiotic-resistance-monitoring

COVID-19 in a community, based on global mapping of research groups active in WBE of COVID-19[5], the current actions occur radically less in Africa than other continents and do not cover at all the countries where ODIN will develop environmental surveillance systems, i.e. Tanzania, DRC, and Burkina Faso.

Antimicrobial agents are currently the recommended option to treat most communicable diseases of bacterial origin. However, concerns are raised about the emergence of multidrug resistant pathogens. Unfortunately, there are no vaccines against most of these pathogens, making it impossible to completely replace antibiotics should a disastrous fall in antibiotic efficacy occur in the next few years. It is therefore crucial to develop other approaches, which notably include close surveillance. For example, researchers from South Africa recently reported surveyal of tuberculosis (TB)-causing mycobacteria from wastewater **(8)**. Indeed, environmental surveillance could serve as a tool for overcoming the underestimation/reporting of TB cases in resource-poor sub-Saharan countries where monitoring of drug-resistant TB (DR-TB) is a major challenge because assays are costly and time-consuming, and laboratories might be ill-equipped. Such missing information greatly compromises outbreak control actions of public health officials who must know exactly what they're facing. DR-TB is only one example of the antimicrobial resistance crisis, a silent, creeping pandemic that endangers the core achievement of modern medicine as antimicrobials lose their effectiveness against common infections[6]. To ensure that these medications will continue to save lives, we need large-scale surveillance systems running in the background to track this silent pandemic as it evolves. Especially for the surveillance of AMR trends over time and space, genomic surveillance based on high-throughput sequencing of environmental metagenomes offers very high value, which has moreover already been tested at a global level **(9)**. Such environmental surveillance programs can thus play a central role in providing currently lacking information of AMR spread, a key element for fact-based awareness raising and response organisation.

State-of-the-art genomic surveillance capacities in sub-Saharan countries
As reviewed by Sims & Kasprzyk-Hordern **(10)**, techniques used for infectious disease surveillance include sentinel surveillance, clinical-based surveillance, questionnaires or surveys, search engine search, mortality and morbidity rates, hospital admission data, prescription rates, human biomonitoring, and wastewater-based epidemiology. A surveillance system needs to have comprehensive data collection capabilities regarding the emergence of new diseases, the re-emergence of old diseases, the threat of imported diseases or pathogens, and the emergence of multidrug or pan-drug resistant organisms. We claim that all this can be achieved by incorporating comprehensive genomic surveillance into environmental surveillance programs with sufficient national coverage.

Continuous advances in high-throughput sequencing technologies have resulted in decreasing costs of sequencing, which can, in turn, revolutionise communicable disease surveillance. While the capacity for producing metagenomics sequence libraries is no longer a challenge, the shortage of bioinformatics expertise to unlock the wealth of acquired information on complex microbial communities in the samples has become a key bottleneck. As metagenomics can provide potentially key information on novel pathogens as well as re-emerging infectious diseases, and on AMR, this highlights the crucial importance of capacity building efforts in bioinformatics processing of the produced metagenome libraries.

Due to the weak health system and inadequate water and sanitation systems, Africa remains highly vulnerable to disease outbreaks that cause severe morbidity, mortality and major economic losses. This highlights the crucial need to strengthen and support the preparedness and response capacity of sub-Saharan countries against infection outbreaks. Importantly, the ongoing COVID-19 pandemic has necessitated the need to establish and reinforce the integration of routine pathogen genomics in health systems in Africa to allow timely detection and response to disease outbreaks, track circulating pathogen variants, keep track of antimicrobial resistance (AMR), and inform control and management strategies. The World Health Organization (WHO) has therefore provided support to laboratory networks within African countries to establish and strengthen capacity for genomic pathogen surveillance, leveraging advanced sequencing technologies to support rapid detection and tracing of circulating SARS-COV2 variants. This support has included technical and financial aids to regional laboratories to sequence at least 5% of collected patient samples at regional reference laboratories.

The Pathogen Genomic Institute (PGI) of the Africa Centres for Disease Control and Prevention (Africa CDC) has set targets to enhance genomic sequencing capacity in all five geographic regions of the African Union (AU), representing all countries on the continent. PGI´s capacity building objectives are to: 1) operationalize the network and support the sequencing of 50,000 SARS-CoV-2 genomes; 2) support AU member states with limited or no sequencing capacity through sample referral; 3) support at least 35 countries and 20,000 samples; and 4) conduct hands-on trainings for at least 100 candidates in support of SARS-CoV-2 sequencing by member states. Although

---

[5] COVIDPoops19 Dashboard | covid19wbec.org
[6] https://www.cdc.gov/drugresistance/biggest-threats.html; https://www.afro.who.int/ResistAMR

these targets have not yet been achieved for financial, technical, and infrastructure reasons, the overall amount of sequencing data in the region has increased and has contributed substantially in shaping the global health response to the pandemic[7]. Despite NGS capacity and technology being sparse and limited in most African countries (exceptions include South Africa, Kenya, Nigeria, and The Gambia), the generated genomic surveillance has nevertheless been critical during the early detection of multiple SARS-CoV-2 lineages in Africa **(11)**. As an example of genomic surveillance capacity, it recently showed how three lineages of SARS-COV2 Omicron variants of concern (VOC), drive the fourth wave of COVID-19 in South Africa **(12)**. Overall, SARS-COV2 genomics data from the region has contributed to the detection of new VOC and (sub)-lineages with multiple mutations at key sites in the spike protein, which decrease antibody neutralisation, thus compromising vaccine efficacy. Increased genomics capacity and its integration into the routine public health response in the African continent is thus critical to support targeted and evidence-based public health interventions; not only for the COVID-19 pandemic response but reaching well beyond to tackle the threats of multiple communicable diseases. As phrased by Dr Nicksy Gumede-Moeletsi, Senior Virologist for Polio Eradication at WHO's Regional Office for Africa in Brazzaville, it remains imperative that "**Routine genetic surveillance should be a part of health systems in Africa**".

Challenges in water and sanitation systems in sub-Saharan countries
Similar to current surveillance system capacity, the capacity to cope with rapid population growth and environmental changes also remains limited in sub-Saharan countries. Emerging population growth and economic water scarcity in urban areas forces people to use untreated wastewater in urban agriculture **(13, 14)**. Moreover, as WWTPs in Sub-Saharan Africa are not able to respond to the increasing population growth, most collected wastewater eventually ends up in the environment untreated or only partially treated. This creates a serious health risk through exposure to waterborne pathogens and antimicrobial resistance determinants and emphasises the urgency of wastewater-based pathogen surveillance. Irrigation water is collected from canals and open street gutters that were originally intended to prevent floods in the wet season **(15)**. Poorly managed or performing pit latrines that may be the only option for excreta disposal, especially in low-income rural and urban regions, pose a risk for disease transmission **(16)**.

The hygienic quality of drinking water is of paramount importance in the prevention of communicable diseases. However, seasonal water scarcity and deficiencies in sanitation might lead to a situation where water of impaired quality has to be used for domestic purposes. The intrusion of human excreta is considered as the most dangerous threat to drinking water quality. In terms of clean water infrastructure, drinking water distribution systems and sewerage systems in rural areas differ greatly compared to urban areas. For drinking purposes, showering and for irrigation of plants, inhabitants of rural areas mostly rely on well water, rivers, streams and other surface waters that often are consumed untreated. Rurality and seasonal variation in water availability are operating as the determining factors in water consumption habits **(17)**. Therefore, surveillance-based sampling of aforementioned water sources could lead to the origin of contamination. Furthermore, the provision of safe drinking-water supply and sanitation requires the use of risk assessment and management frameworks such as Water Safety Planning (WSP) and Sanitation Safety Planning (SSP). WSP, recommended by the WHO Guidelines for drinking-water quality, encompasses all steps in a drinking water supply system from catchment to consumer. SSP supports the implementation of WHO's Guidelines on sanitation and health and Guidelines for the safe use of wastewater, excreta and greywater. SSP provides a framework to protect human health from sanitation-related risks, including from reuse of wastewater in agriculture and aquaculture.

Research methods utilised in ODIN WPs
The ability to rapidly monitor the spread of diseases is key for prevention, intervention and control of the outbreaks of communicable diseases **(10)**. Environmental surveillance is a new epidemiology tool that has potential to act as a complementary approach for current infectious disease surveillance systems and as an early warning system for disease outbreaks. Environmental surveillance, through the analysis of population pooled wastewater and other environmental sample matrices that contain human excreta, can monitor infectious disease and resistance spread comprehensively and in real-time, tracking the emergence of new disease outbreaks to the community level. While environmental testing cannot reliably and accurately predict the number of infected individuals in a community, and while community-level wastewater surveillance at a treatment plant will not capture homes on a septic system or communities and facilities served by decentralised systems, there is an abundance of clinically relevant information to be found in waters contaminated with human excreta that allows wastewater testing to catch outbreaks that the health care system misses **(18)**. Throughout the ODIN work packages, we aim to utilise this new approach of comprehensive genomic surveillance of environmental samples under anthropogenic impact to provide such comprehensive health information on communities.

Traditional wastewater-based epidemiology (WBE) is based on centralised, water-based sanitation systems and

---

excludes wastewaters from shared on-site sanitation facilities, solid waste including faecal sludge from non-flushing on-site sanitation systems, raw/untreated water, and drinking water supply systems in low-income countries (LICs). Wastewater in LICs, if not recycled for plant irrigation, is stored in septic tanks or similar on-site sanitation systems. ODIN will apply the comprehensive approach of Wastewater, Waste, and Water-Based Epidemiology (WWW-BE) proposed by Gwenzi **(19)**. Figure 2 illustrates potential sampling locations for environmental surveillance of communicable diseases in sub-Saharan low-resource settings.

Figure 2. Examples showcasing potential environmental sampling locations in sub-Saharan Africa for surveillance of multiple human pathogens. Sample matrices known to harbour human waste include e.g. wastewaters of hospital building septic tanks, influents and effluents of the wastewater treatment plants, surface waters and even groundwater sources contaminated with human faeces.

A total of six study locations will serve as example cases for the sub-Saharan region (Table 1). ODIN will facilitate and help in setting up the environmental communicable disease surveillance programs and the subsequent genomic epidemiology investigations at these cities. In each participating sub-Saharan country, one large city and one smaller location will be included. The intention is that the experiences gained at these study locations can be later employed in other areas.

Table 1. ODIN study locations and the local actors in charge of environmental sampling and microbiological investigations in Tanzania (TZ), DR Congo (DRC) and Burkina Faso (BF).

Study location	Population estimate	Laboratory	Environmental sampling collaborator	Stakeholders in public health sector
Dar es Salaam, TZ	4,364,541	NIMR (partner 4)	Ministry of Water, Water Institute; www.waterinstitute.ac.tz	National Public Health Laboratory, Ministry of Health
Tanga,TZ	273,332			
Kinshasa, DRC	17,000,000	UNIKIN (partner 5)	Ministry of Health	Ministry of Health
Kisenso, DRC	386,151			
Ouagadougou, BF	3,055,700	IRSS-DRCO (partner 6)	Ministry of Health and Public Hygiene; https://www.insp.gov.bf/	Ministry of Health and Public Hygiene; https://www.insp.gov.bf/
Nanoro, BF	188,837			

ODIN will base its capacity building activities upon existing science, technology and innovation contributions, many of these raised as part of the global response to combat COVID-19 disease. In particular, considerable progress has been made on the systematic development of environmental and genomic surveillance programs for communicable diseases, and the corresponding experiences and lessons learnt will be available through the extensive ODIN partnership and collaborator network. The case studies from African countries discussed above and to be utilised here include e.g. establishment of environmental surveillance, and SACCESS network in sewered and non-sewered communities of South Africa **(20)**. ODIN communication efforts will be aimed at transparency and effectiveness to fully convey the strength of the environmental surveillance method to those health authorities and stakeholders unfamiliar with this type of monitoring of communicable diseases. Once in place, surveillance of local communities can be used to direct community screening and alert medical authorities to potential increases in patient numbers.
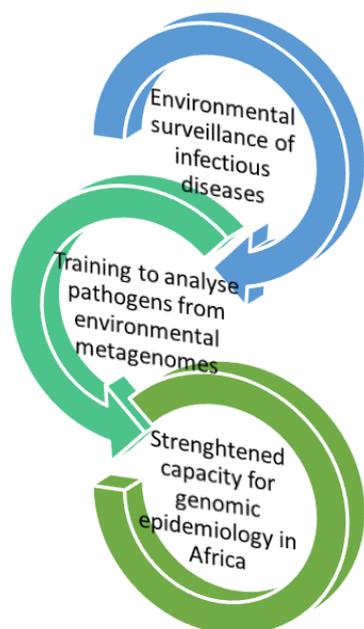
Chart 1. The ODIN concept.

The ODIN consortium partners in the African node and thus the study locations for environmental monitoring and genomic surveillance are located in three out of five Regional Collaborating Centres (RCC) of Africa CDC: Tanzania in Eastern Africa RCC, Democratic Republic of Congo in Central Africa RCC, and Burkina Faso in Western Africa RCC. The planned development of environmental monitoring-based genomic surveillance in these three sub-Saharan countries by African researchers in collaboration with the sequencing laboratories of their National Public Health Institutes (NPHIs) serves as a basis for implementation of the ODIN results in the greater African region, in the European partnering countries, and even globally as the dissemination actions of the project will be open to anyone. As expected outcomes towards better control of poverty-related and emerging/re-emerging infections in sub-Saharan Africa, ODIN will contribute to the following major development needs:

● establishment of genome-based, non-centralized, sewage-environment pathogen surveillance that produces actionable information about the health burden of clinically relevant pathogens and AMR determinants; and

● building local capacity on processing and easy visualisation and dissemination of the acquired high-throughput sequencing results to provide useful and timely genome-based pathogen surveillance information to the health authorities and policy makers.

National and international R&I activities linked with project ODIN

The proposed framework will strengthen national, regional and international capacity in pathogen genomics through linkage with the African Centre for Disease Control (ACDC)/Pathogen Genomic Institute and Pan-African based research networks. At a national level, Public Health Laboratories (NPHL) designated by the Ministries of Health (MoH) in Tanzania (TZ), DR Congo (DRC) and Burkina Faso (BF) will be the immediate beneficiaries from the project through their engagement in laboratory training, and ICT infrastructure development to support computation capacity. Through the partnership with NPHL, the project will enhance the epidemic preparedness and response of the established East African Network mobile BSL3/4 laboratories. The integration of genomics in the mobile NPLH laboratories will increase the capacity for multi-pathogen detection, AMR surveillance, and outbreak investigation to rapidly respond to epidemics and future pandemics. This capacity will support regional efforts to incorporate genomics tools and training during cross-border simulation exercises. The direct partnership with the MoH through NPHL in the respective countries (Tanzania, DR Congo and Burkina Faso) will facilitate sustainable capacity building and integration of genomics data into public health systems for policy impact. Also, the project will feed into the National AMR Coordinating Committees and Technical Working Groups (TWGs) through establishment of population AMR surveillance platforms in the respective countries.

To achieve this, we will implement a One Health Approach (OH) to established communities of practice (COP) in TZ, DRC and BF. The OH COP will consist of multi-disciplinary teams of experts from environmental, animal and public health, including experts in genomics, computational/data science, environmental and clinical epidemiologists, social sciences, as well as water and sanitation experts from respective ministries to support translation through technology transfer, and knowledge sharing focused on genomic surveillance of infectious diseases and AMR. Through the OH COP we will support knowledge sharing, best practices, standardised protocols for WWW-BE, computational tools and training materials beyond public health use. The ODIN project will feed into the Africa CDC through the Pathogen Genomics Institute (PGI) and other established regional and international consortia on pathogen genomics and epidemiology, including the Malaria Genetic Epidemiology Network (Dr Vito Baraka is a member), and the EDCTP Alumni network (Dr Marc Tahita-IRSS, Prof Vivi Maketa-UNIKIN and Dr Vito Baraka-NIMR are members).

Moreover, our project will engage other regional networks through the establishment of a Scientific Advisory Board (SAB) to ensure that our project creates synergies and complementarity with Pan-African and international initiatives that support genomics. SAB members will be identified from NPHL/MoH in respective countries (DRC, Tanzania and Burkina Faso), the Africa Centre for Disease Control/PGI, the Pan-African Bioinformatics Network, Human Heredity and Health in Africa (H3Africa), the Genomic Medicine Training Initiative (AGMT), the African Society for Laboratory Medicine (ASLM), SAMRC Wastewater Surveillance and Research Programme/SUCCESS Project as well as Water Research Commission (WRC) in South Africa. These representative members will officially be engaged at project start. As one member of SAB, Dr Eunice Ubomba-Jaswa, Research Manager at the WRC, will support the project outreach and networking with the thematic areas of source water pollution and protection (including both microbial and chemical emerging contaminants), water-related human health and WASH activities.

At the global level, one the ODIN partners is The Global Health Network (Trudie Lang, Professor of Global Health Research, University of Oxford) that will support a platform for transferring skills, sharing of standardised protocols for genomics surveillance as well as WWE and providing know-how to enable application of genomics in public health systems. TGHN is a vast global community of practice, working to support research where evidence and skills are lacking. This facility has contributed substantially in supporting research during previous (Ebola, Zika, and recently COVID-19) outbreaks in the African region, and works to connect excellence and transfer research know-how across diseases and between countries and organisations.

The Global Health Network is working with several pathogen genomics research networks across the Global South and so working with ODIN will enable further connecting of excellence in this field and ensuring not only that methods and processes are shared across teams working on different pathogens, in different regions and settings but also that genomics teams are connected up with policy makers and the wider ecosystem of health researchers in order that maximum benefit is made from the findings, and also the methods and approaches used.

Integration of expertise/methods from different disciplines to reach success with ODIN
The specific composition and expertise of the different organisations participating in ODIN is described in detail in section 3.2. Briefly, our project calls for the need of individuals highly experienced within molecular biology, sequencing, bioinformatics, teaching, social sciences, medical assessments, engineering, epidemiology, and organisation. On top of these skills, particular knowledge within microbes, water, LMICs and environmental surveillance is permeating the organisation. As can be particularly seen in the PERT chart, all WPs communicate with each other, indicating the true interdisciplinary character of our project, and the need to bring together a consortium with these skill sets in order to have a profound impact on the large and complex societal and medical problem posed by infections. It can be clear that, while no individual research organisation could complete this project, as a consortium we bring together the required unique and complementary knowledge and expertise that will guarantee the success of our project.

Social sciences and Humanities within project ODIN
The WWW-BE approach produces community level information of the spread of infectious disease, where infected individuals are not identified, and is thus considered free of social stigma related to e.g. person-to-person transmission routes of illnesses. One part of ODIN relates to understanding human behavioural and structural social aspects at community levels related to the intervention we intend to put in place. More generally, we will assess local communities' acceptance of the proposed interventions (surveillance of environmental samples and wastewater), and its feasibility and adequacy in the local context. In addition, we will identify potential bottlenecks for implementation at community level.

Gender dimension within the project
This project aims to strengthen capacity and develop a genomic surveillance system of both wastewater and clean water for the timely detection and prevention of epidemics. This is because epidemics significantly increase health condition deterioration and, in concert with poor economic growth, have had a negative impact on the entire population, with greater impact on vulnerable people and women. In sub-Saharan settings, women are known to be on the front lines of their families' and communities' well-being and resilience. As women often are the primary caregivers in the home and have greater domestic responsibilities, capacitating them will result in a greater impact. However, fathers are often the main providers of income, paying for medicines and school expenses. Therefore, the fast recovery of every family member is important.

In our project, we therefore integrate sociobehavioral studies into the lab work, focusing on understanding the pathway of contaminated water amongst communities in order to develop a clear approach to prevent these diseases. Women are known to be the cornerstone of many families due the role and leadership they are bringing in these settings, so capacitating them will result in a great impact. Our proposal will address a critical maternal/women's health issue in Sub-Saharan Africa, i.e., ensuring access to appropriate preventive and curative health technologies and services in epidemics. Although epidemics most likely equally affect both boys and girls, their household management may be different. In addition to being at potential risk of contracting cholera or other epidemic related diseases, women in sub-Saharan countries are most often the only or the most important caretakers of sick children and other sick family members. This puts an additional burden on their daily routines and may interfere with occupational duties. Appropriate diagnostic approaches will speed up treatment and subsequent recovery of patients and alleviates women's care tasks. Furthermore, the outcomes of this research are directly relevant to health programs throughout sub-Saharan Africa and will thus directly benefit women's health. Preventing epidemics through earlier identification of suspected cases in the community and referring them to health centres are key in this proposal and will greatly reduce the many risks associated with an expansion of the epidemic.

Open Science Practices

The guiding principle of ODIN is the availability of all data and code to interested parties as early as possible. ODIN will therefore ensure semi-automation of data processing and dissemination with a strong emphasis on the development of bioinformatic pipelines. ODIN will develop a new interactive dashboard, where all collected data will be available in real time (data streaming). The goal of ODIN is to serve society with providing reliable information sources on time. ODIN will move beyond open access in implementing open science practices while adhering to Findability, Accessibility, Interoperability and Reusability (FAIR) principles. The training programs in ODIN will focus on the improvement of methods leading to better open science practices, and students will be familiarised with the FAIR principles. By default ODIN manuscripts are published in open access journals. If the publication forum is not fully open, the organisation of the corresponding author will pay the extra fees for gold, or at minimum, green level open access. Authors will make pre-prints available via repositories such as Zenodo and BioRxiv, and publishing in journals offering an open peer review process is recommended.

ODIN will have a knowledge hub on TGHN and this will function as an open working space for our network, but also connect us and allow sharing of methods and tools between us and other networks. This is a significant opportunity for driving open science as we will be enabling other groups to use our methods, processes and training. Knowledge hubs such as this have worked well for many years for the EDCTP networks of Excellence, the EU funded Zika consortia and the ALERRT networks, to name just a few of the 63 component and connected knowledge hubs on TGHN.

Research Data Management and Management of other Research Outputs

A dedicated Data Management Team (DMT) consisting of one participant from each work package will be established, whose responsibility will be to create the detailed data management plan within the first 6 months of the project and to periodically review the plan. This team will ensure that our data management policy follows open science practices according to the European Commission's requirements and that ODIN follows the FAIR principles. ODIN will generate 1) raw data in various formats depending on the analysis type (e.g., text, CSV, fatsq, fast5), 2) metadata, 3) processed data to be shared on the ODIN platform (will be developed during the project, WP4), 4) processing code and bioinformatic pipelines (e.g., R scripts, Jupyter notebooks), and 5) end-products (e.g., handbooks, training material, journal articles). Within WPs 2 and 5 we will also re-use data from already existing clinical surveillance systems in sub-Saharan countries. Metadata originating from different geographical locations and laboratories will be collected in a harmonised manner according to established standards, e.g., Minimum Information about any (x) Sequence (MIxS) and Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) Guidelines. All acquired data will be deposited in discipline-specific, community-recognized repositories as early as possible (e.g., Sequence Read Archive, SRA) enabling their findability and long-term storage. Otherwise, generalist repositories will be used (e.g., Zenodo). Developed protocols and applied workflows will also be made freely available through public platforms (e.g., bio-protocols, protocols.io). Emphasis will be placed on reproducibility of research outputs via the publication and sharing of source and analytic code (e.g., using GitHub) under permissive open-source licenses (e.g., Apache2 or MIT). Besides adopting open data management software tools, ODIN will contribute to the development of any needed extensions and will share these as open source via public repositories (e.g., GitHub). Data management software tools developed within our project will also be made freely available under permissive open source licenses. Moreover, all project reports and technical notes will be freely web-accessible (project dashboard and on OpenAIRE document repositories). Enabling better use of data science in LMICs is a core goal of TGHN and therefore ODIN will connect with Global Health Data Science and work with other pathogen genomic projects to share approaches, tools and overcome challenges in the component elements of strong data science in genomics. This will leverage Gates funding to support this objective.

2. Impact

2.1 Project's pathways towards impact

Due to the weak healthcare system and the high burden of infectious diseases, healthcare systems in Africa remain vulnerable to, and ill-prepared for the management of, epidemic and pandemic situations. The overall impact of the ODIN project is to reduce mortality and morbidity to help bring about poverty reduction and economic development on the continent. The ODIN project will impact livelihood by strengthening the public health response to disease outbreaks through flexible surveillance systems. Indeed, there is an urgent need to strengthen resilience and preparedness to public health emergencies of international concern (PHEICs), and to outbreaks of public health relevance (e.g., vaccine preventable diseases) caused by known and unknown pathogens and AMR. The recent emergence and spread of the COVID-19 pandemic has revealed serious capacity gaps to detect and timely respond to outbreaks. To address these challenges, there is a crucial need for development and integration of innovative technologies and approaches to improve local and regional capacity for outbreak investigation, pathogen detection

and appropriate interventions. Advances in pathogen genomic technologies have demonstrated potential to enhance disease surveillance and outbreak investigation, which ODIN aims to leverage to improve outbreaks investigation and response capacity in regions through their integration into public health systems.

The ODIN project will achieve this overall impact by focusing on four targets:

1. Enhanced capacity for pathogen genomics and workforce. The ODIN project, in partnership with MoH/NPHL and research networks in the respective African countries, will contribute to the capacity strengthening of the local laboratories to implement pathogen sequencing capacity (wet lab) and improve the computational skills (dry lab) of key staff at NPHL and the Institutes for epidemic preparedness and response. ODIN will identify trainers of trainees (ToT) from the NPHL and institutes to implement hands-on laboratory training (library preparation, and sequencing), IT/ bioinformatics tools and training, provide technical support and standardised protocols. In addition, our project will establish links between the NPLH, research institutions, and networks to enable knowledge and technology exchange on metagenomes and targeted sequencing of pathogens such as *Vibrio cholerae* and multidrug resistant *E. coli* as case studies. Furthermore, in recognition of the need for bioinformatics skills, the project will ensure the NPHL are installed with the necessary ICT infrastructure to support computational analysis and data visualisation using the ODIN-developed tools. This established capacity can be leveraged to support multi-pathogen sequencing, data analysis, and results interpretation to enable real-time sharing for public health interventions and/or policies;

2. Improved population-based monitoring of Antimicrobial Resistance and interventions. Antimicrobial resistance (AMR) is a serious public health concern and recent evidence suggests that it is escalating in the sub-Saharan region. However, surveillance data on AMR remains limited and is often based on passive reporting of phenotypic laboratory results for specific pathogens isolated from human clinical infection. Currently, the lack of evidence on the dissemination of bacterial resistome profiles, abundance and diversity of AMR genes in most African settings is due to limited genomics capacity. However, this knowledge is crucial to predict and limit the environmental spread of AMR through sanitation systems to reduce the local and global burden of AMR. In the context of the One Health Approach, the consortium will therefore complement existing national, regional and global AMR surveillance by integrating environmental/population-based metagenomic analysis for AMR surveillance into public health systems;

3. Development of low-cost, field-adaptable molecular tools for pathogen surveillance in remote settings: The ODIN project will implement innovative research to support development and optimization of field deployable, low cost and user-friendly molecular tools that can be used in remote settings with limited resources. This will increase access to these surveillance tools at research institutes and the NPHLs, and will improve national and regional epidemic preparedness and response capacity. These tools will be based on the portable Nanopore Oxford Sequencing and Biomeme Franklin™ qPCR devices, allowing multi -target and targeted diagnostic approaches for environmental samples derived from human surroundings (WWW-BE case studies for this project). The development of these tools in combination with complete bioinformatics pipelines, semi-automatic analysis, and the interactive visualisation of pathogen detection enhanced with metadata presentation will further enhance the existing mobile laboratory infrastructure in terms of flexibility to use in remote settings, reduced cost of pathogen detection, and enables the fast dissemination of results. Successful development and evaluation of these tools will impact local and regional disease control programmes and public health networks such as the Africa CDC/PGI. Importantly, the customisable nature of this technology will also allow for the detection of other pathogens of public health relevance, even well beyond the broad scope of ODIN.

4. Capacity building through short and long-term training: In recognition of the need to improve computational capacity in the sub-Saharan region, ODIN will implement a capacity development program that includes short-term and long-term training. Our project aims to fully support six (tentatively three female and three male) Masters in Bioinformatics (two per country) at recognized Universities in Africa, career development opportunities for one Postdoctoral fellow, and organise shorter training and workshops for others. The selected students will have opportunities for 2 to 3 months academic visits at a partner institution in Europe. The MSc students will be linked with regional networks on pathogen genomics and bioinformatics such as the Africa CDC/PGI, Global Health Network for career development opportunities and training. ODIN capacity building will translate to an increased workforce with bioinformatics skills and the required competence to support pathogen genomics in the region for long term sustainability. Further, the extensive ODIN on-line training material will be made freely available for all.

2.2   Measures to maximise impact - Dissemination, exploitation and communication

Africa's health systems are still weak and vulnerable to frequent outbreaks of pathogens and face increasing concerns about antimicrobial resistance (AMR). Therefore, strengthening epidemic preparedness and response capacity is needed to improve regional and global health security, and to protect vulnerable populations. Development and integration of innovative solutions is critical for resilient health systems in the sub-Saharan region region. Advances in genomics technologies and integration into the public health response have demonstrated great potential for

efficient multi-pathogen disease and AMR surveillance. Below we outline our strategy for dissemination, exploitation and communication to maximise the impact of ODIN.

Dissemination of data

Within ODIN we will generate many types of data that will be suitable for dissemination to different targeted institutes / organisations / individuals. All PIs in the project will be responsible for dissemination of data, but the overall strategy for dissemination will be through one of the ODIN partners, The Global Health Network (TGHN), which has vast experience and an established network in that area. Here our knowledge hub will be continuously sharing our outputs in the form of recommendations from our findings, our methods, tools, processes and data. TGHN can issue DOI numbers and so everything we release, such as policy guidance documents, or SOPs, protocols and even training videos will be citable and trackable and so this will also allow us to see where these have been taken up and used by other groups or where recommendations from our research are being accessed and used. Success in dissemination will also be measured through publication in high-tier journals, presentations at conferences, continued funding support to the project's goals, and increased interest in ODIN's goals. We will use the following strategy to maximise project dissemination:

1. *Dissemination to local stakeholders and Ministries of Health*
   Communication with stakeholder's communities and local health management teams, NPHL/Ministries of Health and their technical and implementing partners will be through meetings and local forum presentations at village, district, regional, county and national levels. Research results will be communicated through dissemination meetings to key national level stakeholders and a multidisciplinary technical committee on Emergency preparedness, control programmes, and One Health Community of Practice. Policy documents will be prepared using applicable local language (English in TZ, and French in DRC and BF) targeting country policy makers to facilitate policy adoption and subsequent implementation, as appropriate.

2. *Academic and Scientific beneficiaries*
   Communication of the research outcomes to academic and scientific audiences will be achieved through open-access, peer-reviewed journal publications, articles in science review magazines such as International Innovation, symposia, conferences, and the dissemination workshops to multidisciplinary technical committees on Epidemic Preparedness. Publications will also be disseminated via institutional websites of partners, and to all ODIN partners by e-mail and through the project social media channels. The research will be disseminated also within each of the participating institutions' annual dissemination seminars.

3. *Regional and Global partners*
   ODIN results will be communicated to WHO-AFRO, Africa CDC/PGI, relevant networks (EDCTP networks such as EACCR/CANTAM/WANETAM, and ALLERT) and at international scientific conferences and Symposia for more widespread translation and policy consideration. These will also be shared across all the knowledge hubs on TGHN which can connect these and make them discoverable.

Exploitation of data

During the project a lot of data will be generated, including raw data (sequences), but also reports, articles, methods, scripts, and innovative tools. The consortium will adapt to an, in principle, open access-based ideology, where all data should be available for all participants, as well as outside the consortium. This does by no means indicate that results cannot be exploited, but ensures an efficient outreach strategy:

1. All scientific reports released from the consortium will be open access, allowing everyone to continue building upon our results and implement these elsewhere.
2. All courses and materials (workshops etc) made available will remain freely available on TGHN website, for facilitated knowledge transfer.
3. We will construct a Dashboard, showcasing data from the WWW-BE in real-time, allowing key stakeholders to exploit these data points and take actions (e.g. renovate water supplies due to contamination, initiate vaccine programs, etc).
4. All generated SoPs, methods, etc will be freely available on our website (TGHN), enabling others to exploit these for their own implementation. TGHN is thus a guarantee for the sustainability of these methods, as they will remain on their site even after project end. Each will be issued a DOI number for tracking impact and citations.
5. Certain generated scripts and tools may be considered of extra commercial interest and may then be IP protected and exploited in a commercial way. This commercialisation should, however, not interfere with the basic idea of facilitating the implementation of our systems globally to reduce illness and death due to lack in surveillance capacity, but rather act as a facilitator to enhance such globally implemented systems.

Communication of data

As we are targeting to mitigate a globally relevant research and societal problem, it is imperative that not only stakeholders and researchers have access to the information, but that it reaches the general public as well. We have

significant expertise and experience in data visualisation to generate informative materials that are commonly understood by a broad audience. For project sustainability, it is also imperative that we engage the public, and raise their awareness of, and interest in, our project, to elicit general acceptance of the concept we want to implement. It is thus not enough to have governments and institutes to approve of the project, as we need to build a high public acceptance rate through information, engagements, and various communication channels. Communication success will be measured through small digital surveys on the national level to investigate changes in awareness before and after our communication rounds. Specifically, we will target our audience in the following ways:

1. News articles in regular daily journals describing our concept as well as the results obtained. This will be conducted around halftime as well as when the project is reaching its end. TGHN is administering contacts with local journals and will facilitate this aspect.
2. Popular science journals will also be targeted for a more detailed description of our project, and will reach a slightly different audience compared to the daily newspaper. Similarly, we will target these journals in several continents, during halftime as well as at the end of the project. Target journals will be "*Scientific American*", "*National Geographic*", "*Discover Magazine*", "*The Scientist*" and national popular science journals. The project coordinator will initiate journal contacts, assisted by TGHN.
3. We will further strive to communicate our concept and results in different media, including radio and TV, to target a broad audience not necessarily interested in reading specific journals. All PIs are encouraged to mediate such contacts.
4. Finally, we will launch a website that will convey information from our surveillance results in close to real-time. Such information was highly sought during the COVID-19 pandemic by the general audience, where they can engage with the data and follow changes in prevalence of different microorganisms at different sites. Through all these means we intend to increase the awareness of the general audience, while also building interest and enthusiasm for the project, thus creating an educated audience that will put demands on key stakeholders to implement changes for better surveillance in other regions.

Intellectual Property strategy

Patents protecting innovations can serve several means: both allowing for the commercialisation of a product with limited direct competition, or in order to stop others from releasing products within certain areas. However, for certain innovations it may be more beneficial not to patent, which describes the detailed procedure, and rather maintain the innovation as a trade secret. The overall goal with all of these strategies is to gain commercial advantages, expand the reach of your company, and increase your revenues. Patenting may however also be detrimental in some instances, hampering the uptake of a product or an idea due to lack of possibilities to use it and/or commercial interests aggravating the uptake in LMIC or poor regions. The main purpose of the project is to develop a surveillance method that can be implemented in regions of different economic status, both in Africa and elsewhere. As a research community we are also prone to share our findings through e.g. open access and facilitate the sharing of ideas and results. A commercial interest may benefit the production and scaling of some of the tools developed, but may hamper the implementation of our discoveries if added to other aspects of the project. This is thus a delicate question that we will approach carefully:

1. Protocols, SOPs, articles, and general reports from the study will all be freely available on the ODIN website as well as made available through open access publications. This will provide access to all of our detailed protocols and operating procedures to other researchers and stakeholders that may want to implement similar systems and exploit the findings. The data will be generated by all researchers within the consortium and will be treated as non-confidential.

2. Developed mobile solutions may have a value to protect, in order to facilitate their production by a commercially interested partner and thus exploitation. Such discussions will be handled by ODIN WP3 lead partner NORCE during the development of the tools.

3. Software code for analysing data will be open source; this to increase the global implementation of these methods and thus achieve high societal impact. Certain aspects of data visualisation may be considered protectable and this will be evaluated on a project basis. Dr. Mesuere (UGENT), leading the visualisation part of the project as well as being highly involved in the programming, will advise on the possibilities throughout the project - with the main consideration that the project aims at disseminating our data as well as processes to as broad an audience as possible. If that can be facilitated through protection of key elements, we will take such measures, but our default strategy will be to opt for open source licences for all generated code.

4. Before initiating the project, all PIs in the project will sign a non-disclosure agreement (NDA) detailing processes for joint ownership, royalty agreements, licensing agreements, general handling of confidential material within and outside of the consortium, as well as research data ownership etc. The NDA will be distributed by the coordinator through Lund University Legal Department. Most institutes participating in the consortium have access to free legal advisory boards through their respective organisations, and many also have commercialisation units at the universities, facilitating informed discussions regarding IP.

2.3 Summary
KEY ELEMENT OF THE IMPACT SECTION

SPECIFIC NEEDS

What are the specific needs that triggered this project?

Africa's health systems are still weak and vulnerable to frequent outbreaks of pathogens and antimicrobial resistance (AMR) is an increasing concern. Therefore, strengthening epidemic preparedness and response capacity is needed to improve regional and global health security, and protecting vulnerable populations.

D & E & C MEASURES

What dissemination, exploitation and communication measures will you apply to the results?

Dissemination towards the scientific community and industry: Participating in conferences; arranging workshops directed towards key stakeholders; writing peer-reviewed articles.

Exploitation: Patenting mobile low-cost solutions to facilitate long-term interest in investments in production thereof. Considering patenting developed scripts and visualisation technologies.

Communication towards citizens: News articles in national and international journals describing the ODIN research conclusions, as well as through open seminars to the general audience. Through the Global Health Network we will have several websites with accessible information for the general audience.

EXPECTED RESULTS

What do you expect to generate by the end of the project?

Environmental genomics data and its integration in public health systems generates **early warning signals, epidemic preparedness and response**.

Improved and innovative environmental surveillance to monitor the real pathogen and antimicrobial resistance trends also in remote settings.

Strengthened genomic epidemiology capacity for improved population-based pathogen and AMR surveillance to inform policy makers about the health burden.

Strengthened capacity in genomic and bioinformatics skills for epidemic preparedness and response.

Established **Community of Practice** and networks for genomic surveillance using One Health Approach.

TARGET GROUPS

Who will use or further up-take and benefit from the results of the project?

Communities vulnerable to infectious disease outbreaks and AMR in sub-Saharan Africa.

National Public Health Laboratories under the Ministries of Health and public health research institutions.

Regional research networks and authorities involved in coordination of genomics surveillance and capacity strengthening eg. Africa CDC/Pathogens Genomics Institute.

International organisations and stakeholders supporting epidemic preparedness and response capacity and technology exchange (WHO, and private companies e.g. ONT, Illumina).

OUTCOMES

What change do you expect to see after successful dissemination and exploitation of project results to the target group(s)?

Integration of environmental surveillance for genomic epidemiology in public health systems for infectious disease preparedness and response.

Enhanced genomic surveillance of AMR and its application into public health interventions and programs.

Improved capacity for NGS sequencing and bioinformatics, data handling infrastructure for genomic epidemiology to respond timely on outbreaks of pathogens in the region.

Reduced burden of disease after successful water and sanitation interventions following the **One Health Approach**.

IMPACTS

What are the expected wider scientific, economic and societal effects of the project contributing to the expected impacts?

Scientific: Development of highly adaptable mobile surveillance equipment, as well as adaptable scripts to facilitate visualisation of complex biological data.

Societal: Improved health system resilience and preparedness to respond to pathogens outbreaks, reduce mortality, promote well-being and accelerate poverty reduction (Sustainable Development Goals, SDG 3).

Improved regional and global health security through environmental surveillance of pathogen and AMR prevalence trends.

Impact on sustainability by skilled workforce and innovation to respond to infection outbreakson.

Economic: Cost-effective ways to limit illness and infection through genomic surveillance.

3. Quality and efficiency of the implementation

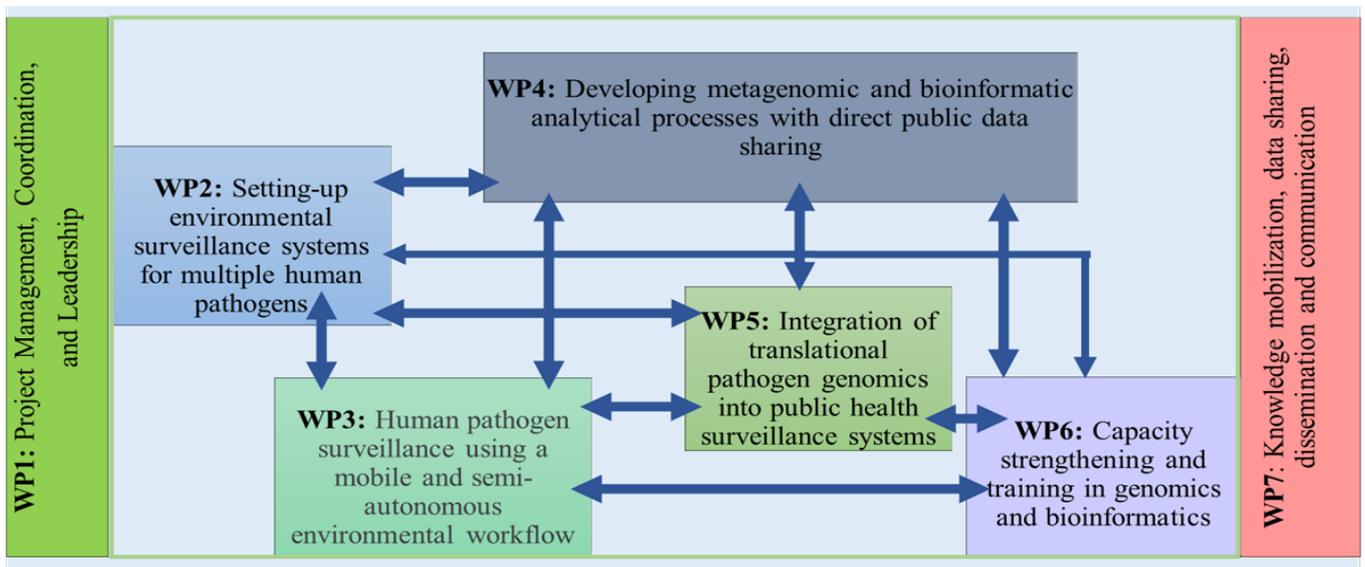## 3.1   Work plan and resources

The overall aim of the project is to build up a reliable, robust, and sensitive surveillance system that is based both on static and mobile surveillance of water. For this purpose, a surveillance system based on WBE and environmental monitoring will be set up to continuously monitor prevalence of different microbes in the sewage and non-sewage environments. Similar detection systems will be evaluated for clean water systems as well. To access remote areas, a mobile, low-cost system will be developed. The generated data will be processed through several analytical workflows / scripts to facilitate generation of reports that will be easily understandable even by laymen. The system will be validated in two translational studies, as well as through their cost-effectiveness and sustainability.

Table 2. Gantt chart for the project.

		2023			2024				2025			2026	
		0-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-30	31-33	34-36
WP1	Task 1.1 Organisation of the consortium		D1.1; D1.2			M1.1							
	Task 1.2 Internal meetings					D1.3				D1.3			D1.3
	Task 1.3 External meetings	M1.3				D1.3				D1.3			M1.2
	Task 1.4 Follow-up on progression of project					D1.4				D1.4			D1.4
	Task 1.5 Establishing Community of Practice, networking, and coordination												D1.5
WP2	Task 2.1 Gathering information on the existing clinical and environmental surveillance systems		M2.1	D2.1									
	Task 2.2 Developing a proposal for a cost-efficient environmental surveillance strategy					D2.2							
	Task 2.3 Preparing guidance and a handbook for microbiological investigations					D2.3							
	Task 2.4 Collecting environmental surveillance data				M2.2			M2.3		M2.4	D2.4		
WP3	Task 3.1 Workflow design and development of protocols for mobile and semi-autonomous						D3.1	M3.1					
	Task 3.2 Operationalization and cost- benefit analysis of mobile surveillance								M3.2; M3.3			D3.2	D3.3;M3.4
WP4	Task 4.1 Integrate with data generation processes					D4.1							
	Task 4.2 Initial data interpretation and quality control					D4.2; M4.1							
	Task 4.3 Centralised data collection								D4.3				
	Task 4.4 Data processing and pipeline							D4.4; M4.2					
	Task 4.5 Develop a dashboard for interactive data visualisation and analysis												D4.5; M4.3
	Task 4.6 Data dissemination												D4.6; M4.4
WP5	Task 5.1. Epidemiological links between presence of waterborne pathogens and clinical						M.5.1; D5.1; D5.2						
	Task 5.2. Translational investigation of species- and strain-dependent outbreaks of cholera								M5.2; D5.5				
	Task 5.3 Water and Sanitation Safety Plans			M5.3					M5.4		D5.3		
	Task 5.4. Cost effectiveness of the developed methods										D5.4	M5.5	
WP6	Task 6.1 Knowledge Mobilisation		M6.1		D6.1								
	Task 6.2 Regional Hubs		M6.2										
	Task 6.3 Teaching, training and career development		M6.3	D6.2	D6.3; M6.4								
	Task 6.6 Convening and connecting resources and excellence								D6.4				
WP7	Task 7.1. Develop Consortium communication, dissemination, and exploitation plan		M7.1	D7.1									
	Task 7.2. Organise project workshops, communication, and dissemination activities				D7.2				D7.2				D7.2
	Task 7.3. Development of the ODIN project webpage, and social media presence			M7.2	D7.3				D7.3				D7.3
	Task 7.4. Coordinate scientific papers and policy to ensure timely publication				D7.4;M7.3				D7.4;M7.3				D7.4;M7.3

Figure 2. Pert chart for the project.

To facilitate our project, one WP will focus solely on project management and coordination (WP1), and one WP (WP7) will focus solely on how these data and reports can be implemented into current systems, reaching key stakeholders, sharing data with public authorities, and building local capacity to run the developed systems. Our project is thus built upon four pillars: 1) research, 2) development and innovation, 3) capacity building, and 4) implementation, communication, and dissemination. Individual work packages are presented in detail below.

Table 3.1a: List of work packages

WP No	Work Package Title	Lead Partner No	Lead Partner	PMs	Start (M)	End (M)
1	Project Management, Coordination, and Leadership	1	LU	43	1	36
2	Setting-up environmental surveillance systems for multiple human pathogens	3	THL	149	1	30
3	Human pathogen surveillance using a mobile and semi-autonomous environmental workflow	2	NORCE	136	1	36
4	Developing metagenomic and bioinformatic analytical processes with interactive visualisation and sharing endpoints	7	VIB	140	1	36
5	Integration of translational pathogen genomics into public health surveillance systems	6	IRSS-DRCO	75	8	36
6	Capacity strengthening, knowledge mobilisation and training in genomics and bioinformatics	4	NIMR + TGHN	73	1	36
7	Dissemination, Communication and Exploitation of the Results	8	TGHN + NIMR	50	1	36

Table 3.1b: Work package descriptions

Work package number	1
Work package title	Project Management, Coordination, and Leadership
Lead Beneficiary	LU (1)

Objectives
- Project Management, Coordination and Leadership

Description of work

The work is administrative in nature, but entails coordination of resources, setting up meetings with external collaborators, submit reports to stakeholders, arrange internal and external meetings, and follow up on the progression of the project. Specifically, tasks 1.1-1.5 will be performed.

Task 1.1 Organisation of the consortium [M01-M36], Partners: all

The consortium is currently involving 10+ PIs according to the organisation structure below, with several more researchers involved in collecting data/material, performing experiments, analysing material, and writing reports. A Project Manager will continuously update the Organisation Chart so that all communications will be done in established routes, simplifying the coordination and reporting of progress within the projects. Briefly, all local hospitals and local government contacts will be directed by a country-specific PI, reporting all the activities within that country. The reports are delivered to one of the two nodes (Europe and Africa), and summarised by the project manager before given to the coordinator. For meetings of broad relevance, it is up to the discretion of the country-specific team leaders to involve the node coordinator or the project coordinator if need be. A Scientific Advisory Board (SAB) will support the project fulfilment and ensure true connections to the other relevant research activities, major stakeholders, health organisations and networks active in sub-Saharan Africa.

Task 1.2. **Internal meetings [M01-M36], Partners: all**

The consortium will have online bi-monthly meetings. All PIs will be invited, as well as other individuals invited by the PIs. The project manager will arrange for the meetings, send out an agenda well in advance and take notes during the discussion. The coordinator will act as chairman during the meeting, discussing the supplied data with the other PIs. A summary of the meeting will be written and submitted to all individuals within the project. For these meetings, the SAB will also be invited to participate.

There will also be on-site meetings, with the first meeting starting month 2 (M2) in Finland. The purpose is dual, both bringing together all of the PIs of the project, as well as having a chance to see a similar surveillance system, facilitating downstream discussions and implementations. The second meeting will be hosted in Tanzania (M14), the third in Burkina Faso (M26), and the fourth in DR Congo (M36) - this to enable visits to the surveillance stations and insights into the implementation, as well as means to host workshops and meetings with stakeholders. Travel restrictions will be considered when the internal meetings are taking place, and if not deemed safe to travel to a specific place, we will either change location, or preferably have the meeting through two hubs (Europe, Africa), to still allow for meetings in person.

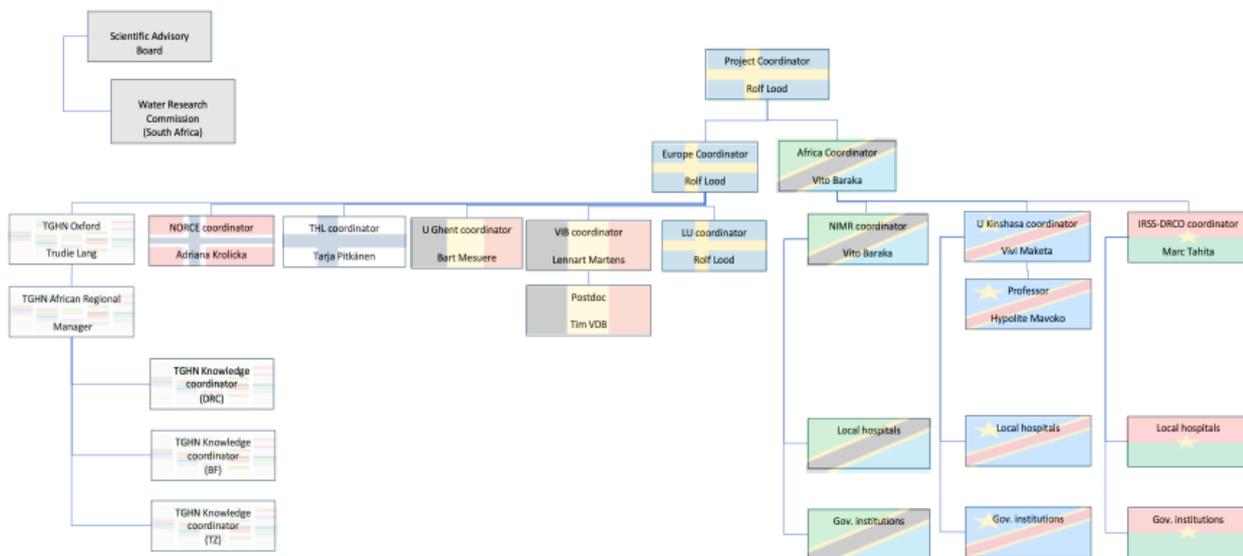**Task 1.3. External meetings [M01-M36]; Partners: LU, TGHN**

Most external meetings are managed on a local level, but initiations of new collaborations, invitations to other organisations to share knowledge, common infrastructure, or new opportunities with impact on the whole consortium will be in consultation with the coordinator. Similarly, additions to the SAB will also be executed through the coordinator, but contacts may be initiated by all members of the consortium.

Task 1.4. Follow-up on progression of project and budget [M01-M36], Partners: all

All project progression will be reported on a bi-monthly basis, before the PI meetings, and major changes, difficulties, etc brought up for discussion. Would there be any significant changes, this will be directly reported to the funding organisation. Budget restrictions and coordination thereof will be followed-up and discussed. Similarly, any significant changes according to plan will directly be communicated to the funding organisation.

Task 1.5. Establishing Community of Practice, networking, and coordination [M01-M36], Partners: LU, TGHN, NIMR

We will establish a community of practice (CoP) on AMR and pathogen genomics using the One Health approach. The project will also establish partnerships within the countries and regions to establish a network that will share expertise by active knowledge sharing, exchanging methods, and sharing best practices. The networking and collaborations will support equitable and sustainable capacity strengthening aiming to facilitate wider multi-pathogen sequencing and bioinformatic analyses that will facilitate timely epidemic and pandemic preparedness and response across the region. We will also develop guidelines, protocols, and standards of the WBE leveraging the advances in genomics. Through support from the ACDC/PGI and TGHN we will ensure these programmes are sustained by building lasting research leadership skills, establishing tailored-made training, in person and virtual across the networks of laboratories in the partner countries and beyond. The physical actions will take place in WP6 and WP7 but will be coordinated on a high level in WP1.

Work package number	2
Work package title	Setting-up environmental surveillance systems for multiple human pathogens
Lead Beneficiary	THL (3)

Objectives

● Develop environmental surveillance systems for multiple human pathogens to supplement the data from other infectious disease surveillance techniques at the selected study sites

● Document as a handbook the analytical workflow for processing environmental samples, generating metagenomic sequence libraries and gathering all information generated for further use among the local health authorities and other stakeholders

● Test in practice the developed environmental surveillance system to define a baseline microbiome content in the different sample matrices at the study sites

● Report the ODIN experiences and advances in infectious disease surveillance reached at the study sites to be further employed in the further years to come and in other sub-Saharan countries

Description of work

In this work package we will set-up environmental surveillance programs for major infectious diseases agents occurring in sewage and non-sewage environments containing or contaminated with human excreta in sub-Saharan countries. The development and testing of the environmental surveillance programs will take place in representative pre-selected study locations in Tanzania, DR Congo and Burkina Faso. During the project planning, we have identified a total of six sites where we will set up these environmental surveillance systems to support the existing sentinel and clinical surveillance of communicable diseases and to enable improvement of the water safety and subsequently human health in the area. We will prepare an environmental surveillance scheme to be utilised in sub-Saharan countries for communicable disease monitoring. We will also prepare a laboratory handbook with detailed analytical workflow for generating microbiological and high-throughput metagenomic sequence information from the collected environmental samples. The usability of the surveillance scheme will be verified at the study sites during a one-year monitoring pilot. The work has been divided into four tasks (2.1-2.4). The analysing of environmental samples does not involve a need for gaining consents since it is not possible to identify the individuals whose excreta is present in the collected environmental samples. Only sampling locations collecting excreta from a large number of individuals (preferably populations more than 20.000 persons) will be included in the study material.

Task 2.1 Gathering information on the existing clinical and environmental surveillance systems [M01-M09], Partners: THL, NIMR, UNIKIN, IRSS-DRCO

In this task, we will identify the relevant stakeholders, investigate the infectious disease surveillance techniques already in place in the sub-Saharan countries and define the main improvement needs in the communicable disease epidemiology at the selected study locations in Dar es Salaam and Tanga in Tanzania, in Kinshasa and Kisenso in DR Congo and in Ouagadougou and Nanoro in Burkina Faso. We are interested in clinical surveillance but also seek information about previous outbreaks of infections that could have had a waterborne transmission route. We will gather information about the coverage of sewerage systems and map the potential wastewater treatment plants that could commit to participate in long-term influent wastewater sampling campaigns. We will also gather the prior knowledge from the study sites about the clean water resources and their hygienic quality.

By using the expertise and connections of partners, local stakeholders will be reached and a workshop arranged. A questionnaire will be launched during the workshop to survey the existing infectious disease surveillance techniques and to identify the specific needs for improvement.

Task 2.2 Developing a proposal for a cost-efficient environmental surveillance strategy [M06-M18], Partners: THL, NIMR, UNIKIN, IRSS-DRCO

Based on the information gathered in Task 2.1, a set of targeted pathogens for each study location will be agreed. The microbial targets are selected based on local needs and by taking into consideration the illnesses that are difficult to detect by the existing clinical surveillance strategies. Candidate pathogens include *Vibrio cholerae* and resistant bacteria such as Extended Spectrum Beta-Lactamase (ESBL) producing *E. coli* and/or carbapenem producing *Enterobacterales* **(21)**, for which methodology applicable to sub-Saharan Africa is existing[8].

Our intention is to develop a surveillance program where as low as possible number of samples for representative

---

[8] https://www.who.int/publications/i/item/who-integrated-global-surveillance-on-esbl-producing-e.-coli-using-a-one-health-approach

surveillance outcome from each study location will be analysed within the sampling campaign covering both dry season and rainy season. We will map in detail the study sites to provide a list of environmental sample matrices to be collected, to identify the sampling locations and to define the optimal sampling frequency for these locations. This work is connected to the actions in WPs 3 and 5.

The anticipated sample matrices to be utilised in the environmental surveillance programs are hospital and community wastewater and other septic systems, surface water and groundwater. The population coverage in each sampling location will be estimated, based on population and sample type: e.g. septic tank, WWTP influent (sewerage sample), street gutter, river, open well, drill well, or tap water. In addition, we will study the possibilities of microbial and chemical markers , such as *E. coli* count, and CrAssphage or pepper mild mottle virus gene copy numbers **(22, 23)** of the collected samples, to produce metrics capable of producing numerical information to be used for population normalisation. Then, the selected markers of human excreta are quantified in each sample **(22)** for normalisation to the amount of human faecal material, as in non-sewered environments, the population size sampled at a specific site is generally unknown.

To overcome difficulties of wastewater based sampling in non-sewered environments, instead of grab samples, we plan to use passive samplers for collecting composite samples over a determined period of time (usually over 24 or 96 hours, **(24)**. The passive samplers **(25)** will be validated for detection of metagenomes and single pathogens, as these allow for sampling over longer timeframes with a better coverage of the underlying population. The water level, flow, consistency and proportion of wastewater in canals and other water bodies is dependent on the season and hence sampling is scheduled in both wet and dry seasons. The environmental surveillance in urban areas focuses on wastewater treatment plant influents, canals, street gutters and stagnant water bodies around cities, but also tap water and well water representing the clean water infrastructure are targeted. Furthermore, especially in rural areas systematic sampling of latrines and septic tanks of hospitals or other institutions offer a possibility to identify circulating and emerging infectious agents quickly and cost-effectively.

Task 2.3 Preparing guidance and a handbook for microbiological investigations [M06-M18], Partners: THL, NIMR, UNIKIN, IRSS-DRCO

This task will be performed in collaboration with clinical microbiology and water laboratories to produce a common operation procedure for conducting ODIN environmental surveillance scheme. The procedure entails guidance for collecting the environmental samples, sample pretreatment protocols for each sample matrix, detailed laboratory methods for microbiological and molecular investigations and protocols for the sequencing of the total nucleic acids from the environmental samples. The preparation of sample metadata tables, the reportable result calculations and reporting of the results will be planned in collaboration with WP4. For the selection of target pathogens defined in Task 2.2, methodological aspects related to the available laboratory network, conducting culture-dependent analytics and molecular microbiology work including metagenomic sequencing of the total microbiome and resistome of the samples can be modified based on the need of individual sites.

To establish integrated environmental multi-pathogen and AMR surveillance system for sub-Saharan countries, we will at first identify the exact capacities at the microbiological expert laboratories in regards of conducting the desired investigations from wastewater samples and agree on the list of microbial targets to be analysed in each study location where the ODIN environmental surveillance system will be tested. We aim to include quantitative analysis of culturable bacteria, namely faecal indicator bacteria *Escherichia coli* and intestinal enterococci (FIBs), *Vibrio cholerae* and ESBL *E. coli* as common targets for each study location and then add to the surveillance program other pathogens and microbial targets as proposed in Task 2.2. We will use metagenomic tools for covering the whole microbiome and the pathogen content of the collected water samples. For metagenomic sequencing of the environmental samples, we will employ commercial nucleic acid extraction kits capable of handling diverse environmental samples and high-throughput Illumina MiSeq and Oxford Nanopore Technologies. Where necessary, we will employ large volume sampling by using dead-end ultrafiltration **(26)** for the clean water resources to collect enough biomass for metagenomic sequencing. The produced sequence libraries will be transferred to WP4 for further processing.

We anticipate that pathogen-specific analytics based on traditional culture-based methods and quantitative (RT-)PCR are crucial to include for validation of the genomic surveillance approaches during the development phase of the metagenomic analysis pipelines in WP4. We foresee that after bioinformatic analysis the genomic surveillance of the produced metagenomic libraries might even reveal some emerging surveillance targets. For the validation purposes, some of the collected environmental samples will be used for inter-laboratory calibration purposes, where all participating laboratories in TZ, DRC and BF analyse the same sample and a sample fraction is shipped to the THL´s reference laboratory in Finland too.

Task 2.4 Collecting the environmental surveillance data [M10-12; M16-30], Partners: THL, NIMR, UNIKIN, IRSS-DRCO

Piloting of the prepared sampling scheme and analytical procedure will be conducted in collaboration with WP3 already at the end of the first study year to gain preliminary information about the sampling locations and their hygienic quality. Further, the first samples will reveal the performance of the proposed methods in the participating laboratories and produce preliminary sequence material to the WP4. Later on [M18], after preparations in tasks 2.1-2.3 have been completed, actual environmental surveillance campaigns at the participating cities will be launched. The outcomes of this task are connected to translational and risk assessment activities in WP5.

We plan to analyse a total of 360 environmental samples to produce an analytical data library to be further deployed in the bioinformatic pipeline development (WP4), epidemiological studies (WP5) and in training and data sharing activities (WP6). We will collect and analyse environmental samples from two distinct locations per country (see Table 1. for the study locations). Tentatively, each location will be monitored over three months in the dry season and over three months in the rainy season. The cornerstone of microbiological water monitoring is enumeration of faecal indicator bacteria (FIB). In cases when FIB's are present in water intended for human consumption, there is a possibility that also waterborne pathogens are occurring.

The microbiological investigations of the collected samples include analysing FIB, selected bacterial pathogen targets by culture-based methods, nucleic acid extraction, selected viral and bacterial pathogen targets by (RT)-qPCR methods and performing metagenomic shot-gun sequencing (and tentatively also 16S rRNA gene amplicon sequencing of bacterial community). Analysing approximately 60 samples per each study location (n=6) is planned in collaboration with partners NIMR (Tanzania), UNIKIN (DRC) and IRSS-DRCO (Burkina Faso) in collaboration with the local stakeholders.

The wastewater findings will be compared to the information gained from the local medical centres about the prevalent communicable diseases. Special attention is given, in collaboration with WPs 4 and 7, for communication and engagement with the local medical doctors, and epidemiologists to guide the interpretation of the wastewater/environmental results and to enable timely and proportionate public health actions.

Work package number	3
Work package title	Human pathogen surveillance using a mobile and semi-autonomous environmental workflow
Lead Beneficiary	NORCE (2)

Objectives

● Develop a mobile surveillance laboratory (MSL) (vehicle-based) for semi-automated molecular genetic tracking of human pathogens in environmental samples

● Update relevant libraries of human pathogen qPCR detection tools using new genomic knowledge of endemic pathogen diversity

● Provide extensive laboratory training of students in the molecular genetic methods used by the mobile surveillance laboratory through 3-5 month research exchanges

Description of work

To deliver efficient infectious disease surveillance, disease prevention and control, as well as public health emergency preparedness there is a strong need for implementing a holistic One Health approach that aims to sustainably balance and optimise the health of societies and the natural environment. Implementation of the environmental and veterinary pillar within the One Health framework is especially crucial in resource- and infrastructure-poor regions such as slum settlements and small villages, which represent a larger interface between human vs livestock / environment and therefore stronger human exposure to contaminated water, food and animal excrements that are carriers of human pathogens.

Researchers and public health professionals today have access to improved and integrated research and surveillance infrastructure that enables the combination of genomic and epidemiology data. These capabilities allow better understanding of infectious disease epidemiology and the development of new, low-cost, easy-to-implement solutions, which can deliver public health interventions for vulnerable populations in low-resource settings. Mobile surveillance will be more suitable for detecting early signs of outbreaks in small villages that are not linked with a common WWTP or small outbreaks that are not detected by community-wide environmental surveillance approach (WP2) because of the signal averaging and silencing at the community level.

The proposed mobile surveillance will include microbial source tracking (MST) i.e. mapping of sources of contamination and potential reservoirs of the outbreaks. In addition to early signaling about detection of genetic markers indicating upcoming disease outbreaks, our approach will provide indications for better interventions since we will aim at identifying the origin of the outbreaks in the surrounding environment. In other words, we will implement the One Health approach and its environmental pillar. Moreover, the integration of resistome analysis into the bioinformatic pipelines is expected. This will directly impact appropriate treatment and interventions.

We are planning to develop surveillance protocols to be used for drinking water sources, water bodies in the vicinity of settlements, seafood, and soil with animal faeces but as well as sewage at the inlet of small and hard to reach small WWTPs. Finally, we are proposing a mobile (bus-based) laboratory that will be reaching small villages/towns over 4-5 months to demonstrate the value of the One Health approach to understand the epidemiological status in a given region. The mobile laboratory will be fully equipped to perform sampling, nucleic acid extraction, PCR/qPCR amplification, sequencing, and data analysis. We are planning to utilise the only currently available portable solution, Oxford Nanopore sequencing technology, for the qualitative characterization of samples in terms of the presence of microorganisms. In parallel, the portable Biomeme Franklin™ Biomeme qPCR device and associated kits will be used for more frequent quantitative diagnostics of selected targets in a variety of samples collected from the surrounding environment. The flexibility of the approach will be the basis of the cost reduction of mobile surveillance having in mind that high-throughput sequencing (HTS) methods are not continually needed. Identification of pathogens based on the sequencing results will provide a way to use the qPCR targeted approach (Biomeme).

In this WP we are aiming to develop flexible solutions that will be tailored to local settings, with particular attention to resource availability. This flexibility will be enabled by a modular configuration consisting of independent functional modules: microbial profiling by Nanopore amplicon sequencing; metagenomic characterization of virulence- and antimicrobial resistance genes using Nanopore sequencing; targeted human pathogen detection using qPCR. Sample turnover and analysis costs will be minimised using single-target flow cells for Nanopore, or qPCR multiplexing. It will thus be possible to customise the approach to the changing local conditions and actual needs by combinations of functional modules as well as module customization. Training of students will be provided within all modules and project tasks.

Task 3.1 Workflow design and development of protocols for mobile and semi-autonomous surveillance [M01-M27]; Partners: NORCE, NIMR, UNIKIN, THL
Task 3.1.1 Collection of relevant samples from Congo, Tanzania and Burkina Faso
The aim of this subtask is to collect samples for environmental workflow development as well as to understand the potential challenges during processing of samples with different origins such as drinking water, water from water bodies in the vicinity of settlements, seafood, WWTPs inlet and soil that will be collected in Congo, Tanzania and Burkina Faso. During the workshop that is planned in WP2, task 2.1 the harmonisation of study locations and sampling schedule (Table 1.) will be achieved for sewage samples. In total, 40-60 environmental samples (n=9, 13-20 per country) will be collected and transported to NORCE.

Task 3.1.2 High-quality DNA purification and long-fragment sequencing analysis
High quality RNA and high molecular weight (HMW) genomic DNA are essential, both for successful Nanopore sequencing but also for downstream metagenomic analysis. Resource-greedy methods such as commercial kits for consistent production of high-quality nucleic acids, however, are poorly suited for rapid, field-based genomic surveillance programmes in resource-limited areas. Low-cost and low-resource solutions designed for field-based nucleic acid purifications proffer a viable alternative for use within a mobile public health surveillance framework. The concept for simple, cartridge-based, eDNA and eRNA extraction kits by Biomeme will therefore be used for the further development. Moreover, the issue with access to laboratory equipment will be taken into consideration. We are aiming at developing a method/ methods for nucleic acid extraction that could result in high quality DNA and RNA both be suitable for quantitative analysis (on the Biomeme qPCR device) and for HTS sequencing (amplicon-based, metagenomics, metatranscriptomics). This solution will significantly reduce the analysis time and cost in the situations when both qualitative and quantitative data will be needed. The strong focus will be brought to the metagenomic and metatranscriptomic approach because very often this is the only manner to track rapidly mutating viral pathogens. Sequencing technology is evolving rapidly, therefore it will be necessary to monitor the development of technological solutions made by Oxford Nanopore technology. This subtask will result in the development of a rapid, simple and robust method for nucleic acid purification from environmental samples that provides sufficient quality for shotgun metagenome library preparation and Nanopore sequencing in the field.

Task 3.1.3 Mobile in situ qPCR detection
Based on output of Tasks 2.1 (*a priori* targets) and 2.4 (new targets) we are planning to generate a general pool of qPCRs assays i.e., the molecular targets that could be narrowed and adapted to individual regions and tuned for instance to endemic diseases. We foresee the need of novel qPCR assays design, in addition to usage of the published ones, with focus on re-evaluating the accuracy of existing assays with target sequence variants identified using metagenomic approaches (Task 2.4). This development and optimization, as well as analysis of some samples from *Task 3.1.1* will be carried out using high-throughput digital PCR prior to implementation on the mobile qPCR device. The selection of targeted pathogens will be highly dependent on WP2 and WP5. Particular attention will be given to detection, quantification and identification of pathogenic *Vibrio* species and their genes responsible for pathogenesis. The qPCR targeted approach will be used (1) to confirm questionable sequencing results and (2) to demonstrate that the pathogen detected by HTS is metabolically active and (3) to limit the costs of analysis. For confirming metabolic activity, RNA based investigation using Reverse Transcription qPCR (RT-qPCR) will be performed.

Task 3.1.4 Improvement of semi-autonomous data processing module
The aim of this task will be to test the prototypes of bioinformatics pipelines developed in WP4. Moreover, the close cooperation between WP4 and WP3 to facilitate the transfer of knowledge between bioinformatics / IT specialists and end users is expected in this task. We will enable acquisition of sample metadata and analysis results into a larger project database. We will also enable semi-automatic data processing and dissemination by combining bioinformatic pipelines with IT scripts and tools. More specifically, semi-autonomous methods to analyse the raw data that are transferred from Franklin Biomeme and stored in Biomeme Cloud will be developed.

Task 3.2 Operationalization and cost- benefit analysis of mobile surveillance [M18-M36]; Partners: NORCE, NIMIR, UNIKIN, THL
Task 3.2.1 Student training for MSL deployment
This task will be closely linked to WP6. Students included in the ODIN educational program will be trained to acquire the necessary knowledge to carry out the tasks proposed in the subtask 3.2. under the supervision of the NORCE personnel and bioinformatic support provided in WP4. Moreover, students will transfer knowledge further to the students in Africa when the mobile laboratory is launched. A 3-5 month internship at NORCE for two MSc students from NIMR Tanzania is planned.

Task 3.2.2 Preparation of mobile surveillance laboratory (MSL)
In this task we aim to identify a suitable vehicle and purchase necessary furniture/furnishings to provide a mobile laboratory bench for sample processing. We intend to equip it with a Nanopore sequencing device and Franklin qPCR Biomeme device. In addition, we will create space for necessary computational and communications stations; waste storage; chemical storage, etc. Finally, we will perform n=10 test deployments to ensure success of all functional components of the mobile surveillance laboratory. Our intention is as well to take advantage of the expertise of currently available mobile laboratory personnel and the vehicle at the National Public Health Laboratory in Tanzania in the case of disease outbreaks.

Task 3.2.2 Operationalization of mobile surveillance
The sampling plan of MLS will be harmonised with the activities in WP2 and WP5, 2-3 time points from the same localisations is planned and the same types of samples will be collected as described in task 3.1.1 i.e soil and water. DNA/RNA purification from samples will be performed using the optimal field-friendly methods identified in Task 3.1.2. In total, it is expected that MLS personnel will collect and analyse 150-180 samples (n=3) in Congo, Tanzania and Burkina Faso (50-60 samples per country), whereof 150-180 samples (n=1, a pooled sample that will represent the collected 3 replicates) will be sequenced by using a metagenomic/metatranscriptomic approach and amplicon sequencing. Additionally,150-180 samples (n=3) will be further analysed by qPCRs on the Franklin Biomeme. qPCR products will be pooled and sequenced to confirm target accuracy and to identify new sequence variants present in environmental pools. The mobile laboratory will visit selected sampling locations as proposed in WP2 and will process the samples from the same source. A comparison of environmental surveillance data from this WP and the clinical data from WP5 will be used to evaluate the relevance of the mobile microbial source tracking approach. This task will allow us to establish a knowledge-based surveillance strategy, and will help to develop recommendations for sampling location and frequency of environmental surveillance efforts.

Task 3.2.3 Cost-benefit analysis of mobile surveillance solutions
The detailed evaluation of the costs and time needed for a sample processing in the combination with the description

of equipment, characteristic of facilities and qualifications and number of personnel will be performed for each variant and module proposed. In addition, the evaluation of the cost-effectiveness and sustainability of the mobile-based analyses of environmental samples that have been developed will be performed. The goal is to achieve the best balance between surveillance coverage and quality while minimising time to decision and requirements for resources and handling expertise.

Work package number	4
Work package title	Developing metagenomic and bioinformatic analytical processes with interactive visualisation and sharing endpoints
Lead Beneficiary	VIB (7)

Objectives

● Setting up bioinformatic processes to analyse both amplicon sequences as well as metagenomes generated from nucleic acids of the environmental samples

● Developing scripts that can facilitate targeted analyses

● Developing scripts that will facilitate generation of easy-to-understand reports from the raw sequence data

● Develop methods for data sharing in real-time

Description of work

Access to appropriate pipelines and well-prepared infrastructure i.e., servers and software, is essential to enable genetic surveillance. Moreover, the developed systems should be as simple as possible, yet as automated as possible to limit costs and maintenance requirements for the long run. This is especially important in resource-poor regions/villages with limited resources and qualified personnel. As a result, WP4 will focus on the development of such data processing and presentation pipelines, building on current state-of-the-art algorithms and interactive and intuitive data visualisations, all packaged in an easily deployable format (Docker) that will facilitate long-term use, and will offer the lowest possible threshold to adoption elsewhere. Throughout, the core objective is capacity building, by ensuring robust solutions that are deployed locally, and through extensive training and associated skills and knowledge transfer (see also WP6).

Task 4.1 Integrate with data generation processes [M1-M12]; Partners: UGENT, VIB, THL

The first key task in this WP is to integrate with the data generation processes established in WPs 2, 3 and 5, to ensure that the acquired data can readily be sent to the centralised processing infrastructure. As data types and associated metadata will vary, each type of data collection (e.g., amplicons, metagenomics data) will be analysed and standard formats and procedures will be established to ensure the consistent collection and transmission of these data and their associated metadata to the central storage system (see Task 4.3 below). Potential logistics limitations in the field will be taken into account, including intermittent or slow connectivity.

Task 4.2 Initial data interpretation and quality control [M1-M12]; Partners: UGENT, VIB

As data is collected, it will be useful to perform initial processing and associated quality control (QC) analysis as close as possible to the time of data acquisition. This task will therefore perform a detailed analysis of feasible processing and QC analysis at each point of data generation, and will provide automated pipelines to execute these steps locally. This will ensure that possible issues with data generation are caught sooner rather than later, and will allow remedial actions to be taken immediately (e.g., re-analysis of the sample while fresh and available). All processing metadata and QC information will become part of the metadata, and will be sent along to the central data storage, allowing longitudinal monitoring of each site in terms of expected data quality. This in turn can be used to inform the local QC pipelines to automatically flag aberrant data. It will also allow all information to be associated with its QC metrics in downstream visualisation (see Task 4.5).

Task 4.3 Centralised data collection [M1-M24]; Partners: UGENT, VIB

The data collected throughout the project (and beyond the end of the project) in the surveillance set up in WPs 2, 3 and 5, will be consolidated in a single, centralised data management environment. This system will be fed by the standardised and automated submission systems that will be built in Task 4.1, after processing and metadata enrichment by initial, local processing as described in Task 4.2. The data storage system will be built as a mixed relational and NOSQL object storage system. The relational data storage will primarily be used for metadata, while the object storage will be used for the actual raw data. For processed results, a flexible approach will be taken that optimises fast data access for interactive data visualisation. Importantly, free and open technologies will be used for the back-end engines of this data storage system, and it will be developed in such a way as to be easy to either

deploy locally, or in a cloud environment. Special attention will be given to optimise storage space requirements, as these can build up over time. Indeed, longitudinal data analysis will likely be relevant, but will probably require some form of specific data analysis to be re-run on original raw data (see also Task 4.4 below). As a result, all data will be maintained in high-availability form as much as possible.

Task 4.4 Data processing and pipeline [M1-M36]; Partners: NORCE, UGENT, VIB, THL

This task will continue throughout the project, as it is expected that new processing requirements will continue to emerge as the system is being used. The overall concept behind the implementation strategy is therefore iterative, in which we will start from initial, automated data processing pipelines for each data type, leading to interactively visualised results (see Task 4.5). By presenting these to experts and users, we will gather requirements for novel analyses, and implement these as well. Over time, we aim to deliver a semi-automated overall pipeline, in which essentially all analyses can run automatically, but in which certain analyses are triggered and/or parameterized by user action. Two major types of analyses are expected: overall analysis of data, which aims to extract trends and descriptive metrics for the acquired data, and targeted analyses in which, say, a given (set of) pathogen(s) is (are) specifically analysed in greater detail. As it is expected that these targeted analyses require somewhat different approaches, they provide a good illustration for the choice of a semi-automated approach. Indeed, the user can, for instance, trigger a targeted search, requiring the provision of a (set of) target(s), along with the data sets to screen. This will then trigger a parameterised pipeline, which may internally change settings on data processing tools based on the specifics of the task, or use different tools (e.g., viral variant callers) based on the specific request. The pipelines themselves will be built on the NextFlow system, for maximal flexibility and reproducibility in other settings.

Task 4.5 Develop a dashboard for interactive data visualisation and analysis [M1-M36]; Partners: NORCE, UGENT, VIB, THL

One of the key outcomes of the ODIN project is the ability to communicate any findings quickly yet effectively to a wide range of users, with a variety of expertises and backgrounds. This notably includes the general public as well. For this, we will build on directly relevant experience and expertise in user interface design, (interactive) data visualisation, and dashboarding at NORCE, UGENT and VIB. Interactive visual analysis, adapted to the needs of this project, will be achieved using an extensive suite of functionalities, including brushing and linking that can be used to explore trends, correlations, and outliers in the data. This will make it possible to better understand geographic and temporal aspects of the collected data, a key element given the context. Usability of the visual analytics will be further strengthened by making it possible to merge multiple related datasets for combined analysis, for instance by combining surveillance data, population data, and environmental data for a given geographic area, or to draw the QC profiles for a certain type of data across various data generation sites. Obviously, the presence of easily accessible and standardised metadata (see Tasks 4.1, 4.2, and 4.3 above) will be critical for these analyses. The dashboard itself will be based on Enlighten-web, a web application developed by NORCE. Enlighten-web is used in several operative web-portals, e.g., the Norwegian portal in the European Plate Observing System infrastructure (EPOS) (https://epos-no.uib.no:444). The Enlighten-web server is based on Python and the Pandas data analysis library. The client utilises WebGL for fully interactive visualisation of large datasets in the client's browser. This dashboard will be linked from The Global Health Network website, and will constitute an online resource for health personnel, scientists and decision makers alike. It will convey the results from environmental monitoring of major communicable disease agents to these key stakeholders, and will facilitate the generation of visualisations that can be used in easily understandable reports for these key stakeholders to be able to take informed, and rapid, decisions. We will also ensure that visualisations are made available to the general public.

As mentioned in Task 4.4, there will be a close link between data visualisation and (pipelined) data analysis, allowing the user to trigger a new data processing pipeline based on the interactive visualisations (e.g., performing a targeted search across multiple sampling sites for a particular pathogen that has newly emerged in a given site).

Task 4.6 Data dissemination [M1-M36]; Partners: NORCE, UGENT, VIB

A key goal for the ODIN project is to ensure that the collected information, code, and associated documentation can be disseminated and re-used as widely as possible. For this, careful attention will be given to ensure the FAIRness (Findability, Accessibility, Interoperability, and Reusability) of the data collected and stored in this project. NORCE, UGENT, and VIB all have extensive relevant experience in data dissemination. NORCE from participation in the ENVRIFAIR project (https://envri.eu/home-envri-fair), UGENT through a variety of projects, including the citizen science VLINDER project and COVID-related response data visualisation and dissemination, and VIB through their creation of the world-leading PRIDE database for proteomics data dissemination, and through very active involvement in, and leadership of, an ELIXIR Community. Moreover, permissive licenses will be used

as default for all content created in this WP (e.g., Apache2 or MIT open source licenses for code, Creative Commons 4.0 BY or BY-SA for content).

Work package number	5
Work package title	Integration of translational pathogen genomics into public health surveillance systems
Lead Beneficiary	IRSS-DRCO (6)

Objectives

● Investigate epidemiological links between surveillance system outputs and clinical situations

● Investigate species- and strain-dependency in cholera outbreaks using developed surveillance systems

● Prepare Water and Sanitation Safety Plans for the study sites and propose risk management approaches to reduce waterborne infection risks

● Estimate cost-efficiency of the developed methods

Description of work

The work will be focused on implementing the findings, by investigating potential correlations between the epidemiological data from the surveillance data and the clinical situation in translational settings. The WP also deals with development of safety plans for drinking water, sanitary inspections as well as evaluation of cost effectiveness of the developed methods.

Task 5.1. Epidemiological links between presence of waterborne pathogens and clinical situation [M07-M21]; Partners: IRSS-DRCO, NIMR, UNIKIN, THL, NORCE, LU

In order to link, and thereby validate, the results from the environmental surveillance system, data on prevalence on selected microbes that can be identified in wastewater (*E. coli*, *Enterococcus*, *Vibrio*, *Shigella*, polio virus, HIV) will be collected from all surveillance points. Several of these microbes are only opportunistic pathogens (*e.g. E. coli* and *Enterococcus*) but can still indicate outbreaks if their prevalence and/or carriage of AMR is affected. The actual collection of microbial data is conducted in WP2.4, as previously described.

Daily hospital admissions at the hospitals and health facilities in regions with WBE surveillance systems will be compiled for individuals suffering from infections, and confirmed laboratory testing data summarised. Health facility collected metadata will be obtained from the health facilities following local ethical and regulatory procedures. We will collect daily data of the confirmed infections, date of specimen collection and date of reporting of the laboratory results. The aggregate case of the laboratory-confirmed positive infection and date will be compiled from daily reports through the responsible local public health officers or epidemiologist. No patient data will be collected other than a sum of the total infections at a location over a set time-period.

Secondly, confirmed hospital cases of infections will prompt an investigation of the source of contamination. Upon increased prevalence of specific microbes in the environmental surveillance systems, several environmental samples will be collected with a mobile laboratory (see WP3) from that region to rapidly allow for identification of the source of contamination, and reporting to governmental institutes for further actions (e.g. sanitising water supplies, vaccination supply, etc) to limit further infections in the general population. The mobile investigations will be conducted in Tanzania during 6 months, and take actions on several (5-10) occasions. Several (5-10) different sites will be analysed from the region of detection, and a report of the identified spread of pathogens submitted to local authorities in order to promptly take actions. The analysis will be based on the methods developed in WP2 and WP3 combined with ddPCR of defined targets for methods validation.

Task 5.2. Translational investigation of species- and strain-dependent outbreaks of cholera [M01-M24]; Partners: IRSS-DRCO, NIMR, UNIKIN, LU, NORCE

Outbreaks of cholera are both episodic and endemic depending on different locations in Africa, but constitute an enormous challenge in terms of hospitalisation and resources, to some extent due to a lack of an efficient and long-lasting vaccine, as well as means to timely identify outbreaks, localise contaminated water sources and implement actions to reduce spread. In this task, we will investigate the clonal nature of episodic and endemic outbreaks of cholera, as well as the identity of the bacterial species conferring such disease in a translational setting. Preliminary data from episodic outbreaks of cholera in Tanzania indicate that not only *V. cholerae* can cause cholera, or cholera-like, symptoms, but also *V. metschnikovii*, a currently rather uncharacterized bacterium. It is therefore critical to understand its role in cholera outbreaks, prevalence, as well as genetic background in order to possibly re-evaluate the treatment regime and prophylaxis given to withstand cholera infections.

Cholera outbreaks identified by increased prevalence of *Vibrio* in environmental surveillance of wastewater will result in an immediate sampling of individual sites to localise the outbreak. We will use our mobile laboratory

(WP3) to drive out and sample individual segments of water sources from different locations with rapid-tests, developed in WP3. Positive samples will result in a more thorough sampling, as discussed below. Suspected outbreaks from sites not connected to our stationary sampling, but reported to our teams through local authorities/clinicians will result in a similar action.

Samples will be collected from water sources at the suspected outbreak site, as well as more distal from the outbreak site. DNA will be isolated, and amplicon sequencing and ddPCR (*Vibrio* species) conducted. Besides initial DNA isolation, the samples will be plated for isolation of *Vibrio* species, analysed with multiplex PCR for species identification, followed by WGS. In total, we will sample three episodic outbreak sites per year during two years (e.g. in total six samplings per country) in Tanzania and Burkina Faso, with an estimated five environmental samples per occasion. For DR Congo (endemic), sampling will be every six months at three different geographic locations, for two years, with five environmental samples from each location and time-point.

Task 5.3 Water and Sanitation Safety Plans [M04-09; M22-33] Partners: THL, IRSS-DRCO, NIMR, UNIKIN, LU

In this task, we will implement a drinking water risk assessment and quality monitoring system in the case study areas occurring in urban and rural areas. The results of the WP2 will act as the baseline information about the occurrence of the selected human pathogens in the study areas and provide a basis for developing Water Safety Plans (WSPs), Sanitation Safety Plans (SSPs) and Quantitative Microbial Risk Assessments (QMRAs) and waterborne risk management approaches. WSPs are used as a systematic approach to ensure water quality and SSP is a step-by-step risk-based approach to assist in the implementation of local level risk assessment and management for the sanitation service chain. Once embedded with QMRAs, the actual numbers of infections as well as the risk values shown as disability adjusted life years (DALYs) can be displayed. In this task, WSPs and SSPs with QMRA will be prepared for a selected set of water supply and sanitation infrastructures in the target city areas. The assessments involve enumeration of FIB's and selected pathogen analysis from the drinking water samples and sanitary inspections at the communities. Tentatively, we plan to complete QMRA for *Vibrio cholerae* during the project life span **(27)**, while WSP and SSP cover the potential waterborne hazards more widely. As an outcome, the number of cholerae cases in different water management scenarios will be calculated to showcase possibilities how the water safety could be maintained in cases it is at an adequate level and improved when there is need for that. Successful implementation of the WSPs, SSPs and hygienic quality monitoring requires commitment from the local health authority and laboratory facilities capable for bacteriological testing of water and this will be mediated through the communication actions of WP7. An open dissemination of the analysis results to the water users is enhanced by developing electronic reporting systems.

Task 5.4. Cost effectiveness of the developed methods [M19-M36]; Partners: UNIKIN, NIMR

In order to assess the cost effectiveness of the proposed methods (Task 5.1-5.3), health economics protocols will be developed and tested in the respective countries (Burkina Faso, DRC, and Tanzania). The proposed method will be compared to the ones used by the national surveillance program using three aspects: i) Intervention costs alone, ii) potential costs/savings to health providers (MoH), iii) potential societal costs/savings, which includes both health providers and households. The use of a standardised cost template in all the three study sites will allow for a comparable setting, and to investigate country-specific challenges.

Work package number	6
Work package title	Capacity strengthening, knowledge mobilisation and training in genomics and bioinformatics
Lead Beneficiary	NIMR (4)

Objectives

- Select and conduct training of trainer (ToT), workshops on genomics and laboratory skills
- Conduct training on bioinformatics, data interpretation and support ICT infrastructures
- Support postgraduate training (MSc bioinformatics) for selected candidates
- Organise workshops, short courses, and webinars on environmental surveillance (PGI and GHN platform)

Description of work

To support capacity building within genomics and bioinformatics, we will develop a program to enhance knowledge transfer, laboratory skills and capabilities, enhanced infrastructure and collaboration of laboratory networks within this area. The WP will support training to enhance skills transfer and competencies in genomics, laboratory skills for genomics work, and bioinformatics and genomic epidemiology and its integration.

This includes training of trainers (ToT), tailor-made training of genomic data analysis and interpretation. Tailor-made training will be provided for enrolled students, project staff, and selected candidates from the institutions in the sub-Saharan countries to strengthen their skills in genomics and data science. Due to the weak and inadequate infrastructure capacity at most institutions in Africa, the WP will also focus on ensuring that laboratory equipment is installed and ICT-infrastructure is to support the strengthening of bioinformatics units. The project will attract at least 4 postgraduate students (MSc candidates) on bioinformatics that will be fully supported by the project (NIMR, UNIKIN, IRSS-DRCO). The students will be integrated in the research project to conduct specific research activities throughout the project period. Teaching will also be scalable and equitable by the approached developed by TGHN such as supported learning, research skills, clubs in laboratories, and other settings and virtual classrooms. These approaches ensure not just that the whole team have access to training by the supporting institution but also benefits by making these opportunities to learn available to other laboratory staff and wider local stakeholders. This will widen the support and understanding of this research and increase the impact, as well as widening the skills base.

Task 6.1 Knowledge mobilisation [M01-M12]; Partners: NIMR, TGHN
We will establish a Knowledge Hub on The Global Health Network platform to facilitate a community of practice on AMR and pathogen genomics. This facility (and the wider components afforded through the platform) will offer longer-term sustainability in which the value and utility of the resources developed across the lifespan of this programme, will be realised beyond the remit of this grant (see task 2.4, 3.2 and 5.2).

Task 6.2 Regional hubs [M01-M12]; Partners: NIMR, TGHN, UNIKIN, IRSS-DRCO
A Regional Coordinator (RC) will be based at each co-applicant institution (Tanzania, DRC, Burkina Faso) and will manage the delivery of the initiatives outlined in this WP. These coordinators will be operationally supported by The Global Health Network team. The coordinators will be key to (i) the embedded and iterative approach of framing the training and resource development around the needs of local partners; (ii) the integration of capacity strengthening within the planned research and its evolution; (iii) the participation of the individual institutions, which will greatly enhance this approach and increase its impact.

Task 6.3 Teaching, training, and career development [M01-M36]; Partners: all
The ODIN partners and wider institutional groups will be enrolled into the TDR Professional Development Scheme (PDS) to track research competency development over time. This is a quantifiable measure of capacity development for an individual but can also be aggregated at team level, institutional level or even network level. All course and training materials created and authored by regional experts will be acknowledged and recognised. The assignment of DOI numbers will allow for greater discoverability, trackable and will ensure these assets are citable. This will enable recognition for the contributors and a good measure of impact.

Task 6.3.1 Educational activity review and development
We will perform a comprehensive review of all available training and educational activities related to the bioinformatics and sequencing of WBE pathogens. This includes a comprehensive and iterative review (performed periodically) of currently available and upcoming guidelines for the prevention, detection and treatment of specified infectious diseases. In addition to the research programme that underpins this proposal, the results from this exercise can inform targeted course content and prioritisation of workshops.

Task 6.3.2 Shared online learning
A catalogue of generic and specialist research skills eLearning will be made freely available through TGHN Training Centre[9]. Across our study teams and partner institutions we will deliver the WHO-TDR *Essential Curriculum for Research*. This framework was developed in partnership with The Global Health Network involving 7,000 participants and resulting in 13 key modules including i) project, team and finance management ii) governance and regulation ii) community engagement; and iii) research uptake, alongside all the critical components needed to run safe, ethical and high-quality studies.

Task 6.3.3 Course development
A comprehensive training course will be developed in the field of bioinformatics. The component custom-made modules will be hosted open access within TGHN Training Centre. Additionally, these modules will be made available through the in-house online learning environments within each institution. The online courses will also be integrated with complementary in-person training sessions, to deliver the necessary practical skill elements.

Task 6.3.4 Regional Workshops
In person, virtual and hybrid skills-based workshops will be organised offering training and knowledge exchange

---

[9] https://globalhealthtrainingcentre.tghn.org/elearning/

in genomic surveillance, bioinformatics and sequencing. Sessions will also hone key aspects of the research ecosystem (e.g. ethics, research administration) to enable a wider number of local and regional health facilities to develop capabilities (ca 10 workshops in total).

Task 6.3.5 Site exchange / Twinning program
Institutional level mentoring will be established through a twinning programme. Sites will be paired to share and coach each other in the various logistical aspects of implementation. This will be defined across the course of the programme, and sites paired accordingly (Project 1-2 site exchanges per year).

Task 6.4 Convening and connecting resources and excellence [M04-M36]; Partners: TGHN
Task 6.4.1 Support setting up of infrastructure for data handling, analysis, reporting, and communication
Leveraging the existing collaboration with Africa CDC, ALERRT and WANETAM local data scientists and bioinformatics personnel will be trained in Tanzania, DRC and Burkina Faso. Data tools built from the ALERRT network and hosted through The Global Health Network will be shared with the ODIN study team including assistance in infrastructure instalment, maintenance and troubleshooting. Moreover, The Global Health Network will link the data science community that develops from this project with the Global Health Data Science [10]Community of Practice for further support and knowledge sharing within our knowledge hub and across the platform (including the new research resources gateway.

Task 6.4.2 Resource Bank
The articles, templates, downloadable toolkits and guidance notes generated through this programme of work will populate a resource bank accessible through the Hub and also discoverable through other hubs (such as EDCTP, ALERRT and the resources Gateway). Key data management and ICT support tools[11] and others developed through the ALERRT network will be leveraged to support the objectives of WP4. This will further integrate the DM groups working across this consortium with those in ALERRT to cascade active learning and high value resources. The research embedded within this programme will also be used to create 'study profiles[12]' published on the hub. This includes protocols and SOPs, which can be readily downloaded and modified by others to raise standards and speed up research.

Task 6.4.3 Set up a collaboration mechanism with other international networks
We will seek to connect and collaborate with The World Health Organization Regional Offices for Africa (WHO Afro), EDCTP, PANDORA-ID-Net, ALERRT, CANTAM, EACCR, WANETAM, the WHO Special Programme for Research and Training in Tropical Diseases (WHO TDR), the Sanger Institute, Water Research Commission (WRC) South Africa, and the Africa CDC pathogen Genomics Initiative (PGI). Nominated researchers and executive board members from the aforementioned networks will be asked to serve on the scientific advisory board/executive steering committee for ODIN (WP1). In addition to creating synergies, collaborations will also assist in sharing and standardising research tools that can benefit the activities under WP4 (for more details see task 6.5).

Task 6.4.4 Pathfinder project
The research that underpins the programme of work outlined within this application will form the host studies. We propose implementing an 'add-on' study that will map, track and record the steps taken across the course of each host study, from conception to completion. Using a mixed quantitative and qualitative methodology, this 'pathfinder project' will enable us to determine the specific barriers and challenges faced in implementing and achieving the host studies set out. Each step will be documented, and various metrics captured (e.g. for time delays incurred) to create a publishable 'process map'. Integral to the 'add-on' study, where challenges are encountered, solutions will be actively identified through the wider expertise of The Global Health Network international research community. The premise is that solutions likely already exist, in other disease areas or settings. Using the global health research community, we can identify such excellence and connect this so other teams can benefit immediately.

As a result of the Pathfinder project the process map and all the associated tools, resources, guidance, training materials and recommendations that are used will be shared. The protocols will be made open access through the Hub to any group who would like to take up the opportunity of running a methodology pathfinder add-on study alongside their programme.

Work package number	7
Work package title	Dissemination, Communication and Exploitation of the Results
Lead Beneficiary	TGHN (8)

Objectives

The overall objective of WP7 is to ensure that results of the study are communicated in a regular and timely manner to key stakeholders and they translate into policy for improvement of public health systems in the respective countries (TZ, DRC and BK). The work package will also ensure the results are exploited to have impact beyond the project lifespan, supporting uptake of the genomics technologies and its integration in epidemic preparedness and response.

● Develop a consortium plan for communication, dissemination and exploitation which will identify key stakeholders and scientific Advisory Board (SAB), and ensure the engagement mechanisms during the study duration.

● Ensure there is regular and smooth communication within the consortium members and also with the SAB, key stakeholders such as NPHL/MoH stakeholders, networks, regional/international agencies

● Ensure visibility and timely communication between consortium members and the public through newsletters, the knowledge hub on TGHN, TV/Radio, and social media platforms

● Ensure the project results and disseminated to the scientific forums via established channels, networks, stakeholders, and local communities

Description of work

Efficient communication and engagement with local stakeholders, regional/global policy makers, academicians and the general public is a priority in all ODIN activities. The ongoing study processes and findings will be continously communicated with local level stakeholders including the study communities, respective Ministries of Health and surveillance programmes/national public health institutions, and the research partner institutions, and with other national and international/regional partners and technical agencies, such as WHO-AFRO, Africa CDC, technical agencies, (genomic-) surveillance networks and the general public. We will use different communication channels and strategies to reach each target audience, as outlined in our D&E&C plan (see section 2.2). Specifically here, we shall be using the ARCH facility within TGHN which supports translation of research findings into recommendations for uptake and application. ODIN will be case study in this facility and thereby providing the opportunity for sharing with many partners how to optimize use of genomics findings and work with policy makers to provide usable tools for translation and uptake

Task 7.1. Develop Consortium communication, dissemination, and exploitation plan [M01-M12]; Partners: all

To ensure that the study findings are promoted to have policy impact and exploitation generated knowledge is well communicated to the partners. Task 7.1 will entail development of the study communication and dissemination plan, and be executed during the first quarter FY1 of the project. The plan will be discussed and approved by the consortium members as well as the Scientific Advisory Board (SAB) to be established. The engagement of the Consortium members and SAB will ensure that there is consensus on the plan and responsibilities for each partner are clear:

Task 7.2. Organise project workshops, communication, and dissemination activities [M01-M36]; Partners: TGHN

The WP leader will ensure there is good communications between the ODIN partners and the stakeholders involved, the NPHL/Ministries of Health and technical working groups and partners in TZ, DRC and BF. The WP7 leader will ensure there is arranged ad hoc meetings to address implementation challenges as it may be necessary during the course of the study. For the dissemination, the WP7 leader will liaise with local PIs to ensure the local forum presentations at village, district, regional national/Ministries of Health levels are conducted applicable.

Task 7.3. Development of the ODIN knowledge hub, and social media presence [M01-M36]; Partners: all

The leader of the WP7 will ensure the knowledge hub for the study is designed, populated with resources, and launched 6 months after the study kick off. The hub contains information regarding the study, consortium members and frequent updates of the Consortium activities such as, workshops, meeting, conference participation, training opportunities organised during the course of the study. Importantly, it also connects this study with key programmes such as ALERRT, the EDCTP networks of excellence and Africa CDC. It also facilitates immediate access to the tools and training our teams need in elements from community engagement to data science, and can bring our

training to this 'one stop shop'. The WP leader will also ensure media presence to ensure the project visibility is increased including TV and local radio. The online communication to the general public will be via ODIN project account managed by the WP leader on social media platforms (Twitter, Facebook, YouTube and ResearchGate etc). The online media/platforms will communicate project key milestones and engagement in key events such International Day of Epidemic Preparedness, engagement with the Emergence and Epidemic Preparedness Technical Working Group EDCTP organised forums and events, Africa CDC Annual International Conference on Public Health in Africa (CPHIA) and workshops etc.

Task 7.4. Coordinate scientific papers and policy to ensure timely publication [M07-M36]; Partners: all
The project WP7 leader will coordinate with the other WP leaders to ensure timely communication of the scientific finding through open source and high impact journals, newsletters, which will contribute towards disseminating the results of the study and facilitate sharing with scientific and academic community. Policy briefs will be prepared towards the end of the project and in a local language to ensure key findings and recommendations translate into decisions at local and regional levels. The timing and decisions on publications in peer reviewed scientific journals, contributions to conferences, and information to health authorities will be managed by the WP7 leader.

Summary of Expected Outcomes from Work Packages

Task No	Expected Outcome (EO)	Activities supporting EO	Expected Impact[1]
1.1	An open, inclusive consortium with members from different countries, continents, and cultures striving to solve a common goal	WP1	8
1.2	Four physical meetings between the researchers within the project that will facilitate research exchange and sustainability within the project	WP1	8
1.5	A highly diversified network of researchers, institutes, stakeholders, and companies that will ensure the longevity and sustainability of the program	WP1	5,7,8
2.1	Relevant stakeholders in charge of communicable disease epidemiology and water supply and sanitation are brought together to create a common understanding of the existing challenges in controlling poverty related and emerging/re-emerging infectious diseases. The stakeholders in TZ, DRC and BF will better realise the possibilities the genomic epidemiology based on environmental surveillance could offer and the findings are available for dissemination also elsewhere.	WP2, WP7	1,5,8
2.2	Ministries of Health in TZ, DRC and BF will consider the uptake of the developed cost-efficient environmental surveillance strategy including the sampling schemes with good population coverage for genomic surveillance with aim to control the poverty related and emerging/re-emerging infectious diseases	WPs 2-7	1,2,5,6
2.3	The laboratories of National Public Health Institutes (NPHIs) will deploy the ODIN laboratory handbook including SOPs for microbiological investigations and sequence library production from environmental samples in their future genomic epidemiology work	WP2-3, WP6-7	2,3,4,5,7
2.4	Environmental surveillance data produced in TZ, DRC and BF serves as an example for other countries in the African Union to employ such monitoring programs to strengthen the public health actions against poverty related and emerging/re-emerging infectious diseases in sub-Saharan countries	WPs 2-7	1,2,8
3.1	Workflow for mobile and semi-autonomous surveillance	WP3	1,2,5,8

3.2.1	A mobile surveillance laboratory (MSL) for semi-automated human pathogen surveillance in environmental samples	WP3	1,5
3.2.2	Recommendations how the instruments, protocols and bioinformatic pipelines can be used after the end of this grant, and how the costs can be covered by other institutes and governments in a model system	WP3	1,2,8
3.2.3	A comparison of the surveillance approach outcomes as described in WP2 and the mobile approach	WP3	1
4.1	Integrated data collection and submission from local data generators to a central data management system.	WP2, 3, and 5	2,5,7,
4.2	Established local data processing and QC analysis to ensure data quality and efficient use of samples.	WP2, 3, and 5	2, 5, 7
4.3	A centralised data collection and management infrastructure	WP4	2,5,7,8
4.4	A semi-automated data analysis infrastructure, consisting of a series of automated pipelines that can be parameterised and called by the user.	WP4	2,5,7,8
4.5	An interactive and flexible, real-time data visualisation dashboard	WP4	1,2,5,7,8
4.6	FAIR and reusable data, code, and documentation	WP4	3,4,5,6,7,8
5.1.1	Validation that WBE surveillance systems correlate with actual clinical situations and thus is a good proxy for health status.	WP5	1
5.1.2	A report debriefing the cost-efficiency of the suggested workflow based on a stationary and mobile lab-investigation to timely report outbreaks of (general) infectious diseases.	WP5	2
5.2.1	A better understanding on the species- and strain- dependent nature of cholera outbreaks in sub-Saharan Africa, allowing for an epidemiological-based recommendation for prophylaxis and treatment regimes for cholera in different regions.	WP5	1, 3
5.2.2	A report debriefing the cost-efficiency of the suggested workflow based on a mobile lab-investigation to timely report outbreaks of cholera.	WP5	2
5.3.1	QMRAs provide a clear tool to showcase clear visuals on the numbers of illness cases before and after the proposed interventions of the water and sanitation systems and are used to increase awareness of public health consequences of faecal contamination of water.	WP5	1,2,4,7
5.3.2	Uptake of the water and sanitation management actions proposed in WSPs and SSPs offer possibilities for improvement of public health at the study location and they can be used as examples in other locations too.	WP5	1,2,4,5
5.4.1	A report debriefing the cost-efficiency of the suggested workflow based on a mobile lab-investigation to timely report outbreaks of (general) infectious diseases.	WP5	1,2,3,5,7
5.4.2	Recommendation for prophylaxis and treatment regimes for cholera in different regions.	WP5	1,2,5
6.1.1	Generation of a knowledge hub that will facilitate exchange of information in a sustainable way.	WP6	7,8

6.3.1	Provide a rich source of free and open supporting materials for implementing genomic surveillance systems.	WP6	7,8
6.6.1	Increased collaborations with other established networks and leverage input and data from peer networks with the aim to speed up evidence generation and improve research efficiencies within genomic surveillance.	WP6	3,4,5,6,7,8
7.1	Increased dissemination and Exploitation of the results for policy impact	WP7	2
7.2	Improved communication between consortium members and stakeholders	WP7	8
7.3	Increased project visibility via webpage and Social media Platforms	WP7	2
7.4	Increased dissemination of the results through Scientific conference and forums and other academic beneficiaries	WP7	3,5,8
7.5	Strengthening the linkage between project and policy via SAB	WP7	4,7,8

[1]_Expected impacts are listed below, and numbers given above:_

1. Achieve SDG3 "Ensure healthy lives and promote well-being for all at all ages" in sub-Saharan African countries.
2. Provide evidence for informed health policies and guidelines within public health systems in sub-Saharan Africa and at international level.
3. Strengthen clinical research capability in sub-Saharan Africa to rapidly respond to emerging epidemics.
4. Enable a regulatory environment that can ensure effective development, delivery, and uptake of new or improved safe health technologies guaranteeing that trials in sub-Saharan African countries meet international standards.
5. Enable countries in sub-Saharan Africa to better understand pathogen epidemiology and support effective public health monitoring through integration of genomics and epidemiology.
6. Increase cost effectiveness of public investment through collaboration of funders of clinical trials in the area of infectious diseases in sub-Saharan Africa.
7. Strengthen health systems to ensure uptake of effective health technologies and innovations.
8. Enhance sustainable global scientific collaboration in health research and international cooperation across sub-Saharan Africa.

Table 3.1c: List of Deliverables

Number	Deliverable name	Work package number	Lead partner	Type	Dissemination level	Delivery month
D1.1	Kick of meeting report	WP1	LU	R	PU	6
D1.2	Consortium Data Management Plan	WP1	LU	DMP	PU	6
D1.3	Minutes of Annual Progress and SAB meetings	WP1	LU	R	PU	13,25,36
D1.4	Yearly progress reports	WP1	LU	R	PU	13,25,36
D1.5	Signed contract by key stakeholders for continued financial support / maintenance of the surveillance systems	WP1	LU	R	PU	36
D2.1	Report of the existing surveillance and development needs in TZ, DRC, BF drafted as a manuscript	WP2	THL	R	PU	7
D2.2	Report entailing the protocol for environmental pathogen surveillance in sub-Saharan countries	WP2	THL	R	PU	16

D2.3	Laboratory handbook (SOPs) for pre-treatment of environmental samples, pathogens analytics and high-throughput sequencing	WP2	THL	R	PU	16
D2.4	Manuscript(s) describing the ODIN environmental surveillance campaign results and lessons learnt drafted	WP2	THL	R	PU	28
D3.1	Webinar records describing processing pathways as training video. Laboratory handbook (SOPs) for mobile surveillance workflow transferred to Zenodo	WP3	NORCE	R	PU	19
D3.2	Manuscript on methodological developments for mobile pathogen surveillance solutions	WP3	NORCE	R	PU	35
D3.3	Final report on cost-benefit analysis of mobile surveillance solutions	WP3	NORCE	R	PU	36
D4.1	Report on integration with data generation processes	WP4	VIB	R	PU	12
D4.2	Report on local data processing and QC analysis	WP4	VIB	R	PU	12
D4.3	Report on the central data collection and management infrastructure	WP4	VIB	R	PU	24
D4.4	Report on the fully operational data processing pipeline	WP4	VIB	R	PU	24
D4.5	Report on the final interactive data visualisation dashboard	WP4	UGENT	R	PU	36
D4.6	Report on all implemented data dissemination approaches and data FAIRness	WP4	NORCE	R	PU	36
D5.1	Report on the links between presence of waterborne pathogens and clinical situation	WP5	IRSS	R	PU	23
D5.2	A report on the impact of WBE surveillance as proxy for health status	WP5	IRSS	R	PU	23
D5.3	A report about cost-effective water quality monitoring and water safety interventions	WP5	IRSS	R	PU	33
D5.4	A report of the cost-efficiency of the suggested workflow based on a mobile lab-investigation to timely report outbreaks of cholera	WP5	IRSS	R	PU	30
D5.5	Policy dissemination plan for each country on prophylaxis and treatment regimens for cholera for each country	WP5	IRSS	R	PU	26
D6.1	An online knowledge transfer hub to support the community of practice (COP) convened in WP1, that tracks, measures and reports the successful uptake of the COP	WP6	NIMR	DEC	PU	12
D6.2	Review of current existing material for education within bioinformatics and sequencing	WP6	NIMR	R	PU	9
D6.3	A package of workshops for stakeholders, students, and clinicians	WP6	NIMR	DEC	PU	12
D6.4	Signed contracts with key stakeholders within the regions for continued scientific and financial support	WP6	NIMR	R	PU	27
D7.1	Dissemination and exploitation plan	WP7	TGHN	R	PU	12
D7.2	Conference presentations (e.g. EDCTP/CDC)	WP7	TGHN	DEC	PU	12,24,36

D7.3	Activity reports of ODIN webpage and social media	WP7	TGHN	DEC	PU	12,24,36
D7.4	Peer reviewed papers, newsletters and policy briefs	WP7	TGHN	R	PU	12,24,36

Table 3.1d: List of milestones

Miles tone	Milestone name	WPs	Due mont h	Means of verification
M1.1	SAB consisting of 5+ organisations covering national and international stakeholders	WP1	15	Signed contracts
M1.2	Confirming continued (financial) support for surveillance systems by regional and local stakeholders in all three countries	WP1	36	Signed contracts
M1.3	Effective study start in respective countries	WP1	3	Reports from countries
M2.1	Stakeholder workshops arranged and the questionnaire launched at TZ, DRC and BF	WP2	6	Report from workshops
M2.2	Performance of laboratory methods checked and metagenome sequence library production tested with environmental samples at TZ, DRC, BF	WP2	12	Summary report submitted to coordinator
M2.3	First round of sampling conducted, target pathogen and metagenome libraries produced	WP2	21	Validation report
M2.4	Second round of sampling and analytics completed, metadata and result files finalised	WP2	27	Report
M3.1	Processing pathways (workflow) developed for mobile and semi-autonomous surveillance and entered the best practices.	WP3	18	Report submitted to Knowledge Hub
M3.2	Students trained and ready for deployment	WP3	21	Certificate of studies
M3.3	Fully functional MSL, tested and ready for deployment. Established route and plan how to broadcast systematically form MSL route	WP3	21	Verification of function; report
M3.4	Benefits of the novel mobile surveillance solutions demonstrated and documented to stakeholders	WP3	38	Report, Movie
M4.1	Local data generators' bioinformatics needs met	WP4	12	Report
M4.2	Central data collection, management, and processing operational	WP4	24	Report
M4.3	Interactive dashboard complete	WP4	36	Report
M4.4	FAIRness of data, code, and documentation ensured	WP4	36	Report
M5.1	Validation report on the environmental surveillance systems	WP5	21	Report
M5.2	Policy dissemination plan drafted for surveillance of *V. cholerae*	WP5	24	Report
M5.3	First drafts of the water and sanitation safety plans (WSP, SSP) drafted and the information needs identified	WP5	9	Report
M5.4	WSP and SSP content updated and QMRA initiated based on the results from the 1st sampling round	WP5	24	Report

M5.5	Cost effectiveness database available	WP5	33	Country reports
M6.1	Launch of a Knowledge Hub (website)	WP6	6	Website
M6.2	Identification, hiring, and placing of regional coordinators in TZ, DRC and BF	WP6	6	Signed contracts
M6.3	Launch of an Educational Platform	WP6	6	Website
M6.4	First held Workshop with key stakeholders	WP6	12	Workshop report
M7.1	Dissemination, Communication and Exploitation Plan drafted	WP7	9	Consortium plan
M7.2	ODIN webpage and the social media accounts established	WP7	9	Website
M7.3	Results dissemination to multidisciplinary stakeholders	WP7	12, 24 36	Publication, Progress reports

Table 3.1e: Critical risks for implementation

Description of risk; level of (i) likelihood and (ii) severity: Low/Medium/High	WP(s) involved	Proposed risk-mitigation measures
Existing/new travel warnings hamper the organisation of on-site meetings, research visits and participation of partners on sampling activities at the study locations; (i) High, (ii) Medium	1,2,3,5,7	Careful planning of the on-site visits. Arrangement of virtual meetings, virtual visits and the use of two meeting hubs - one in European node and one in African node. With having six study locations, the potential difficulties to travel in one location does not necessarily affect all travelling. Consecution of the laboratory work by partners at NIMR, IRSS-DRCO and UNIKIN who have strong laboratory expertise, supported with connections to NPHIs and experience on hosting international visitors.
Delays in delivery of consumables and equipments due to the global shortage; (i) High, (ii) Medium	2,3,5	Ordering of all laboratory consumables early. Seeking for alternative suppliers where necessary. Possibility to perform part of the analysis in partnering consortium laboratories in case of major delays.
Issues with lcoal data processing and QC generation; (i) Low; (ii) Low	4	It may be that there is not enough compute power available at a given data generator to perform initial data processing and/or QC analysis of these data. In those cases, this task will be outsourced to the central data collection facility, and the system will be developed such that this process is automated, and responds with QC metrics as soon as possible, so that the benefit of fast QC feedback to the data generator remains preserved.
Delays in setting up centralised IT infrastructure in Africa; (i) Medium; (ii) Low	4	Supply-chain problems and local logistics issues may delay the set-up of the local server infrastructure for the centralised data collection and management. To mitigate any delays, the same server infrastructure will be set up at VIB, and this can be used already during the project. As this centralised system will be built as a deployable unit (Docker technology), the switch to a local installation will be relatively straightforward, and its exact timing of installation will not be critical.

Unwanted variations in execution of laboratory procedures; (i) Medium, (ii) Low	2,3,5	Although each country has its own surveillance system, harmonisation of metadata and production genomic sequence libraries is desired. To avoid major differences in execution of the work, and its subsequent outcomes, specific standard operating procedures (SOPs) will be put in place at the participating laboratories and study sites to minimise this risk. Cross-checking of data and repeating experiments in other partner laboratories will be done to confirm study results if needed.
The consortium fails to deliver due to a major deviation from the approved work plan; (i) Low, (ii) Low	1	Project builds on already existing collaborations which is a great advantage for the success of the project. The partners of this consortium are complementary and have already been successfully working together in different consortia for many years on various projects. These projects have been efficiently managed, resulting in major policy recommendations, numerous joint publications and delivered several PhD students.

Table 3.1f: Summary of staff effort (PM, Person Month)

	WP1	WP2	WP3	WP4	WP5	WP6	WP7	Total PMs per Participant
LU (1)	**22**				19	2	2	45
NORCE (2)			**26**	14	3			43
THL (3)	1	**39**	8	17	9	3	2	79
NIMR (4)	6	20	22	6	6	**21**	15	96
UNIKIN (5)	6	38	40	6	6	30	18	144
IRSS-DRCO (6)	6	52	38	28	**32**	9	10	175
VIB (7)	1		2	**54**		6	3	66
TGHN (8)						21	**4,2**	25,2
UGENT (9)	1			15		2		18
Total Person Months	43	149	136	140	75	94	54.2	

Table 3.1g: 'Subcontracting costs' items

Partner	Cost (€)	Description of tasks and justification
3/THL	40 000	In WP2, 15 000 € is reserved for gaining information about the water and sanitation infrastructures at the study locations and for defining the population coverage of each sampling location. In WP5, 25 000 € is reserved for a consultancy company to facilitate the Water Safety Planning at the clean water resources and Sanitation Safety Planning at the study locations.
4/NIMR	20 000	Subcontracting will be used for hiring of vehicles for field work during the duration of the project, as well as hiring microbiological equipment and general laboratory equipment maintenance fees for the duration of the project.
5/UNIKIN	20 000	The amount of 20,000€ covers the car rental fees 12,000€ for vehicles during the field work and the 8,000€ maintenance contract that includes cost shared with the ongoing project at the Laboratory. The maintenance contract will be requested by the ODIN project as participation fees to cover the cost of microbiology equipment maintenance as well as laboratory general equipment maintenance project duration.
6/IRSS-DRCO	20 000	Subcontracting will be used for hiring of vehicles for field work during the duration of the project, as well as hiring microbiological equipment and general laboratory equipment maintenance fees for the duration of the project.
8/TGHN	258 750	Face to Face Workshops - €33,750 – These costs are to support the development and delivery of face-to-face capacity development workshops in Tanzania, DRC and Burkina Faso. We will liaise with expert organisations locally in each location to assist in the development and delivery of these workshops using a tried and tested model that The Global Health Network has successfully operated in other regions globally. Each

		workshop will cost an estimated €3,750 and we plan to deliver one workshop in each region per year resulting in nine face to face workshops across the duration of the award. These workshops will hone key aspects of the research ecosystem (e.g. ethics, research administration) to enable a wider number of local and regional health facilities to develop capabilities. Regional Knowledge Coordinators – €225,000 – these costs are to place Regional Coordinator (RC) (€75,000 per coordinator) at each co-applicant institution (Tanzania, DRC, Burkina Faso) who then will manage the day-to-day delivery of the initiatives outlined in this WP 6. These coordinators will be operationally supported by The Global Health Network team as such collaboration agreements will be implemented between the University of Oxford and each institution.

Table 3.1h: **'Purchase costs' items (travel and subsistence, equipment and other goods, works and services)**

2/NORCE		
	Cost (€)	**Justification**
Travel and subsistence	49 400	Workshop, WP2-WP3:12600 (travel x 3 person: 12150€, meals (e.g. 5 meals for 3 ppl x 30€): 450€), Attendance to Annual Meeting x 1 person, WP3: 6000€, research stay for 2 masters students from TZ, DRC and BF institute for 2-3 months c/o NORCE, WP3: 22 800€ (accommodation: 5 400€, per diem: 14 400€, cost per travel 2 student return: 3000€), conference attendance WP3: 2 000€, Attendance to Annual Meeting x 1 person WP4: 6 000€.
Equipment	10 500	Depreciation costs for: Nanopore for mobile lab: 750€, Biomeme for mobile lab: 7 500€, Small Equipment for mobile lab: 2 250€
Other goods, works and services	229 140	Consumables, WP3: Filters: 3250€, other general lab consumables: 7340€, Nanopore Oxford products: 72840€, qPCR (or dPCR): 20884€, products form Biomeme: 26070€, Mobile laboratory 27500€, Publishing: 3 000€, Mobile Lab 27500€, data storage 10000€ Biomeme Web API License (Annual):4800€ , WP4: Linux server for visualization and analysis, 24 000€; Gitlab, WP4: 4800€, Audit:3 000€
Total	289 040	
3/THL		
	Cost (€)	**Justification**
Travel and subsistence	46 800	Travel of the project personnel to the consortium meetings (WP1, 2 000€) and stakeholder workshops (WP2 and WP7, 3 000€), researcher visits to the African partner laboratories to set-up the environmental sampling schemes and analytical operation procedures for multiple pathogens (WP2, 25 000€), and site visits during preparation of water and sanitation safety plans (WP5, 16 800€).
Equipment	3 000	Purchasing of the passive samplers to be tested in reference conditions prior employment to the study locations (3000€).
Other goods, works and services	62 000	Laboratory consumables to conduct inter-laboratory validation tests in the reference laboratory of THL (WP2; 18 000€), this entails plasticware, membrane filters, ultrafiltration units, nucleic acid extraction kits, primers and probes, qPCR kits, culture media and Illumina kits. Funds for the metagenomic sequencing service provision are part of costs reserved for services (25 000€). In addition, we reserve funds for sample shipping costs (14 000€) and other other costs that may arise when sampling strategy and sanitation interventions are planned (5 000€).
Total	111 800	
4/NIMR		
	Cost (€)	**Justification**
Travel and subsistence	36 000	Travelling to yearly internal meetings, workshops, and conferences for 1 PI and two researchers.
Equipment	20 000	Pump for wastewater, fridges for sample storage, software and laptops
Other goods, works and services	202 000	This funding is dedicated for DNA isolation, library preparation, and metagenome / amplicon sequencing of several hundreds samples, based on the cost of ca 450€ per metagenome sample. The cost further includes sampling costs (wastewater, well water, clean water), reagents / consumables for microbiology cultures, and identification using API gallery and BD BBL systems. Further, the academic fees and other costs related to

		university training for 2 MSc in bioinformatics (40 000€) is included.
Total	258 000	

5/UNIKIN

	Cost (€)	Justification
Travel and subsistence	40 000	Travelling to yearly internal meetings, workshops, and conferences for 1 PI and two researchers, as well as funding for MSc student travels.
Equipment	20 000	Pump for wastewater, fridges for sample storage, software and laptops
Other goods, works and services	162 000	This funding is dedicated for DNA isolation, library preparation, and metagenome / amplicon sequencing of several hundreds samples, based on the cost of ca 450€ per metagenome sample. The cost further includes sampling costs (wastewater, well water, clean water), reagents / consumables for microbiology cultures, and identification using API gallery and BD BBL systems. Further, the academic fees and other costs related to university training for 2 MSc in bioinformatics (40 000€) is included.
Total	222 000	

6/IRSS-DRCO

	Cost (€)	Justification
Travel and subsistence	54 000	The estimation for International travels (36 600€) is made based on the guidance of institutional rules: ca. 2 000€ per ticket per travel and an average daily cost of 100-150€ for a maximum of one week stay/trip, with the local PI with two team members attending 2 international scientific meetings during the course of this project, in addition to the consortium meetings. Local travels (17 400€) will be needed to cover the travel costs either from Nanoro to Ouagadougou for sampling, meeting or at IRSS headquarters for administrative purposes. In addition, the sum requested covers the travel fees for local collaborators from the MoH (Institut National de Santé Public and Ministry of water and agriculture) in Nanoro.
Equipment	20 000	The total sum requested for the small additional equipment to conduct this study (20 000€). Although core essential equipment is available, e.g. a powerful pump for the filtration of the wastewater, fridges for pathogens and samples storage, software, and printers are needed.
Other goods, works and services	162 000	This funding (162 000€) is dedicated for DNA isolation, library preparation, and metagenome / amplicon sequencing of several hundreds samples, based on the cost of ca 450€ per metagenome sample. Also sampling costs (wastewater, well water, clean water), reagents / consumables for microbiology cultures, and identification using API gallery and BD BBL systems are included. Further, the academic fees and other costs related to university training for 2 MSc in bioinformatics (40 000€) are covered.
Total	236 000	

7/VIB

	Cost (€)	Justification
Travel and subsistence	10 000	Travel of project personnel to yearly consortium meetings (WP1, €3 000) and for on-site training visits (WP6, €7 000).
Equipment	120 000	Funding (120 000€) will be used to establish the necessary server infrastructure for the centralised data management and storage, and analysis pipelines. To ensure fail-over and availability, the system will be set up in duplicate (always two identical servers, which will synchronise their data holdings between them), with an uninterruptible power supply to allow controlled shutdown in case of power outage. All storage will be duplicated in RAID-1 configuration and hot-swappable disks will be used.
Total	130 000	

8/TGHN

	Cost (€)	Justification
Travel and subsistence	15 000	This is the cost for travel and subsistance for 1 site exchange between research groups per year as part of the twinning programme so that research groups can learn from each others organisations on the ground activities and operations. This consists of two international trips per year (€1,500 per flight, €1,000 for accomodation, subsistance and local traveler trip), 6 trips for the total project.
Other	42 500	These costs consist of €37,500 for digital development, hosting, maintenance and

goods, works and services		trouble-shooting for the project hub within The Global Health Network and €5,000 in year 3 of the award for auditing fees as required by the EC. The digital development costs of €12,500 per year are the standard costs required to develop and launch the project hub on The Network and then the standard costs for maintaining the hub through out the course of the project, this includes all functionality updates, troubleshooting, site and usage reports, hosting of materials, tools and resources and all interactivity with other hubs and resources across the platform.
Total	57 500	

The consortium costs do not include items in other cost categories or in-kind contributions and thus Tables 3.1i and 3.1j are not presented.

3.2 Capacity of participants and consortium as a whole

The consortium is built up from two general nodes. In simplified terms, the European node will contribute with methodological expertise regarding genomics surveillance and bioinformatics tools and developments, and the African node will contribute with local knowledge of medical needs. Overarching these two nodes is The Global Health Network that, through its extensive network, will facilitate implementation of methods and contacts with key stakeholders on a local, regional, and global level. To support the proposed work, the consortium has existing connections to National Public Health Institutes (NPHIs) and their public health laboratories in Tanzania, DR Congo and Burkina Faso (e.g. Institut National de Santé Publique in Burkina Faso). **Further, the project has full confirmed support from both Africa CDC and Africa PGI**, and also ODIN Scientific Advisory Board (SAB) will support the work. The main PIs in this consortium are described below.

Rolf Lood, Lund University (LU), Sweden (Associate Professor)
Associate professor within experimental infection medicine, with 15+ years of professional experience working with molecular and genomic aspects of microbes. A strong focus the last few years has been on investigating prevalence of pathogens and resistance genes in wastewater and clean water in LMICs (Chile, Tanzania). The study was coordinated by Dr. Lood in a Nordic collaboration (JPIAMR 3rd JPI-EC-AMR), and a Chile-Sweden collaboration (ACCESS Sweden-Chile AMR). Dr. Lood has had 20+ MSc, PhD, and postdocs supervised in his lab. Dr. Lood holds several patents within biotechnological and biopharmaceutical solutions, with several being in clinical trials and/or already developed into products and sold from international companies.
Specific part in the project: Dr. Lood's main part in the project is leadership, management and coordination. Based on his experience in coordinating similar projects before, as well as acting as the VP of R&D for a global biotechnology company, Dr. Lood has the experience to conduct that role. Further, Dr. Lood will mainly assist with targeted genomics through ddPCR.
Available resources for the project: Dr. Lood has assigned 25% of his time for this project. Further, methods for ddPCR as well as primers for several pathogens and resistance genes have been designed and validated in earlier projects. The lab has a BSLII lab for both aerobic and anaerobic bacteria, as well as all modern molecular biology equipment needed for microbiology work. A ddPCR dedicated for this project is available in Dr. Lood's lab.

Adriana Krolicka, Norwegian Research Centre (NORCE), Norway (Senior Researcher)
Senior researcher with 17 years of professional experience holds a PhD in molecular ecology. Adriana's scientific work addressed the biodiversity in aquatic environments, application of monitoring tools and environmental DNA. Dr. Krolicka has been involved in projects leading to semi- or automation of environmental DNA aiming on tracking molecular signatures of pollution. She has extensive experience with microbial indicators and implementation of biological components on robotized genosensors.
Specific part in the project: Dr. Krolicka will lead WP3. The NORCE personnel will be involved in the preparation of the protocols for mobile environmental surveillance strategy (WP3) as well as in the preparation of the ODIN portal for data streaming in real time (WP4).
Available resources for the project: NORCE will support the project with experienced personnel in molecular ecology, field work and IT data processing. NORCE molecular laboratories are well-equipped with modern equipment and laboratory robots. Senior Researcher Jessica Louise Ray in the Molecular Ecology and Paleogenomics group at NORCE will assist Dr. Krolicka with field campaign design and coordination, sampling methods and surveillance implementation, and data analysis in WP3. Andrea Bagi, PhD in the Marine Ecology group at NORCE will be available for the project to contribute with expertise on Nanopore sequencing technology and associated data analysis solutions (WP3).

Senior Researchers Jeremy Cook and Tor Langeland in the Digital System group at NORCE will participate in WP4. They have long experience with developing web-based data portals providing advanced visual analytics functionalities. For this purpose, NORCE has developed Enlighten-web, which is a web application for online interactive visual analysis of large datasets. Enlighten-web is used in several operative web-portals, e.g., the Norwegian portal in the European Plate Observing System infrastructure (EPOS[13]). Enlighten-web will be re-used in the proposed dashboard for visualisation and analysis of data. The dashboard will be an online resource for health personnel, scientists and decision makers.

Tarja Pitkänen, The Finnish Institute for Health and Welfare (THL), Finland (Associate Professor)
Assoc. Prof. Tarja Pitkänen, PhD (ORCID 0000-0002-7591-9148) is a Group Leader at Water Microbiology Laboratory in Expert Microbiology Unit of THL´s Health Security Department, and has over 20 years' experience on molecular water microbiology and water quality investigations. As co-affiliation, Dr. Pitkänen work part-time as Assoc. Prof. (tenure track) in the field of environmental health at the Helsinki One Health position of University of Helsinki, Faculty of Veterinary Medicine, Department of Food Safety and Environmental Health. She has published 51 peer-reviewed scientific articles, has supervised 2 PhD theses and currently supervises 6 doctoral students. Currently Dr. Pitkänen leads a national consortium project WastPan[14] developing wastewater-based surveillance as a pandemic preparedness tool.
Specific part in the project: Dr. Pitkänen leads Work Package 2 "Setting-up environmental surveillance systems for multiple pathogens" in this consortium. Her research team will participate in research tasks related to planning of the surveillance systems and in performing the verification of the environmental monitoring programmes with the aim to support the projects capacity building and training aspects.
Available resources for the project: The Finnish Institute for Health and Welfare, THL, is one of the leading public health organisations in Europe (www.thl.fi). The research areas include environmental microbiology, virology, bacteriology, molecular epidemiology, antimicrobial surveillance, toxicology, epidemiology and chemistry, water safety, indoor and ambient air quality and risk assessment. THL has all the expertise and equipment needed for state-of-the-art water and public health research with a special emphasis on microbes causing infectious diseases. THL has 60 years experience in wastewater-based poliovirus surveillance and conducts weekly national wastewater surveillance for COVID-19, including next-generation high-throughput sequencing of SARS-CoV-2 variants. Also whole genome sequencing and data analysis with sophisticated bioinformatic tools is part of national surveillance of numerous viruses and bacteria. THL will offer modern facilities and has skillful personnel in microbiology, bioinformatics and epidemiology that support the proposed research and the laboratory work will follow a quality assurance system accredited according to standard ISO 17025. THL maintains, based on the communicable disease law, the National Infectious Disease Registry, related microbial strain collection, and the antimicrobial resistance surveillance database (Finres). THL's water research is part of the Kuopio Water Cluster research infrastructure in Pohjois-Savo region, Finland.

Vito Baraka, National Institute for Medical Research (NIMR), Tanzania (PhD)
Dr Vito Baraka (https://orcid.org/0000-0001-9694-9293) is a postdoctoral researcher at NIMR Tanzania with a background in Molecular Epidemiology of infectious diseases. His research applies molecular tools and approaches to solve practical challenges in infectious diseases control. His current projects include development and molecular monitoring of malaria chemoprevention strategies in different vulnerable populations, tracking and mapping spread AMR/drug resistance infections using genomics tools to inform local and regional policies and strategies. Other projects include diagnostic studies of SARS-COV2 and surveillance of SAR-COV2 variants (EDCTP RIA2020EF) and EU/EDCTP Career Development Fellowship (CDF). He has extensive research network collaboration in the region, South-South as well as in the North. At the institute, he participates in the several regional and international genomics networks MalariaGEN consortium, and worked in collaboration with Prof Rolf Lood (Lund University, Sweden) on a AMR project (Manuscript submitted). He has published several papers in peer reviewed journals (26 papers, H-index 11) and is involved to support postgraduate students (MSc and PhD))**.**
Specific part in the project: He will be involved in the co-implementation of the Capacity strengthening and training in genomics and bioinformatics (Work Package 6) and Knowledge mobilisation, data sharing, dissemination and communication (Work Package 7). Through his previous research, network and experience in coordinating similar projects he has the capacity to ensure the targets are timely achieved and results are translated to the key stakeholders for public health action.
Available resources for the project: Dr Vito Baraka will coordinate the study locally in Tanzania to ensure

---

deliverables are timely achieved. NIMR will support the laboratory infrastructure to support conventional and real-time PCR analysis, library preparation and purification and Sequencing using Mk1C Oxford Nanopore Technology (ONT). Also, the institute will support samples purification, processing laboratory for microbiological analysis, including culture and sensitivity testing, purification and detection of resistance genes. Through linkage with the Ministry of Health/Public Health Laboratory, we will support the use of a mobile Lab for sampling and analysis which has key equipment of BSL3/4 containment foldable glove box in which viruses and bacteria can be inactivatedReal time PCR for molecular diagnosis.

Vivi Maketa, University of Kinshasa (UNIKIN), DRC (assoc. professor)
Vivi Maketa has been involved in research, particularly clinical trials, for over a decade. For the past several years, she has been the national coordinator of a trial vaccine, an open-label, phase 2 study to assess the immunogenicity and safety of prophylactic vaccination of healthcare providers by administration of a heterologous Ebola vaccine regimen in the Democratic Republic of Congo funded by IMI. She is currently involved in two other vaccine trials. One is an open-label effectiveness study of a typhoid conjugate vaccine in Kisantu, Democratic Republic of Congo (TyVECO), targeting children under 16 years of age. The second is a multicenter, randomised, double-blind trial, phase 2b to evaluate the safety and immunogenicity of Janssen Ad26COVS1 and Novavax NVX-CoV2373 COVID-19 vaccines for a homologous and heterologous booster in adolescents and adults aged 12-64 years with and without HIV infection in 3 African countries (Kenya, Democratic Republic of Congo and Rwanda). Assoc. prof. **Hypolite Muhindo Mavoko**, who is Principal Investigator in the EBOVAC3 vaccine trial (EBL2007), to evaluate the Safety and Immunogenicity of a prophylactic vaccination of healthcare providers by Administration of a heterologous prime-boost vaccine against Ebola in the Democratic Republic of the Congo, Principal Investigator in the PYRAPREG trial, a multicenter, randomised, open label study on Efficacy and Safety of a newly registered ArtemisininBased Combination (Pyronaridine - Artesunate - PYRAMAX®) for the treatment of uncomplicated malaria in African pregnant women, and Co-Investigator in a prospective evaluation of a rapid diagnostic test to screen for gambiense human African trypanosomiasis and diagnose *Plasmodium falciparum* malaria in the Democratic Republic of the Congo will contribute to ODIN.
Specific part in the project: Vivi Maketa and Hypolite Muhindo Mavoko will be involved in the Work Package 5 "Integration of translational pathogens genomics into public health surveillance system" of the consortium, as well as WP2-3 for sampling and analysis.
Available resources for the project: Vivi Maketa has experience in research in the field of infectious and neglected tropical diseases, capacity building, daily management and supervision of the research projects. She will be the study coordinator through the University of Kinshasa in DRC. She will ensure the link between the University of Kinshasa and the National laboratory of the Ministry of health that has capacities in molecular biology namely in PCR thermal cycle, Real time PCR and sequencing. Hypolite Muhindo Mavoko has built his experience in the field of infectious and neglected tropical diseases for more than 12 years.

Marc Christian Tahita, CNRST/IRSS-DRCO, Burkina Faso (PhD)
Dr Marc Christian is an Associate Researcher with more than 10 years of experience in infectious diseases working at the Clinical Research Unit of Nanoro (https://crun.bf/). He holds a PhD in Medical Sciences at the University of Antwerp. Dr Tahita has a large experience in setting up, managing and coordination of many studies including multicenter studies. His research activity now is focused on the epidemiology of resistance for malaria with artemisinin combination treatments and antibiotics. He is a research fellow of the EDCTP (Career Development Fellowship) and Fondation Mérieux. Dr Tahita has practical experience in design, conduct, analysis and monitoring of malaria intervention trials and has experience of the evaluation of regulatory issues. Next to these research activities, he has strong skills in Quality Management System and is ISO 15189 and GCLP auditor. Currently, Dr Tahita is leading the Clinical Laboratory of the Clinical Research Unit of Nanoro carrying out the malaria vaccine R21. He has authored or co-authored more than 40 refereed articles in international journals and conferences (https://orcid.org/0000-0003-2158-0182). He has supervised 4 medical students (2 medical doctors and 2 pharmacists) and involved also in the training and supervision of masters and PhDs students.
Specific part in the project: Dr. Tahita leads Work Package 5 "Integration of translational pathogens genomics into public health surveillance system" of the consortium. This work package group will help to develop and capacitate public health laboratories for routine surveillance systems of key pathogens.
Available resources for the project: Dr Tahita vast experience in preparation and conduct of studies will help to set up and conduct the study in Burkina Faso. He will be the coordinator of this study in the country. The platform of the Clinical research Unit of Nanoro/CNRST/IRSS-DRCO will be used to achieve this present proposal. In terms of capacity, the site has a molecular biology laboratory (PCR and real time PCR) and also genomics capacity with one (1) Illumina MiSeq. The reference laboratory from the Ministry of Health also has genomic capacities with two (2)

Illumina MiSeq.

Lennart Martens, Flanders Institute for Biotechnology (VIB), Belgium (Senior Full Professor)

The CompOmics group headed by Prof. Martens, who has more than 20 years experience in cutting edge bioinformatics solutions for managing, processing, and disseminating large scale omics data, specialises in state-of-the-art omics data analysis and management, and has long-standing expertise in (i) developing bioinformatic software, (ii) (global) data dissemination, and (iii) data standardisation. First of all, the group has developed many popular bioinformatics software tools and algorithms over the past several years (e.g., https://www.compomics.com/#web-applications), resulting in over 265 published, peer-reviewed papers, of which 60 were published in the last five years alone. Notably, these tools are released as open-source under permissive licences (e.g., MIT and Apache2; https://compomics.github.io/). The CompOmics group thus has highly relevant expertise in developing production-grade bioinformatics tools, and can quickly build computational pipelines relevant for the project. Indeed, the CompOmics group has already built the Pladipus grid engine **(28)**, ProteoCloud **(29)**, and MetaProteomeAnalyzer **(30)**, to name but a few such large-scale data analysis pipelining systems. Moreover, **Dr. Tim Van Den Bossche**, who will actively participate in this project, already built automated connections between several popular, independent software packages **(31)**, contributed heavily to the development of a multi-omics pipeline for integrated host and microbiome analysis of non-model organisms **(32)**, and co-developed bioinformatic tools such as MegaGO **(33)** and Pout2Prot **(34)**. Furthermore, the group has proven experience in large-scale genome analysis and annotation projects, including the re-use of public data, with LNCipedia as the prime example **(34)**, along with a study to validate a large collection of small open reading frames **(35)**. Second, the CompOmics group also has highly relevant expertise in global data dissemination. Most notably, in 2003, Prof. Martens designed and built the PRIDE repository for the global dissemination of proteomics data, which to this day remains by far the world's largest and most popular data repository with over 21.000 data sets (**36**). And, last but not least, the CompOmics group is very active in developing community data standards in the context of the Human Proteome Organisation's Proteomics Standards Initiative (HUPO-PSI). For example, the group was instrumental in designing mzML, the community standard for mass spectrometry data, and qcML, the first exchange format for quality control metrics from mass spectrometry experiments **(37)**. Notable recent examples where Dr. Van Den Bossche actively contributed include the Universal Spectrum Identifier (USI) for mass spectra **(38)**, and MAGE-TAB-Proteomics, a standardised format to describe sample metadata and their relationship with dataset files **(39)**.

Specific part in the project: With our relevant expertise in the development of bioinformatic tools, pipelines, and management of computational infrastructure, we will contribute primarily to WP4 ("Developing metagenomic and bioinformatic analytical processes with direct public data sharing"), but will also assist in the (bio)informatics sections in WP3 and WP6. In addition, we will assist with bioinformatics training in WP7.

Available resources for the project: The CompOmics group of Prof. Martens possesses two local cluster computing systems. The first runs Kubernetes and consists of four high-end servers with dual AMD CPUs with 128 cores and 2TB of RAM, each. The second is a 40-server cluster with dual Intel Xeon CPUs with 32 cores and 32GB of RAM each. A total of 250TB of storage is available for these systems. The CompOmics group also has free access to the Flemish SuperComputer (VSC) Tier-1 and Tier-2 resources.

Bart Mesuere, Ghent University (UGENT), Belgium (assistant professor)

Bart Mesuere is a professor in Big Data Science at the Computational Biology lab at Ghent University. Prof. Mesuere has 10+ years of experience in building open-source data-analysis tools in several fields of research. A common theme in these tools is that they were built using the highest software development standards and that they make extensive use of dashboards and interactive data visualisations to communicate results. Prof. Mesuere was the initiator of the Unipept project which is a collection of software tools for the taxonomic and functional characterization of meta-omics samples. In Unipept, particular care was given to create a set of novel visualisations to efficiently convey the results to non-technical users. Prof. Mesuere was also responsible for the data strategy and dashboard development of the VLINDER project. In this citizen science project, weather data is collected from all over Flanders to improve high-resolution forecasting models and study the urban heat island effect. The collected data is also made available to the public both as open data as well as through an interactive dashboard. While initially set up using a research grant, VLINDER is now entirely self-supported because the public dashboard helped convince policy makers and stakeholders of the value of the collected data. More recently, Prof. Mesuere was involved in the communication on the evolution of the COVID-19 pandemic in Belgium using clear data visualisations and dashboards. He also advised the government during the subsequent vaccination campaign and was made responsible for the official dashboard giving insight into the progress of the campaign.

Specific part in the project: Given the experience in developing (biological) data-analysis tools and creating data visualisations and dashboards for experts and the general public alike, Prof. Mesuere will be mainly involved in WP4

(Developing metagenomic and bioinformatic analytical processes with direct public data sharing). In addition, his long-standing experience in developing an online training platform and teaching data visualisation courses will help with WP6 (Capacity strengthening and training in genomics and bioinformatics).

Available resources for the project: Ghent University is a partner of the Flemish Supercomputer Centre. This means the university has access to significant compute and cloud resources, both at TIER1 (regional) and TIER2 (university) level.

Trudie Lang, The Global Health Network (TGHN), University of Oxford, United Kingdom (Full professor)
Professor Lang brings a unique blend of experience from both the pharmaceutical industry and academia, focusing her career on research capacity and methodology in LMICs. Initially, she devoted her first 12 years of her career working in trial development and management in supervising a large Tropical Medicine portfolio at GlaxoSmithKline, working mainly on Malaria and Helminths. She then transferred to Oxford where she set up the clinical trial facility in the KEMRI-Wellcome programme in Kilifi, Kenya. She then founded The Global Health Network to support research capacity development, process improvement, research systems and management, aligned with all medical services, large public sectors and governments.

Specific part in the project: Africa CDC in collaboration with The University of Oxford have established a new Global Health Network Africa as a regional leadership and coordination centre. This centre takes forward the proven mechanism for open and agile connecting of research teams, efforts and outputs across the region. The program comprises a comprehensive research and data science element within the regional leadership hubs in Africa and works closely with sister hubs in Asia and Latin America and the UK. To fulfil The Global Health Network's strategic development plan; which is to shift the leadership to the Global South, the regional coordination centre is working closely with coordination teams with EDCTP Networks of excellence and other regional partners to drive its strategic goals. This successful approach for sharing know-how and giving visibility and engagement to the research team is highly scalable and extendable by design, and so through this award, as part of WP6 lead, we can take this facility to full impact through a regional focus that will establish regional leadership in Burkina Faso, Democratic Republic of Congo and Tanzania.

Available resources for the project: We will establish a Knowledge Hub on The Global Health Network platform to facilitate a community of practice on AMR and pathogen genomics and support the community with tools, training and guidance to assist research capacity development. The Global Health Network has developed over a decade to become established as a trusted, neutral and well-regarded facility for knowledge mobilisation in Global Health research. It works because it has two highly connected elements that bring global health organisations and the health research community in LMICs to the platform. One side is the knowledge exchange hubs[15]. The newly established knowledge hub will be linked to other more than 60 of these and whilst they are highly linked and interconnected (for maximum knowledge mobility) each is a community of practice for a health research topic or network such as MESH[16] (community engagement in research), or ARCH[17] (taking research finding into practice), or CEPI's space for Epidemic Preparedness and Innovation[18], and WHO's Epidemic Ethics[19]. The other element is the provision of research tools, training and guidance to support research capacity development and the conduct of quality research. These are made available through the capacity building[20] side which we support along with our partners, and is how it all began. The platform has been accessed over 50 million times and has over 500,000 active users who share their best practices. We have had over 3.5 million health research skills training courses taken within our health research skills training facility[21] where we maximise discoverability and access by translating the courses and supporting materials. We have developed several health research capacity-building resources which will be made available for this community, such as SiteFinder[22], the essential curriculum for research[23], the professional development scheme[24] that we run with TDR, and the tools, templates and guidance that researchers access every day through the resources gateways[25] in the main hub and also discoverable through the different network hubs, such as EDCTP knowledge

[15] https://hub.tghn.org/members/

[16] https://mesh.tghn.org/

[17] https://arch.tghn.org/

[18] https://epi.tghn.org/

[19] https://epidemicethics.tghn.org/

[20] https://tghn.org/

[21] https://globalhealthtrainingcentre.tghn.org/

[22] https://sitefinder.tghn.org/

[23] https://globalresearchmethods.tghn.org/methodology-projects/essential-research-skills-training-curriculum/

[24] https://globalhealthtrainingcentre.tghn.org/pds/about/

[25] https://hub.tghn.org/resources-gateway/

Composition of the consortium and overlapping skills

The consortium brings a broad experience within microbiology, genomics, bioinformatics, engineering, data science and visualisation skills in general, and within microorganisms in particular. The methodological approaches and specific research topics are, however, highly complementary, and add value to the consortium, enabling us to tackle larger research questions together. We also bring a broad experience in societal interaction, be it through commercialization of products, teaching, reaching global audiences on social media, implementing new clinical practices, or writing white papers. Specifically, Dr. Pitkänen's vast experience in designing and managing surveillance systems will be of importance for WP2, as will Dr Krolicka's and her team's experiences to develop low-cost solutions for genomic surveillance. Dr Martens and Dr Mesuere have an impressive resumé in developing bioinformatics tools to facilitate data-rich studies, and to visualise these data in a widely comprehensible manner. Together with NORCE and their team of experienced bioinformaticians and engineers they will be able to develop novel means to present data interactively in real time. Dr. Tahita, together with Dr. Maketa and Dr. Mavoko, bring their clinical experiences of tropical infections to facilitate collection of clinical lab results and samples for our translational studies, facilitating the targeted genomics work by Dr. Lood for identifying species- and strain-dependency in cholera outbreaks. Dr. Baraka will facilitate a capacity building program through NIMR and his extensive network of researchers at other institutes in sub-Saharan African countries. Finally, Prof. Lang and her team at the Global Health Network will facilitate connections between researchers, clinics, institutes, organisations and governments. Working with TGHN is leveraging a vast amount of resources, the digital platform and providing strong value for money for many aspects of our plans. We will be using the digital platform funded by Bill & Melinda Gates Foundation and therefore there is no specific web development costs, only the development of our specific content and coordination of our information. Further, we are leveraging the coordinators already in place across Africa and bringing these together with our teams. This will enable rapid progress, as well as bringing a sustainable future for our work. Within the consortium we strive to have an even gender distribution, resulting in four women and five men constituting the main PIs in the proposal. It should, however, be stressed that all individuals were selected based on their high competence, and not on their respective gender. It is thus clear that the team's competences are highly complementary, and ideally suited to achieving our ambitious project objectives.

Access to critical infrastructure to implement project activities

The NIMR/National Public Health laboratories, National Institute of Public Health laboratories/IRSS- Burkina Faso and UNIKIN/National Reference Laboratory will provide the critical infrastructure required to implement the studies in Africa. This will include equipment for the collection of clean and sewage samples, tablets for epidemiological, demographic, and geographical information and GPS location. Also, the institutes will provide laboratory access for sample storage (refrigerator 4-8°C, -20°C, -40°C, −80 °C, dewars for Storage of Liquid Nitrogen (LN2)) as it may be required. For samples processing, equipment for DNA extraction (centrifuges, heating blocks etc.) is available at the respective sites and will be made available to the study. Regarding sequencing/metagenomic analysis, facilities for library preparation (conventional PCR and Real-Time PCR) and sequencing are available at NIMR (MinION-MK1C, Oxford Nanopore Technology, UK) and the NPHL (ION torrent, Thermo Fisher Scientific). IRSS Burkina Faso have Illumina MiSeq platform (Illumina Inc.) and DRC-UNIKIN/National Reference Lab have MiniON sequencer (Oxford Nanopore) that will be accessible for the project. Also, the Mobile Laboratory will be available for the project from the NPHLs, equipped with BSL3/4 containment foldable glove box. All institutions have infrastructure for basic microbiological investigation to support culture and sensitivity assays.

Earlier interactions between the consortium members

Several of the members in the consortium have interacted before. Dr. Lood has conducted several studies with Dr. Baraka (manuscript under revision) and they have submitted joint grant proposals. Dr. Lood has further submitted to NordForsk (the Nordic Council of Ministers) a grant proposal together with Dr. Pitkänen and Dr. Krolicka earlier this year under the Societal Security and Antimicrobial Resistance call. In addition, Dr. Tahita has a long-term research collaboration with Dr. Baraka and Dr. Maketa since more than 7 years in the framework of multicenter projects. Currently, they are carrying out a project on diagnostic studies and surveillance of SARS-COV2 variants (EDCTP RIA2020EF) in the 3 respective countries: Burkina Faso, DRC and Tanzania. Dr. Mesuere is a former

---

[26] https://edctpknowledgehub.tghn.org/

[27] https://dac-trials.tghn.org/

[28] https://alerrt.tghn.org/about/network-networks/

postdoc of Dr. Martens, and they frequently publish together. As such they have extensive knowledge about each other's capacity and expertise. Several members of the Global Health Network have interacted with Dr. Baraka as well in different networks. All together, these earlier successful interactions facilitate the proposed work.

Commercial involvement

The ODIN project does not include any binding forms of industrial/commercial involvement. However, we are aiming to remain in tight collaboration with companies such as Biomeme and Oxford Nanopore. We believe our findings can be further used by companies working on molecular-based portable solutions. More specifically, the protocols, bioinformatic pipelines and scripts developed in WP2, WP3, WP4 and WP5 can be further used by companies for improvement and acceleration of sample processing and data analysis. This interest from industry may not only come from parties working on pathogen surveillance solutions but also from a wide spectrum of industries such as O&G and aquaculture that are planning or considering using environmental DNA for monitoring purposes. It should, however, be noted that the ambition of the project is to develop sustainable, open-source solutions that can easily be adapted in all sub-Saharan (and worldwide) countries with a high success rate.

References

(1) Ashbolt, N.J., Toxicology, 2004. 198(1-3): p. 229-38.
(2) Lin, B., J. Hui, and H. Mao, Biosensors, 2021. 11(7).
(3) Wang, M., et al., Small, 2020. 16(32): p. e2002169.
(4) Bull, R.A., et al., Nat Commun, 2020. 11(1): p. 6272.
(5) Whittle, E., et al., mSphere, 2022. 7(1): p. e0096421.
(6) Konechnyi, Y., et al., Clin Case Rep, 2021. 9(4): p. 2236-2244.
(7) Hovi, T., et al., Epidemiol Infect, 2012. 140(1): p. 1-13.
(8) Mtetwa, H.N., et al., Heliyon, 2022. 8(2): p. e08910.
(9) Hendriksen, R.S., et al., Nat Commun, 2019. 10(1): p. 1124.
(10) Sims, N. and B. Kasprzyk-Hordern, 2020. Environment international,139: p. 105689.
(11) Tessema, S.K., et al., Lancet Microbe, 2020. 1(6): p. e227-e228.
(12) Tegally, H., et al., Nat Med, 2022.
(13) Bougnom, B.P., et al., Environmental Research, 2019. 168: p. 14-24.
(14) Mateo-Sagasta, J., et al., https://doi.org/10.1016/j.envres.2018.09.022
(15) Nitiema, L., et al., Journal of Environmental Protection, 2013. 4: p. 575-584.
(16) Nakagiri, A., et al.,BMC Public Health, 2016. 16(1): p. 120.
(17) Edokpayi, J.N., et al., Water, 2018. 10(2): p. 159.
(18) Diemert, S. and T. Yan, Appl Environ Microbiol, 2019. 85(10).
(19) Gwenzi, W., Science of The Total Environment, 2022. 806: p. 150680.
(20) Tlhagale, M., et al., 2022. Journal of Water and Health, 20(2): p. 287-299.
(21) Blaak, H., et al., Science of The Total Environment, 2021. 776: p. 145925.
(22) Sabar, M.A., R. Honda, and E. Haramoto, Water Research, 2022. 221: p. 118827.
(23) Hsu, S.Y., et al., Biomarkers Selection for Population Normalization in SARS-CoV-2 Wastewater-based Epidemiology. medRxiv, 2022.
(24) Habtewold, J., et al., Environ Res, 2022. 204(Pt B): p. 112058.
(25) Schang, C., et al., Environmental science & technology, 2021. 55(15): p. 10432-10441.
(26) Smith, C.M. and V.R. Hill, Appl Environ Microbiol, 2009. 75(16): p. 5284-9.
(27) Ahmed, J., et al., Int J Environ Res Public Health, 2020. 17(8).
(28) Verheggen, K., et al., Journal of Proteome Research, 2016. 15(3): p. 707-712.
(29) Muth, T., et al., Journal of Proteomics, 2013. 88: p. 104-108.
(30) Muth, T., et al., Journal of Proteome Research, 2015. 14(3): p. 1557-1565.
(31) Van Den Bossche, T., et al., Journal of Proteome Research, 2020. 19(8): p. 3562-3566.
(32) Muñoz-Benavent, M., et al., NAR Genomics and Bioinformatics, 2020. 2(3).
(33) Verschaffelt, P., et al., Journal of Proteome Research, 2021. 20(4): p. 2083-2088.
(34) Volders, P.-J., et al., Nucleic Acids Research, 2012. 41(D1): p. D246-D251.
(35) Olexiouk, V., et al., Nucleic Acids Research, 2015. 44(D1): p. D324-D329.
(36) Perez-Riverol, Y., et al., Nucleic Acids Research, 2021. 50(D1): p. D543-D552.
(37) Walzer, M., et al., Molecular & Cellular Proteomics, 2014. 13(8): p. 1905-1913.
(38) Deutsch, E.W., et al., 2021. 18(7): p. 768-770.
(39) Dai, M., et al., Nat Commun, 2021. 12(1): p. 5854.