D3.4: Final Solutions for Healthcare Verticals Use of 5G

Revision: v.1.0

<table>
<thead>
<tr>
<th>Work Package</th>
<th>WP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>Task 3.3</td>
</tr>
<tr>
<td>Due rate</td>
<td>M42 – 30 November 2022</td>
</tr>
<tr>
<td>Submission date</td>
<td>30 Nov 2022</td>
</tr>
<tr>
<td>Deliverable lead</td>
<td>Telenor</td>
</tr>
<tr>
<td>Version</td>
<td>1.0</td>
</tr>
<tr>
<td>Authors</td>
<td>Per H Lehne (editor) (Telenor), Mikko Uitto, Antti Heikkinen (VTT), Linda de Goederen, Fons de Lange, Njin-Zu Chen, Ewout Brandsma, Henk Huijgen, Paul Gruijters, Elia Sarta, Gijs Neerhof (Philips), Havard Solvin, Mohammad Reza Sajadi, Ole Jakob Elle, Henrik Brun, Nilojan Gnanakulasekaran, Ali Khalegi, Hemin Qadir (OUS), Haibin Zhang (TNO), Donal Morris (RedZinc), Nicolas Deparis (CEA), Vera Stavroulaki (Wings)</td>
</tr>
<tr>
<td>Reviewers</td>
<td>Panayiotis Verrios (ACTA), George Fourtinias (OCC), Ruth Anderson (OCC), Jarno Pinola (VTT)</td>
</tr>
</tbody>
</table>

Abstract

This deliverable describes the phase-3 (final) trials of the healthcare use cases in 5G-HEART. These phase-3 trials contribute to the milestone MS5 of the project.

Keywords

5G, healthcare, trials
Disclaimer

The information, documentation and figures available in this deliverable, are written by the 5G-HEART (5G HEalth AquacultuRe and Transport validation trials) – project consortium under EC grant agreement 857034 and does not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

<table>
<thead>
<tr>
<th>Project co-funded by the European Commission in the H2020 Programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of the deliverable:</td>
</tr>
<tr>
<td>Dissemination Level</td>
</tr>
<tr>
<td>PU</td>
</tr>
<tr>
<td>CI</td>
</tr>
<tr>
<td>CO</td>
</tr>
</tbody>
</table>

Copyright notice

© 2019 - 2022 5G-HEART Consortium

¹ R: Document, report (excluding the periodic and final reports)
DEM: Demonstrator, pilot, prototype, plan designs
DEC: Websites, patents filing, press & media actions, videos, etc.
OTHER: Software, technical diagram, etc.
EXECUTIVE SUMMARY

This deliverable describes the final solutions used in the phase-3 trials of the healthcare vertical use cases of 5G-HEART, which contribute to milestone MS5.

During the final step of the three phased approach, the focus of the work has been towards conducting trials on 5G infrastructures.

The three healthcare use cases, H1 – “Remote Interventional Support”, H2 – “Automatic Pill Camera Anomaly Detection”, and H3 – “Vital-sign Patches with Advanced Geo-location” are further split into nine subcases. The paragraphs below give a summary of the achievements for each of the nine subcases.

For the subcase H1A – “Educational Surgery”, improved solutions for remote learning, consultancy, and remote attendance of surgical and other medical interventions have been explored by leveraging 5G technology. The setup has been tested by VTT using the 5GTN testbed in Oulu. The focus has been on delivering a real-time feed from an operating room to students in the classroom or at home for educational purposes. The initial target was to provide a 360° video streaming platform which was to be tested and run on 5GTN. Progressive insights led to the inclusion of single lens video streaming with focused cameras. This way the application latency was reduced, and improved video quality could be provided to better visualize essential parts of the surgical procedure. Two separate activities have been performed during the project course. The first activity involved completing the 5G Standalone (SA) setup and comparing, testing, and refining the network configuration to have optimal support for the subcase. This involved evaluating and configuring the uplink channel, since this determines the video quality originating from the content source. It became clear that single lens cameras have a lower capturing delay than the 360° cameras and are more optimized for low latency streaming. The application performance has been evaluated in terms of Motion-to-Photon (MTP) latency – also referred to as end-to-end (E2E) latency – and network KPIs for 5G have been measured. A baseline for latency providing the E2E user experience is 200 ms. During the evolution of the setup, performance has improved from 500 ms using the 360° camera over LTE, through 200 ms for using a single lens camera on the 5G Non-Standalone (NSA), to less than 180 ms for using single lens camera over 5G SA, which is only 12 ms more than performance over LAN. Also, it has been demonstrated that the uplink latency depends on the UL/DL frame structure in the network configuration, showing that higher bitrates with no latency increase can be achieved by changing the frame ratio from 1/4 to 3/7. The second activity involved RedZinc and Oulu University Hospital carrying out trials with RedZinc’s wearable equipment platform, with the aim to improve remote consultancy and education, especially during COVID-19. During the experiments, this platform was used in medical tutorials and the remote students were able to attend the tutorials from home with the ability to see the medical procedure from the trainer’s point of view which provides an optimal view of the procedure. The wearable video tests with the Oulu University Hospital showed that 89 % agreed that wearable video can be used for training and teaching that would otherwise not be possible, while 82 % felt that wearable video could replace traditional face-to-face teaching wholly or partly.

For the subcase H1B CHD – “Remote Ultrasound – Congenital Heart Disease (CHD)”, the use of teleguided ultrasound technology, possibly further enhanced with immersive telepresence technology, has been investigated. This could enable local paediatric cardiologists to arrive at a correct diagnosis in case of suspected CHD under the guidance of a remote expert, without the new-born or expert cardiologist having to travel to conduct the examination. Three technical setups have been used in the evaluations: (1) the commercially available Philips EPIQ/Collaboration Live system, (2) an experimental digital multi-stream ultrasound system based on Digital Navigation Link (DNL), and (3) an experimental AR/VR system enabling remote and local doctors to cooperate in a 3D setting. The first setup has been used for clinical evaluations at Oslo University Hospital (OUS). Specifically, a team of a remote expert and a local doctor can effectively decide whether a patient with suspected CHD needs to be transported to a central hospital or not. This solution was appreciated by both professionals as well as by the patient’s parents. However, the feedback from the clinicians was too limited in volume to be of statistical significance. Another clinical evaluation showed that the setup could be useful for
inexperienced students to conduct a rudimentary evaluation of cardiac function when guided by an expert. In addition, this setup has been used for 5G performance testing on the 5G-VINNI facility. The second setup was used for 5G performance testing on the 5Groningen facility. For the third setup, three experimental campaigns were performed. All involved were experienced paramedics and/or emergency doctors. The findings are based on qualitative evaluation from the clinical partners. The aim was to test and evaluate ultrasound image viewing in AR using e.g., HoloLens 2 or Oculus Quest 2. All the aforementioned experiments were performed either using Wi-Fi technology or wired setups. The experiments with the second and third setups were conducted in close cooperation with 5G-TOURS (another ICT-19 project) as well as with 5G-EINDHOVEN (a local cooperation in Eindhoven involving VodafoneZiggo, Ericsson, the local hospital and ambulance services). Combining the experiences from the three technical setups highlighted that the suitability of 5G strongly depends on the chosen technologies at user / application level. The EPIQ/Collaboration Live system, tested with and at OUS, does not put a high demand on network performance and can be supported by today’s 4G networks. Digital multi-stream ultrasound image and data sharing clearly requires a true 5G network configuration. Telepresence using 3D image capturing and AR visualization technology poses strong requirements on latency and bandwidth beyond the network performance demonstrated by the testbed. However, in this case the application technology is also not mature enough. Nevertheless, it has potential to further enhance the clinical collaboration experience when utilizing B5G/6G technology. Additional learnings from the experiments for this subcase highlighted the need for new business models and for modification of the existing clinical workflows which will slow down adoption.

For the subcase H1B Robot – “Remote Ultrasound – Robotics”, the objective was to develop a teleoperated robotic system. A master-slave configuration was developed for performing comprehensive ultrasound examinations of the heart, with the controlling expert sonographer sitting at a remote place, manipulating the robot via a 5G connection. Developing such a platform poses several challenges. Firstly, the visualization of the ultrasound probe over the patient’s body must be addressed. It has been tested using three 2D cameras, a wearable camera, and a 360° camera. Secondly, the usable workspace of the robotic system must satisfy the cardiologist’s demand during the examination. This has been optimized by testing three different probe holders, concluding that a 30° angle holder provided the best performance. Thirdly, a comprehensive risk and ethical assessment was required to be allowed to use the system on healthy volunteers. Based on this, approval was given by The Norwegian Medicines Agency to OUS, and testing was performed on 23 healthy volunteers. The system consists of three parts: the master side, the slave side and the communication link. The master side is typically located at the central hospital where the expert cardiologist works and it comprises a master robot, a haptic device, and a monitoring system. The slave side can be at a remote location, e.g., a local hospital or a general practitioner’s office, where a local doctor and/or a medical assistant is present together with the patient. The communication link consists of three sub links: a high-priority link carrying the real-time ultrasound image stream, a low-priority link carrying ambient audio/video between the master and slave locations, and a data link for establishing the bidirectional connection between the haptic device and the slave robot. Clinical testing was done via wired communications, since the 5G connectivity was not available until later in the trial period. A clinical evaluation was done by assessing the ability to create high quality images of the heart, in which important cardiac structures can be identified and whether it is possible to make precise measurements comparable to the gold standard examination as used in clinical practice. A complete protocol of the heart, consisting of 15 different images was done, and the image quality was scored by two-blinded fully trained sonographers from the paediatric echo-lab at OUS. Post examination tenderness among patients was also investigated. In total, the clinical evaluation of the robotic system did not reach the quality of the reference method. However, in the light of this being a preliminary clinical test, there is potential for further development and improvement, making it a very promising direction. Final system verification over Telenor’s 5G test network has been performed measuring the round-trip latency over the robot manipulation link, showing the feasibility of doing robotic ultrasound examinations over 5G networks.

For the subcase H1C – “Paramedic Support”, we have studied the sharing of wearable audio/video (A/V), patient vital parameters and ultrasound images, between a paramedic and a remote Chief Medical Officer (CMO), enabling the CMO to see the same patient context as the paramedic. This can help to
accelerate diagnosis and improve decision making. For this subcase, TNO and RedZinc have tested this scenario on different network setups. The remote CMO can be located anywhere and access the live A/V feed on different types of devices, like a laptop, tablet, or smartphone. The early tests were done using a legacy 4G network (phase-1), before moving the trials to an indoor 5G SA network in Groningen (phase-2). The final trials (phase-3) were performed in realistic outdoor conditions, using an outdoor 5G SA network in Helmond, including an edge server for low-latency applications. The phase-3 trials were a combined case with the transport vertical case T2S1 “Smart junctions and network assisted & cooperative collision avoidance (CoCa)” to test how RAN slicing could separate traffic and improve performance. In the setup, three network slices were implemented: one for video traffic (H1C), one for the Cooperative Connected and Automated Mobility (CCAM) traffic (T2S1), and one general purpose slice. Field trials were done with a vehicle were the CCAM traffic was used to obtain traffic light priority at a road junction, simultaneously with the paramedic’s A/V traffic and background traffic. The wearable A/V feed was realized using RedZinc’s BlueEye platform designed for first responders, a cloud-based service where the servers can be deployed regionally, or even physically on the premises of the regional ambulance service. Testing and verification have been focused on how video quality can be guaranteed while sharing the same network resource with the CCAM traffic and general-purpose traffic using network slicing with different priorities. The tests were performed with and without network slicing. While the CCAM traffic (highest priority), and the video traffic (2nd priority) were live, the general-purpose traffic was emulated using a roadside unit. The focus was on the uplink video traffic throughput, where the use of slicing showed a significant performance improvement. Without slicing, the general-purpose traffic was reduced to 50 % resource assignment (7.5 Mb/s), while the video throughput was lowered to 1.5 Mb/s because of the sensitivity of the algorithm towards the competition of the general-purpose traffic. With slicing, the video traffic was given higher priority and reached a throughput between 10 Mb/s and 15 Mb/s. The CCAM traffic, which had the highest priority, but involved little data did not affect the video traffic significantly. The study has validated that the use of real-time video and vital-signs data significantly improves the assessment effectiveness in emergencies. Furthermore, it has validated the role of 5G as the enabler for real-time video communication between a paramedic and a CMO, not just in almost ideal conditions but also in more realistic and practical conditions with significant mobile traffic. 5G network slicing can guarantee sufficient radio resources for the delivery of the A/V stream and vital-signs data for the ambulance service.

For the subcase H1D – “Critical Health Event”, the focus has been on planning the use of RedZinc’s BlueEye solution in extended trials with Oslo Ambulance Service’s Urban Search and Rescue/ Chemical, Biological, Radiation, Nuclear and Explosives (USAR/CBRNE) team. Video is increasingly being used in different settings, as it enhances situational awareness and leads to improved outcomes. Multiple pre-hospital, clinical applications can be supported where a paramedic or ambulance technician may need support. Moving away from the use of manikins and actors, the next step is to perform a larger pilot in a production environment, supporting real patient environments. The USAR and CBRNE operations can be enhanced with more ‘eyes in the field’ for supervision, support, and expert opinions. In the case of USAR, the incident commanders play a critical role as they establish objectives and supervise implementation. Similarly, the remote commanders are not able to assist the disposal technician at the front in absence of visual details. A secure mobile edge computing (MEC) environment has been developed, complying with GDPR security. Data is AES-256 encrypted, and TLS 1.2 and TLS 1.3 is used updated with secure ciphers. Multiple levels of security have been implemented, including the encryption of control and data planes, the use of a private APN to direct traffic to a secure, non-Internet environment, and the deployment of sovereign data servers located inside the region and geographical boundaries of the ambulance service. At the end of the 5G-HEART project, the secure service is under deployment, and full-scale testing is planned to be done as a continuation of the work outside the project. The use case proposes a trust-based approach to engage with medical users, offering integrated, trust-based solutions.

For the subcase H2A – “Automatic Pill Camera Anomaly Detection”, the aim has been to demonstrate the feasibility of video streaming over a 5G network from an endoscope capsule, receive and analyse the streamed video at the network edge, and provide feedback to the capsule, all with acceptable latency. Imaging technology in a capsule no bigger than a pill can provide a diagnosis within hours, furthermore,
AI is playing an increased role in technology development of clinical practices. The potential of AI to automate tasks is the motivation to combine precancerous pathology detection and novel advances in wireless capsule endoscopy into an easy-to-use alternative that can potentially save thousands of lives and improve the lives of patients and clinicians. The test setup comprises a wireless capsule endoscope (WCE) which uses backscatter technology to transmit live video to a body reader, or wireless capture system (WCS). The WCS relays the video stream over the 5G network to a hospital medical cloud, which can be deployed at the network edge, to reduce latency. In the cloud, real-time image processing and AI is used to detect polyps and provide feedback to the capsule to adjust the sensor settings such as light intensity, spectrum, and camera resolution. A convolution neural network (CNN) encoder-decoder network has been developed by OUS. This is adapted from AlbuNet34, which is a UNet-like architecture. In the test setup, the endoscope streaming has been simulated from a computer and sent to the AI server in the cloud. Quantitative evaluation of the performance of the proposed detection method has been done by calculating sensitivity (recall) and precision using well-known medical parameters: true positive, false positive, true negative and false negative. The ASU-Mayo clinical dataset with 20 specific videos for training was used for the evaluation. The backscatter system between the WCE and WCS is developed by OUS to support high data rates up to 16 Mb/s. Practical testing has been done for 8 Mb/s and 10 Mb/s. Telenor’s 5G-VINNI testbed was used in the early phase to measure end-to-end latency on different transmission protocols (UDP-RTP, TCP-RTSP, and TCP-HTTP). The measurements shows that UDP-RTP is the fastest e2e protocol with a measured latency of 46.74 ms, while the two TCP protocols performs from 240 ms to 470 ms. From the trials, it was clear that capsule endoscopy combined with deep learning and fast communication can lead to a paradigm shift in clinical activities in these settings. With a phantom-based experiment, the concept has been proven and video streaming from a capsule to 5G network has been validated.

For the **subcase H3A** – “Vital-sign Patch Prototype”, a single-use, Direct-to-Cloud, vital-signs patch prototype has been developed by Philips. This is a smart band-aid that measures a patient’s vital-signs 24/7 and communicates these directly to the cloud through the cellular network. This way doctors can keep a tab on their patients, no matter where they are. The scope of the work in 5G-HEART relates to efficiently, securely, and reliably uploading vital-signs data to the Philips Cloud, the key challenges being battery life and coverage. To investigate these challenges, an automated test framework was used for rigorous evaluation of the impact of different IETF protocols, protocol options, and 3GPP features and standards on energy usage and coverage. The test framework interoperaled with live commercial LTE-M and NB-IoT networks in different locations. Specifically, three topics have been investigated: (1) the selection and optimization of upload protocols to reduce energy consumption, (2) the feasibility of NB-IoT for improved coverage, and (3) the feasibility of firmware over the air updates. For all topics, the ability to operate within a very limited energy budget was the key issue investigated. Another concern was how to integrate the proposed solutions into the Philips Cloud. In relation to the first topic, observe that this use case involves the upload of very small payloads. Therefore, the impact of connection setup on overall energy consumption dominates over payload transfer. All protocols and options target a reduction of the number of messages sent and of their payload, in turn, leading to energy reduction. For the protocol optimization investigation, three different protocols have been tested over LTE-M: “plain UDP” (to provide a baseline), CoAP over DTLS, and HTTPS. Under poor coverage, HTTPS uses 4.7x more energy than “plain UDP”, and 1.6x more energy than CoAP over DTLS. With regards to HTTPS, also the improvements enabled by TLS session resumption have been quantified, 14-33% energy reduction is possible. Some further improvements in protocol efficiency may be expected leveraging recent developments in IETF such as QUIC, but the law of diminishing returns will apply and support by modem and cloud vendors is still unclear. In relation to the second topic, coverage extension using NB-IoT instead of LTE-M promises 15 dB to 20 dB better coverage, considering that CE Mode B is not deployed in practice. This is due to the massive repetitions employed by NB-IoT, implying exceptionally low data rates. A more lean-and-mean protocol based on CoAP is proposed. Practical coverage improvement in the tests proves to be around 6.5 dB, far from the (15-20) dB promise, probably due to network operator implementation limitations. Finally, in relation to the third topic, the energy consumption for a single upload is insignificant compared to the total energy budget, making a single firmware update over the operational life of the patch very well feasible.
For the **subcase H3B** – “Localizable Tag”, the objective is to evaluate the feasibility of a reliable, low cost, low power localization technique to complement the cloud connectivity of H3A. Therefore, the research work focuses on providing a lightweight and accurate radio-localization feature on wearable health monitoring patches. GNSS modules with low power consumption are currently commercialized for IoT devices. However, since the current mMTC application demands very stringent power consumption, these modules, even the ones with low power consumption, do not meet the requirements. Therefore, complementary location method should be specified for narrowband transmissions, notably for 5G- NB-IoT evolutions. The aim is to demonstrate the feasibility of localizing a low power wearable patch over a 5G network, only using received base station radio metrics, all with acceptable or predictable accuracy. For this subcase, a new “phase-coherent multi-channel” approach has been proposed by CEA that addresses uplink tag packets by sequentially transmit their narrow-band channels over the whole band, to create the much-required bandwidth for accurate localization. For the tests, the focus is on urban environments, where the radio propagation conditions result in strong multipath components. The field trials have been done in the city of Grenoble using CEA’s proprietary infrastructure where radio metrics have been collected into a database using crowdsourcing. The infrastructure comprises 6 base stations based on NB-IoT similar radio performances LPWA operating in the 868 MHz ISM band. These radio parameters are like NB-IoT operating in e.g., band 20 (800 MHz). The database contains more than 1 million samples which has been exploited to predict a Localization Accuracy Map (LAM). The network-based localization uses multilateration techniques, which is based on Time of Arrival (ToA) measurements from multiple base stations. The resulting localization accuracy depends on the ToA measurements, and it is clear that the propagation conditions contribute significantly (Line-of-Sight (LOS) vs Non-Line-of-Sight (NLOS)), as well as the base station position. Increased base station density and deploying them at the highest possible locations should be considered to improve the LOS conditions. A very good matching (< 12% difference) has been observed between the localization accuracy prediction and the measured one. In very dense urban condition, this accuracy is less than 300 m.

For the **subcase H3C** – “Aquaculture Remote Health Monitoring”, a remote health monitoring system has been tested. It has been designed, amongst others, to mitigate aquaculture workers’ hazards and risks. Workers in the aquaculture industry will be equipped with wearable devices (smart watches) measuring major vital signs such as heart rate, electrocardiogram (ECG), and oxygen saturation (SpO2). Two different test cases have been addressed. The first is the use of wearables to collect vital-signs measurements. These are further analysed using WINGS’ STARLIT platform, which includes a dashboard offering notification management and video calls on demand. The second test case is to use so-called “smart glasses” for live video streaming, to connect supervisor caretakers within the aquaculture area to remote medical experts. The subcase is concurrent to the 5G-HEART aquaculture scenario A1. The solution comprises the needed intelligence for the identification of issues and health emergencies, like vital signs being out of range, trending towards potential problematic situations, or patient leaving a pre-defined geographical area. Corresponding notifications are generated, as well as a dashboard to provide health care professionals with visualization of health monitoring data, notifications, and alerts. A deep convolutional neural network has been developed to analyse the ECG signal obtained from the smart watch. This is based on ResNet and has been tested using open data repositories from the MIT-BIH Arrhythmia Database. Similarly, a deep convolutional neural network for analysis of SpO2 signals has been developed and tested using the University College Dublin open data repository. Experiments have been performed using the 5G-EVE platform in Athens, but also the commercial network. Accuracy, sensitivity, specificity, and precision have been tested for ECG and SpO2 measurement analysis and show values close to or exceeding the targets. The overall system performance validation shows latency averages of 27 ms, while the user experience UL/DL throughput was 37 Mb/s and 134 Mb/s, respectively. The usability score was 3.5 on a scale from 1 (worst) to 5 (best). In total, the tests confirm that both the intelligence (AI) and the use of 5G can improve health care for workers in remote, potentially hazardous environments.

The overall conclusion from the implementation work is that the 5G-HEART project has been able to address and implement experimental setups for nine different subcases within the healthcare vertical. Even though not all setups have been connected and tested using 5G connectivity, the feasibility has
been proven for all, thanks to a committed cooperation between expert groups from the telecom area, the health technology industry, and clinical experts. Some of the subcases can be close to commercialization, provided that the stakeholders manage to forge the right partnerships. Others are still in an early exploratory phase and require further research and development of a partnership model.
TABLE OF CONTENTS

EXECUTIVE SUMMARY .............................................................................................................. 3
TABLE OF CONTENTS ............................................................................................................... 9
LIST OF FIGURES .................................................................................................................... 13
LIST OF TABLES ....................................................................................................................... 16
ABBREVIATIONS ....................................................................................................................... 17
1 INTRODUCTION .................................................................................................................... 20
  1.1 Use cases and phase-3 trials overview ............................................................................. 20
  1.2 Organization of this deliverable .................................................................................... 21
2 SUBCASE H1A: EDUCATIONAL SURGERY ........................................................................ 22
  2.1 Description and motivation ............................................................................................ 22
  2.2 Final setup ...................................................................................................................... 22
    2.2.1 Network architecture ............................................................................................... 23
    2.2.2 User application architecture .................................................................................. 23
    2.2.3 Hardware components ............................................................................................ 24
    2.2.4 Software components ............................................................................................ 24
  2.3 Testing and verification ................................................................................................. 25
    2.3.1 Methodology ............................................................................................................ 25
    2.3.2 List of key performance indicators .......................................................................... 26
    2.3.3 Measurement and testing tools .............................................................................. 26
    2.3.4 Final results ............................................................................................................. 27
  2.4 Recommendations ......................................................................................................... 29
3 SUBCASE H1B CHD: REMOTE ULTRASOUND – CONGENITAL HEART ................................ 30
  3.1 Description and motivation ............................................................................................ 30
  3.2 Final setup ...................................................................................................................... 30
    3.2.1 Network architecture ............................................................................................... 31
    3.2.2 User application architecture .................................................................................. 31
    3.2.3 Hardware components ............................................................................................ 34
    3.2.4 Software components ............................................................................................ 35
  3.3 Testing and verification ................................................................................................. 35
    3.3.1 Methodology ............................................................................................................ 35
    3.3.2 List of key performance indicators .......................................................................... 35
    3.3.3 Measurement and testing tools .............................................................................. 36
    3.3.4 Final results ............................................................................................................. 36
  3.4 Recommendations ......................................................................................................... 41
4 SUBCASE H1B ROBOT: REMOTE ULTRASOUND – ROBOTICS ........................................... 43
4.1 Description and motivation.................................................................43
4.2 Final setup ..........................................................................................43
4.2.1 Network architecture ........................................................................44
4.2.2 User application architecture ............................................................44
4.2.3 Hardware components .......................................................................46
4.2.4 Software components .......................................................................46
4.3 Testing and verification .......................................................................46
4.3.1 Methodology ....................................................................................46
4.3.2 List of key performance indicators ....................................................47
4.3.3 Measurement and testing tools ..........................................................47
4.3.4 Final results ......................................................................................48
4.4 Recommendations ..............................................................................52

5 SUBCASE H1C: PARAMEDIC SUPPORT ..................................................53
5.1 Description and motivation.................................................................53
5.2 Final setup ..........................................................................................53
5.2.1 Network architecture ........................................................................53
5.2.2 User application architecture ............................................................54
5.2.3 Hardware components .......................................................................55
5.2.4 Software components .......................................................................55
5.3 Testing and verification .......................................................................56
5.3.1 Methodology ....................................................................................56
5.3.2 List of key performance indicators ....................................................58
5.3.3 Measurement and testing tools ..........................................................58
5.3.4 Final results ......................................................................................58
5.4 Recommendations ..............................................................................60

6 SUBCASE H1D: CRITICAL HEALTH EVENT .............................................61
6.1 Description and motivation.................................................................61
6.2 Final setup ..........................................................................................61
6.2.1 Network architecture ........................................................................61
6.2.2 User application architecture ............................................................62
6.2.3 Hardware components .......................................................................63
6.2.4 Software components .......................................................................63
6.3 Testing and verification .......................................................................64
6.3.1 Methodology ....................................................................................64
6.3.2 List of key performance indicators ....................................................65
6.3.3 Measurement and testing tools ..........................................................65
6.3.4 Final results........................................................................................................................65
6.4 Recommendations..................................................................................................................65
7 SUBCASE H2A: AUTOMATIC PILL CAMERA ANOMALY DETECTION........................................................................................................................67
7.1 Description and motivation....................................................................................................67
7.2 Final setup ............................................................................................................................67
7.2.1 Network architecture .......................................................................................................69
7.2.2 User application architecture ..........................................................................................69
7.2.3 Hardware components ......................................................................................................71
7.2.4 Software components .......................................................................................................72
7.3 Testing and verification ........................................................................................................72
7.3.1 Methodology ....................................................................................................................72
7.3.2 List of key performance indicators ..................................................................................78
7.3.3 Measurement and testing tools .......................................................................................78
7.3.4 Final results .....................................................................................................................79
7.4 Recommendations ...............................................................................................................81
8 SUBCASE H3A: VITAL-SIGN PATCH PROTOTYPE .................................................................................82
8.1 Description and motivation....................................................................................................82
8.2 Final setup ............................................................................................................................82
8.2.1 Network architecture .......................................................................................................82
8.2.2 User application architecture ...........................................................................................83
8.2.3 Hardware components .....................................................................................................83
8.2.4 Software components .......................................................................................................83
8.3 Testing and verification ........................................................................................................83
8.3.1 Methodology ....................................................................................................................83
8.3.2 List of key performance indicators ..................................................................................83
8.3.3 Measurement and testing tools .......................................................................................83
8.3.4 Final results .....................................................................................................................84
8.4 Recommendations ...............................................................................................................102
9 SUBCASE H3B: LOCALIZABLE TAG .........................................................................................105
9.1 Description and motivation....................................................................................................105
9.2 Final setup ............................................................................................................................105
9.2.1 Network architecture .......................................................................................................106
9.2.2 User application architecture ...........................................................................................107
9.2.3 Hardware components .....................................................................................................112
9.2.4 Software components .......................................................................................................112
9.3 Testing and verification ........................................................................................................113
9.3.1 Methodology .................................................................................................................... 113
9.3.2 List of key performance indicators ................................................................................... 116
9.3.3 Measurement and testing tools .......................................................................................... 116
9.3.4 Final results ...................................................................................................................... 119
9.4 Recommendations ............................................................................................................ 120

10 SUBCASE H3C: AQUACULTURE REMOTE HEALTH MONITORING ................ 121
10.1 Description and motivation ............................................................................................... 121
10.2 Final setup ....................................................................................................................... 121
10.2.1 Network architecture ........................................................................................................ 122
10.2.2 User application architecture ............................................................................................ 122
10.2.3 Hardware components ...................................................................................................... 123
10.2.4 Software components ....................................................................................................... 123
10.3 Testing and verification .................................................................................................... 126
10.3.1 Methodology .................................................................................................................... 126
10.3.2 List of key performance indicators ................................................................................... 126
10.3.3 Measurement and testing tools .......................................................................................... 127
10.3.4 Final results ...................................................................................................................... 127
10.4 Recommendations ............................................................................................................ 128

11 CONCLUSION .............................................................................................................. 129
REFERENCES ............................................................................................................................. 135
LIST OF FIGURES

Figure 1 Overall architecture for H1A .................................................................23
Figure 2 The different components forming the experienced motion-to-photon (MTP) latency ...25
Figure 3 Motion-to-Photon (MTP) evaluation setup in 5G laboratory in Oulu .......................26
Figure 4 Essential 5GTN components for the trial in H1A: (a) 4G and 5G indoor small cells, (b) Intel NUC mini pc as video encoder, (c) TeleWell 5G SA modem ...........................................26
Figure 5 Detailed evaluation and functional architecture for phase-3 trials in H1A ...............27
Figure 6 User experienced motion-to-photon (E2E) latency for H1A trial evaluation for baseline LAN and 5G SA ..........................................................28
Figure 7 The effect of different video bitrate to the E2E latency ........................................28
Figure 8 Hardware architecture view of the 3D capturing and AR visualization application ....31
Figure 9 EPIQ (left) with USB webcam (middle top), remote control (middle bottom) and EPIQ screen (right) ..................................................................................................................32
Figure 10 Schematic view of 3D telepresence with context of local HCP (a) and Remote Expert (b) ..................................................................................32
Figure 11 Reference layout without 5G RAN ..................................................................33
Figure 12 Viewing US images with the HoloLens 2. Holding up the hand helps improve contrast on the images ..........................................................38
Figure 13 Setup of the user experience test of the 3D telepresence with GGD Eindhoven ...39
Figure 14 Captured scene as seen by the Remote Expert ..................................................39
Figure 15 HCP exploring the local role ..........................................................................40
Figure 16 Remote expert and local HCP collaborating .....................................................41
Figure 17 Ultrasound robot system architecture ................................................................44
Figure 18 Experimental Setup for Robotic Ultrasound Examination using Telenor’s 5G test network at OUS ..........................................................45
Figure 19 Experimental Setup for Robotic Ultrasound Examination over local Wi-Fi. One of the CPEs are used only as a Wi-Fi router ..................................................45
Figure 20 Spread in differences between left ventricular fractional shortening from the robotic examination compared to the reference ...........................................48
Figure 21 Proportion of cardiac structures deemed as assessable by the two independent rating sonographers. Right: Robot examination. Left: Reference examination ................................................................49
Figure 22 Ultrasound image quality rating for the robot compared to the reference ............49
Figure 23 Ultrasound image usability assessment .............................................................50
Figure 24 Experimental results of measuring round-trip time (RTT) ..................................51
Figure 25 5G SA network architecture (H1C, phase-3) ....................................................54
Figure 26 User application architecture for RedZinc's video streaming (H1C) ....................55
Figure 27 BlueEye application on an Android smartphone .................................................56
Figure 28 BlueEye hot desk ...........................................................................................56
Figure 29 Test scenario of the phase-3 trials (H1C and T2S1 jointly) ................................57
Figure 30 Observed uplink throughput of “All vehicle traffic” (H1C + T2S1) and general purpose traffic (Case A) .................................................................................................................................59
Figure 31 Observed uplink throughput of “Video traffic” (H1C), CCAM traffic (T2S1) and general purpose traffic (Case B)........................................................................................................60
Figure 32 BlueEye MEC architecture .................................................................................................................................62
Figure 33 Control and Data planes in BlueEye MEC ....................................................................................................................64
Figure 34 Full backscatter system demonstration setup ........................................................................................................68
Figure 35 End-to-End connection setup for video streaming via 5G network ........................................................................69
Figure 36 Patient side system and capsule connection to the 5G network ................................................................................70
Figure 37 Image processing and AI functionalities ..................................................................................................................71
Figure 38 Feedback loop and control process of the WCE ........................................................................................................71
Figure 39 End-to-End connection setup for video streaming via 5G VINNI network ................................................................72
Figure 40 AI-based anomaly detection and localization consisting of two paths: an encoder and a decoder ..................................................................................................73
Figure 41 Concept of the integrated antenna inside WCE for RF backscatter application (US20210059526A1 [18]) ................................................................................................................75
Figure 42 Drawing of on-body antenna array for RF backscatter with the transmitter and receiver array on a flexible substrate to set on an on-body surface (WO2021048303A1 [19]). 75
Figure 43 Prototype of the backscatter WCE with integrated electronics and the electrode antenna ..................................................................................................................75
Figure 44 Backscatter system includes RF transmitter of controllable power up to 500 mW, the receiver system and clock data recovery system ..................................................................76
Figure 45 Array antenna configuration for RF backscatter in the reader system ........................................................................77
Figure 46 Overview of the colon capsule system level design ..................................................................................................77
Figure 47 Test setup using streaming video file and phantom human simulator ........................................................................78
Figure 48 The full AI-based automatic polyp detection is developed and fully trained ................................................................78
Figure 49 Architecture for H3A ..............................................................................................................................................83
Figure 50 Block-diagram of automated test framework for H3A ..............................................................................................84
Figure 51 Initial and resumed TLS handshake with session ID resp. session tickets (indicative, simplified) (H3A) ........................................................................................................85
Figure 52 Distribution of total energy as function of attenuation for Plain UDP (H3A) .................................................................89
Figure 53 Distribution of total energy as function of attenuation for CoAP over DTLS (H3A) ... 89
Figure 54 Distribution of total energy as function of attenuation for HTTPS (H3A) ......................................................................89
Figure 55 Comparison of total energy between protocols at good coverage (H3A) ........................................................................89
Figure 56 Comparison of total energy between protocols at poor coverage (H3A) ........................................................................90
Figure 57 Successful transfers vs. attenuation (H3A) ..................................................................................................................94
Figure 58 Highest attenuation at which transfer still possible (H3A) ..........................................................................................95
Figure 59 RSRP as function of attenuation for NB-IoT (H3A) .......................................................................................................95
Figure 60 RSRP as function of attenuation for LTE-M (H3A) .................................................................................................96
Figure 61  Maximum UL repetition factor for LTE-M (top) and NB-IoT (bottom) vs. attenuation (H3A) ...............................................................................................................................................96
Figure 62 Maximum DL repetition factor for LTE-M (top) and NB-IoT (bottom) vs. attenuation (H3A) ...............................................................................................................................................97
Figure 63 Energy consumption vs. attenuation; HTTPS / OAuth 2.0 vital signs upload over LTE-M (H3A) ..........................................................................................................................................97
Figure 64 Energy consumption vs. attenuation; CoAP vital signs upload over LTE-M (H3A)..............................................................................................................................................................................98
Figure 65 Energy consumption vs. attenuation; CoAP vital signs upload over NB-IoT (H3A)...............................................................................................................................................................................98
Figure 66 Energy as function of payload; CoAP vital signs upload over NB-IoT (H3A)...............................................................................................................................................................................99
Figure 67 RSRP levels for subsequent iterations; CoAP upload over NB-IoT (45 & 200-byte payloads) (H3A) .......................................................................................................................................99
Figure 68 - Time of arrival (ToA) measurements ........................................................................ 106
Figure 69 Network architecture used for the field trial.............................................................. 107
Figure 70 LPWAN Tags used for field trials................................................................................ 108
Figure 71 Crowd sourcing ............................................................................................................ 108
Figure 72 Field trials: Base Station deployment setup ................................................................. 108
Figure 73 Base stations installed on the roof of the cable-car arrival station at Bastille # 2 (left) and Chamrousse # 3 (right), overhanging the city of Grenoble, France ............................................. 109
Figure 74 Base Station Heatmap Spectral Scan for different channel (Frequency in MHz) versus time, BCC@CEA #1, observation period = 30 days..............................................................................................110
Figure 75 Localization Algorithm webpage outputs: Grenoble City explored area during field test (top), Predicted Accuracy with CRLB (middle), Position Error CDF (bottom)....................... 112
Figure 76 LOS and NLOS propagation in urban condition .......................................................... 114
Figure 77 Field Trial visited cells in the city of Grenoble............................................................ 117
Figure 78 Overview of LOS analysis (Matlab raytracing tool).................................................... 118
Figure 79 ToA errors empirical CDF for different base stations according to their channel conditions .................................................................................................................................................. 120
Figure 80 Aquaculture remote health monitoring subcase architecture for H3C ...................... 122
Figure 81 WINGS STARLIT Remote Health Monitoring login page........................................ 124
Figure 82 WINGS STARLIT Remote Health Monitoring patients view .................................... 124
Figure 83 WINGS STARLIT Patients detailed personal information. ....................................... 125
Figure 84 WINGS STARLIT Notifications and alerts management........................................... 125
Figure 85 WINGS STARLIT Patients location monitoring (geofencing).................................... 126
LIST OF TABLES

Table 1 Healthcare subcases: 5G enablers, phase-3 trial locations and involved parties ........20
Table 2 H1A HW components. ...................................................................................................24
Table 3 H1A SW components. ..................................................................................................24
Table 4 Target KPI values for H1B CHD ..............................................................................36
Table 5 Mean fractional shortening measured on M-modes acquired by robot compared to reference examination.................................................................48
Table 6 Blood flow Doppler velocities measurements with robot and reference examination. ....50
Table 7 Round-trip-time (RTT) averages for 5G compared to Wi-Fi.................................52
Table 8 Revised target KPI values for H1C ..............................................................................58
Table 9 End-to-End latency measurements for different transmission protocols...............80
Table 10 Data reduction for full vs. resumed handshake for vital-signs patch use case, handshake only.........................................................87
Table 11 Data reduction for full vs. resumed handshake for vital-signs patch use case, full upload (H3A).................................................................87
Table 12 Energy gains of session resumption for vital-signs patch use case (H3A) ..........87
Table 13 Parameter settings for protocol comparison on Sequans GM02s (H3A) .............88
Table 14 Comparing the Nordic to the Sequans results (indicative) (H3A).....................91
Table 15 Firmware download energy & duration for different test sites (H3A).................102
Table 16 LPWAN LoRaWAN vs 5G NB-IoT comparison......................................................106
Table 17 Base Station location ...............................................................................................109
Table 18 Intelligence components performance validation results..................................127
Table 19 Assessment of total test emphasis of the different healthcare subcases in 5G-HEART. .................................................................129
Table 20 Progress-to-plan for final (phase-3) trials .............................................................130
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3GPP</td>
<td>3rd Generation Partnership Project</td>
</tr>
<tr>
<td>4G</td>
<td>4th Generation wireless systems</td>
</tr>
<tr>
<td>4K</td>
<td>3,980x2160 pixel resolution</td>
</tr>
<tr>
<td>5G</td>
<td>5th Generation wireless systems</td>
</tr>
<tr>
<td>5G-HEART</td>
<td>5G Health Aquaculture and Transport validation trials</td>
</tr>
<tr>
<td>5GTN</td>
<td>5G Test Network</td>
</tr>
<tr>
<td>5QI</td>
<td>5G QoS Identifier</td>
</tr>
<tr>
<td>6DOF</td>
<td>6 Degrees of Freedom</td>
</tr>
<tr>
<td>AES</td>
<td>Advanced Encryption Standard</td>
</tr>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>AoA</td>
<td>Angle of Arrival</td>
</tr>
<tr>
<td>AR</td>
<td>Augmented Reality</td>
</tr>
<tr>
<td>AV</td>
<td>Aortic Valve</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>AWS</td>
<td>Amazon Web Services</td>
</tr>
<tr>
<td>BEST</td>
<td>Battery Efficient Security for very low throughput MTC devices</td>
</tr>
<tr>
<td>CBRNE</td>
<td>Chemical, Biological, Radioactive, Nuclear and Explosive (materials/agents)</td>
</tr>
<tr>
<td>CCAM</td>
<td>Cooperative Connected and Automated Mobility</td>
</tr>
<tr>
<td>CCM</td>
<td>Counter with CBC MAC (AES)</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CDP</td>
<td>Customer Data Platform</td>
</tr>
<tr>
<td>CE</td>
<td>Coverage Enhancement</td>
</tr>
<tr>
<td>CHD</td>
<td>Congenital Heart Defect</td>
</tr>
<tr>
<td>CINR</td>
<td>Carrier to Interference and Noise Ratio</td>
</tr>
<tr>
<td>CIoT</td>
<td>Cellular IoT</td>
</tr>
<tr>
<td>CMO</td>
<td>Chief medical officer (Ambulance care)</td>
</tr>
<tr>
<td>CNN</td>
<td>Convolutional Neural Network</td>
</tr>
<tr>
<td>CoAP</td>
<td>Constrained Application Protocol</td>
</tr>
<tr>
<td>COTS</td>
<td>Common Off-The-Shelf</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Coronavirus Disease 2019</td>
</tr>
<tr>
<td>CPE</td>
<td>Customer Premises Equipment (5G modem)</td>
</tr>
<tr>
<td>CRLB</td>
<td>Cramer-Rao Lower Bound</td>
</tr>
<tr>
<td>CW</td>
<td>Continuous Wave Doppler</td>
</tr>
<tr>
<td>D2C</td>
<td>Direct-to-Cloud</td>
</tr>
<tr>
<td>DAS</td>
<td>Distributed Antenna System</td>
</tr>
<tr>
<td>DDoS</td>
<td>Distributed Denial of Service (attack)</td>
</tr>
<tr>
<td>DL</td>
<td>Downlink</td>
</tr>
<tr>
<td>DNL</td>
<td>Digital Navigation Link (Philips)</td>
</tr>
<tr>
<td>DNS</td>
<td>Domain Name System</td>
</tr>
<tr>
<td>DTLS</td>
<td>Datagram Transport Layer Security</td>
</tr>
<tr>
<td>EC</td>
<td>European Commission / Enhanced Coverage</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>EDT</td>
<td>Early Data Transmission (3GPP Release-15 feature)</td>
</tr>
<tr>
<td>EGD</td>
<td>Esophagastroduodenoscopy</td>
</tr>
<tr>
<td>eMBB</td>
<td>enhanced Mobile Broadband</td>
</tr>
<tr>
<td>EPC</td>
<td>Evolved Packet Core (4G)</td>
</tr>
<tr>
<td>FOTA</td>
<td>Firmware Over the Air</td>
</tr>
<tr>
<td>FP</td>
<td>False Positive</td>
</tr>
<tr>
<td>FPGA</td>
<td>Field-Programmable Gate Array</td>
</tr>
<tr>
<td>GDOP</td>
<td>Geometric Dilution of Precision</td>
</tr>
<tr>
<td>GDPR</td>
<td>General Data Protection Regulations (EU)</td>
</tr>
<tr>
<td>gNB</td>
<td>Gigabit Node B (5G)</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite System</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile Communications</td>
</tr>
<tr>
<td>H.264</td>
<td>MPEG-4 AVC (video codec)</td>
</tr>
<tr>
<td>HAS</td>
<td>HTTP Adaptive Streaming</td>
</tr>
<tr>
<td>HAZMAT</td>
<td>Hazardous Materials</td>
</tr>
<tr>
<td>HCP</td>
<td>Health Care Professional</td>
</tr>
<tr>
<td>HD</td>
<td>High-Definition</td>
</tr>
<tr>
<td>HSV</td>
<td>Hue, Saturation, Value</td>
</tr>
<tr>
<td>HTTP(S)</td>
<td>Hypertext Transfer Protocol (Secure)</td>
</tr>
<tr>
<td>HW</td>
<td>Hardware</td>
</tr>
<tr>
<td>IAM</td>
<td>Identity Access Management (AWS)</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communications technology</td>
</tr>
<tr>
<td>IEEE</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>IETF</td>
<td>Internet Engineering Task Force</td>
</tr>
<tr>
<td>IF</td>
<td>Intermediate Frequency</td>
</tr>
<tr>
<td>IoMI</td>
<td>Internet of Medical Implants</td>
</tr>
<tr>
<td>IoT</td>
<td>Internet-of-Things</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol / International Protection</td>
</tr>
<tr>
<td>ISM</td>
<td>Industrial, Scientific, and Medical (band)</td>
</tr>
<tr>
<td>KPI</td>
<td>Key Performance Indicator</td>
</tr>
<tr>
<td>LAA</td>
<td>LOS Analysis Area</td>
</tr>
<tr>
<td>LAM</td>
<td>Location Area Map</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>LoRa</td>
<td>Long Range</td>
</tr>
<tr>
<td>LOS</td>
<td>Line of Sight</td>
</tr>
<tr>
<td>LPWA</td>
<td>Low-Power Wide-Area</td>
</tr>
<tr>
<td>LTE</td>
<td>Long-Term Evolution</td>
</tr>
<tr>
<td>LTE-M</td>
<td>LTE Machine Type communication</td>
</tr>
<tr>
<td>LTS</td>
<td>Long-term Support (Ubuntu)</td>
</tr>
<tr>
<td>LVOT</td>
<td>Left ventricular outflow tract</td>
</tr>
<tr>
<td>M(V)NO</td>
<td>Mobile (Virtual) Network Operator</td>
</tr>
<tr>
<td>MAC</td>
<td>Message Authentication Code</td>
</tr>
<tr>
<td>MCL</td>
<td>Maximum Coupling Loss</td>
</tr>
<tr>
<td>MCU</td>
<td>Microcontroller Unit</td>
</tr>
<tr>
<td>MEC</td>
<td>Mobile Edge Computing</td>
</tr>
<tr>
<td>MF-PDoA</td>
<td>Multi Frequency-PDoA</td>
</tr>
<tr>
<td>MLE</td>
<td>Maximum Likelihood Estimation</td>
</tr>
<tr>
<td>M-mode</td>
<td>Motion mode</td>
</tr>
<tr>
<td>mMTC</td>
<td>massive Machine Type Communications</td>
</tr>
<tr>
<td>MP</td>
<td>Megapixel</td>
</tr>
<tr>
<td>MPEG</td>
<td>Moving Picture Experts Group</td>
</tr>
<tr>
<td>MPV</td>
<td>MPEG Elementary Stream Video</td>
</tr>
<tr>
<td>MQTT</td>
<td>Message Queue Telemetry Transport</td>
</tr>
<tr>
<td>MTC</td>
<td>Machine Type Communications</td>
</tr>
<tr>
<td>MV</td>
<td>Mitral Valve</td>
</tr>
<tr>
<td>NB</td>
<td>Narrow Band</td>
</tr>
<tr>
<td>NB-IoT</td>
<td>Narrow-Band IoT</td>
</tr>
<tr>
<td>NLOS</td>
<td>Non-Line of Sight</td>
</tr>
<tr>
<td>NR</td>
<td>New Radio (5G)</td>
</tr>
<tr>
<td>NSA</td>
<td>Non-standalone (5G architecture)</td>
</tr>
<tr>
<td>NUC</td>
<td>Next Unit of Computing</td>
</tr>
<tr>
<td>OBU</td>
<td>On Board Unit</td>
</tr>
<tr>
<td>OSM</td>
<td>Open Street Map</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computer</td>
</tr>
<tr>
<td>PKI</td>
<td>Public Key Infrastructure</td>
</tr>
<tr>
<td>PSM</td>
<td>Power Save Mode</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>PTP</td>
<td>Precision Time Protocol</td>
</tr>
<tr>
<td>PW</td>
<td>Pulsed Wave Doppler</td>
</tr>
<tr>
<td>QoE</td>
<td>Quality of Experience</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality-of-Service</td>
</tr>
<tr>
<td>QUIC</td>
<td>Quick UDP Internet Connections</td>
</tr>
<tr>
<td>RAI</td>
<td>Release Assistance Indication</td>
</tr>
<tr>
<td>RAN</td>
<td>Radio Access Network</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RGB</td>
<td>Red, Green, Blue</td>
</tr>
<tr>
<td>RPM</td>
<td>Remote Patient Monitoring</td>
</tr>
<tr>
<td>RRC</td>
<td>Radio Resource Control</td>
</tr>
<tr>
<td>RSRP</td>
<td>Reference Signal Received Power</td>
</tr>
<tr>
<td>RSRQ</td>
<td>Reference Signal Received Quality</td>
</tr>
<tr>
<td>RSSI</td>
<td>Received Signal Strength Indicator</td>
</tr>
<tr>
<td>RTP</td>
<td>Real-time Transport Protocol</td>
</tr>
<tr>
<td>RTSP</td>
<td>Real Time Streaming Protocol</td>
</tr>
<tr>
<td>RTT</td>
<td>Round Trip Time</td>
</tr>
<tr>
<td>SA</td>
<td>Standalone (5G architecture)</td>
</tr>
<tr>
<td>SIM</td>
<td>Subscriber Identity Module</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>SOC</td>
<td>System on Chip</td>
</tr>
<tr>
<td>SpO2</td>
<td>Peripheral Oxygen Saturation</td>
</tr>
<tr>
<td>SW</td>
<td>Software</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TDD</td>
<td>Time Division Duplex</td>
</tr>
<tr>
<td>TDoA</td>
<td>Time Difference of Arrival</td>
</tr>
<tr>
<td>TLS</td>
<td>Transport Layer Security</td>
</tr>
<tr>
<td>ToA</td>
<td>Time-of-Arrival</td>
</tr>
<tr>
<td>ToD</td>
<td>Time-of-Departure</td>
</tr>
<tr>
<td>ToF</td>
<td>Time of Flight</td>
</tr>
<tr>
<td>TR</td>
<td>Tricuspid regurgitation</td>
</tr>
<tr>
<td>TV</td>
<td>Tricuspid valve</td>
</tr>
<tr>
<td>UART</td>
<td>Universal Asynchronous Communication Protocol</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
<tr>
<td>UL</td>
<td>Uplink</td>
</tr>
<tr>
<td>UPF</td>
<td>User Plane Function</td>
</tr>
<tr>
<td>URLLC</td>
<td>Ultra-Reliable Low-Latency Communications</td>
</tr>
<tr>
<td>US</td>
<td>Ultrasound</td>
</tr>
<tr>
<td>USAR</td>
<td>Urban Search and Rescue</td>
</tr>
<tr>
<td>USB</td>
<td>Universal Serial Bus</td>
</tr>
<tr>
<td>VHDL</td>
<td>Very High-Speed Integrated Circuit (VHSIC) Hardware Description Language</td>
</tr>
<tr>
<td>VPN</td>
<td>Virtual Private Network</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual Reality</td>
</tr>
<tr>
<td>WCE</td>
<td>Wireless Capsule Endoscopy</td>
</tr>
<tr>
<td>WCS</td>
<td>Wireless Capture System</td>
</tr>
<tr>
<td>WebRTC</td>
<td>Web Real-Time Communication</td>
</tr>
<tr>
<td>Wi-Fi</td>
<td>Wireless Fidelity</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

This deliverable describes the phase-3 (final) trials of the healthcare vertical use cases of 5G-HEART. These phase-3 trials contribute to the milestone MS5 of the project.

Phase-1 was strongly focused on gathering clinical requirements, developing prototype systems, and verifying the feasibility using a number of legacy networks (4G, Wi-Fi, cable, LAN, etc.). In phase-2, the focus gradually shifted towards using 5G networks to validate the network’s capabilities to serve the use cases. In phase 3, the final solutions have been implemented and tested, to large degree using 5G technology.

Even though 5G networks are now being rolled out massively, implementing and testing advanced healthcare vertical use cases using 5G has not been trivial. Due to a combination of test platform instabilities and problems, and the effects of a two-year period of COVID-19 restrictions, a few subcases have not been tested over 5G. Instead, these have been tested either using commercial 4G/5G networks or alternative wireless and wired technologies. In those cases, the feasibility assessment of enabling those subcases over 5G is based on analysis of their network requirements compared to the 5G network capabilities as experienced during other experiments.

1.1 Use cases and phase-3 trials overview

Table 1 shows an overview of the nine healthcare subcases addressed by 5G-HEART, including their primary 5G enablers, the location of the final trials, as well as the involved parties. The emphasis in this document is on the complete technical description of the final solution for testing.

Clinical test cases and motivations are only briefly mentioned, the reader is referred to D3.2 [1] and D3.3 [2] for further details. Generally, only results from clinical and application trials are included in this document. The trial results related to 5G network performance are documented in D6.4 [3].

<table>
<thead>
<tr>
<th>Subcase</th>
<th>Primary 5G enablers</th>
<th>Location phase-3 trials</th>
<th>Subcase owner and partners</th>
<th>External collaborators phase-3 trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1A Educational surgery</td>
<td>eMBB</td>
<td>Oulu, Finland</td>
<td>VTT, RedZinc, Oslo University Hospital</td>
<td>Oulu University Hospital</td>
</tr>
<tr>
<td>H1B Remote ultrasound examination – CHD</td>
<td>eMBB, URLLC</td>
<td>Oslo, Norway</td>
<td>Philips, Oslo University Hospital, Telenor</td>
<td></td>
</tr>
<tr>
<td>H1B Remote ultrasound examination – Robot</td>
<td>eMBB, URLLC</td>
<td>Oslo, Norway</td>
<td>Oslo University Hospital, Telenor</td>
<td></td>
</tr>
<tr>
<td>H1C Paramedic support</td>
<td>eMBB coverage</td>
<td>Groningen, the Netherlands</td>
<td>TNO, Philips, RedZinc</td>
<td>Ambulancezorg Groningen, 5Groningen</td>
</tr>
<tr>
<td>H1D Critical health event</td>
<td>eMBB</td>
<td>Oslo, Norway</td>
<td>RedZinc, Oslo University Hospital, Telenor</td>
<td></td>
</tr>
<tr>
<td>H2A Automatic pill camera anomaly detection</td>
<td>URLLC</td>
<td>Oslo, Norway</td>
<td>Oslo University Hospital, Telenor</td>
<td></td>
</tr>
</tbody>
</table>
Subcase | Primary 5G enablers | Location phase-3 trials | Subcase owner and partners | External collaborators phase-3 trials
--- | --- | --- | --- | ---
H3A Vital-sign patch prototype | mMTC | Eindhoven, the Netherlands | **Philips**, Oslo University Hospital, Telenor |
H3B Localizable tag | mMTC localisation | Grenoble, France | **CEA**, **Philips** |
H3C Aquaculture Remote Health Monitoring | mMTC | Greece | **WINGS** | External Medical Expert |

### 1.2 Organization of this deliverable

The remainder of this deliverable is organized as follows. Each of the Chapters 2 -10 contains a detailed description of the phase-3 (final) solution for one of the healthcare subcases respectively, including a recap of motivations, clinical (test) cases where applicable, final test setups, testing and validation, and recommendations:

- Chapter 2: H1A – Educational surgery
- Chapter 3: H1B CHD – Remote ultrasound examination – Congenital Heart Disease
- Chapter 4: H1B Robot – Remote ultrasound examination – Robotics
- Chapter 5: H1C – Paramedic support
- Chapter 6: H1D – Critical health event
- Chapter 7: H2A – Automatic pill camera anomaly detection
- Chapter 8: H3A – Vital-sign patch prototype
- Chapter 9: H3B – Localizable tag
- Chapter 10: H3C – Aquaculture Remote Health Monitoring

Concluding remarks are given in Chapter 11.
2 SUBCASE H1A: EDUCATIONAL SURGERY

2.1 Description and motivation

The “Educational Surgery” subcase aims to find improved solutions for remote learning, consultancy, and remote attendance of surgical and other medical interventions, which can be supported by 5G technology. The main initial target of H1A starting from the project proposal was to provide a 360° video streaming platform to be tested and run on the 5G test network (5GTN), relying purely on the latest mobile 4G/5G connectivity. The advantage of having mobile flexibility for setting the educational video streaming session in mobile networking environment enables various healthcare use cases in different medical context ranging from the depicted educational surgery in fixed hospital facility, to the ambulance field scenarios possibly within a moving vehicle. All such scenarios have the common need; remote participation to the live streaming session either for learning the treatment procedure or even consulting in real-time, which requires adequate video quality and low network latency. During the era of COVID-19, such remote techniques have proven to possess valuable technology advantage.

The developed, tested, and experimented video streaming platform in 5GTN focused on delivering a real-time feed from an operating room to students in the classroom or home for educational purposes. During platform development, the findings have led to include single lens video streaming as well – even with multiple cameras – to reduce application latency and improve video quality in scenarios where dedicated camera stream from a certain area or angles can provide essential information for the end user. Such capability especially in terms of quality of experience is not often adequate according to the capabilities observed from the 360° equipment.

During phase-3, H1A has conducted two separate activities. Firstly, VTT has completed the 5G SA setup in 5GTN and compared, tested and refined the network configuration to have optimal support for this subcase. The focus has been especially on evaluating and configuring the uplink channel, since it defines the video quality originated from the sender’s camera from the field. Additionally, the live streaming setup has evolved more towards low latency streaming, which has clear benefits for several 5G use cases even at scenarios requiring also high throughput i.e., high video resolution and quality. Therefore, single lens cameras have lower capturing delay and are more optimised for low latency streaming for optimising the playback performance. Secondly, RedZinc, Oulu University Hospital with VTT as assistance role have carried out trials with the RedZinc wearable equipment. The aim has been in helping remote consultancy and education, especially during COVID-19.

In this chapter for H1A, we showcase the final setup and results both for the 5G SA evaluation conducted in the 5G TN laboratory in Oulu, as well as outcome from the Oulu University Hospital (OYS) trials. The comparative measurements done in phase-1 (LTE) and phase-2 (5G NSA) are reported in D3.2 [1] and D3.3 [2]. The results gathered at phase-3 do not only validate the streaming platform and KPI performance, but verbal assessment of the trial outcome is also provided from the Oulu University Hospital trials. Thus, the actual quality of service and -experience from video related experiments is easier to validate by the experience of the audience. Furthermore, the extensive laboratory results with live video streams with congested uplink revealed how network delay behaves when the uplink capacity reaches its limitation. The results especially from the SA networking are divided between this document and D6.4 [3].

2.2 Final setup

This section explains the final setup for the experiments and tests conducted during phase-3. The H1A has been divided into two parts during phase-3: 1) 5G SA tests utilizing the /single lens setup carried out in 5G facilities at VTT, and 2) Oulu University hospital trials utilizing RedZinc wearable camera equipment. The first trial mentioned has been focusing on a research-oriented approach while the latter one evaluates the developed system more in the actual hospital context.
2.2.1 Network architecture

Phase-3 contains SA capability in the core and indoor small cells of which one was deployed in the trials. Open5GS core was used in the evaluation, because its usage in band 78 together with the SA small cells provided better performance in terms of latency and throughput compared to the other core options, which are introduced in detail in D6.4 [3] alongside with more detailed description of 5GTN. The phase-3 emphasis was on evaluating the SA uplink channel performance against low latency streaming with congested network scenario. Furthermore, different network configuration parameters were tested during the evaluation focusing especially on the different variations for the data slot ratio, which defines the UL/DL share.

In the Oulu University trial with RedZinc equipment, commercial LTE and in-hospital Wi-Fi were utilized in the education premises for streaming the wearable video feed from doctor’s viewpoint to the BlueEye platform. Akamai cloud server in Stockholm was utilized as the video streaming platform through which the video clients (students) have access to the video stream from the hospital.

2.2.2 User application architecture

The application testing was carried out in 5GTN test facility, in Oulu. The architecture depicted in Figure 1 was deployed in the H1A scenario through the project. Different camera options were utilized in different phases to have solid evaluation and measurements. The injection of sensor data into the video stream as an overlay was tested but left out of the final setup since the use case did not include any actual sensor provider. Multicamera setup was introduced in phase-3 alongside with 360° camera for having dedicated, high-quality viewpoints for the medical operations. The detailed SA network evaluation in phase-3 was carried out using single lens cameras. During phase-2 the output quality of 360° camera was determined only as adequate level. The final demonstration taking place in November will include use of 360° as well as multicamera setup functioning together.

The hospital/ambulance operations include three main components: high quality camera capable for live streaming, video encoder for packetizing the video into transmittable format, and 5G SA capable access point for connecting the video encoder into network as well as providing near real-time transmission to the core network. The edge server in the core network was responsible for receiving the input stream,
transcoding the feed into predetermined representation, and for serving the output towards the video clients. During phase-3, mobile devices (smartphones and laptops) were used as the remote client sites to enable easier setup for measurements.

The wearable video camera setup used in the Oulu University Hospital trials follows similar overall approach depicted in Section 5. The wearable camera user is located at the hospital operational room with the BlueEye application running in the smartphone connected to the local Wi-Fi or mobile network. The video is streamed from the BlueEye application to the cloud server to be served to the clients through the platform.

2.2.3 Hardware components

The essential HW components regarding the 5GTN camera setup, network, and player equipment are listed in Table 2. The HW specifications for the wearable video system can be found from Section 5.2.3.

<table>
<thead>
<tr>
<th>Video Encoder inc. Camera</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Camera HW</strong></td>
<td>Logitech Brio 4K Ultra HD Pro</td>
</tr>
<tr>
<td><strong>Camera-encoder connection</strong></td>
<td>USB 3.0</td>
</tr>
<tr>
<td><strong>Encoder HW</strong></td>
<td>Intel NUC Core i7-8809G @ 3.1 GHz</td>
</tr>
<tr>
<td><strong>Edge server</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Streaming server HW</strong></td>
<td>Intel Core i9-9900K @ 3.60 GHz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Network</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mobile network</strong></td>
<td>5G SA</td>
</tr>
<tr>
<td><strong>EPC core</strong></td>
<td>Open5GS</td>
</tr>
<tr>
<td><strong>5G Small Cell</strong></td>
<td>Indoor gNB: Band n78 TDD @ 60 MHz</td>
</tr>
<tr>
<td><strong>5G Frequency</strong></td>
<td>3.5 GHz</td>
</tr>
<tr>
<td><strong>Access point</strong></td>
<td>TeleWell 5G USB modem</td>
</tr>
<tr>
<td><strong>5G QoS</strong></td>
<td>5QI= 9 has been used as specified in [4]</td>
</tr>
</tbody>
</table>

Table 2 H1A HW components.

2.2.4 Software components

The essential SW components regarding the video encoder and video player equipment are listed in Table 3. The SW specifications for the wearable video system can be found from Section 5.2.4.

<table>
<thead>
<tr>
<th>Edge server</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OS</strong></td>
<td>Ubuntu 18.04 LTS</td>
</tr>
<tr>
<td><strong>Streaming server</strong></td>
<td>rtsp-simple-server</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Video Encoder</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Encoder / streaming server OS</strong></td>
<td>Ubuntu 20.04 LTS</td>
</tr>
<tr>
<td><strong>Encoding SW</strong></td>
<td>FFmpeg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Video player</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UE OS</strong></td>
<td>Ubuntu 20.04 LTS</td>
</tr>
<tr>
<td><strong>Video player application</strong></td>
<td>MPV media player</td>
</tr>
</tbody>
</table>
2.3 Testing and verification

This chapter illustrates the final results from the application perspective according to the phase-3 trials.

2.3.1 Methodology

The evaluation methodology was divided into two parts. The first part evaluates the application performance and overall visual quality in terms of Motion-to-Photon (MTP) latency, which defines the reaction latency starting from the camera capture ending to visual display as a reflection to the movement. The second part was the main emphasis in the H1A and 5GTN, which was measuring the network KPIs in 5G low latency video streaming. These results are presented in detail in D6.4 [3]. There are several components in the low latency video streaming chain and each component adds a certain amount of delay to the chain and the sum of the delays forms the End-to-End (E2E) latency. These are visualized in Figure 2.

![Figure 2 The different components forming the experienced motion-to-photon (MTP) latency.](image)

We set up the pilot environment in VTT’s 5G laboratory in Oulu following the architecture in Figure 1 with more detailed version presented in Figure 5. Figure 3 shows the actual setup and devices in the pilot measurements for evaluating both the application performance (motion-to-photon) as well as network KPIs.

The single lens camera (Logitech Brio) in the front captures the digital clock with millisecond accuracy from the middle screen and outputs the raw video stream into Intel NUC mini-PC (Figure 4 (b)), which encodes the video into H.264 and encapsulating it into RTSP transport format. The encoder is connected to the 5G network (SA in phase-3) via USB modem (Figure 4 (b)), which interconnects with the 5GTN indoor 5G small cells (Figure 4 (a)) and video stream is sent via uplink to the edge server functioning as a video streaming server. This edge server is physically located in 5GTN core network. Finally, the video client (Figure 3, the device on the left) with MPV video player receives the video from the edge server and plays the content on a display. Additionally, we also had one laptop (right side on Figure 3) with Qosium network measurement tool for monitoring the network KPIs.

The actual evaluation occupied 1-2 VTT persons for assessing video quality to see possible momentary degradations, and to gather network KPIs from separate runs to get an averaged result.
2.3.2 List of key performance indicators

From the application perspective, only motion-to-photon (MTP) is reported here. The extensive results related to network KPIs are presented in D6.4 [3].

2.3.3 Measurement and testing tools

Figure 5 shows the detailed architecture from the evaluation trials. The roles of the different components were described earlier in Section 2.3.1. The role of extra components in this figure, such as iPerf3 related functionalities, PTP, and Qosium will be introduced in detail in D6.4 [3] since they are more related for testing and evaluating the network performance accurately.
We measured the user experienced E2E latency or motion-to-photon latency with the help of digital clocks with millisecond accuracy. This latency is the time it takes for an image that the camera captures from screen 1 to be encoded, transmitted through 5G SA and displayed on screen 2. On screen 1 we have a running time in ms which the camera captures. We used a Linux terminal command:

```bash
while true; do echo -ne "'date +%H:%M:%S:%N'"; done
```

to show the time on screen 1. The video player plays the streamed video on screen 2 that is placed next to screen 1 as shown in Figure 3. We recorded a slow-motion video at 240 FPS of the two screens using the modern smartphone camera application. The E2E latency is calculated from the recorded video using the frame by frame forwarding feature in the MPV player. From the recorded video we calculated how many frames it takes until the same time value identified on screen 1 is seen on screen 2. For instance, if it takes 40 frames when screen 1 time is displayed on screen 2 the end-to-end latency is 1/240 x 40 = 167 ms. Therefore, the precision of the measurement method is 4.2 ms (which equals for 1/240 FPS).

### 2.3.4 Final results

The baseline for the E2E user experience video latency was set as 200 ms, which is one of the common thresholds in video streaming [5]. Higher E2E latency usually degrades the subjective video quality, namely as quality of experience (QoE). With the initial solution and verification using the 360° camera over HTTP adaptive streaming (HAS), the E2E latency was over 500 ms in LTE without considering the video buffer [1]. In phase-2, we evolved the system by introducing single-lens cameras and 5G NSA, and we achieved slightly under 200 ms. However, in phase-2, the already only slightly congested UL channel caused problems in the 5G NSA network, which we resolved and improved in phase-3 trials with 5G SA using optimized video streaming solution and proper network configuration.

Figure 6 shows the final results regarding E2E latency using 5G SA, and LAN as a reference. Both UDP and TCP for the transmission were tested. The user experience is under 180 ms, which is only ~12 ms higher than corresponding LAN. As a conclusion, the difference yields from the network used.
Figure 6 User experienced motion-to-photon (E2E) latency for H1A trial evaluation for baseline LAN and 5G SA.

Figure 7 shows how different (in this case increasing) bitrate affects the experienced E2E latency using UDP and TCP with different UL/DL slot ratio in the 5G SA network configuration. With the video bitrates under 40 Mb/s both the 1/4 and 3/7 frame configuration result in packet lossless states, but for 1/4 ratio the latency exceeds 200 ms. As a conclusion, optimised UL/DL data slot ratio configuration enables higher data rates especially for the UL and guarantees satisfying QoE for the end users.

The Oulu University Hospital trials were carried out during several courses and took place in the department of Children and Youth Clinic. The questionnaire at the end of the courses provided the following feedback:

- **89%** of the 26 respondents agreed that BlueEye wearable video can be used for teaching or training that would otherwise not be possible and
- **82%** of the 27 respondents felt that BlueEye could replace traditional face-to-face teaching wholly or partly.
2.4 Recommendations

H1A reached its main target KPIs (5G UL/DL throughput & delays) from the network point of view and video streaming optimization for E2E latency was achieved during the project. The use of 360° camera does not necessarily provide decent quality in case sharp video quality is needed and therefore use of multiple single lens cameras may lead to better QoE also providing possibilities for lower E2E latency. Furthermore, the use of live streaming for educational and consultancy purposes can lower the boundaries between physical and virtual attendance and enable modern ways for remote attendance, which are important especially during exceptional times, such as COVID-19, when physical attendance is not either recommended or even possible.

According to the trials in H1A, utilization of mobile networks and especially 5G can enable a flexible set-up for live streaming. Naturally, UL is usually more restricted than e.g., in Wi-Fi networks, which means that live streaming from the field might need flexible video adaptation to the available channel. From the application perspective in the actual facilities, the ease-of-usability is important to have where the streaming can be started easily by the medical staff.
3  SUBCASE H1B CHD: REMOTE ULTRASOUND – CONGENITAL HEART

3.1 Description and motivation

Congenital heart disease (CHD) is rare amongst new-borns (0.8 %), the potential impact of poor or late diagnosis and hence wrong or late treatment can be huge for babies, their parents and indirectly society. A correct diagnosis at the local hospital is vital for timely treatment decisions. It is however not easy for regular paediatric cardiologists at local or regional hospitals to recognize CHD through ultrasound examination because they simply lack the experience of recognizing malformations on the echocardiographic images due to the low incidence and large variation of congenital cardiac malformations. Specialized paediatric cardiologists who are skilled in conducting and interpreting echocardiographic images can diagnose CHD quickly and reliably but in a country like Norway, there are only a hand full of them. The use of teleguided ultrasound technology, possibly further enhanced with immersive telepresence technology could enable local paediatric cardiologists to come to a correct diagnosis in case of suspected CHD under the guidance of a remote expert, without the new-born or expert cardiologist having to travel to conduct the examination. A correct diagnosis can help to initiate timely and appropriate treatment and can also avoid unnecessary travel of new-borns in case CHD can be ruled out. Especially in countries like Norway with a big part of the population living in rural areas, high-quality tele-sonography may reduce travel, improve time-to-diagnosis, and generally improve the quality of care, while lowering its cost. The preliminary overall set-up to investigate the use of teleguided ultrasound is elaborately described in D3.2 [1] and comprises the following individual set-ups each covering different aspects of the CHD use case:

**Teleguided ultrasound based on the commercially available EPIQ/Collaboration Live system**
- Technical evaluation: the 5G network setup to enable teleguided ultrasound is described in D6.4 [3], Section 3.6, under test case 2 and will therefore not be extensively described in this document.
- Usability/clinical evaluation: The aspects of the clinical exploration are described in this document under "EPIQ/Collaboration Live system".

**3D capturing and visualization with AR**
This is a parallel experiment to explore enhanced guidance of the ultrasound probe positioning through 3D capturing and visualization using augmented reality (AR) glasses and supporting technology to realize 3D telepresence. This exploration is described in this document under “3D capturing and visualization”.

**Digital Navigation Link (DNL) streaming for digital multi-stream ultrasound image sharing**
This experiment was focused on deriving network KPIs for multi-channel ultrasound data streaming, without clinical and usability evaluation. Because of the technology nature of this experiment, the setup and results are described in D6.4 [3], Section 3.6, test case 1 and only occasionally referred to in this document as “DNL streaming”.

3.2 Final setup

This section explains the final (phase-3) set-up of the experimental CHD use case. The various underlying setups related to CHD are in accordance with the earlier proposal in D3.3 [2] Section 3.2 and include the 5G-VINNI network in Oslo, the cart-based ultrasound solution for remote collaboration and the configurations for enhanced ultrasound streaming and 3D capturing and visualization. Note that especially in this final phase, Philips has leveraged several partnerships to be able to create sufficient critical mass especially around the further developments related to 3D telepresence and DNL streaming. For these, in addition to 5G-HEART, developments have taken place within the scope of the 5G-TOURS healthcare vertical (Rennes) and within a private cooperation between Ericsson, Vodafone and the local hospital and ambulance service in Eindhoven.
3.2.1 Network architecture

The two 5G testbeds that have been established in 5G HEART for the H1B CHD case focused on (1) experiments with the Philips EPIQ/Collaboration on the 5G facility in Oslo University Hospital (OUS) and (2) experiments with digital multi-stream ultrasound (DNL streaming) on the 5Groningen facility. Both setups and their related architecture are described in D6.4 [3].

For the 3D capturing and visualization implementation, different network variants have been identified. The required location flexibility for the Point-of-Care and Remote Expert side drives the specific network configuration since each side can be either connected to the 5G RAN or to the 5G Core. Upload of 3D (point cloud) data from the Point-of-Care location via the RAN is the most critical in terms of bandwidth (Figure 8).

The actual 3D telepresence demonstrator has been implemented with a wired network to serve as a lab/reference set-up, aiming to gain insights on usability rather than on network performance.

For this set-up, specific priority and quality settings were defined and trialled for the three different streams: 3D data (streaming of point clouds), ultrasound images (using WebRTC video) and probe position (WebRTC plain text). All data streams were set up as peer-to-peer.

3.2.2 User application architecture

EPIQ/Collaboration Live system

The user application architecture for the remote ultrasound set-up consists of a cart-based ultrasound machine [6] that can be equipped with the Collaboration Live software and service [7] to enable remote guidance during ultrasound examination. Besides audio/video communication and the streaming of ultrasound images to a remote expert, Collaboration Live provides remote control of the EPIQ through a UI that mimics the control panel of the machine itself (see Figure 9).
The configuration is extensively described in D3.3 [2] Section 3.2.2. and has not been subject to change for the clinical studies that continued in 2022.

3D capturing and visualization

One development focused on enhancing and enriching telepresence involves moving from 2D video to 3D video. This 3D telepresence architecture is described below. Figure 10 illustrates how the work settings are perceived by the users. The left side (a) shows the context of the local health care professional (HCP). The ultrasound display and virtual probe (in green) are augmented images. The right side (b) shows the context of the Remote Expert. The red elements are the 3D capture of the physical part of the Point of Care setting. An extra virtual probe and ultrasound display, in green, are added as virtual objects. The Remote Expert is seeing and interacting with all red and green elements in a virtual reality (VR) setting. Using this setup, the local HCP and the Remote Expert have a shared 3D view on the scene. The local HCP has a virtual ultrasound screen and a virtual probe displayed by the extended reality (XR) headset. Guidance with this virtual probe is given by the Remote Expert, who manipulates a virtual object in the virtual scene presented to him. Both can see the ultrasound image data and can place it in a convenient spot in the work environment.
Relative to the 2D collaboration use case, three major components change:

- The local HCP uses an XR headset, to enable spatial interaction. User interface elements such as video displays, clinical data displays and control elements of the Ultrasound equipment can be virtualized.
- The 2D video capture is replaced with 3D video capture. This is done with a set of depth cameras, allowing direct capture of a 3D image. By using multiple cameras, multiple sides of an object can be captured, to minimize occlusion.
- The Remote Expert views the scene with an XR or VR headset. He can interact with a virtual probe to provide guidance to the local HCP. He can place any additional information at his discretion in his work environment. Interaction can be done in several ways – this demonstrator uses handheld controllers.

This effectively creates a combined physical/virtual workspace. The local HCP works in a physical space which is captured in 3D. He uses an AR interface to inspect the ultrasound data as well as the visual instructions from the Remote Expert. The Remote Expert in turn sees the local scene in 3D, together with the ultrasound images. The hand controllers enable him to interact with virtual objects, some of which are reflected back to the local HCP. This closes the feedback loop between the two users.

The resulting hardware setup is shown in Figure 11 below. The Point of Care side setup consists of a 3D capture setup, an ultrasound probe with processing and connectivity, and an XR display for the local HCP. The Remote Expert side provides 3D visualization and interaction. For this, 3D capture equipment is added to the Point of Care equipment setup. Optimal viewing and interacting with 3D video put the Remote Expert in an XR or VR setup.

![Figure 11 Reference layout without 5G RAN](image)

On the Point of Care side, the 3D capture is done with multiple Microsoft Azure Kinect DK cameras recording both colour and depth information. The data from the cameras is fused to create a 3D reconstruction of the infant, the ultrasound probe, the hands of the practitioner and a part of the patient’s surroundings during the ultrasound exam. The use of multiple cameras addresses occlusion and viewpoint selection issues, removing the effort required from the local HCP to keep a proper view on the work area. The 3D perspective decreases the cognitive load required to understand the geometry of the situation of the local patient, as well as making the geometry easier to interpret and work with. The combination of multiple cameras requires calibration between the cameras. The resulting 3D representation is streamed to the Remote Expert.

The XR interface is implemented on a HoloLens 2. It visualizes the ultrasound video stream, as well as the virtual probe controlled by the Remote Expert. The virtual ultrasound stream can be placed at any desired location, enhancing the ergonomics of the local HCP. Proper visualization of the probe requires calibration of the XR device relative to the cameras.
Additional audio and video devices may be present but have not been included in the current test setup. On the side of the Remote Expert, the streamed 3D representation is rendered for the Remote Expert in 3D. This can be either XR, using a HoloLens 2, or VR, using an Oculus Quest 2, depending on the exact test or evaluation that is required. Both variants have been implemented. In both cases, a host PC renders the scene. The rendered images are streamed to the headset by holographic remoting (HoloLens) or over a tether (Quest).

For user testing, VR was used to facilitate the interaction, and minimize the training time needed to let new users operate the system. Remote rendering over holographic remoting also incurs a latency penalty that degrades the XR user experience substantially. Interaction in VR is done exclusively with the handheld controllers.

Next to the latency, additional timing aspects arise from the additional data streams and the interplay with the use case. As multiple streams are joined, these need to be synchronized. The ultrasound and 3D cameras have different capture rates. They also run on separate systems with no hardware synchronization between them. Synchronizing the channels properly can require an additional buffer, increasing the latency experienced by the end user. As the two users work together, the quality of the experience is affected by the roundtrip time, which now includes the local processing time as well (visualization, user reaction time, capture and reencoding). Jitter on all of these may require extra buffers, leading to further increase in end user latency. One of the goals is to get insight into the trade-offs and boundary conditions between these factors, and their impact on the usability of the system.

### 3.2.3 Hardware components

The hardware components used for the H1B CHD subcase are described below, covering both the basic set-up for teleguided ultrasound, as well as the set-up for enhanced telepresence using AR/VR:

**EPIQ/Collaboration Live system**
- Philips EPIQ (7G, hardware revision B.0) ultrasound machine
- Huawei 5G CPE Pro 2 (2x)
- USB camera (regular webcam)
- Laptop

**3D capturing and visualization**

*Local paediatrician*
- Azure Kinect DK (3x)
- HoloLens 2
- USB camera (any regular webcam, e.g., embedded in laptop or Logitech StreamCam)
- Lumify Ultrasound probe (USB). All three types are used: S4-1 broadband phased array, C5-2 broadband curved array and L12-4 broadband linear array
- Huawei 5G CPE Pro 2 or any other suitable device (3x) – in case of RAN deployment
- Laptops/PCs, minimal Intel Core i7

*Remote expert*
- Visualization – either:
  - Screen based
  - HTC Vive (VR headset)
  - HoloLens (AR headset)
3.2.4 Software components

The software components used for the H1B CHD subcase are described below, covering both the basic set-up for teleguided ultrasound as well as the set-up for enhanced telepresence using AR/VR.

**EPIQ/Collaboration Live system**
- Version 6.0 software for EPIQ with Collaboration Live enabled
- Reacts software on laptop and server in Canada (e.g., version 3.15.2.3630)

**3D capturing and visualization**
- HoloLens application for the local paediatrician
- 3D data processing/communication application
- Remote expert application

3.3 Testing and verification

The evaluation of the EPIQ/Collaboration Live system in this final stage has been a continuation of the network test and usability evaluation tracks that were already running last year based on this system. In addition, the further technology developments around multi-stream ultrasound data (DNL streaming) and 3D telepresence with AR have been shown to clinical and/or network partners for first evaluation.

3.3.1 Methodology

For the evaluation of the usability of remote ultrasound, two studies were initiated and conducted by OUS, involving medicine students for study 1. and neonates for study 2. Furthermore, the 3D telepresence been evaluated in a qualitative manner with clinical partners from the 5G TOURS and 5G Eindhoven consortia.

3.3.2 List of key performance indicators

At the beginning of the project, the user requirements associated with the CHD case have been converted into an initial set of network KPIs. These are listed in D2.2 [8], Section 4.2.1 (Table 20). The performance tests for the various elements (EPIQ/Collaboration Live, DNL streaming and 3D AR enhancement for immersive telepresence) that were conducted in the earlier phases of the project have resulted in updated insights on differentiated network requirements for each of the elements, as described in D3.3 [2], Section 3.3.2. The overview of updated network requirements for each of the elements is shown in Table 4 below. Measurements addressing selected KPIs for the digital multi-stream ultrasound and the EPIQ/Collaboration Live setup are reported in D6.4 [3].
Table 4 Target KPI values for H1B CHD

<table>
<thead>
<tr>
<th>Network requirements</th>
<th>initial target values</th>
<th>Target for EPIQ/Collaboration Live</th>
<th>Target for multi-stream US</th>
<th>Target for immersive AR/VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL throughput</td>
<td>High: 200 Mbps</td>
<td>Min: 6 Mbps</td>
<td>Min: 30 Mbps</td>
<td>Min: 600 Mbps</td>
</tr>
<tr>
<td>UL throughput</td>
<td>High: 200 Mbps</td>
<td>Min: 9 Mbps</td>
<td>Min: 60 Mbps</td>
<td>Min: 600 Mbps</td>
</tr>
<tr>
<td>Broadband connectivity/peak data rate</td>
<td>both DL and UL: 100 &lt; Medium ≤ 1000 Mbps</td>
<td>100 Mbps</td>
<td>100 Mbps</td>
<td>Min: 100 Max: 1000 Mbps</td>
</tr>
<tr>
<td>Latency requirements</td>
<td>5 &lt; Medium &lt; 25 ms</td>
<td>Max: 200 ms</td>
<td>Max: 20 ms</td>
<td>Max: 100 ms</td>
</tr>
<tr>
<td>Reliability</td>
<td>High: 99.99999%</td>
<td>Min/Max: 99 / 99,9999%</td>
<td>Min/Max: 99 / 99,999%</td>
<td>Min/Max: 90 / 99,999%</td>
</tr>
<tr>
<td>Mobility</td>
<td>Stationary</td>
<td>Stationary</td>
<td>Stationary</td>
<td>Stationary</td>
</tr>
<tr>
<td>Location accuracy</td>
<td>Low &gt; 25 m</td>
<td>Max: 25 m</td>
<td>Max: 25 m</td>
<td>Max: 25 m</td>
</tr>
<tr>
<td>Connection (device) density</td>
<td>Low: 5 devices</td>
<td>Max: 1 device/ m²</td>
<td>Max: 1 device/ m²</td>
<td>Max: 1 device/ 10 m²</td>
</tr>
<tr>
<td>Interactivity</td>
<td>100 &lt; High ≤ 1000 transactions/s</td>
<td>Min/Max: 100/1000 transactions/sec</td>
<td>Min/Max: 100/1000 transactions/sec</td>
<td>Min/Max: 100/1000 transactions/sec</td>
</tr>
<tr>
<td>Area traffic capacity</td>
<td>200 Mbps/m²</td>
<td>10 Mbps/ m²</td>
<td>100 Mbps / m²</td>
<td>200 Mbps/m²</td>
</tr>
<tr>
<td>Security/privacy</td>
<td>High: Confidential</td>
<td>Confidential</td>
<td>Confidential</td>
<td>Confidential</td>
</tr>
</tbody>
</table>

3.3.3 Measurement and testing tools

The enhancement of remote ultrasound with 3D capturing and visualization to enable immersive telepresence as explored in the joint effort for 5G HEART and 5G TOURS has been evaluated on technical feasibility and usability but since this was an explorative initial set-up, no network measurements have been logged.

3.3.4 Final results

Usability of the EPIQ/Collaboration Live system

The clinical studies on the use of teleguided ultrasound that were conducted in the earlier phases of the 5G HEART project as part of a PhD study within OUS have been further continued in 2022. The setup as described in D3.3 [2] Section 3.3.4 has remained unchanged. The number of study subjects has increased from eleven to thirteen students and from three to fifteen new-borns, respectively. These numbers are still too low to draw statistically significant conclusions, but they do provide some first directions.

Experiment 1: guided cardiac ultrasound (inexperienced students)

This experiment addresses the following research question:

Is it possible to conduct remotely guided ultrasound examinations of the heart with inexperienced students and how do the presently available features of Collaboration Live improve the quality and time of the examination?

Thirteen students have been included so far. For the first image of the ultrasound protocol, the students were divided into two groups. One group was guided through a smartphone videoconference call, while the other group was guided through Collaboration Live. The image quality that resulted was significantly better in the Collaboration Live group, but the examination time in the smartphone group was...
significantly shorter. The shorter examination time in the smartphone group might have led to lower image quality, therefore there might not be a clear-cut difference between both remote guidance methods. A study design with a set examination time to get the best image quality possible, would have been appropriate.

For the remainder of the examination, all students were guided by Collaboration Live, to obtain five more images. These images were compared with those obtained from manual echocardiography by an expert. The images resulting from remote guidance were of significantly lower quality than those of the expert. Even though only three out of the 52 images were rated as unusable, 60% of them were rated as of good quality. However, the mean examination time was six times longer for remote guidance, compared to manual echocardiography by the expert. We have also compared the ability to delineate specific cardiac structures, where the remote guidance images showed poor agreement to what structures the expert could clearly visualize. Lastly, we have compared measurement of left ventricular function between the images from the students and the expert. The results showed, approximately, only 3% deviation between remote guided students and expert images, which is acceptable in clinical use, though the variation coefficient is around 15%, falling short of the clinically acceptable standard for measurement variation.

Although one should be cautious in drawing conclusions, the results show a trend that remote guidance of novice echocardiographers could be useful to do crude evaluation of cardiac function.

Experiment 2: guided cardiac ultrasound on neonates (relatively inexperienced doctor)

This experiment addresses the following research question:

Is remotely guided ultrasound examination of the heart of new-born babies feasible? Is the image quality good enough to evaluate the necessity of transfer from a local hospital to a paediatric heart centre?

Fifteen new-borns from the neonatal intensive care unit of Oslo University Hospital (OUS) have been recruited in the experiment as patients. These neonates have been examined by a responsible clinician prior to the experiment and a treatment plan was already in place. Following this, a remote guidance team – consisting of a relatively inexperienced doctor and a remote expert paediatric cardiologist – has examined the patients with cardiac ultrasound, with the help of the remote guidance software. The remote guidance team assessed whether transfer of the neonate was needed and was blinded to the diagnosis prior to examination. Only in one case, the new-born was proposed to be moved to the paediatric heart centre, when it was not necessary. In the other fourteen cases, the clinician expert and the remote guidance team agreed with regards to patient transport. Out of the fourteen, six cases required need for transfer to a paediatric heart surgery centre and all six of these were identified in the remote guided examinations. The sample size is too small to make significant conclusions, but from our experience this method is worth further exploration and academic work as it seems that the remote guidance method can differentiate between critical congenital heart disease and other heart diseases/healthy hearts with good certainty. The evaluation of the parental experience has been good, with 100% of the participant’s parents, saying they would consider remote guided examination again.

Usability of 3D capturing and visualization

The findings from the 3D capturing and visualization evaluation are based on qualitative evaluation with clinical partners from the 5G TOURS and 5G Eindhoven consortia.

November 2021 experiments (Eindhoven)

In November 2021, two experienced paramedics and two emergency doctors were present in Eindhoven to test US image viewing in AR. The live US video stream was viewed in the HoloLens, on a fixed position in front of the local HCP. As connectivity requirements were low, local Wi-Fi was used for the US video stream. The local HCP was imaging an abdominal phantom, as shown in Figure 12. The Remote Expert had access to both the US stream, as well as to a video feed from the HoloLens 2 main camera. Participants switched roles between local HCP and Remote Expert.
The test case covers a Remote Expert giving guidance to the local HCP. In this initial test, the local HCP is asked to give an impression on the general use of the virtualized US image and the communication thereof.

Figure 12 Viewing US images with the HoloLens 2. Holding up the hand helps improve contrast on the images

The following feedback was obtained:

- The size of the ultrasound image was appreciated. It was perceived as three times as big compared to a laptop.
- The camera of the HoloLens was better, and there was less glare visible.
- Contrast on the ultrasound image was lacking when looking at a bright background. A dark backing improves this, as illustrated in Figure 12.
- The probe was almost never in the camera frame of the HoloLens. The ability to reposition the US video feed might improve that.

**April 2022 experiments (Eindhoven)**

In April 2022, another test was run in Eindhoven on the 3D Telepresence/AR setup in fully wired configuration, as illustrated in Figure 11. The US probe was not active, as there was no phantom to capture from. Instead, a dummy US image loop was played. The 3D video capture was done with two Azure Kinect cameras capturing at 640x576.

Two experienced emergency responders, both with experience in doing ultrasound exams, were alternatingly taking the roles of local HCP and Remote Expert. The task they were asked to execute is for the Remote Expert to guide the local sonographer in positioning the probe. Figure 13 shows the participants executing the test. The participants were close enough to each other that they could conduct a normal voice conversation. Figure 14 shows the interface on the PC mirroring the expert view in the VR headset.
The following feedback was obtained:

- The 3D capture detail is currently insufficient. On a scale from 1 (unreadable) to 10 (Excellent), it was rated a 5.
- The marker on the side of the virtual ultrasound probe should be clearly visible. This refers to the orientation markers (notch and line) that are placed on one side of the physical device.
- Positioning the virtual probe was easy for the participants.
- The registration of the probe in the AR glasses was incorrect, making absolute guidance impossible. However, relative movements were usable.
- It would be useful if also the patient could look along.
Discussion on deployment of a 3D telepresence solution

After the evaluation, there was a short discussion on adoption of a telepresence system such as this one. The reimbursement system in the Netherlands is judged as not ready for remote assistance. In the current system, the transport is only fully reimbursed if the patient arrives at hospital i.e., the patient must physically travel by ambulance to the hospital. For patients that are treated on site, only a token compensation is awarded.

The opportunity for support with 3D telepresence is to prevent incoming patients in the hospital by being able to deliver patient care on site. This should result in more efficient care. However, since hospitals do not get properly compensated for remote assistance, they would get a cut in income. Also, allocating staff to provide remote support and work in the field would result in a reduction of people working in the hospital, resulting in more work pressure at the hospital.

May 2022 experiments (Rennes)

In May 2022, a usability evaluation was done at the BCOM premises in Rennes, using a 3D Telepresence/AR setup in fully wired configuration as illustrated in Figure 11. An abdominal phantom was available as ultrasound target, allowing for live US images. Attempts to run the system over the RAN at BCOM (a 5G TOURS facility) failed. The test was executed with three Azure Kinect cameras capturing at 640x576.

An experienced emergency responder, with extensive experience in doing ultrasound exams, was guiding a non-expert. To allow an evaluation of the local HCP interface, the roles were also switched. The Remote Expert was asked to guide the local sonographer in positioning the probe. Figure 15 and Figure 16 illustrate the execution of the test. An abdominal phantom was available as ultrasound target.

![Figure 15 HCP exploring the local role](https://b-com.com/en)
The following feedback was obtained:

In Rennes, the usability of this enhanced set-up was deemed good, although the technical maturity of the 3D telepresence is relatively low, the potential added clinical value for both local HCP and Remote Expert is recognized. A few details on use cases were discussed:

- With the fixed cameras, this system could be installed inside ambulances. Not all ambulances need it though, so a flexible mounting system would be needed. The system should be set up such that the staff does not need to be involved with calibration or configuration.
- Other possible applications are education and operating rooms. The 3D viewing enables demonstration of physical gestures and movements that are hard to convey properly in 2D.
- Nowadays, many ORs are already fitted with cameras for other purposes; an enhanced camera system may easily allow addition of the 3D telepresence capability.

Conclusions on usability

It should be noted that the feedback gathered from clinical partners during the various experiments related to remote ultrasound is too limited in volume to be of statistical significance, yet one may conclude with some caution that there seems to be appreciation for teleguided ultrasound. The enhancement with 3D capturing and visualization is well appreciated as a concept. The current execution of showing the probe in AR to obtain higher accuracy in probe guidance is however technically not mature enough to really meet usability expectations and requirements. The visual quality/resolution of the 3D probe representation is too low and there many operational steps required to get the setup to work.

As a sidenote, it was also mentioned that implementation of teleguided ultrasound will require modification of the existing clinical workflows, which will require organizational commitment from the healthcare institutions and their willingness to change established practices.

### 3.4 Recommendations

The experiments have made clear that the suitability of a particular 5G network configuration strongly depends on the chosen technologies at user/application level. The EPIQ/Collaboration Live implementation at OUS that is built on a particular instantiation of WebRTC uses data and image compression techniques that do not put a high demand on network performance and can even very well
be supported with today’s 4G networks. From that perspective, 5G remains a promising, but not critical, enabler for remote ultrasound support.

However, digital multi-stream ultrasound image and data sharing – as has been explored with the DNL experiments – clearly requires a true 5G network configuration, with best performance when the expert application and the WebRTC services are running on the edge (see D6.4 [3]).

At the other end of the scale of 5G enabled applications, the 3D telepresence experiments that use 3D image capturing and AR visualization techniques bring along bandwidth and latency requirements that exceed the network performance demonstrated in the 5G-VINNI and 5Groningen networks. At the same time, the experiments made clear that these kinds of immersive technologies themselves are still insufficiently mature, even though they have the potential to further enhance the clinical collaboration experience and fully exploit the power of B5G/6G.

We have also observed that clinical partners are positive yet cautious in their reflections on possible adoption of remote collaboration technologies. Developments in these fields may therefore be driven by other industries than healthcare.

Overall, we see that for teleguided ultrasound, a full-fledged immersive application along with the supporting 5G (or beyond) network configuration both have a long way to go before they will reach maturity. At the same time, the first instantiations for remote collaboration are already technically feasible on today’s LTE networks. It is therefore recommended to follow a paired step-by-step approach whereby network and application maturity in other industries may be the pointers for future opportunities in healthcare.
4 SUBCASE H1B ROBOT: REMOTE ULTRASOUND – ROBOTICS

OUS and Telenor are collaborating to investigate the feasibility of robotic-assisted remote ultrasound examination which is targeted at cardiac sonography for adult patients. The research focuses on developing subsequent technical setups from both the technical and clinical standpoint.

4.1 Description and motivation

Ultrasound of the heart is a complex task, demanding substantial experience in the healthcare professional to correctly perform and interpret the examination. In smaller healthcare centres, the availability of these trained and experienced healthcare professionals is scarce, especially after workday hours. Patients with signs and symptoms warranting ultrasound examination of the heart, appear at healthcare centres across the globe at a steady pace. This creates a problem of meeting a demand for ultrasound examinations of the heart compared to the supply of healthcare workers competent of properly performing the examination. With this background the objective of this use case is to develop a teleoperated robotic system in term of a master-slave configuration for performing comprehensive ultrasound examinations of the heart, with the controlling expert sonographer sitting at a remote place, manipulating the robot via 5G connection.

The robotic system has been developed as it is specified in D3.2 [1] over ethernet connection. Testing the developed robotic platform presented the following challenges.

First, the visualization of the ultrasound probe over the patient's body has been researched to determine the best course of action. We evaluated the system's performance for the visualization using three 2D cameras (to provide a perspective visualization), a wearable camera, and a 360° camera. We have concluded that using three 2D cameras (it is possible to use two cameras as well) is the best method for visualizing the probe over the patient's body, because it provides the best flexibility to focus on the probe location and orientation.

The second challenge was the dexterous workspace of the robotic system to satisfy the cardiologists demands during the comprehensive examination of the heart. To address this challenge, three ultrasound probe holders were 3D printed, and the holder with a 30° angle off the axis of the robot's end-effector provided the highest performance.

Finally, because the robot is directly operating the examination over the patient's body, a comprehensive risk and ethical assessment was required to get approval from The Norwegian Medicine Agency3 (NOMA) to enable us using the system on healthy volunteers. We conducted a very thorough risk assessment and filed the necessary paperwork. Finally, NOMA authorized testing the system on 23 healthy volunteers to assess the efficacy of the robotic system in taking ultrasound images of hearts from diverse perspectives.

D3.2 [1] Sections 3.1.2 and 3.1.2.1 provide a detailed discussion of the motivation and business validation, respectively.

4.2 Final setup

The robotic system has been developed as it is specified in D3.2 [1], section 3.2.3.

---

3 https://legemiddelverket.no/English
4.2.1 Network architecture

The network architecture for final verifications is the same as for the H1B CHD case, and is described in D6.4 [3].

4.2.2 User application architecture

In this use case, the developed robotic solution is classified as a bilateral teleoperation system which satisfies the requirement of telepresence to capture ultrasound images in both short and long-distance teleoperated system paradigm. The teleoperation system primarily involves human-in-the-loop control, which means that a human operator (in the H1B subcase a cardiologist) controls the slave robot from the master site. The system consists of three main parts: Master site, Slave site and Communication link as shown in Figure 17.

![Ultrasound robot system architecture](image)

Figure 17 Ultrasound robot system architecture

The Master site can be viewed as the central hospital that houses the expert cardiologist, master robot, and monitoring system for the slave sites. The Slave site can be considered as a remote area that houses the patient, the slave robot, the ultrasound machine, the local doctor and/or the medical assistant and the medical monitoring devices. The connection between the master and slave sites consists of three sub-links. The sub-link with the highest priority is responsible for streaming the real-time ultrasound image, while the sub-link with the lowest priority is responsible for streaming video/audio between master and slave sites. Another data link is responsible for establishing a bidirectional connection between the haptic device and the slave robot.

The primary components of the robotic architecture are hardware architecture and software architecture. Each piece of hardware used in the robotic system's design has its own software layer. The link between these layers and the various layers of the control algorithm is managed by the software architecture. D3.2 [1] Sections 3.2.3.1 provide the detailed information about the user application architecture. In the clinical experiment, the real-time control data and the High-level management data were transmitted over the Ethernet level, the ambient video/audio streamed over the wireless hospital network and the ultrasound video transmitted over a HDMI cable.

We are interested in evaluating the performance of the robotic system in terms of round-trip time (RTT) when a UDP connection is established between the master and slave sites via 5G. For comparison, measurements over local Wi-Fi connectivity are also done as a baseline.

Figure 17 illustrates the experimental setup for robotic ultrasound examination over Telenor’s 5G test network, where the master and slave sides are in two different rooms at OUS. The master site is in room 1 with the haptic device and master site controller, and the master site's controller is connected to CPE
1. The slave site is in room 2 with the slave robot, ultrasound machine, and slave site controller, where the slave site's controller is connected to the CPE 2.

Figure 19 shows the comparison setup for measuring over local Wi-Fi, where one of the CPEs is used only as a Wi-Fi access point and router.
4.2.3 Hardware components

The hardware components of the ultrasound robot setup consist of:

- Slave manipulator: a UR5 robot arm manipulator with 6 degrees of freedom (DoF) and 5 kg payload.
- Slave Robot internal controller: The controller box has digital and analog inputs/output connected to the main controller through an ethernet connection.
- Slave site’s main controller: A PC running Linux 2.6.38 and Xenomai 2.6.0.
- Force torque sensor: The Gamma SI-65-5 from ATI Industrial Automation with 6 DOF.
- Ultrasound machine: A GE Vivid E 95.
- Video conferencing cameras.
- Haptic device: Phantom Omni” from Sensable Technologies with 6 DoF.
- Master site main controller: A PC running Linux 4.9.90 and Xenomai 3.1.

4.2.4 Software components

The following software has been used:

- LabVIEW Interface
- Robotic Controller
- Haptic Device Controller
- Xenomai real time framework 3.1 and 2.6.0
- Linux Kernel versions 2.6.38 and 4.9.90

4.3 Testing and verification

The robotic system has been tested both from the clinical perspectives and verified to work over 5G. The clinical evaluation of the remote ultrasound robot was done by assessing the ability to create high quality images of the heart, depicting important cardiac structures, and making precise measurements compared to gold standard examination, used in clinical working life. System verification over 5G has been done by measuring round-trip latency as seen from the master side.

4.3.1 Methodology

Clinical trials for assessing examination quality

23 male healthy volunteers were included into the study. All volunteers were first examined using the remote-controlled robot by one doctor and then a reference examination was done by another doctor.

A simplified but complete protocol of the heart examination was performed. The study protocol contained 15 images: Parasternal long axis view 2D and colour doppler of the aortic valve and of the mitral valve (1 and 2), parasternal short axis view 2D at the mitral valve/papillary muscle level (3) and M-mode (motion mode 4), apical four chamber view with left and right ventricle focussed 2D image (5 and 6), colour doppler of the mitral valve and tricuspid valve (7 and 8), pulsed wave (PW) Doppler measurement of the mitral valve (9), continuous wave (CW) Doppler interrogation of the tricuspid regurgitation (10), apical 5 chamber colour interrogation of the left ventricular outflow tract and across the aortic valve (11), PW and CW of the left ventricular outflow tract and the aortic valve (12 and 13), sub xiphoidal view abdominal aorta long axis colour doppler and pulsed wave Doppler (14 and 15).

The examinations were limited to 45 minutes. All measurements were means calculated from three consecutive beats. The diastolic measurements of the M-mode were done on the top of the R waves of the ECG and the systolic measurements done on the point of smallest systolic left ventricular cavity. The M-mode measurements were all quality checked by a consultant cardiologist (HB). The image
quality was scored by in a double-blind study fully trained echo-technicians from the paediatric echo-lab of Oslo University Hospital. The images were scored by questionnaire on a scale from 0 to 3 where 0 is “not usable quality”, 1 is “poor quality”, 2 is “moderate quality” and 3 is “good quality”. There were also a set of dichotomous questions whether the images could be used to evaluate function of valves, identify significant valvular regurgitation, obstruction, or measure velocity of the tricuspid regurgitation. All projections were timed from when the probe touched the skin to when the image was stored on the ultrasound machine.

Following the reference examination, the participants were asked if they experienced any post robotic examination tenderness the first day after examination. They were also asked to compare the probe pressures from the two examinations and tell if they were equal or if there was exerted more pressure during one of the examination types.

The clinical tests were done via wired connection because the 5G testbed was unavailable at the time. The software made to manipulate the robot worked without delay.

**System round-trip latency performance measurements.**

The operator at the slave site (room two) configures high-level management data such as robot initialization and the mode of the slave’s robot controller using a LabVIEW user interface. To achieve the hard real-time requirements outlined in D2.2 [8], the master site control loop in this setup runs at 700 Hz, while the slave site controller operates at 100 Hz. These two controllers are connected by a UDP connection across the network, which also serves as the real-time data link between them. To measure the round-trip latency, a UDP server is implemented in the master site controller, while the client is implemented in the slave site controller. The server’s current time (master controller) is sent to the client as part of the robotic real-time data (manipulation commands) through the forward link, and the same timestamp is included in the force-feedback data and sent back to the server via the backward link. When a timestamp is received by the server, it is compared to the master controller current time to determine the RTT.

### 4.3.2 List of key performance indicators

The following application KPIs has been evaluated:

- Precision and variation coefficient of the measured left ventricular shortening fraction.
- Proportion of important cardiac structures sufficiently imaged for anatomical and functional evaluation.
- Image quality.
- Precision of measured velocities of blood flow.
- Examination time
- Patient experience
- Round-trip (end-to-end) latency

### 4.3.3 Measurement and testing tools

The ultrasound probe is a GE M5Sc-D phased array 1.5 MHz to 4.6 MHz probe connected to a GE E95 ultrasound machine (GE Healthcare, Boston, USA) used to create the ultrasound images. The measurements were done using EchoPAC software (GE Healthcare, Boston, USA).

The slave site controller is running the Xenomai real-time framework in order to have direct access to the system hardware, and the UDP client and the slave site controller is executing in separate threads to reduce the latency. Likewise, the UDP server and the master site controller are both running in different threads.
4.3.4 Final results

Clinical trials for assessing examination quality

Left ventricular shortening fraction

Table 5 lists the results from measuring the M-modes acquired by the robot compared to the reference examination.

Table 5 Mean fractional shortening measured on M-modes acquired by robot compared to reference examination.

<table>
<thead>
<tr>
<th></th>
<th>Mean (CI)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robot FS (%)</td>
<td>34.0 (32.2-35.9)</td>
<td>4.2</td>
</tr>
<tr>
<td>Reference FS (%)</td>
<td>36.7 (34.2-39.2)</td>
<td>5.6</td>
</tr>
<tr>
<td>Difference (%)</td>
<td>2.7 (P=0.0599)</td>
<td>6.3</td>
</tr>
</tbody>
</table>

The standard deviation of the difference between robotic fractional shortening (FS) and reference FS was 6.3 with a range of deviation between -15.4 % and 6.4 %, corresponding variation coefficient of 17.2 % (relative value). Fractional shortening from robot-acquired M-mode images were 2.7 % lower compared to the reference, barely missing statistical significance. The spread in differences between left ventricular fractional shortening from the robotic examination compared to the reference is shown in a Bland Altman plot, Figure 20.

Assessability of cardiac structures

The proportion of cardiac structures depicted with usable quality for assessing anatomy and function is shown in Figure 21. Two independent sonographers evaluated the assessability of cardiac structures. In the robot examination 22 % of total cardiac structures were evaluated not assessable by both sonographers, compared to 5.543 % in the reference case.
Ultrasound image quality

Image quality for the robotic examination was a little above 1.5 and the reference examination had an average image quality of a little over 2.5, see Figure 22.
Blood flow velocity measured by Doppler interrogation:

The Doppler velocities measured during the echocardiographies show mean angle differences higher than 25° for all measurements, except from E wave velocity of the mitral valve, which did not reach statistical significance. The results are shown in Table 6.

Table 6 Blood flow Doppler velocities measurements with robot and reference examination.

<table>
<thead>
<tr>
<th>Doppler Measurement</th>
<th>Robot (m/s)</th>
<th>Reference (m/s)</th>
<th>Difference (m/s)</th>
<th>Corresponding angle(°)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MV CW E</td>
<td>0.76 (0.70-0.83)</td>
<td>0.77 (0.70-0.84)</td>
<td>0.004 (-0.056-0.067)</td>
<td>9.2</td>
<td>0.87</td>
</tr>
<tr>
<td>MV PW A</td>
<td>0.44 (0.41-0.47)</td>
<td>0.49 (0.43-0.53)</td>
<td>0.045 (0.001-0.09)</td>
<td>26.1</td>
<td>0.06</td>
</tr>
<tr>
<td>LVOT PW</td>
<td>0.96 (0.88-1.03)</td>
<td>1.07 (0.99-1.15)</td>
<td>0.11 (0.06-0.17)</td>
<td>26.2</td>
<td>0.0005</td>
</tr>
<tr>
<td>AV CW</td>
<td>1.06 (0.98-1.14)</td>
<td>1.25 (1.17-1.33)</td>
<td>0.19 (0.14-0.24)</td>
<td>32.0</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Examination time:

Robotic echocardiography was 1.92 times slower compared with manual echocardiography (p=0.0000). Mean robotic examination time was 26.4 minutes and manual echocardiography 14.1 minutes. None of the examinations reached the time limit of 45 minutes, the longest robotic examination lasted for 43.9 minutes.

Patient experience:

Two participants rated physical probe pressure during the examinations as similar, one participant rated the pressure during the reference examination to be heavier than with the robot. 18 participants rated pressure from the robot to be greater compared to the reference examination and two participants reported tenderness the day after examination.

Concluding remarks from preliminary clinical testing:

The robotic system for performing remote controlled ultrasound was tested on healthy male volunteers and the result is thus, in a global medical perspective, not widely generalizable. To evaluate the technical system itself it is a more suitable test design. Compared to earlier prototypes of robots used for remote ultrasound examination of the heart this robot has a big dexterous workspace, although it could not completely cover all the examination locations on the volunteer’s chest, without the volunteer moving on the bench. The robot base also needed manipulation along the vertical axis to reach the different locations and angles necessary for complete examination. We had some difficulties with the robot going
into singularity mode, sometimes warranting a reset. The full examination could still be performed in all cases, except for one case, where the examination was completed the following day.

The force feedback worked quite well, and all the volunteers thought that the robot examination was bearable. One volunteer stopped the robot once due to uncomfortable pressure being applied.

The clinical evaluation of the robotic system shows that it did not reach the quality of the reference exam. The measurement of fractional shortening showed variation coefficient of above 15 %. Clinically acceptable variation coefficient is 5 % to 10 %. The assessability of structures of approximately 80 % will not be deemed appropriate as a clinical test.

In the light of this being a preliminary clinical test and the potential for further development of the system and adding additional improving features, it is still promising.

**System round-trip latency performance**

Ten experiments were conducted to assess the latency performance of the robotic system in the teleoperation configuration across a UDP connection utilizing 5G or Wi-Fi networks. Five trials were conducted over the 5G test network, while five others were conducted over Wi-Fi. For each experiment, we manipulated the robotic system and recorded the round-trip delay for the first 10000 packages. The average round trip latency for the five experiments using 5G compared to the five experiments using Wi-Fi is shown in Figure 24. Table 7 provides the results for the experiments conducted using Wi-Fi and 5G in terms of the maximum, minimum, and average latency. The target latency requirement was defined in D2.2 [8] to be between 5 ms and 25 ms, while the average 5G latency round-trip latency was measured to be 40.05 ms. Target latency is a one-way latency definition, while RTT is two-way, we can conclude that the system verification has met the target.
Table 7 Round-trip-time (RTT) averages for 5G compared to Wi-Fi.

<table>
<thead>
<tr>
<th>Connection</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Average</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wi-Fi</td>
<td>24.69 ms</td>
<td>1.62 ms</td>
<td>2.92 ms</td>
<td>1.61 ms</td>
</tr>
<tr>
<td>5G</td>
<td>73.7 ms</td>
<td>26.81 ms</td>
<td>40.05 ms</td>
<td>4.73 ms</td>
</tr>
</tbody>
</table>

4.4 Recommendations

This use case has developed a system for remote controlled robotic ultrasound examination. Looking at the clinical results the system did not match up to the quality of a manual, reference exam. Having in mind that this was the first clinical test of a prototype, we managed to create many good and usable ultrasound images of the heart. The manipulation of the robot worked well, but the dexterous workspace and ability to make small positional adjustments was a challenge. A well performed ultrasound examination of the heart is all about the small and precise movement of the probe to get the best acoustic window and image quality to make the correct assessment of the patient.

Final verification over Telenor’s 5G test network has demonstrated the feasibility of the complete end-to-end system, also showing that it is possible to achieve low latency, sufficient to provide a real-time experience for the users.

Further research and development are certainly warranted and needed.
5 SUBCASE H1C: PARAMEDIC SUPPORT

5.1 Description and motivation

Ambulance services are striving towards healthcare with higher quality and possibly better cost-effectiveness. Increasingly more and more ambulance services wish to improve pre-hospital triage, and for this purpose it is important to make correct and timely decisions in emergencies [9].

The Chief Medical Officer (CMO) of an ambulance service may be remotely consulted for decision-making in rare or difficult emergency situations. At present, this is only possible via an audio connection (telephone) or through a dedicated communication channel via the dispatcher. The CMO has thus no visual feed of the patient information for more effective assessment. Therefore, the ambulance crew often makes conservative decisions (over-triage), propagating the risk and confidence in diagnosis/treatment to the emergency department. This results in a high percentage of cases that are brought to hospital while treatment in situ would suffice [10]-[12], resulting in lower availability of emergency services for more acute cases as well as unnecessary costs and burden on the healthcare system. If a paramedic can use wearable video-audio and share a patient’s vital parameters or even ultrasound images, the remote CMO can see the same patient context that the paramedic sees, help accelerate diagnosis and improve decision-making. This will lead to accelerated patient treatment, improved outcomes, reduced travel, cost savings and higher efficiencies.

In the phase-1 trials (see D3.2 [1]), we have validated the H1C subcase using a legacy 4G network (see D3.2). In the phase-2 trials (see D3.3 [2]), we have further validated the H1C subcase using an indoor 5G SA network. We have concluded that promising network and application-level performance of the used 5G SA network partially originates from its good deployment and operation conditions: an indoor network in a largely isolated warehouse, and with little competing mobile traffic.

This chapter includes a description of the final (phase-3) trials performed by 5G-HEART partners TNO and RedZinc, in a more realistic ambulance scenario. The phase-3 trials took place in Helmond, the Netherlands, using an outdoor 5G SA network of 5Groningen. In the final (phase-3) trials, we combined H1C and T2S1 (see D4.4 [13]) subcases in validating how both subcases can jointly empower ambulance services from both medical and transport perspectives, facing competition of regular data traffic in using the radio network resource. The single 5G SA network was used to support three network slices, one for video traffic (H1C), another for Cooperative Connected and Automated Mobility (CCAM) traffic (T2S1) and a last one for general purpose traffic (regular data). This exemplifies how a single 5G network may simultaneously support multiple (vertical) applications via network slicing.

5.2 Final setup

This section explains the final (phase-3) setup of the experiments involving both H1C and TS21 use cases, with more focus on the relevance for the H1C subcase. Aspects which are more relevant for the TS21 use case are addressed in D4.4 [13].

5.2.1 Network architecture

The used 5G SA network consists of the following major components:

- An operator-grade outdoor 5G SA base station with carrier frequency of 3.65 GHz (TDD), a maximum carrier bandwidth of 100 MHz, TDD downlink/uplink frame structure of DDDSU (“S”-slot setup 10:2:2) and sub-carrier spacing of 30 kHz. The base station is deployed nearby an intersection of the highway A270 in Helmond, the Netherlands, with two directional antennas which aim to optimize network coverage along the highway.

- A local edge server, located in Helmond, which hosts a User Plane Function (UPF) of the 5G core network and delay-sensitive applications (e.g., the CCAM applications for the T2S1 subcase).
• The remaining of the 5G core network, located remotely in TNO office in the Hague, the Netherlands.
• (RAN) Network slicing. Network slicing is a technology used to create multiple virtual networks within a given physical network, which can extend across both the Radio Access Network (RAN) as well as the core network for end-to-end QoS provisioning. In this study, we focus on RAN network slicing implemented in the used 5G SA base station. By introducing RAN slicing to the network, the network operator has direct control over the radio resource distribution between configured slices. This enables the guarantee of certain QoS and capacity requirements posted by certain applications, even in an overloaded network.

Figure 25 shows the 5G SA network architecture used in the phase-3 trials of H1C. More description of the used 5G SA setup is given in D6.4 [3].

Figure 25 5G SA network architecture (H1C, phase-3)

5.2.2 User application architecture

Figure 26 shows a user application architecture for the video-audio streaming between a paramedic and the CMO of an ambulance service, based on the award-winning BlueEye platform designed for first responders [14]. Systems architecture is shown in more detail in the following section on use case H1D.

The paramedic wears a BlueEye camera headset connected with smartphone for mobile connection. The camera can capture point-of-view video with a resolution of 1920 x 1080. The wide-angle camera of 158° ensures the remote CMO sees the ‘complete picture’. The camera headset was connected to a 5G Android smartphone via a Universal Serial Bus (USB) connector. Here USB is preferred over wireless connectivity (e.g., via Bluetooth or Wi-Fi), for reduced latency, improved reliability, and minimised power consumption and weight of the headset, which directly benefits the usability and its wearing comfort. Security is also enhanced using a wired connection.

The smartphone is further connected (via a 5G network) to two servers in the control and user planes respectively; one server oversees application management and control, while another server is dedicated to user-plane data (video-audio etc.) delivery. For countries where ambulance care is regionally organised, such as in the Netherlands, one video-delivery server could be deployed per region (as shown in Figure 26). In practice, these servers may be physically located on the premises of the regional ambulance services, or at another place where the requirements of security, privacy and performance can be guaranteed. A local server could avoid extra transmission latency and jitter due to the Internet while potentially addressing the security and governance issue. However, due to practical limits, in this study an existing remote server located in Frankfurt was used.

The remote CMO, who can be located anywhere (home, office, hospital or on the move) can access the live video-audio feed (and vital data if any) using different types of devices (e.g., laptop, iPad, smartphone). Typically, a single dashboard is desired on the device to display the received video-audio and vital data, with which the CMO can easily switch between different types of patient information. Different device types obviously may pose different requirements regarding video resolution and the
usability of dashboards. In this study, we assume the remote CMO is on the move and uses a smartphone while receiving the video-audio streaming.

Figure 26 User application architecture for RedZinc’s video streaming (HIC)

5.2.3 Hardware components

User equipment at the side of the CMO:

- Smartphone iPhone 12, assuming that the CMO is on the move while receiving the video-audio streaming.

User equipment at the side of the paramedic:

- Camera headset of RedZinc, which was connected to an Android smartphone (Samsung Galaxy S20) with a USB 3.0 connector. Further details can be found in D3.3 [2].
- Smartphone Samsung S20 with USB 3.0 interface.
- 5G SA-compatible on-board unit (OBU): Fibocom FM150-AE + APU2 platform (an embedded Linux device) + Netgear Nighthawk M5 MR5200 (WiFi 6 router).

User equipment for network KPI measurements:

- Fibocom FM150-AE + APU2 platform.

5.2.4 Software components

The software components used for the HIC subcase are the following:

- Android application, via which the paramedic can connect with the video servers (Figure 27).
- Custom management services in the Cloud.
- BlueEye hot desk, which can be used by the CMO to receive the video streaming (Figure 28).
5.3 Testing and verification

Testing and verification have been focused on how video quality can be guaranteed while sharing the same network resource with CCAM traffic (for the T2S1 subcase) and general purpose (regular data) traffic via network slicing, in a more realistic ambulance scenario where the ambulance for a large portion of the time is on the move (in comparison to phase-2 trials with an indoor network in a largely isolated warehouse and with little competing mobile traffic, see D3.3 [2]).

5.3.1 Methodology

Figure 29 depicts the test scenario of phase-3 trials, showing how the ambulance drives to the scene of emergency (Steps 1-3), the paramedic examines the patient when arriving at the scene and in case necessary consults the remote CMO (Step 4), and how the ambulance drives towards the hospital with the patient on board (Steps 5-7). The used 5G SA station is also depicted in the figure, noted as “gNB”, near the intersection. Pragmatically a TNO vehicle is used to emulate a real ambulance.
Each of the steps is described as follows:

- **Step 1:** at a parking lot on the highway, the OBU of the vehicle is connected to the 5G SA network;
- **Step 2:** after being successfully connected to the 5G SA network, the vehicle starts driving toward the intersection;
- **Step 3:** when approaching the intersection, the OBU automatically asks for a green-light priority (T2S1); when a green-light priority is granted, the driver is informed of the current phase-timing of the traffic light and the optimal speed to drive; and the vehicle turns right at the intersection accordingly and drives to the scene of the emergency;
- **Step 4:** after examining the patient at the scene, the paramedic decides to consult the remote CMO by starting a video-audio streaming (H1C);
- **Step 5:** in consultation with the remote CMO, it’s decided to bring the patient to the hospital; the vehicle leaves the scene with the patient on board; the video-audio connection may be kept active on the way to the hospital, keeping on monitoring the patient and surrounding conditions (H1C).
- **Step 6:** the vehicle drives back to the highway, and when it approaches the intersection, it again asks for a green-light priority (T2S1);
- **Step 7:** the vehicle turns left at the intersection and drives on the highway; the scenario ends when the OBU of the vehicle is out of coverage of the 5G SA network (note: there is only one 5G SA base station in the 5G test network used).

![Figure 29 Test scenario of the phase-3 trials (H1C and T2S1 jointly)](image_url)

Three (RAN) network slices may be configured, one for video traffic (H1C), another for CCAM traffic (T2S1) and a last one for general purpose traffic (regular data). The CCAM slice is assigned the highest priority in using radio resources, considering the critical nature of the CCAM traffic (green-light priority requests) in minimizing the time spent on the way from the scene to the hospital. The video (or medical) slice is assigned the 2nd priority, while the general-purpose slice is assigned the lowest priority. At a certain moment of scheduling at the base station, traffic with a higher priority will be first scheduled. By configuring slices with different priorities, we effectively allow all available radio resources to be used by the slice with the highest priority. In case that a certain resource is not used by a higher-priority slice, at a certain scheduling moment, it can be used by a lower-priority slice.

The general-purpose traffic is emulated by using a road-side unit (RSU) near the intersection, with a relatively short distance from and good radio condition with the base station. The RSU is configured to send uplink data with throughput up to ~15 Mb/s. As we have observed, the camera of RedZinc may generate video-audio stream with throughput no more than 15 Mb/s in the uplink, and the uplink CCAM traffic of T2S1 (asking for green-light priority) consumes very little data throughput (see D4.4). To be
able to see a significant effect on the performance, when enabling or disabling the slicing configuration, it is necessary to reduce the carrier bandwidth to 20 MHz. This emulate a real practice in a 5G commercial network where a significant part of the overall bandwidth (e.g., of 100 MHz) is reserved or used by other applications.

### 5.3.2 List of key performance indicators

The key performance indicators associated with the H1C subcase are summarised in Table 8. During the course of work, we have observed that the RedZinc camera can generate higher uplink, and acceptable average latency is higher, up to 350 ms, compared to the original assumptions, as described in D3.3 [2], section 5.3.2.

<table>
<thead>
<tr>
<th>Network requirements</th>
<th>Target values</th>
</tr>
</thead>
<tbody>
<tr>
<td>User experienced DL throughput</td>
<td>1 &lt; Medium ≤ 15 Mb/s ( a )</td>
</tr>
<tr>
<td>User experienced UL throughput</td>
<td>1 &lt; Medium ≤ 15 Mb/s ( a )</td>
</tr>
<tr>
<td>Broadband connectivity / peak data rate</td>
<td>100 &lt; Medium ≤ 1500 Mb/s ( a )</td>
</tr>
<tr>
<td>Latency requirements</td>
<td>Low &lt; 350 msec ( b )</td>
</tr>
<tr>
<td>Reliability</td>
<td>Medium: 99.999%</td>
</tr>
<tr>
<td>Mobility</td>
<td>50 &lt; Medium ≤ 160 km/h</td>
</tr>
<tr>
<td>Location accuracy</td>
<td>Low &gt; 25 meters</td>
</tr>
<tr>
<td>Connection (device) density</td>
<td>Low: &lt; 10 devices in and around the ambulance</td>
</tr>
<tr>
<td>Interactivity</td>
<td>Low ≤ 1 transactions/s</td>
</tr>
<tr>
<td>Area traffic capacity</td>
<td>&gt; 1, ≤ 15 Mb/s/m² ( a )</td>
</tr>
<tr>
<td>Security / privacy</td>
<td>High: Confidential</td>
</tr>
</tbody>
</table>

\( a \): as we have observed, the camera of RedZinc can generate uplink throughput up to about 15 Mb/s, larger than the original estimate of 10 Mb/s per video stream.

\( b \): according to the feedback of ambulance professionals (see D3.3 [2]), an average latency of about 350 ms is acceptable.

### 5.3.3 Measurement and testing tools

For the phase-3 trials, the following measurement and testing tools are used:

- Ping for measuring the round-trip time (RTT) at the network layer.
- iPerf3 for measuring network layer throughput.

### 5.3.4 Final results

At the reduced bandwidth of 20 MHz, the measured maximum downlink (uplink) throughput is 119 (15.7) Mb/s.

In this section, we compare the achieved results with and without network slicing enabled (results of some more configuration options are captured in D4.4 [13]):

- Case A: no slice configured, the video traffic (H1C), the CCAM traffic (T2S1) and the general-purpose traffic is scheduled by the base station with the same priority.
- Case B: three slices configured with different priorities, where the CCAM slice is assigned the highest priority, the video slice the 2nd highest priority and the general purpose slice the lowest priority.

For each of the cases, the vehicle performed four runs following the route and steps as described in Figure 29 and Section 5.3.1. Each run takes between 250 s and 300 s. For the sake of simplicity, in the
context below we show and compare the results of one single run of each of the cases. Note that the same/similar behaviour in terms of throughput patterns has been observed among the four runs of a certain case.

Figure 30 shows the observed uplink throughput performance of Case A along the route. “All vehicle traffic” refers to the sum of video traffic and CCAM traffic sent from the vehicle. At the start (Steps 1-3 of the whole route), “All vehicle traffic” only consists of CCAM traffic with low data transmission demand, and therefore the general-purpose traffic could utilize most of the available radio resource (in this case the reduced bandwidth of 20 MHz) and reach uplink throughput around 15 Mb/s. When the video traffic is started (at Step 4 and being kept on until the vehicle runs out of the coverage), “All vehicle traffic” is dominated by the video traffic and competes with the general-purpose traffic in the use of the available radio resource. The results show that on average the general-purpose traffic is assigned about 50% of the radio resource and reaches uplink throughput around 7.5 Mb/s. The video traffic is assigned the other 50% of the radio resource. As observed, the adaptive behaviour of the video algorithms is quite sensitive to the competition of the general-purpose traffic in using the radio resource, and it lowers the video throughput to around 1.5 Mb/s. The video algorithms could be improved so that the potential of a much higher throughput can be exploited (up to around 7.5 Mbps as we would expect, considering the amount of radio resource available to the video traffic), and accordingly higher video quality.

![Figure 30 Observed uplink throughput of “All vehicle traffic” (H1C + T2S1) and general purpose traffic (Case A)](image)

Figure 30 shows the observed uplink throughput performance of Case A along the route. “All vehicle traffic” refers to the sum of video traffic and CCAM traffic sent from the vehicle. At the start (Steps 1-3 of the whole route), “All vehicle traffic” only consists of CCAM traffic with low data transmission demand, and therefore the general-purpose traffic could utilize most of the available radio resource (in this case the reduced bandwidth of 20 MHz) and reach uplink throughput around 15 Mb/s. When the video traffic is started (at Step 4 and being kept on until the vehicle runs out of the coverage), “All vehicle traffic” is dominated by the video traffic and competes with the general-purpose traffic in the use of the available radio resource. The results show that on average the general-purpose traffic is assigned about 50% of the radio resource and reaches uplink throughput around 7.5 Mb/s. The video traffic is assigned the other 50% of the radio resource. As observed, the adaptive behaviour of the video algorithms is quite sensitive to the competition of the general-purpose traffic in using the radio resource, and it lowers the video throughput to around 1.5 Mb/s. The video algorithms could be improved so that the potential of a much higher throughput can be exploited (up to around 7.5 Mbps as we would expect, considering the amount of radio resource available to the video traffic), and accordingly higher video quality.

Figure 30 shows the observed uplink throughput performance of Case B along the route. In this case, the video traffic and CCAM traffic are explicitly separated since they use two different slices with different priorities. Similar to Case A, at the start (Steps 1-3), the general-purpose traffic could utilize almost all of the available radio resource and reach uplink throughput around 15 Mb/s. When the video traffic is started at Step 4, it has higher priority in using the radio resource than the general-purpose traffic (note again that the amount of the CCAM traffic is so little that it almost has no effect in radio resource assignment although with higher priority). Therefore, the video traffic can reach much higher uplink throughput (with higher video quality), at about 10 Mb/s to 15 Mb/s at Steps 4 to 6. During this period, some general-purpose traffic could still be scheduled using (remaining) radio resource not utilised by the video traffic. When the vehicle drives away from the base station, the throughput of the video traffic drops although it still consumes most of the radio resource. When the vehicle runs further away from the base station, the received signal strength at the vehicle is so low that no video traffic can be transmitted. In this case, the general traffic is (again) assigned most of the radio resources.
5.4 Recommendations

The study has validated, in collaboration with ambulance professionals, that the use of real-time video and vital data significantly improves the effectiveness of assessment in emergencies, compared to the current audio-only communications. In addition, a few implementation challenges were encountered during the trials and recommendations have been made to address these challenges (see D3.3 [2]).

The study has further validated the role of 5G in enabling the use of real-time video communication between the paramedic and the remote CMO, both in (more or less) ideal (D3.3, with an indoor 5G SA network, with little competing mobile traffic) and more practical conditions (This report, with an outdoor 5G SA network, on-move, with significant competing mobile traffic). The study has shown that 5G network slicing is able to guarantee sufficient radio resource for the delivery of video-audio streams and vital data of the ambulance service.

It is recommended to continue similar trial-based investigation with a commercially deployed 5G network with more challenging network conditions and varying environments, including investigation of certain network features (e.g., slicing, relaying) which are expected to be able to improve network and radio resource availability and performance for ambulance services. A holistic approach may also include business model study (see the first study in D7.4 [15]), since the aimed network availability (i.e. anywhere and anytime) and performance are not only about 5G network capabilities but also about business policies of stakeholders (e.g. who should invest to guarantee network availability and performance required by ambulance services).
6 SUBCASE H1D: CRITICAL HEALTH EVENT

6.1 Description and motivation

Video is being increasingly used in different settings as it enhances situational assessment and leads to improved outcomes. BlueEye Handsfree is an innovative wearable video solution, which facilitates transmission of real-time point-of-view video from one location to another. BlueEye Handsfree has a large number of applications across industries such as Telemedicine, Tele-maintenance and Public Safety.

The key benefits are
- Expedite patient treatment
- Avoid unnecessary conveyance
- Accelerate pre-hospital decision-making
- Improve patient outcomes
- Save costs and improve efficiencies
- Eliminate contagion risk
- Provide instant remote assistance or expert advice

Multiple pre-hospital clinic applications can be supported, where a paramedic or ambulance technician may need support such as: Paramedic doctor support; Paramedic buddy support; Pneumothorax; Stroke, Heart Attack; Trauma; Burn Victim; Brain Injury; Special needs; Frail/Elderly; Disaster; Mass Casualty; Training.

Further details are given in Deliverable D3.2 [1] and D3.3 [2] and an application example is shown in use case H1C in the Chapter 5. In this section we focus on the necessary secure architecture to deliver the service.

6.2 Final setup

This section explains the final (phase-3) setup of the experimental use case with a particular focus on secure architecture set up.

6.2.1 Network architecture

Following the pilots in January 2020 and June 2021 with Oslo Ambulance Service the Urban Search and Rescue/Chemical, Biological, Radiation, Nuclear and Explosives (USAR/ CBRNE) team in Oslo Ambulance Service wish to use BlueEye connecting paramedic team with remote supervisors in real patient environment. To date the pilots have used manikins and actors and the USAR/ CBRNE team have requested that production environment be put in place so that real patient scenarios can be supported. There is wide support from the USAR/ CBRNE team to go to the next level. The USAR/ CBRNE team consists of approximately 20 paramedics and 6 supervisors.

The USAR Scene commander cannot see the visual details of the incident that a team is involved in. This is applicable to a wide range of USAR and CBRNE cases.

In the USAR and CBRNE cases, time is of essence and the decision making needs to be precise and quick to contain loss and save lives. There is little to no scope for delayed and/or incorrect decision making as it may result in irrecoverable losses. Hence the system needs to be such that there is little to no scope for human error. As well as saving a pre-hospital patient the supervisor has responsibility to ensure that the paramedic team is detracted safely from the incident.
The USAR and CBRNE operations can be enhanced with more ‘eyes on the field’ for supervision, support, and expert opinion. But the critical nature of these operations and the associated costs prohibit deploying many people on the field.

In the case of USAR, the incident commander plays a critical role as they establish objectives and supervise implementation. A key challenge for the incident commander is to keep an eye on the team working and make quick decisions in tricky situations which are hampered in the absence of visual details for the incident commander. Similarly, in the CBRNE cases, the remote commander is not able to assist the disposal technician at the front in the absence of visual details.

In both cases, the victims need to be attended to by the medical team and the critical cases need to be escalated quickly to save lives.

The Mobile Edge Computing (MEC) architecture is shown in Figure 32.

![Figure 32 BlueEye MEC architecture](image)

### 6.2.2 User application architecture

The BlueEye application contains the following features:

- Live secure video unidirectional broadcast to a remote clinician with bi-directional live continuous audio
- Multiple concurrent support clinicians supported in a single event
- Web-based remote Clinician hot-desk with support for multiple devices
- High-Quality Snapshots (available on request)
- BlueEye App audio toggle between loudspeaker for patient interaction by remote clinic and earpiece for a private paramedic to clinician interaction
- Adjustable video settings with quick pre-set based on location brightness
- Adaptive video quality based on network conditions
- Poor network condition mode based on still images
- Geolocation with reverse geocoding to retrieve location and address
• Secure access with two-factor authentication support with configurable rules
• Configurable inactive session closure for extra security
• Private app auto-update process outside the scope of app stores for extra security and privacy

6.2.3 Hardware components

The user hardware contains:
• Samsung A52 (or similar) 5G android based smartphone
• Smartphone Body Dock Case for Peter Jones Klickfast connector series
• Peter Jones Klickfast connector
• BlueEye Handsfree Camera (Model 268)
  o 2Mpixel 1080p high definition camera
  o Max image resolution 1920x1080
  o 1/6" Color CMOS 1080p Camera
  o Camera Image Area(mm): 2.7288x1.5498
  o Pixel Size(um): 1.4x1.4
  o Dynamic range: 73.3 dB @ 15.5x gain
  o Sensitivity: 553 mV/(Lux-sec)
  o Lens Options
    ▪ 93° Narrow-Angle Lens Option; H=81°90'; V=50°18'; D=92°82'
    ▪ 158°Wide Angle Lens Option; H=132°; V=65°; D=158°
  o 1 meter USB cable to smartphone
  o BlueEye Quick change camera shoe.
• Options
  o Arm mounted smartphone case
  o Velcro-based helmet mount

6.2.4 Software components

Software components are:
• User application for wearable camera
• WebRTC enabled browser
• BlueEye application stack
  o BlueEye Application Management
  o Video Router
  o Media Relay

These are configured in the scenario shown in Figure 33
6.3 Testing and verification

The methodology used was to develop a secure MEC environment.

6.3.1 Methodology

The baseline applications BlueEye architecture complies with GDPR security and the latest security standards. All BlueEye data is AES-256 encrypted. BlueEye uses TLS 1.2 and TLS 1.3 and is always updated with secure ciphers only. This includes:

- All communications are encrypted with HTTPS.
- Video server signalling happens in HTTPS.
- With standard WebRTC encryption, video images are secured.
- No video, audio or patient information is stored; therefore a breach of data storage is not possible.
- The following encryption software is used
  - RSA2048, SHA256 Hashing, TLS
  - Apache commons libraries, Apache Tomcat for TLS
  - Sensitive data such as passwords is hashed
  - Standard SSH, TLS 1.3 mechanisms
  - TLS 1.3
  - TLS_AES_128_GCM_SHA25
  - TLS_AES_256_GCM_SHA38
  - TLS_CHACHA20_POLY1305_SHA256
- Periodical update of ciphers as they are superseded.

The approach is to extend this using security features of 5G networks.
6.3.2 List of key performance indicators

Multiple levels of security are defined:

- Control and data plane encrypted
- Private APN so that 5G radio traffic is directed to a secure non-internet environment
- MPLS/BGP VPN with private addressing decoupled from public internet
- Sovereign data servers with servers located inside the region and geography boundaries of the ambulance service. This approach avoids Schrems II GDPR issues which are associated with non-EU region hyper-scalers.

6.3.3 Measurement and testing tools

The secure service is under deployment at the date of this deliverable and end-to-end testing has not yet occurred.

6.3.4 Final results

Due to confidentiality, the final architecture cannot be shown in this report. The final architecture includes:

- Paramedic wearable video equipment
- Browser access based on WebRTC standards
- Secure Network consisting of the below elements
  - 4G/5G wireless access
  - Private APN
  - Back haul to VPN built MPLS/BGP
  - Sovereign data servers in MEC
  - VPN gateway to enable browser access to video

6.4 Recommendations

Based on the collaboration between RedZinc and Telenor, a trust-based approach is proposed for engagement with medical users offering integrated trust-based solutions. The key ingredients of a trust package for emergency services are listed below.

**Application**

- Mobile Video Tele-medicine applications
- Connected sensors including medical monitoring and other sensors

**Customer Private network**

- Private APN
- MPLS BGP VPN

**Security**

- Encrypted application data and control plane
- Routing of radio traffic to private network via Private VPN
- Private IP addressing via MPLS/BGP
- Virtual network functions inside sovereign MEC data centres

_Public 4G/5G_
- National 5G coverage
- Fall back to 4G public wireless if 5G not available
- Campus 4G/5G (for hospital resuscitation rooms etc)
- National roaming

_5G Slicing_
- Radio resources in a slice allocated to customer user equipment
- Quality of Service
  - QCI-2 & QCI-4 Option for quality class indicators
  - Guaranteed bearer GBR option
  - Static or Dynamic resource allocation (N5 interface)
- Traffic Profiling
  - eMBB (enhanced Mobile Broadband) for video applications
  - mMTC (massive Machine Type Communications) for medical sensors
7  SUBCASE H2A: AUTOMATIC PILL CAMERA ANOMALY DETECTION

7.1 Description and motivation

The imaging technology, in a capsule no bigger than a pill, can provide a diagnosis within hours. Known as a colon capsule endoscopy. Traditional endoscopies mean patients need to attend hospital and have a tube inserted whereas the new technology means that people can go about their normal day. Using wireless technology to connect medical implantable devices to a cloud, provides a high impact on monitoring health information. Colon cancer is the second most common cause of cancer mortality for both men and women.

Artificial intelligence (AI) has played an increasing role in the technological development of clinical practice. Together with the numerous medical data, the evolution of technology has led to recent advances in AI using deep learning in the medical field. Computer-aided diagnosis (CAD) systems using esophagogastroduodenoscopy (EGD) and colonoscopy images have become a vigorous research field. Especially deep learning, and convolutional neural networks (CNN) have been very successful to advance computer vision and image processing of natural images databases. The potential of AI to automate tasks such as precancerous pathology detection and novel advances in endoscopy such as wireless capsule endoscopy (WCE) motivate us to combine these technologies as an easy-to-use alternative that can potentially save thousands of lives and improve the lives of patients and clinicians. We implement a high data rate wireless capsule endoscope (WCE) bridging the video data to the cloud by means of a 5G network, enabling data processing, AI analysis for decision making support and feedback control from the cloud. The WCE uses a novel approach for high data rate wireless communication, in which the wireless backscatter is applied to download video data from the capsule device.

An external body reader device Wireless Capture System (WCS) is used to transmit the RF signal and read the video signal in real time. The WCS system will bridge to a 5G network for transmission of capsule image data to our medical cloud. Real-time image processing is conducted by using cloud computing and artificial intelligence (AI) methods, with the aim of automatic, fast, and reliable diagnosis and detection. Real-time communication, image processing and analysis with low latency protocols enable adjustment of the capsule’s sensor setting such as light intensity, spectrum, and camera resolution to obtain high-quality images for reliable detection of polyps. This culminates in an Internet of medical implants (IoMI) for future medicine. In addition, the bridging between the implant and the 5G network enables the application of cloud-based AI analysis for accurate therapy, diagnosis, and decision support.

AI is defined as the use of computers and technology to simulate intelligent behaviour and critical thinking comparable to that of a human being. The ever growing need to provide high-quality and cost-efficient global healthcare has resulted in a corresponding expansion in the development of computer-based and robotic healthcare tools that rely on artificially intelligent technologies. Capsule endoscopy is one of the branches of gastroenterology that can benefit the most from the application of this type of technology. Indeed, the use of AI in this field shows great promise and capsule endoscopy can serve as a steppingstone for the broader application of AI in endoscopy and gastroenterology.

7.2 Final setup

The aim of this subcase is to demonstrate the feasibility of video streaming over a 5G network from an endoscope capsule, receive and analyse the streamed video at the network edge, and provide feedback to the endoscope capsule, all with acceptable latency.

The endoscope capsule camera operates using unique backscatter communications. The capsule camera streams the video frames through the 5G network to the network edge where an AI based anomaly
detection model is ready to process the received frames. The detection model is an efficient and accurate deep learning-based algorithm built from CNN blocks.

Wireless video capsule endoscopy is used to stream the captured high data rate (10 Mb/s to 20 Mb/s) videos to the 5G network. The video acquisition is performed by using a NanoEye® camera sensor that transmits the raw video data of 12 Mb/s to the capsule onboard micro processing unit. The raw data controls the capsule antenna reflection coefficient that is sensed by an on-body array antenna using RF backscatter. The external backscatter unit converts the radar data to digital data corresponding to the camera information. The output of the external reader device is serial data that is synchronized with the recovered clock information. The reader output voltage level follows low voltage differential signalling (LVDS) voltage levels. The reader data is bridged to a custom-designed FPGA in which the data is converted to a standard UDP LAN protocol. The interconnection from this stage follows standard interface connections.

Once the video frames are received, our AI-based anomaly detection algorithm commences the detection and localization in real-time. If a polyp, or a suspicious region has been identified, a signal with instructions is sent back to the capsule camera through the same end-to-end 5G network. The capsule camera adjusts its parameters according to the instructions sent by the deep learning model, to obtain frames with the best lighting and resolution possible. In the demonstration of the system in the phantom experiment, the setup as shown in Figure 34 is prepared. The streaming video from first PC has a data rate of 10 Mb/s is sent to a microcontroller unit and data formatting to control the capsule antenna. Further it consists of the reader system with array on-body antennas, the reader RF system, FPGA modem and second PC for displaying the detected data.

![Figure 34 Full backscatter system demonstration setup.](image)

. The video data is created from a pre-recorded video from a clinical colonoscopy, to be able to test the full system operation. The video data is played to a capsule mock-up and the backscatter data is extracted with the reader and bridged to the FPGA modem. The data can either be transmitted to the 5G modem or directed through LAN cable or Wi-Fi to a second computer where the video is played in the second receiver PC. In the network implementation, the video data is analysed to detect the polyps. The project follows the implementation of a camera in a capsule endoscope or a benchtop model.
7.2.1 Network architecture

The 5G-VINNI network in Oslo – and the installation located at the OUS buildings – was used for the technical evaluation in the period until June 2020. Further testing was done using Telenor’s commercial network, using mainly 4G.

![Image of network architecture](image)

Figure 35 End-to-End connection setup for video streaming via 5G network

Figure 35 shows an overall implementation of the backscatter based capsule endoscope system in conjunction with 5G network deployment and features. The system-level development is illustrated in the picture, where a capsule is manufactured that can stream a pre-recorded video to the capsule’s internal backscatter switch system.

7.2.2 User application architecture

The aim of this subcase is to demonstrate the feasibility of video streaming over a 5G network from an endoscope capsule, receive and analyse the streamed video at the network edge, and provide feedback to the endoscope capsule, all with acceptable latency.

The endoscope capsule camera operates using unique backscatter communications. The capsule camera streams the video frames through the 5G network to the network edge where an AI based anomaly detection model is ready to process the received frames. The detection model is an efficient and accurate deep learning-based algorithm built from CNN blocks. After receiving the video frames, Developed AI-based automatic detection algorithm detects and localizes in real-time. Suspicious polyp, or a region has been identified, a signal with instructions is sent back to the capsule camera through the same end-to-end 5G network. The capsule camera adjusts its settings as per to the instructions sent by the deep learning model, to obtain frames with the best lighting and resolution possible.

The real-time transmission of information to the cloud is essential and can be used to control the imaging parameters of the WCE online, with the help of a feedback loop from the deep learning model to generate high quality frames for reliable detection of gastric disease such as polyps, bleeding, etc. It is also possible to alternate the stream rate of the WCE based on the received frames and the analysed data to save energy of the implant sensory system. Once the data package sent by the WCS is received at the inference location (cloud and/or edge), our polyp detection algorithm commences the detection and localization of polyps in real-time. If a polyp or a suspicious region has been identified, a signal with instructions is sent back to the WCE through the same end-to-end virtual network. The WCE then adjusts its parameters according to the instructions. Figure 36 demonstrates the patient side system and the swallowable capsule connection to the 5G network. The patient is wearing the WCS around the waist. The WCS must be worn during the entire wireless endoscopy. After the clinician has controlled the correct positioning of the WCS the patient swallows the WCE. The WCE travels down the digestive system and the image transmission to the WCS begins when it reaches the large intestine. The WCE transmits continuously the images to the WCS which is bridged through the 5G modem to the Edge computer and to the cloud respectively over the 5G network.
Figure 37 shows the image processing and AI functionalities. The image stream is fed into the edge node and/or medical cloud where the polyp detection algorithm analyses in real time the images looking for polyps. The cloud keeps a copy of all occurrences during each capsule endoscopy and the only difference between the cloud and the edge node is where the computing or inference happens, in the case where the WCS sends the image stream to the edge node, the inference happens in the edge computer, commonly in the place of the experiment. When no edge node is involved, inference happens directly on the cloud.

Figure 38 shows the feedback loop and control process of the WCE. The WCS receives feedback from the edge/cloud and in case the probability of a polyp in a frame is high, instruction with image capture optimal parameters, such as lighting or FPS capture is sent to the WCS. This results in better images from the region of interest, images that will help the clinician in charge get a clearer picture to determine if a polyp is present or not.

Figure 36 Patient side system and capsule connection to the 5G network.
7.2.3 Hardware components

The hardware system near the patient consists of two parts; one is a swallowable capsule with a camera and the supporting electronics: a low power microcontroller, battery, integrated antenna and a backscatter switch that is controlled by the camera data source. An on-body transceiver system supports the wireless communication with the capsule and includes a set of array antennas connected to an RF transmitter that sends a single tone carrier signal and receive the RF reflections of the capsule, a receiver antenna that is connected to a patented receiver system with zero-IF setting and extracts the video data and the synchronized clock data. The receiver is bridged to a custom designed modem that converts the data format to standard LAN (UDP) interface for streaming. The LAN connection can be bridged to any standard modem such as Wi-Fi or 5G modem to interconnect the system to a local or larger network. The following hardware components are used for the verification and demonstration.

- Two CPEs from Telenor (Huawei CPE Pro 2)
- A laptop as the Client PC
- A desktop as the Server PC equipped with NVIDIA RTX 3090
- FPGA PYNQ- Z1 boards (LAN interface)
- RF Backscatter system (transceivers, control unit, power supply, antenna interface)
- On body antenna array of 6 elements to cover the colon area
• Wireless capsule endoscope (includes camera or a dummy data source, battery to power camera, backscatter switch and capsule antenna)

7.2.4 Software components

The software component of the polyp detection prototype consists of several layers of convolutional neural network (CNN) designed and trained for automatic detection of precancerous polyps in colonoscopy images and videos, both retrospectively and in real-time. The model is being adapted for automatic review of capsule endoscope footage to produce control feedback if anything suspicious comes to sight during capsule endoscopy.

The following software components have been used to achieve the target.

• Python programming
• PyTorch for the development of the deep learning algorithm
• OpenCV library for
• Socket package for video streaming
• Pickle package to encrypt the data
• Transmission Control Protocol (TCP), e.g. Hypertext Transfer Protocol (HTTP), and Real Time Streaming Protocol (RTSP), and User Datagram Protocol (UDP), e.g. Real-time Transport Protocol (RTP) for video streaming,
• VHDL for streaming

7.3 Testing and verification

7.3.1 Methodology

End-to-End connection setup

Figure 39 shows the overall 5G connection establishment to measure the end-to-end video streaming latency. In this setup, we use a laptop PC to simulate the endoscope camera streaming videos to the AI server at the edge of the network. The streaming PC is connected to the 5G-VINNI via a CPE, and the AI server is connected to the same 5G network via another CPE. Which were used to establish the point-to-point connection. The AI server at the network edge contains the required processing hardware to analyse the videos in real-time and generate feedback control signals to the capsule reader. The capsule endoscope receives the feedback signals and changes the capsule position, the light intensity, video quality, and/or the frame rate, or uses a hyperspectral imaging sensor to analyse the detected polyps closely and seamlessly. We use different transmission protocols for the video streaming such as UDP-RTP, TCP-RTSP, and TCP-HTTP.
AI-based polyp detection model

Figure 40 presents our proposed method to detect polyps in a single manner in videos. The method is developed based on a 2D CNN encoder-decoder network. The 2D encoder-decoder networks are originally developed for single image analysis and do not incorporate temporal information among neighbouring frames when they are applied for video analysis. This property together with the CNN vulnerability make the 2D encoder-decoder networks produce unstable output predictions for consecutive frames contaminated with a lot of false detection outputs. In this section, we provide a detailed description of our proposed method to make a CNN-based encoder-decoder network a more stable and reliable polyp detector suitable for video analysis. We incorporate temporal information from previous frames to analyse the current frame. We concatenate extracted features from \( N \) previous frames \( f_{xy}(t-n), n \in [1,2,\ldots,N] \) with the extracted features of the current frame \( f_{xy}(t) \) in the bottleneck layer (latent space).

**Neural network architecture**

In this work, we adapt AlbuNet34 for our polyp detection model as shown in Figure 40. AlbuNet34 is a UNet-like architecture consisting of two paths: the contracting path and expanding path. The contracting path (encoder part) takes in an input image frame \( f_{xy}(t) \) and progressively extracts abstract features. The expansive path (decoder part) interprets the extracted features and enables precise localization. AlbuNet34 uses ResNet34 pre-trained on Imagenet dataset for the encoder part. We choose AlbuNet34 because it combines the advantages of both UNet-like architectures and residual learning. In addition, AlbuNet34 is designed from fully convolutional neural networks (F-CNN) predicting outputs in a single shot feed-forward manner which makes them eligible for real-time implementation.

The first block of the encoder is a kernel of size \( 7 \times 7 \) with stride 2 followed by a max-pooling layer with stride 2. The rest blocks of the encoder consist of repetitive residual blocks. In every residual block, the first convolution operation is applied with stride 2 to provide down sampling, while the rest convolution operations are applied with stride 1. We apply a \( 2 \times 2 \) max-pooling operation on the final output feature maps of the final residual block of ResNet34. The result of this max-pooling operation is stored in the bottleneck layer (the latent space). We add this bottleneck layer to facilitate the incorporation of temporal information from consecutive frames.
The decoder part consists of several decoder blocks, each block is concatenated with the corresponding encoder block. In every decoder block, an up-sampling operation is applied to up sample the feature maps by 2, followed by two padded convolution operations of a kernel of size $3 \times 3$ with stride 1 followed by a rectified linear unit (ReLU). The first block of the decoder starts with interpreting the abstract features stored in the bottleneck layer. To generate the final output, we apply a $1 \times 1$ convolution operation followed by the $\tanh$ activation function. Our detection model generates an output image which has the same resolution as the input image frame $f_{xy}(t)$, with the predicted 2D Gaussian shapes $\hat{Y}_{xy}(t)$.

**Concatenation of consecutive features**

It has been shown that the same CNN-based detector can miss the same polyp appearing in the neighbouring frames due to changes in the light conditions, appearances, inherent noises, blurriness, etc. This is due to the vulnerability of CNN to small perturbations. We solve this problem by concatenating the feature maps of a series of consecutive frames extracted by the encoder part. Neighbouring frames are closely related to each other and thus their extracted feature maps should be closely similar and contain complementary information. The elegant structure of UNet-like architectures facilitates this feature concatenation in the bottleneck layer. This way we can incorporate temporal information among neighbouring frames into the detection model.

We store the feature maps from $N$ previous frames $f_{xy}(t - n), n \in [1, 2, 3, \ldots, N]$ and concatenate them with the feature maps extracted from the current frame $f_{xy}(t)$ in the bottleneck layer. We set the number of activation maps in the bottleneck layer to be 256 maps. We equally divide the bottleneck layer into $N$ slots of 256/$N$ activation maps based on the number of previous frames involved. For instance, when only one previous frame $f_{xy}(t - 1)$ is incorporated, 128 feature maps are extracted from the current frame $f_{xy}(t)$ and 128 feature maps from $f_{xy}(t - 1)$. However, when the result of this division is a floating number, we round it up to an approximate number. For instance, when we consider two previous frames $f_{xy}(t - n), n \in [1, 2]$, the result of 256/3 is 85.33, thus we use 86 maps for each frame, resulting in 258 feature maps in the bottleneck layer.

This concatenation of feature maps helps combine complementary information from a series of previous frames $f_{xy}(t - n), n \in [1, 2, \ldots, n]$ with the current frame $f_{xy}(t)$, reducing the effect of a small perturbation and/or change that may pop up in the current frame $f_{xy}(t)$ and fool the CNN-based detector. Therefore, this concatenation strategy helps the CNN-based encoder-decoder network improve its accuracy and produce more stable detection outputs for a series of neighbouring frames.

**2D Gaussian shapes for polyp regions**

A CNN-based encoder-decoder network can be more efficient in detecting polyps when it is trained on 2D Gaussian shapes as the ground-truth masks instead of using binary masks. We train our detection model to predict a 2D Gaussian shape, $\hat{Y}_{xy}(t) \in [0,1]^{W \times H \times 1}$, for a polyp region in an input RGB image frame at time $t$, $f_{xy}(t) \in R^{W \times H \times 3}$, where $w$ is the width and $H$ is the height of both $f_{xy}(t)$ and $\hat{Y}_{xy}(t)$.

We transform the provided binary ground-truth masks, $X_{xy}(t) \in \{0,1\}^{W \times H \times 1}$, to 2D Gaussian ground-truth masks, $\hat{Y}_{xy}(t) \in [0,1]^{W \times H \times 1}$. The 2D Gaussian ground-truth masks are meant to reduce the impact of the outer edges during training and force the network to learn the surface patterns of different polyps more efficiently.

Using 2D Gaussian shapes can also help to realize polyp detection in real-time. We can generate all detected bounding boxes directly from the predicted 2D Gaussian shapes without the need for computing non-maximum suppression (NMS) to eliminate overlapping bounding boxes. At the inference time, we use the strength of the predicted 2D Gaussian shapes as the confidence values of the detected bounding boxes.

© 5G-HEART Consortium 2019-2022
boxes and calculate the two size-adaptive standard deviations (\(\sigma_x\) and \(\sigma_y\)) for the size of the detected bounding boxes.

Moreover, exhibit the feasibility of video streaming from a wireless capsule endoscope to a 5G network that operates using our unique technology with RF backscatter for high data rate wireless communication with deep medical implants [16],[17]. RF backscatter is based on radar approach that permits remote reading of the WCE information. This means that the active transmitter is removed from the WCE that results in significant power saving since streaming high data rate video signals with an active transmitter continues > 100 mW that is not feasible to support with the current battery technology for several hours additionally high-quality video images are necessary for accurate detection of the polyps and their classification. To realize an efficient backscatter system, the capsule antenna should be designed in an effective way in which it plays a crucial role in the system operation; in this project we have designed and integrated the antenna into the WCE prototype, the antenna generates a large radar cross section (RCS) by extending the antenna virtual size beyond its physical dimension. The antenna design is a part of OUS patent US20210059526A1 [18]. Figure 41 shows the antenna drawing with the details.

The antenna uses an active ultra-low power RF switch that alters the antenna RCS for an efficient modulation of the incident electromagnetic (EM) wave transmitted from outside the body. The antenna design considers the specific conditions of wave propagation in the biological environments and antenna loading with the lossy tissues. Polarization diversity using bistatic on-body reader antennas is used for communicating with the implant device. The on-body antennas can direct EM energy to the capsule device for improving the backscatter link performance. The on-body antenna is a part of OUS patent WO2021048303A1 [19] with a drawing as shown in Figure 42.

Figure 41 Concept of the integrated antenna inside WCE for RF backscatter application (US20210059526A1 [18])

Figure 42 Drawing of on-body antenna array for RF backscatter with the transmitter and receiver array on a flexible substrate to set on an on-body surface (WO2021048303A1 [19]).
In principle, the RF backscatter is a RADAR communication technique that uses the RF wave reflections from a target (Capsule) where the reflections are controlled by the data streamed from the capsule’s camera sensor. Using this approach, the capsule’s onboard communication system is reduced to a single switching gate that in principle consumes near to zero power (10 pJ/bit). Using this approach, the capsule’s price becomes extremely cheap, makes it a good case for mass production for a low-cost screening, low power system that can be operated for a longer time. By saving power, the image quality and the illumination can be improved by allocating the saved battery resources; also, other types of sensors can be used in the capsule. The antenna array can be worn to realize the near-field radar system with multiple transmitters and receivers to support the capsule localization and best signal fineing process can be handled. The array antenna can combat with signal fading that would occur due to different rotations and orientations that the capsule would take place in the pathway inside the large/ small intestine. The capsule of length 2.5 cm and diameter less than 1 cm is considered in the demonstration process as shown in Figure 43.

Figure 43 Prototype of the backscatter WCE with integrated electronics and the electrode antenna

An in-house backscatter radio system is developed to support the high data rate system up to 16 Mb/s, where we test it for 8 Mb/s and 10 Mb/s. The backscatter reader is a quite complicated subsystem with controlling units for antennas, received signal sensing, RF power controlling, coupling reduction circuits etc. The system is working at 434 MHz, where only a tone signal emission is the radiation in the environment and the data signal level is below the spectral limits that cannot be detected in a couple of centimetres from the body surface. Therefore, frequency allocation/permission routines that are limiting factors to enter the product to the market is not limiting for our technology. The reader system is mounted inside a shielded box as shown in Figure 44. The system supports up to 6 transmitter and 6 receivers antennas, in total 12 antennas that can be selected by an RF switch and using the communication protocol of the system. Figure 45 shows the array antenna configuration on the body worn, the selection of the active transmitter and receiver antennas are decided by the backscatter radar box based on the quality of the decoded data and the image reception, to provide reliable video streaming.

Figure 44 Backscatter system includes RF transmitter of controllable power up to 500 mW, the receiver system and clock data recovery system.
The output of the backscatter system can be arranged to display on a screen, or it has the main feature to be connected to an in-house FPGA/SOC based system to stream the data via UDP LAN standard to any wire-based network receiver or wireless interface to a Wi-Fi or 5G modem with LAN input. The data can be processed at the network edge, and the related polyp detection algorithms can be applied. Figure 46 shows the overall system level design. The swallowable capsule inside the intestine is detected by the RF backscatter reader system, in which the video data is recovered, and an indicator shows the validity of the signal reception. Once the data signal is validated, the data decoding is applied by the backscatter system and the formatted serial data is streamed. The LAN interface converts the serial data and clock data to standard UDP data protocol with high rate and low latency. To this point the system latency is below 1m sec. The LAN connection to 5G modem is followed and the video is streamed to the cloud.
We have developed a full working system in which the data is generated via dummy data (streamed from a computer) to replace the video from an endoscopy capsule scenario (see Figure 47). For this aim, we have implemented a micro-controller with LAN interconnection, an open-cv program is used to read a video file and stream the video to the microcontroller using the data format that can be read with the backscatter system in its data format. This part is to emulate a camera module in a capsule but with the possibility to stream a video file from pre-recorded endoscopy system to test the system and processing algorithms in the network edge for edge computing and AI application. This benchtop model is a very important stage to take in the project since realizing a standard capsule with all permission to use in human with possible patients with polyps is time consuming approach and is out of the scope of the project. Figure 47 shows the test setup we developed to stream a video file from the first computer via backscatter capsule to the second computer where we use a phantom to simulate the human body, a capsule mock-up with switching feature for backscattering that the switch is controlled by the data from PC/microcontroller. The video can be fed to the FPGA/LAN interface and can be streamed to 5G modem or second local computer for monitoring purposes or for local processing.

![Figure 47 Test setup using streaming video file and phantom human simulator.](image)

### 7.3.2 List of key performance indicators

The following application KPIs have been measured:

- End-to-end video streaming latency (ms) over the 5G-VINNI network from a client to an AI server,
- AI-based polyp detection performance in terms of sensitivity, specificity, and precision.
- Speed of the AI-based detector per frame.

### 7.3.3 Measurement and testing tools

We measured the End-to-End latency in milliseconds. The End-to-End latency is the time that a single frame requires to travel from the streaming device (the laptop PC or pillcam) to the AI server at the network edge. It is noteworthy to mention that the mean to read End-to-End latency include the processing time required by the AI model to handle the received frame.

The output of the proposed method is a set of bounding boxes around the suspected regions in the input frame. The detected bounding boxes are either true or false alarms. To quantitatively evaluate the performance of the proposed method, we calculate sensitivity (recall) and precision using well-known medical parameters:

- true positive (TP): a true detected bounding box around a positive region in the input frame,
- false positive (FP): a false detected bounding box around a negative region,
- true negative (TN): a true detection output for a negative frame in which no bounding box is detected.
- false negative (FN): false detection output where a polyp is missed in a positive frame.

**Sensitivity** measures the ratio of TPs to the total number of polyps in the test set,

\[
Sensitivity = \frac{TP}{TP + FN} \times 100,
\]

while **precision** measures the ratio of TPs to the total number of detected bounding boxes including FPs,

\[
Precision = \frac{TP}{TP + FP} \times 100,
\]

**specificity** measures the ratio of actual negative frames that are correctly classified,

\[
Specificity = \frac{TN}{TN + FP} \times 100,
\]

**Dataset and training details**

In this study, we used four publicly available datasets of still images and videos:

- **CVC-ColonDB**: This is a dataset of 300 still images extracted from 15 colonoscopy videos, each with a unique polyp (15 unique polyps in total). The images have a resolution of 574 x 500 pixels.

- **CVC-ClinicDB**: This is a dataset of 612 still images extracted from 29 colonoscopy videos, each with at least a polyp. There are 31 unique polyps presented 646 times in the 612 images with a pixel resolution of 384 x 288.

- **ASU-Mayo Clinic**: This is a dataset of 38 colonoscopy videos. 20 videos are assigned as a training set while the other 18 videos are assigned as a testing set. Because of the copyright license, we could not get access to the 18 testing videos. The 20 training videos consist of 10 positive videos with a total of 5402 frames (3846 polyp frames) and 10 negative videos with a total of 13500 frames.

- **CVC-ClinicVideoDB**: This is a dataset of 18 colonoscopy videos, each with a different polyp. This dataset comprises 11954 frames (10025 polyp frames). The resolution of the frames is 268 x 576 pixels.

During training, a video dataset is needed to capture the temporal patterns among neighbouring frames. In a video, the neighbouring frames look closely similar. If there are not enough diverse frames in the training videos, our detection model may easily get overfitted on the training frames. To train our detection model and avoid this phenomenon, we use the two datasets of still images namely CVC-ColonDB and CVC-ClinicDB alongside the dataset of videos namely CVC-ClinicVideoDB. We built our final training dataset by mixing the frames of the videos and the still images. Whenever a still image is encountered during training, we count it as its previous frames as well.

As mentioned before, the encoder part uses ResNet34 initialized with ImageNet pre-trained weights. In contrast, we randomly initialize the network parameters of the decoder part. To clean the final training dataset, we apply several simple pre-processing methods to the input images.

1. Image cropping: to remove the canvas around the informative part of the images.
2. Image resizing: by changing the image resolution to 512x512 because the pretrained ResNet34 accepts this image resolution.
3. Image normalization: by converting the pixel values from [0, 255] to [0, 1], subtracting them from the mean, and dividing them by standard deviation both pre-calculated from the ImageNet dataset.

To add further image-level diversity through depth and scale, we apply several image augmentation methods on the fly e.g., rotation, vertical, horizontal flips, random zoom-in (up to 25 %), and zoom-out (up to 50 %), and colour augmentations in HSV space. To keep the balance between large and small polyps and avoid biasing, we apply less zoom-in compared to zoom-out because the training dataset contains more large polyps than small ones.

We randomly split the training dataset into training (85 %) and validation (15 %) subsets. We use Adam optimizer with a batch size of 10 and a learning rate of \(1 \times 10^{-4}\) to train the model for 20 epochs. We change the learning rate to \(1 \times 10^{-5}\) to train the model up to 60 epochs. We use the validation subset to choose the learning rate decay strategy and the number of epochs.
Finally, we use the mean squared error (squared L2 norm) between each element in the input ground-truth image frame at time $t$, $Y_{xy}(t)$, and the output image frame at time $t$, $\hat{Y}_{xy}(t)$ with the predicted 2D Gaussian shapes.

$$L2 \text{ loss} = \frac{1}{M} \sum_{i}^{M} [Y_{xy}(t) - \hat{Y}_{xy}(t)]^2$$

where M is the batch size. We choose L2 norm loss function because it can significantly (quadratically) penalize large errors. This property of the L2 norm makes it favourable especially for the prediction of 2D Gaussian shapes which are normally distributed around a mean value.

7.3.4 Final results

End-to-End latency measurements

Table 9 demonstrates the performance of three transmission protocols used for video streaming. The RTP which is based on UDP is the fastest, end-to-end latency 46.74 ms. This is because UDP is a connectionless protocol meaning there is no acknowledgement signal sent back to the sender when a package is received.

We tested two different connection oriented protocols, HTTP and RTSP. When we used the HTTP for the video stream over the 5G network. We experienced a latency delay of $\sim$200 ms. HTTP is a TCP which is connection-oriented, and a connection between client and server is established before data can be sent. The server must be listening for connection requests from clients before a connection is established. Three-way handshake, retransmission, and error-detection adds to reliability but lengthens latency. RTSP demonstrated even worse performance, $\sim$ 450 ms latency.

It is noteworthy to mention that the End-to-End latency also includes the processing time required by the AI-based detection model per frame which was measured to be 10 ms on NVIDIA RTX 3090.

Table 9 End-to-End latency measurements for different transmission protocols

<table>
<thead>
<tr>
<th>Protocols</th>
<th>Latency [ms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDP-RTP</td>
<td>46.74</td>
</tr>
<tr>
<td>TCP-RTSP</td>
<td>466.48</td>
</tr>
<tr>
<td>TCP-HTTP</td>
<td>240.16</td>
</tr>
</tbody>
</table>

The AI-based model detection results

To quantitatively evaluate our proposed method, we used the ASU-Mayo clinic dataset, more specifically the 20 videos that were originally assigned for training purposes by the authors. The 10 positive videos were used to compute the performance of the proposed method in terms of sensitivity and precision. In contrast, the 10 negative videos were used for the evaluation of specificity. We present our results in curves to facilitate the visualization of the performance evaluation when information from previous frames is incorporated with the current frame.

Results on positive videos: Figure 48 shows the sensitivity and precision measurement setup of the proposed method for all four scenarios. When the current frame is examined alone, AlbuNet34 can provide high sensitivity (91.27 %) but struggles to offer the same level of performance for precision (67.21 %) due to the generation of a substantial number of FPs.

Wireless video capsule endoscopy is used to stream the captured high data rate (10 Mb/s to 20 Mb/s) videos to the 5G network. The Video will be captured using a NanoEye® camera sensor that transmits the raw video data of 12 Mb/s to the capsule onboard micro processing unit. The raw data controls the capsule antenna reflection coefficient that is sensed by an on-body array antenna using RF backscatter. A custom-designed FPGA which is bridged to the system in which the data is converted to a standard UDP LAN protocol. The interconnection from this stage follows standard interface connections.
7.4 Recommendations

The exponential development of the usefulness of AI in capsule endoscopy requires consideration of its medium- and long-term impact on clinical practice. Indeed, the advent of deep learning in the field of capsule endoscopy, with its evolutionary character, could lead to a paradigm shift in clinical activity in this setting.

The subcase is to demonstrate the feasibility of video streaming over a 5G network, all with acceptable latency from an endoscope capsule, receive and analyse the streamed video at the network edge, and provide feedback to the endoscope capsule with the help of unique backscatter communications. The capsule camera streams the video frames through the 5G network to the network edge where an AI based anomaly detection model is ready to process the received frames. The detection model is an efficient and accurate deep learning-based algorithm built from convolutional neural network (CNN) blocks.

In reference to Figure 47 of the demonstration of phantom experiment, the video data is created from a pre-recorded video from a clinical colonoscopy, to be able to test the full system operation. The system is completely working. The team has achieved the required level of progress in this discipline.
8 SUBCASE H3A: VITAL-SIGN PATCH PROTOTYPE

A single-use, Direct-to-Cloud, vital-signs patch prototype has been developed by Philips. The key challenges encountered are battery lifetime and cellular coverage. Exploring and addressing these challenges is Philips’ main contribution within the scope of this subcase.

8.1 Description and motivation

A single-use, Direct-to-Cloud (D2C), vital-signs patch is a smart band-aid that measures a patient’s vital signs 24/7 and communicates these directly through the cellular network. This way doctors can keep a tab on their patients, no matter where they are, potentially supported by AI algorithms running in the cloud. The concept developed is targeted at post-surgery monitoring. It can be used to monitor patients on the general ward and subsequently when they return to the comfort of their home.

This patch gathers vital signs once every five minutes but uploads it to the (AWS based) Philips Cloud less frequently, for example, only once every two hours. The latter leads to 140 to 150 bytes of binary payload. However, depending on the scenario, a more interoperable data format may be desired which enlarges the payload to about 1.7 kB. The available energy budget for a single upload operation is about 3 mWh.

The key challenges in developing such a device relate to battery life (i.e., minimizing energy expenditure) and coverage (i.e., still being able to upload vital signs data when the cellular network conditions are poor). Therefore, within the scope of 5G-HEART, experiments have been conducted to address these challenges and propose solutions. Specifically, during phase-3, the following three topics have been investigated:

- Selection and optimization of upload protocols, in relation to Research Question 1:
  Which protocols and protocol options could help reduce energy consumption for data upload by D2C wearables and how would they best be integrated with the Philips Cloud?

- Feasibility of NB-IoT for improved coverage, in relation to Research Question 2:
  Could the deep coverage modes of NB-IoT help address the coverage concerns for the patch, while staying within its limited energy budget and if so, what would be the implications for integration into the Philips Cloud?

- Feasibility of firmware over-the-air update, in relation to Research Question 3:
  Would it be possible to perform a single firmware over-the-air (FOTA) update for the patch, while staying within its limited energy budget?

The first two topics are a continuation of the work from phase-2. The additional results and insights obtained during phase-3 are reported in this deliverable. The last topic has been started in phase-3 and is also reported upon in this deliverable. For more background on the requirements and challenges, refer to the corresponding Section 8.1 of D3.3 [2].

8.2 Final setup

This section explains the final (phase-3) setup of the experimental use case. The same setup has been used as in phase 2 and as reported in Section 8.2 of D3.3 [2].

8.2.1 Network architecture

As explained in Section 8.2.1 of D3.3 [2], the (prototype) patches connect to live commercial LTE-M networks, using SIM cards with national and international roaming, to enable operation across the world.
8.2.2 User application architecture

A patient is equipped with a sensor that sends vital-signs “Direct-to-Cloud” via the cellular network. The hospital accesses the data from this cloud. The scope of the work within 5G-HEART relates to – efficiently, securely, and reliably – uploading vital-signs data to the Philips Cloud. This aspect is detailed out in Figure 49. Refer to the corresponding Section 8.2.2 of D3.3 [2] for more details.

![Figure 49 Architecture for H3A](image)

8.2.3 Hardware components

Different modems available on the market have been used for the experiments. The reader is referred to Section 8.3 for specifics.

8.2.4 Software components

The modem firmware is part of the modem vendor’s offering. The application firmware running on the MCU is being developed in-house by Philips, but less relevant for the experiments described in Section 8.3 (other than being the subject of FOTA).

8.3 Testing and verification

The experiments conducted, and the results obtained, continue to address the three research questions listed in Section 8.1. This is a logical continuation of phase-2 and for earlier experiments and results the reader is referred to the corresponding sections in D3.3 [2].

8.3.1 Methodology

An automated test framework was used to rigorously evaluate the impact of different protocols, protocol options, and radio access technologies (i.e., LTE-M vs. NB-IoT) on live commercial networks. See Section 8.3.3.

8.3.2 List of key performance indicators

As discussed in Section 8.3.2 of D3.3 [2], energy consumption and coverage are the key KPIs addressed for this subcase. The target values for the other project KPIs are not considered challenging in any way for today’s cellular networks.

8.3.3 Measurement and testing tools

An automated test framework, as depicted in Figure 50, has been developed in house to enable rigorous testing of different (IETF) protocols and protocol options on different modems, utilizing different 3GPP standards and features. Specifically, the framework enables repeated tests (for statistical significance) of energy consumption and peak power of the device, while simulating different coverage scenarios of the network. Testing is done on live commercial networks. Multiple instances of the test framework
have been deployed in several locations in the world. This allows experiments under different coverage conditions and using different MNO networks. For more details refer to Section 8.3.3 of D3.3 [2].

![Figure 50 Block-diagram of automated test framework for H3A](image)

### 8.3.4 Final results

#### Selection and optimization of upload protocols

This topic addressed the first research question: “Which protocols and protocol options could help reduce energy consumption for data upload by D2C wearables and how would they best be integrated with the Philips Cloud?”.

To a large extent this question has already been addressed in phase-2, as described in the corresponding section of D3.3 [2]. However, two further explorations were conducted: (1) assessing the benefits of TLS session resumption across load-balanced nodes and (2) drawing in CoAP over DTLS, as a third protocol, in addition to the comparison between “Plain UDP” and HTTPS/OAuth that was previously reported in D3.3 [2]. The measurements for both these sub-topics were conducted on Sequans GM02s.

These two explorations are discussed below, followed by a more overarching conclusion on the overall work conducted with regards to this topic, to answer the first research question.

#### TLS Session resumption

##### Introduction

TLS session resumption [20] is a well-known mechanism to reduce the overhead of the TLS handshake (HTTPS = HTTP/TLS/TCP/IP). The session keys negotiated for a first upload can be reused for the next ones, at least until they expire, saving a handshake as well as the exchange of certificates. These certificates can be several kilobytes in size, potentially dwarfing the size of the actual vital-signs payload. Two variants exist:

- **Session resumption with session identifiers [20]:** the client and the server each store the session data – parameters, cipher suite and session keys – on their own behalf. For the initial session, the server creates a unique session ID and returns it to the client in the Server Hello message. The server will remember the session data for each client session it had, identified by this session ID. If, for a subsequent session, the client fills in that very same session ID in the Client Hello, the session data for this (now resumed) session can be reused and a shorter handshake suffices.

This mechanism is also known as stateful session resumption because the server is required to maintain some state. A typical server will not remember those session ID’s and data forever and will eventually purge them to avoid scalability issues (e.g., after 10 minutes). This may particularly be a problem for devices like vital-signs patches that upload data only once every few hours. The
client will get confirmation of the session resumption being successful if it receives the same session ID in the Server Hello that it sent in the preceding Client Hello. If it sees a different, new, session ID instead, the server could not resume, and a full four-way handshake will follow to establish a new session. The client can also choose not to resume a session by filling in an empty session ID in the Client Hello message. Also, in this case the full, four-way handshake will follow.

- Session resumption with session tickets [21]: in this case the client also stores the session data on behalf of the server. Therefore, the server does not have to retain any state about the session, which is therefore called stateless session resumption. Therefore, session tickets may have a significantly longer lifetime than session ID’s (e.g., 12 or 24 hours). At the end of the initial, four-way handshake, the server returns all session data to the client as an encrypted blob called a session ticket. Only the server has the key to decrypt it. If, for a subsequent session, the client includes this session ticket in the Client Hello, the server and client can resume the session using a two-way handshake without certificate download. One drawback of session tickets over session ID’s is the fact that session tickets contain more data than session ID’s (e.g., 105 bytes instead of 32 bytes). However, the impact on energy consumption is expected to be minimal.

Figure 51 (simplified view) illustrates a first session starting with a full, four-way, handshake, followed by a second resumed session with a shortened, two-way, handshake that also foregoes the certificate download. As can be seen, the exchanges with session ID and session ticket are similar but, as mentioned above, the session ticket is larger, yet avoids state retention by the server.

![Figure 51 Initial and resumed TLS handshake with session ID resp. session tickets (indicative, simplified) (H3A)](image)

Note that the server includes a timestamp in the session tickets allowing it to expire them after some amount of time for security reasons. So, even with session tickets it may not be possible to resume a session after a certain amount of time.
Load balancing and TLS session resumption

As the term suggests, load balancing enables a cloud service provider to balance the load of requests over a multitude of physical servers (i.e., nodes) offering scalability and availability. AWS offers three different types of load balancers [22]:

- **Classical Load Balancers (CLB):** operate at OSI layers 4 and 7. The IP address of the server changes over time.
- **Application Load Balancers (ALB):** operate at OSI layer 7. The IP address of the server changes over time.
- **Network Load Balancers (NLB):** operate at OSI layer 4. The IP address of the server remains the same within a single availability zone. Any AWS region can contain multiple availability zones.

The type of, and extent to which, session resumption is supported by AWS services depends on the type of load balancer deployed [23]:

- Classic Load Balancers support session ID-based session resumption but don't support session ticket-based session resumption. Session caching is supported at the node level. This means that if a client connects to node B using the session ID received from node A, then the handshake reverts to a full handshake. After that, a new session ID is generated by node B.
- Application Load Balancers support both session ID and session ticket-based session resumption. Both session IDs and session tickets are supported at the node level. This means that if a client connects to node B using the session ID or session ticket received from node A, then the handshake reverts to a full handshake. After that, a new SSL session ID and session ticket are generated by node B.
- Network Load Balancers support only session tickets for session resumption. Resumption using session tickets is supported at the regional level. Clients can resume TLS sessions with a Network Load Balancer using any of its IP addresses.

Review of phase-2 experiments

During phase-2 an initial experiment was conducted to assess the potential benefits of TLS session resumption, that is, its potential for energy reduction. This experiment involved the use of session IDs (not session tickets), as well as a dedicated proxy server to avoid load balancing. Session tickets were avoided because the modem available at that time only supported session IDs. Load balancing was avoided because the Philips cloud services did not support load balancing across nodes at the time. The use of the proxy meant that all communication was terminated in a single node i.e., a single physical server. This proxy server subsequently forwarded the requests to the (load balanced) Philips, Amazon-based, cloud servers. The conclusion, albeit with many disclaimers, was that a resumed handshake would reduce the energy consumption by 19 % relative to a non-resumed one. Refer to the corresponding section in D3.3 [2] for further details.

However, a more realistic scenario, would involve communication with a load balanced cloud server using session tickets. Note that the typical lifetime of session IDs (e.g., 10 minutes) is too small to be meaningful for the vital-signs patch use case with its e.g., two-hourly uploads, whereas the typical lifetime of a session ticket (e.g., 12 or 24 hours) would be very useful in this context.

Experiments with session tickets and load balancing

During phase-3 a new cloud endpoint became available that did support session resumption across load-balanced nodes (at least within a single availability zone, which was sufficient) and at the same time the Sequans GM02s modem got support for TLS session resumption with session tickets. Therefore, a new set of experiments was conducted to evaluate abovementioned more realistic scenario.

These experiments involved uploading vital signs data to this new cloud endpoint, each upload having a 1.7 kB payload. Note that standards-compliant data formatting was used instead of just sending the 140-150 bytes raw binary payload (as described in Section 8.1 of D3.3 [2]), leading to a larger payload.
Table 10 shows that the amount of data transferred during the resumed handshake itself is reduced by a factor 9, compared to the full handshake. This is mainly because downloading the certificates is skipped (and obviously dependent on the size of the certificate chain). Table 11 shows the same comparison, yet relative to the total amount of data exchanged (i.e., for both connection setup and transfer of the payload). The total amount of payload is reduced by 54 % (i.e., 46 % remaining). Obviously, there will still be a full handshake when the 24-hour ticket expires. However, even in case of a two-hourly upload, this affects only one out of every 12 uploads, still leading to an overall data reduction of 49.5%.

Table 10 Data reduction for full vs. resumed handshake for vital-signs patch use case, handshake only.

<table>
<thead>
<tr>
<th></th>
<th>handshake only</th>
<th>full handshake</th>
<th>resumed handshake</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upload</td>
<td>807</td>
<td>503</td>
<td>62%</td>
<td></td>
</tr>
<tr>
<td>Download</td>
<td>5911</td>
<td>266</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6718</td>
<td>769</td>
<td>11%</td>
<td></td>
</tr>
</tbody>
</table>

Table 11 Data reduction for full vs. resumed handshake for vital-signs patch use case, full upload (H3A).

<table>
<thead>
<tr>
<th></th>
<th>full upload</th>
<th>full handshake</th>
<th>resumed handshake</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upload</td>
<td>3797</td>
<td>3493</td>
<td>92%</td>
<td></td>
</tr>
<tr>
<td>Download</td>
<td>7235</td>
<td>1590</td>
<td>22%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11032</td>
<td>5083</td>
<td>46%</td>
<td></td>
</tr>
</tbody>
</table>

Table 12 shows the energy gains observed during the experiments. The energy and the energy gains are shown relative to just the TLS handshake (row “TLS”), as well as relative to the entire upload operation at IP-level, that is excluding the attach and detach (row “All”). These experiments were executed on four different instantiations of the test framework, each being in a different location. This meant that they were connected to different MNOs and that they experienced different signal conditions, as exemplified by the RSRP values reported.

Table 12 Energy gains of session resumption for vital-signs patch use case (H3A)

<table>
<thead>
<tr>
<th>Eindhoven, Netherlands</th>
<th>Aachen, Germany</th>
<th>Dover, US</th>
<th>Bothell, US</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of sessions not resumed</td>
<td>Full TLS</td>
<td>Resume TLS</td>
<td>Savings</td>
</tr>
<tr>
<td>15%</td>
<td>21%</td>
<td>15%</td>
<td>24%</td>
</tr>
<tr>
<td>TLS [mWh]</td>
<td>0.18</td>
<td>0.10</td>
<td>44.4%</td>
</tr>
<tr>
<td>All [mWh]</td>
<td>0.71</td>
<td>0.57</td>
<td>19.7%</td>
</tr>
<tr>
<td>RSRP [dBm]</td>
<td>-101</td>
<td>-106</td>
<td>-112</td>
</tr>
</tbody>
</table>

As these were early experiments with both the new cloud endpoint and the new modem feature, some bugs were still present in the system. These bugs caused a relatively high number (15 % to 24 %) of the sessions not being resumed, even though the session tickets were still valid. Although in the meantime most issues have been resolved, the experiments have not been repeated as the results could easily be extrapolated.

In conclusion, the use of session resumption reduces the energy needed for the vital-signs upload by 11 % to 27 % despite aforementioned bugs. Correcting for those, one would expect to get gains of 14 % to 33 %, for this particular use case.
Observe that the phase-2 experiments reported a 19% reduction relative to the handshake alone, whereas the current results (row “TLS”) show gains of 26% to 44% (not even correcting for the bugs). However, as written, the phase-2 results were obtained with another modem (u-blox R5) and many disclaimers applied at the time.

HTTPS / OAuth 2.0 vs. CoAP over DTLS vs. UDP

Considering that CoAP over DTLS involves fewer handshake than HTTPS – and that typical deployments favour the smaller ECC certificates over RSA certificates – it is worthwhile to assess its potential gains in energy consumption. Therefore, the earlier comparative measurements presented in Section 8.3.4 of D3.3 [2] – see under “Plain UDP vs. HTTPS/OAuth (and other protocols) at varying coverage” – have been repeated by adding CoAP over DTLS as a third option, considering that in the meantime at least some modems do support it properly. For these experiments, the Sequans GM02s modem has been selected instead of Nordic nRF91 which was used for the earlier measurements.

Note that OAuth 2.0 is a client authentication mechanism. It involves a client (device) first contacting a dedicated Identity and Access Management (IAM) cloud service, providing a username and a password to it, and getting a token in return. This token grants permission for accessing a particular cloud service to perform the desired operation, such as uploading some vital signs. That is, the client can subsequently contact that particular cloud service to upload its vital-signs data, while authenticating itself to it by means of the token. This way identity and access management is decoupled from any other cloud service. Note that both service accesses (i.e., the one to IAM and the subsequent one to the particular cloud service) require a full HTTPS handshake.

The parameters used for each of the candidate protocols in the experiment are shown in Table 13. Client authentication (via IAM or otherwise) is not considered in the evaluation. Furthermore, session resumption is not used, neither for CoAP over DTLS nor for HTTPS. Note that having session resumption would have reduced the relative differences between the options.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Certificates</th>
<th>Type</th>
<th>Chain length</th>
<th>Total size (bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain UDP</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>150</td>
</tr>
<tr>
<td>CoAP over DTLS</td>
<td>ECC</td>
<td>1</td>
<td>872 B</td>
<td>150</td>
</tr>
<tr>
<td>HTTPS (/OAuth 2.0)</td>
<td>RSA</td>
<td>4</td>
<td>4.3 kB</td>
<td>1619</td>
</tr>
</tbody>
</table>

Whereas the HTTPS protocol option is intended to be cloud-native and interoperable, the CoAP over DTLS option targets to obtain security and reliability in the most efficient way (e.g., it assumes a binary payload of only 150 bytes), yet while still using PKI to avoid complexities in deployment (i.e., no pre-shared keys). In this case, a self-signed certificate has been used leading to a “single-shackle” chain.

A total of 36 runs were executed on the test framework (see Section 8.3.3). Each run involved a series of measurements at increasing attenuation levels, starting without attenuation and then stepwise increasing the attenuation in 3 dB steps until reaching a total of 27 dB (the 3 dB and 6 dB attenuation levels were skipped). The base level RSRP was about -95 dBm for all the runs, so at deepest attenuation the signal level was around -125 dBm. Each measurement in the series involved executing all three protocols in sequence, preceded by a single attach, and followed by the corresponding detach; that is: attach, UDP, CoAP over DTLS, HTTPS, detach. In other words, the protocol measurements shared the same attach/detach. The energy and duration of attach/detach was included in the corresponding totals for each of the three protocols. This interleaving approach was taken to minimize the spread in the results. Note, that measurements were conducted on a live commercial network (Vodafone NL on the High Tech Campus in Eindhoven) and that this naturally leads to some spread in results.

Figure 52, Figure 53, and Figure 54 show the measurement results for the respective protocols. The distribution of the total energy consumed for the upload is shown for each of the attenuation levels. As
explained above, the total energy includes attach, setting up the (D)TLS connection (where applicable), uploading the data and receiving the acknowledgement, tearing down the connection (where applicable) and detach. The number of runs that have been successful at a particular attenuation level is indicated by the n on the horizontal axes.

Figure 52 Distribution of total energy as function of attenuation for Plain UDP (H3A)

Figure 53 Distribution of total energy as function of attenuation for CoAP over DTLS (H3A)

Figure 54 Distribution of total energy as function of attenuation for HTTPS (H3A)

Figure 55 and Figure 56 compare the three candidate protocols in terms of energy consumption at good (0 dB attenuation, RSRP ~ -95 dBm) and at poor coverage (24 dB attenuation, RSRP ~ -120 dBm)
respectively. Note that the second comparison was not shown for 27 dB, as too few of the CoAP over DTLS and HTTPS uploads (i.e., only 3 resp. 1 out of 35) were successful for the results to be significant.

![Figure 55 Comparison of total energy between protocols at good coverage (H3A)](image)

At good coverage, the average energy spent per (successful) upload is 0.32 mWh for UDP, 0.47 mWh for CoAP and 0.54 mWh for HTTPS. This means that HTTPS uses 1.7x times more energy than UDP and that CoAP over DTLS uses 1.5x more energy than UDP. HTTPS uses just 1.15x more energy than CoAP over DTLS.

![Figure 56 Comparison of total energy between protocols at poor coverage (H3A)](image)

At poor coverage, the average energy spent per (successful) upload is 0.75 mWh for UDP, 2.14 mWh for CoAP and 3.51 mWh for HTTPS. This means that HTTPS uses 4.7 times more energy than UDP and that CoAP over DTLS uses 2.9 times more energy than UDP. HTTPS uses 1.6 times more energy than CoAP over DTLS.

In conclusion, CoAP over DTLS could be an interesting alternative to HTTPS, considering that poor coverage provides the critical design criteria for D2C wearables (i.e., it should also work for patients that find themselves in poor coverage most of the time). However, further optimization of HTTPS (such as session resumption, and shorter certificate chains) will reduce the relative gains.

When trying to compare these results to those of the earlier experiments on Nordic nRF91 (described in D3.3 [2]), one should realize that in the earlier experiments each HTTPS / OAuth 2.0 upload involved accessing IAM, whereas for the current experiments IAM accesses have been omitted (i.e., the current experiments are more about just “HTTPS” than about “HTTPS / OAuth 2.0”). In other words, the results of the Sequans experiments emulate OAuth 2.0 token lifetime extension, which is discussed in Section 8.3.4 of D3.3 [2] (see under “OAuth token lifetime reduction”). Specifically, a 42 % energy reduction is estimated for good coverage and 33 % for poor coverage.
Table 14 shows a very indicative comparison of the results obtained for the two modems. Note that comparing the results of these two experiments should be done with great care. Firstly, they were conducted some two years apart, during which network deployments and conditions, modem firmware, and also the test framework itself have changed considerably. Secondly, obviously, the modems have different characteristics. Thirdly, the extrapolations for session resumption are based on yet other experiments. The energy figures for Nordic are taken from the 0 dB and the 24 dB rows of Figure 53 of D3.3 [2]. The energy figures for Sequans can be found in the text above.

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Modem</th>
<th>(Extrapolated) upload energy (mWh)</th>
<th>HTTPS vs. UDP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>UDP</td>
<td>HTTPS</td>
</tr>
<tr>
<td>Good (0 dB att.)</td>
<td>Nordic nRF91</td>
<td>0.09</td>
<td>0.37 - 42% = 0.21</td>
</tr>
<tr>
<td></td>
<td>Sequans GM02s</td>
<td>0.32</td>
<td>0.54</td>
</tr>
<tr>
<td>Poor (24 dB att.)</td>
<td>Nordic nRF91</td>
<td>0.17</td>
<td>1.49 - 33% = 1.00</td>
</tr>
<tr>
<td></td>
<td>Sequans GM02s</td>
<td>0.75</td>
<td>3.51</td>
</tr>
</tbody>
</table>

Considering all these disclaimers, the relative differences between both protocols are still in the same ballpark; about a factor two at good coverage and a factor 5-6 at poor coverage.

**Overarching conclusions on upload protocol optimizations**

Secure and reliable upload of data via the cellular network involves the following steps:

1. Wake up the modem and attach to cellular network (or resume existing connection).
2. Establish (or resume) secure IP connection to the cloud (e.g., DNS, TCP, and TLS).
3. Transmit data and wait for acknowledgement (e.g., HTTPS Request + Response).
4. Tear down secure IP connection.
5. Detach from cellular network (or go to PSM).

Or in other words, when analysing energy usage, three aspects (or layers) should be distinguished:

A. The cellular connection (steps 1+5).
B. The secure IP connection (steps 2+4).
C. The actual (reliable) data transfer over those connections (step 3).

For a use case involving the upload of just a small amount of payload, such as is the case for the post-surgery vital-signs patch, the impact of connection setup (A+B) on energy usage is quite significant and dominates that of actual data transmission (C). Similarly, the impact of IP connection setup (B) on data cost can be quite significant. The focus of the current research has been on optimizing IP connection setup and teardown (B). Although experimenting with more advanced cellular features (A) has been considered in the beginning of the project, the unavailability (that is unreachability) of experimental facilities caused by the COVID-19 travel restrictions have prohibited such experiments. In addition, also modem vendors were lacking in support for those features. Furthermore, already early in the project it became clear that the anticipated use case was more-or-less feasible leveraging existing commercial LTE-M deployments.

Addressing research question 1, a wide variety of (mostly) IETF protocols and protocol options offering secure IP connectivity have been presented but only few of those are of practical significance today, considering their support by modem and cloud vendors. However, deploying those will already lead to significant savings and applying them should be considered low-hanging fruit.

Particularly, the new cloud endpoint with proper support for session resumption (i.e., ticket-based and across nodes), combined with an extended OAuth 2.0 token lifetime, leads to a very substantial reduction
in handshakes (mostly due to token lifetime extension), as well as to a substantial reduction in data communicated (mostly due to session resumption).

Further gains could be obtained by reducing the certificate chain size. However, the law of diminishing returns applies, considering the use of session resumption and token lifetime extension. These optimizations will be more impactful for use cases that only update data very infrequently (e.g., daily), as in those cases session resumption and token lifetime extension will be less impactful (or not at all).

Going beyond the protocols and protocol options supported by cloud vendors – i.e., opting for dedicated, home-grown endpoints – does not seem to be worth the effort. For example, CoAP over DTLS can still provide some gains over HTTPS, but easier to deploy protocol optimizations, like session resumption, short certificate chains and, hopefully soon, TLS 1.3, will further diminish those gains.

Promising protocol developments on the horizon, such as QUIC, could in the future lead to further improvements, although the law of diminishing returns will apply here as well. It is therefore recommended to monitor future developments in this respect by cloud providers and modem vendors.

As far as the cellular layer (A) is concerned, PSM+RAI will eventually become available for LTE-M (currently it is only available for NB-IoT) and will enable further optimization in that layer.

Feasibility of NB-IoT for improved coverage

This topic addressed the second research question: “Could the deep coverage modes of NB-IoT help address the coverage concerns for the patch, while staying within its limited energy budget and if so, what would be the implications for integration into the Philips Cloud?”. Note, that LTE-M is the default RAT choice for the patch prototype, i.e., a comparison between NB-IoT and LTE-M has been conducted to assess potential gains in deep coverage.

During phase-2 an initial assessment was done, concluding that attach to an NB-IoT network is possible at -12 dB relative to attach to an LTE-M network while still staying within the 3 mWh energy budget for a single upload. Also, it was noticed that RSRP levels below -140 dBm were not possible for NB-IoT, even though the maximum number of repetitions (2048) was not actually used. Refer to the corresponding section in D3.3 [2] for further details.

However, these experiments did not involve the actual data transfer. Therefore, additional experiments have been conducted to compare a very lean-and-mean protocol over NB-IoT (CoAP with out-of-band established keys) to a full-blown cloud-native protocol over LTE-M, considering the very low bandwidths offered by NB-IoT at deep coverage levels. Also, the impact of payload size on energy usage has been assessed for this lean-and-mean protocol over NB-IoT. Note that a summary of most of these results have also been presented at the EuCNC 22 conference [24], as part of the project’s dissemination activities.

Lean-and-mean protocol

NB-IoT is expected to have 15 to 20 dB better coverage than LTE-M, considering that its CE Mode B is not deployed in practice [25]. Due to massive repetitions (up to 2048 in CE2 compared to 32 for LTE-M’s CE Mode B), such deep coverage implies extremely low data rates and consequently protocol overhead becomes a major concern. Given the large number of repetitions required in bad coverage situations, this would imply that either the battery will run out very quickly or that the protocol will time out before that. Also, observe that operators discourage using HTTP-based protocols over NB-IoT as being too chatty relative to its limited bandwidth.

Consequently, a much more lean-and-mean protocol is needed. Nevertheless, integrity, authenticity, confidentiality, and reliability remain hard requirements. As such authenticated encryption (e.g., AES-CCM) and data acknowledgement (with retransmission) are unavoidable. However, easy cloud integration and device provisioning could be compromised in favour of deep coverage, while still fitting the power budget. Therefore, CoAP, a lean-and-mean, UDP-based protocol is proposed, comprising the following elements:
- One uplink CoAP confirmable message containing the vital signs.
- One downlink CoAP ACK for end-to-end acknowledgement.
- VPN (edge) deployment to avoid DDoS attacks.
- Binary-encoded vital-signs data.
- Reduced sampling rate if acceptable to the application (i.e., less than once per 5 minutes).
- In-factory distribution of (symmetric) encryption keys.
- AES-CCM for confidentiality/integrity/authenticity (using a minimum acceptable MAC size).

The purpose of the current research is therefore to assess the potential for coverage improvement by using NB-IoT in combination with this very simple CoAP protocol. It should be regarded as a first feasibility study to assess whether this approach is worth pursuing any further.

Specifically, it should be noted that actual deployment of this protocol requires the development of a dedicated backend service (for protocol and content encoding conversion), involving challenges with regards to authentication, integrity, and confidentiality, all accomplished by TLS when using HTTPS / OAuth 2.0. In other words, the potential gains in coverage will not come for free. Further technical and business evaluation will have to show whether such gains would be worth the effort. Furthermore, the current analysis is based on comparing the alternatives in a fixed setting, connecting to a live network, where both RATs are served by the same base station. Actual gains in the field will also depend on the relative deployment of LTE-M vs. NB-IoT in a certain geography, but these aspects are left for future work.

From a hardware perspective, using NB-IoT instead of – or in addition to – LTE-M is very straightforward, considering that most – if not all – LTE-M modems on the market support NB-IoT as well.

**Experimental setup**

To summarize the above, the aim of these experiments is to assess whether the deep coverage modes of NB-IoT can help address coverage concerns for D2C wearables, while respecting their very constrained energy budgets. This is done by comparing three alternatives:

1. HTTPS / OAuth 2.0 vital signs upload over LTE-M.
2. CoAP vital signs upload over LTE-M.
3. CoAP vital signs upload over NB-IoT.

Specifically, for CoAP the uplink packet comprises a 10-byte CoAP header followed by a 35-byte payload, meant for carrying AES-CCM encrypted, subsampled, application data and MAC. The downlink ACK contains a 4-byte CoAP header. As mentioned above, the encryption keys must be established out-of-band, as the use of full-blown DTLS would substantially increase protocol overhead. This solution is functionally comparable to CoAP/DTLS with pre-shared keys but involves just 2 instead of 4-6 message exchanges. A full-blown DTLS solution with in-band key agreement would require server authentication with certificates and lead to substantially more overhead. The 35-byte payload allows for some 20 bytes of actual vital-sign data, necessitating subsampling to only once per 30 minutes, instead of the desired once per 5 minutes.

In contrast, the HTTPS / OAuth 2.0 baseline option assumes a standards-compliant payload comprising 1.7 kB, which is used as parameter in these experiments.

For each of the alternatives, upload operations are attempted to a commercial LTE-M / NB-IoT network in an alternating manner, while stepwise increasing the level of attenuation, until the point where the operation fails because of the signal conditions becoming too poor. For each of these upload operations, the energy consumption of the modem is measured as well. It is important to note that NB-IoT (user plane) and LTE-M are accessed via the same base station. The automated test framework (see Section 8.3.3) is used to execute the experiments using a Nordic nRF91. The latter has been selected because it is very energy-efficient and supports easy switching between LTE-M and NB-IoT.
Subsequently, a second round of experiments is conducted to investigate the impact of payload size on energy consumption for NB-IoT CoAP. Specifically, the energy consumption of sending 200 bytes of payload, compared to sending 45 bytes of payload (including the CoAP header), is measured for different coverage levels. The 200 bytes of payload correspond to the situation where vital-signs are still sent every 5 minutes (i.e., no subsampling).

**Measurement results**

Figure 57 shows the results of the experiments. Specifically, it shows 22 runs, each run comprising a sequence of transfers (e.g., vital signs uploads) at increasing levels of attenuation, for each of the three successive alternatives. The attenuation level is initially increased in 10 dB steps, followed by 2 dB steps for attenuation levels above 20 dB for LTE-M and above 40 dB for NB-IoT. The transition between the two step sizes has been determined experimentally during earlier runs. Note that the experiments for the three alternatives have deliberately been interleaved to average out slow variations in RF conditions.

The diagram clearly demonstrates that NB-IoT is able to withstand higher levels of attenuation than LTE-M but also that LTE-M CoAP is better capable of coping with attenuation than LTE-M HTTPS / OAuth 2.0. The latter is caused by the fact that at higher levels of attenuation the probability of transmission failures increases, giving short transfers a higher probability of being successful than long ones. Specifically, it has been observed that for HTTPS / OAuth 2.0 often the attach succeeds but that it fails at either the first or the second HTTPS transaction (OAuth involves two HTTPS transactions, one for authentication followed by another one for application data exchange, vital-signs upload in this case). Figure 58 summarizes the results in a box plot, showing an average gain in link budget of 10.5 dB for NB-IoT CoAP over LTE-M HTTPS / OAuth 2.0, of which 6.5 dB can be attributed to NB-IoT vs. LTE-M (i.e., the RAT), while the less chatty protocol (CoAP vs. HTTPS / OAuth 2.0) contributes 4 dB.
The 6.5 dB attributed to the RAT is substantially lower than the 15 dB to 20 dB quoted above. This could relate to the way NB-IoT is deployed, or more specifically to a lower bound to RSRP and/or the so-called deployment, mode as discussed below. However, also note that the experiments were conducted in a life network, a largely uncontrolled setting, leading to variations in results.

Further analysis of actual RSRP levels reveals that those below -140 dBm are not observed for NB-IoT (see Figure 59), while they are – albeit sporadically – for LTE-M (see Figure 60). This may be the result of the cell configuration not permitting UEs to camp on the cell with RSRP values lower than -140 dBm (i.e., q-RxLevMin-r13 and q-RxLevMinOffset setting in SIB messages [26],[27]). A simple experiment on another provider’s network also shows that attach is not possible below -140 dBm. As a sidenote, observe that LTE-M experiences lower RSRP at a similar attenuation level (i.e., RSSI). This can be explained by the difference in the number of resource blocks.
Furthermore, at maximum attenuation, the repetition factor for LTE-M maxes out at the expected 32 for CE Mode A for uplink (and 16 for downlink). However, for NB-IoT the maximum observed was 32 for uplink and 512 for downlink, whilst CE Level 2 supports up to 2048 repetitions (see Figure 61 and Figure 62). The number of runs that have been successful at a particular attenuation level is indicated by the n on the horizontal axis. Attenuation levels with a success rate below 75 % are coloured yellow to indicate that too many samples are missed to be useful in practice. Although this boundary seems rather arbitrary, in various RPM scenarios some missed samples are acceptable.
Figure 62 Maximum DL repetition factor for LTE-M (top) and NB-IoT (bottom) vs. attenuation (H3A)

It is expected that with 2048 repetitions an additional 6 dB could be gained, getting closer to the predicted 15-20 dB. In addition, the operator uses guard-band deployment, while standalone deployment offers the largest link budget [28].

Figure 63, Figure 64 and Figure 65 show the energy consumption for the three respective alternatives as a function of the attenuation level (the n and the yellow colouring are consistent with the figures above).

Figure 63 Energy consumption vs. attenuation; HTTPS / OAuth 2.0 vital signs upload over LTE-M (H3A)
At their respective (and different) highest attenuation levels both LTE-M HTTPS / OAuth 2.0 and NB-IoT CoAP max out at an average energy consumption of around 3 mWh, whereas for LTE-M CoAP this is less than 1 mWh. For the vital-signs patch, these numbers equal the available energy budget. However, for ‘ideal’ NB-IoT deployments (i.e., standalone deployment without limiting q-RxLevMin-r13 and q-RxLevMinOffset settings) energy consumption is likely to exceed that budget.

It should be noted that the modem is switched off between successive transfers, necessitating an attach operation for each one of them. The combination of PSM and Release Assist Indication (RAI) would allow the device to avoid the attach operation, reducing the energy consumption at the expense of a small sleep current. This means that for NB-IoT CoAP energy consumption can be further reduced. Current LTE-M deployments based on Release 13 do not support RAI yet, this may change if RAI, based on Release 14, is deployed by operators and modem vendors.

Figure 66 depicts the results of the second round of experiments; the payload sensitivity experiments for CoAP over NB-IoT. For the 45-byte as well as the 200-byte case (both numbers include the 10-byte CoAP header), 25 upload operations were attempted for each attenuation level. These iterations were equally spread over a 24-hour period to also assess the impact of time-of-day. Although a few messages got through at 50 dB attenuation, most were not successful. Therefore, this is not shown in the graph. Also, at 40 dB to 48 dB – most notably 48 dB – some messages did not arrive (at 48 dB attenuation of the 25 messages, 16 arrived for the 45-byte case and 12 for the 200-byte case). Next to the total energy usage for the upload operation (red and orange graphs), also the contribution of the CoAP message exchange (i.e., without attach/detach) are shown (blue and green graphs). Also note that significant variation was experience in between different measurements; a consequence of using a live public network.
When increasing the payload size from 45 to 200 bytes, a significant increase in energy consumption is observed only for attenuation levels of 40 dB (RSRP ≤ -119 dBm) and above. At 48 dB attenuation (RSRP = -137 dBm), an increase of 28% can be noticed (red line vs. green line). As expected, this difference can be attributed to the CoAP message exchange. Obviously, for larger payloads this difference will become larger but at least subsampling of observations (e.g., sending 30-minute instead of 5-minute averages) could be avoided.

Using DTLS with server authentication would likely still be challenging (considering a significant increase in the number of message flights and the amount of data exchanged), although elliptic curve cryptography, short certificate chains, root-certificate caching and session resumption could help alleviate the problem (see also Section 8.3.4 of D3.3 [2] for an explanation of some of these concepts). More work would be needed to assess the potential of such an approach.

Figure 67 shows again that RSRP levels below -140 dBm are not observed for any of the 2x25 iterations for each of the attenuation levels (including 50 dB). Another interesting observation is that at certain level of attenuation, generally higher levels of RSRP are seen in the middle part of the graph. These iterations were conducted outside office hours at the High Tech Campus in Eindhoven. In other words, the absence of other users on the mobile network improves performance and furthermore seems to make it more predictable.
Conclusions & observations

In conclusion, providing a preliminary answer to research question 2, NB-IoT combined with CoAP, a lean-and-mean, UDP-based upload protocol seems a promising alternative to LTE-M with the cloud-native HTTPS / OAuth 2.0 protocol suite, as it could improve the link budget by about 10 dB, while not exceeding the 3 mWh energy budget for vital-signs upload. Although minimizing payload size is a key design criterion for acceptable performance in deep coverage, it is not necessary to subsample vital-signs observations (i.e., a 200-byte payload would still be acceptable).

However, this approach requires a considerable investment in dedicated backend solutions and in-factory key distribution. This is necessitated by the requirement to have a secure and reliable, end-to-end, Direct-to-Cloud connection, something which comes for free with cloud-native solutions.

More work would be needed to reach a more definitive conclusion. Specifically, field tests should be conducted to include the influence of actual network deployments (i.e., NB-IoT vs. LTE-M). That is, additional coverage assessments in hospitals of devices positioned close to the body are needed. The performance of the LTE-M HTTPS / OAuth 2.0 device should be compared to a similar one based on NB-IoT CoAP. In this context, it should also be noted that attenuated RF is not necessarily fully representative of poor coverage, as under poor coverage also the CINR (Carrier to Interference plus Noise Ratio) may be affected. Another quality measure to consider is RSRQ (Reference Signal Received Quality). Deployment of the (remotely controlled) test framework in a location with actual poor coverage would be a complementary approach to field tests in addressing those shortcomings.

Other topics for future research are the impact of PSM and RAI on (at least) NB-IoT energy consumption, the performance of (optimized) DTLS as security layer underneath CoAP, as well as further studies into the influence of deployment modes and the q-RxLevMin-r13 and q-RxLevMinOffset settings. The latter investigations should involve MNOs to better understand why these limitations are deployed and whether it would be technically and commercially feasible to adapt those deployments for further extended coverage. The BEST initiative [29] might provide an affordable alternative to in-factory key-distribution. The EDT feature could be considered for further reducing communication overhead. It is even conceivable to use non-IP communication via an operator’s CDP, although that will likely compromise the possibilities for roaming. Finally, for the longer-term future, 5G Reduced Capability (RedCap) UEs might be considered as well [30].

Feasibility of firmware over-the-air update

This topic addressed the third research question: “Would it be possible to perform a single firmware over-the-air (FOTA) update for the patch, while staying within its limited energy budget?”.

The use case of FOTA update has very different characteristics from the upload use case addressed by the other two topics; the ability to receive a large amount of data, once over the battery lifetime of the device, as opposed to transmitting many small amounts of data over such lifetime.
Requirements

Any connected device needs to be prepared to receive firmware updates, over-the-air, to fix bugs and patch security issues. The possibility to upgrade functionality over the lifetime of the device may be desirable as well. Direct-to-Cloud wearables are no different in this respect. In many cases, MNOs demand that the firmware of the modems connected to their networks can be updated, and modem vendors provide this functionality accordingly.

Therefore, the firmware for both the application processor (called MCU in the sequel), as well as the cellular modem may need to be updated. Specifically for the patch, this needs to be done at least once during its (battery) lifetime, considering that it may have been on a shelf for a long time before getting activated. As the active life of the patch is fairly short (maximum one month), a single update may suffice.

The prototype patch MCU firmware, i.e., its application code, is 320 kB in size. In the current implementation, it is fetched in 4 kB blocks from the server, using separate HTTP GET calls. This is done because it needs to be transferred from the modem to the MCU in small chunks over the UART interface due to a modem limitation. However, more optimal solutions can be imagined.

The Sequans GM02s modem that was used for these experiments can update its own firmware independently and incrementally. The assumption for this evaluation has been that only a single incremental update of 450 kB in size will be needed over the lifetime of the patch. In contrast to the MCU firmware update, this image can be fetched with a single HTTP GET call.

Experimental setup

For this experiment, multiple instances of the test framework, as described in Section 8.3.3, have been used. On each instance, a test sequence comprising both MCU FOTA and Modem FOTA has been conducted. Such a test sequence is typically repeated twenty times in the experiment to obtain some statistical significance (again, measuring on live networks means that significant spread in results is observed). Both energy consumption and duration are measured for each phase, i.e., MCU FOTA and Modem FOTA. Note that attach is not included in the numbers reported.

Measurement results & conclusions

Table 15 shows the results of the experiments which were run on thirteen different sites. Dover Cambridge and Dover Openlab are physically in the same, US-based location, but using different cellular networks. HTC34 and HTC37 are two of the Philips office buildings on the High Tech Campus in Eindhoven. The other office locations are also Philips offices. Furthermore, some of the tests were performed in the home offices of Philips employees and a drive test has been performed.

For each site, the test sequence involved twenty iterations, except for the drive test which involved 65 iterations. However, for some of the test sequences, one of the iterations failed. If it failed during earlier phases (e.g., attach), FOTA was not even attempted. Therefore, in these cases (Dover Openlab and HTC34) only 19 iterations are reported in total. However, if it failed during FOTA itself (mostly just for modem FOTA, not for MCU FOTA), only the number of successful iterations is decreased by one. Note, in the latter case, the energy consumption and duration of the failed attempt is still accounted for in the average (and may have led to a somewhat higher value).

Median RSRP values are reported for each site to get an impression of the level of coverage. Note, however, that during the drive test RSRP varied considerably, approximately between -48 dBm and -115 dBm.
Table 15 Firmware download energy & duration for different test sites (H3A)

<table>
<thead>
<tr>
<th>Test site</th>
<th>Avg. download energy [mWh]</th>
<th>Avg. download duration [s]</th>
<th>Median RSRP [dBm]</th>
<th>Iterations total/successful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MCU</td>
<td>Modem</td>
<td>MCU</td>
<td>Modem</td>
</tr>
<tr>
<td>Aachen (office)</td>
<td>9.2</td>
<td>5.5</td>
<td>133.7</td>
<td>76.2</td>
</tr>
<tr>
<td>Dover Cambridge (office)</td>
<td>11.5</td>
<td>7.9</td>
<td>126.3</td>
<td>88.3</td>
</tr>
<tr>
<td>Dover Openlab (office)</td>
<td>12.3</td>
<td>7.9</td>
<td>126.8</td>
<td>79.1</td>
</tr>
<tr>
<td>Bothell (office)</td>
<td>9.2</td>
<td>5.7</td>
<td>111.1</td>
<td>71.8</td>
</tr>
<tr>
<td>HTC37 Production (office)</td>
<td>8.5</td>
<td>6.9</td>
<td>116.9</td>
<td>89.9</td>
</tr>
<tr>
<td>Stockholm (office)</td>
<td>6.4</td>
<td>4.3</td>
<td>80.4</td>
<td>53.0</td>
</tr>
<tr>
<td>HTC34 Development (office)</td>
<td>11.0</td>
<td>10.1</td>
<td>121.0</td>
<td>113.0</td>
</tr>
<tr>
<td>Eindhoven (home)</td>
<td>9.5</td>
<td>5.9</td>
<td>150.2</td>
<td>113.4</td>
</tr>
<tr>
<td>Tilburg (home)</td>
<td>7.7</td>
<td>4.5</td>
<td>101.3</td>
<td>64.9</td>
</tr>
<tr>
<td>Best (home)</td>
<td>8.3</td>
<td>5.1</td>
<td>108.4</td>
<td>68.1</td>
</tr>
<tr>
<td>Helmond (home)</td>
<td>12.1</td>
<td>8.3</td>
<td>151.5</td>
<td>101.1</td>
</tr>
<tr>
<td>Gemert (home)</td>
<td>6.5</td>
<td>3.7</td>
<td>78.8</td>
<td>45.6</td>
</tr>
<tr>
<td>Eindhoven region (drive test)</td>
<td>8.8</td>
<td>6.2</td>
<td>111.2</td>
<td>81.3</td>
</tr>
</tbody>
</table>

Turning to the results in the table, MCU FOTA takes between 6.4 mWh and 12.3 mWh, while Modem FOTA takes between 3.7 mWh and 10.1 mWh. Note that the level of coverage (RSRP) varies widely, with lower RSRP values leading to higher energy usage (as expected). Although MCU FOTA involves less data (320 kB) than Modem FOTA (450 kB) it still takes more energy, due to chunking the download in 4 kB blocks.

On the total energy budget of the patch these numbers are insignificant. In other words, answering research question 3, it is indeed possible to perform a single firmware over-the-air update for the patch, while staying within its limited energy budget.

### 8.4 Recommendations

Small wearables, such as vital-signs patches, that are connected Direct-to-Cloud via the cellular network, enable patients to be connected wherever they go, without the need for an additional device (gateway, phone), and without going through cumbersome network setup procedures. As such, this new class of devices will improve patient and staff experience, as well as outcomes, because patients can focus on their recovery, while safeguarding an uninterrupted stream of vital-signs data.

However, this new class of devices comes with its own challenges in fulfilling the needs of long battery life and ubiquitous coverage. Even though LTE-M and NB-IoT have enabled small form factor, energy-efficient modems, these still require significantly more energy than, for example, a Bluetooth-based solution. Just consider that the distance to be covered is several orders of magnitude larger (i.e., kilometres to a base station, instead of a few meters to a phone). For the same reason, coverage is more challenging as well.

#### Battery life

The opportunities for battery life extension have been studied extensively and address different layers:

a) **Cellular network features**: although 3GPP is gradually adding energy-saving features to its standards, LTE-M and NB-IoT in particular, the take-up of those features by MNOs and modem vendors is not a given and may take some time. For these reasons, these features have also not been explored within this project. Furthermore, not all these features (e.g., BEST, EDT) may co-exist very well with the fairly complex IP protocols required by our secure cloud solutions. However,
Release-14 and/or Release-16 RAI may help enable PSM for LTE-M. This will reduce energy consumption by replacing the attach procedure by the more energy-efficient resume procedure. Nevertheless, eventually it is up to the modem vendors to properly integrate the RAI feature into their protocol stacks. How that will unfold remain to be seen.

b) **Secure IP protocols:** the secure IP-protocol suite typically supported by cloud providers (i.e., HTTP/TLS/TCP/IP) originates from PC-based browsers, and its overhead dwarfs the actual payload transmission, at least for use cases that involve infrequent uploads of a modest amount of data like the vital-signs patch. Many developments within IETF aim to reduce this overhead by minimizing both the number of handshakes, as well as the amount of data sent, when setting up a secure connection. However, only few of those are firmly adopted by both cloud providers and modem manufacturers. Nevertheless, OAuth 2.0 token lifetime extension and ticket-based session resumption (across load-balanced nodes) already enable a significant reduction in energy consumption (and data cost) for the patch use case. Further gains could possibly be enabled by opting for smaller certificate chains. Towards the future, the adoption of TLS 1.3, CoAP over DTLS, and QUIC by both cloud providers and modem manufacturers should be monitored, although the law of diminishing returns will apply.

c) **Payload transport:** as mentioned above, the overhead of setting up a secure IP connection tends to dwarf the actual transmission of payload, at least for the patch use case. Furthermore, even firmware-over-the-air (FOTA) updates do not have a significant impact on battery life, considering that these are only expected once over the (battery) lifetime of the device. Therefore, to some extent, using a somewhat less efficient encoding, to achieve better interoperability, may be acceptable.

In summary, for the vital-signs patch readily available solutions address its requirements quite well. In particular, the stated KPI target of ~3mWh per upload has been proven feasible without compromising cloud-nativeness. Also, a single FOTA update can well be catered for within the energy budget. However, for future more demanding use cases monitoring developments by both cloud providers and modem vendors is recommended. As far as the latter are concerned, multiple considerations play a role when selecting a modem, for example: proper support of RAI for LTE-M that is well integrated into the selected protocol suite, and the choice of protocol and protocol options itself (e.g., ticket-based session resumption or, for example, TLS 1.3).

**Coverage**

Despite the promises for deep coverage, in practice, an LTE-M wearable will experience similar or worse coverage than an LTE-phone. Its single, body-worn, small form factor antenna will not be offset by the coverage enhancement features, as they are deployed for LTE-M. This may affect the uninterrupted stream of vital-signs data, which could be problematic, particularly for in-hospital scenarios. The latter could be addressed with infrastructural upgrades. However, the cost of a full-blown DAS or small-cell installation cannot be borne by the wearable’s proposition alone and requires an ecosystem approach [31]. Alternatively, operator-approved repeaters are an option, as outlined in [24]. In any case, it is a good idea to opt for an M(V)NO that offers national roaming, such that any available cellular network, in any environment (hospital, home or on-the-move) can be put to good use.

Coverage enhancements are more strongly deployed for NB-IoT than for LTE-M. Therefore, opting for NB-IoT as an alternative – or complementary – radio access network could help improve coverage, especially there where easy-to-use, infrastructural upgrades are not an option (e.g., in homes). However, experiments have shown the gains to be relatively modest, at approximately 10 dB, while the need for very dedicated, non-cloud-native, endpoints will make this solution cumbersome and costly. These limited gains can be explained by MNOs favouring spectral efficiency over an ultimate deep coverage experience. Note, however, that a hybrid LTE-M / NB-IoT solution – even if not optimized for deepest coverage, that is for very low signal levels – may help address markets where one of the two technologies is either not, or very sparsely, supported. That is, NB-IoT might help provide for a single global solution, also considering that most modems on the market offer support for both LTE-M and NB-IoT.
In summary, the anticipated KPI target of 15 dB to 20 dB coverage gain for NB-IoT relative to existing LTE-M deployments can only be achieved partially (6.5 dB for the RAT itself plus 4 dB for the simplified upload protocol). Furthermore, these benefits will only be achieved at the expense of sacrificing cloud-nativeness, which may very well not be worth the effort. Solutions for better indoor LTE-M coverage in hospitals, as proposed in [24] and [31], seem to be a more feasible approach. Nevertheless, mobile network operators should consider whether supporting RSRP values below -140 dBm for NB-IoT deployments would be worthwhile to better support small wearable and other battery-powered devices.
9 SUBCASE H3B: LOCALIZABLE TAG

9.1 Description and motivation

Wearable health monitoring has emerged as an effective way for improving the quality of life of the patient as it provides seamless remote diagnoses and monitoring.

The subcase, H3B Localizable Tag, is based around the same scenario as Subcase H3A. Its objective is to evaluate the feasibility of a reliable and low cost, low power localization technique to complement the cloud connectivity of H3A. The research work hence focuses on providing a lightweight and accurate radio-localization feature on wearable health monitoring patches. GNSS modules with low power consumption are currently commercialized for IoT devices. Nonetheless, the current mMTC application is expected to demand a very stringent power consumption. Thus, complementary location methods should be specified for narrowband transmissions, notably for 5G-NB-IoT evolutions.

As stated in Section 8.2.1 of D3.2 [1], narrowband communication systems such as NB-IoT and LTE-M are not suitable for accurate ranging [32]. A new “phase-coherent multi-channel” approach has been proposed that addresses these shortcomings by combining (narrow-band) channels over the whole band, to create the much-needed bandwidth for accurate localization. This approach has been evaluated in an indoor environment and compared to simulated performance. The ranging precision is not degraded and consistent with simulated performance when propagation occurs in quasi-Line-of-Sight (LOS), even for short-range indoor environment. In an outdoor environment (in the CEA Campus), 30m accuracy was measured under multipath propagation environment with the proposed Multi-Frequency-Phase Difference of Arrival (MF-PDoA) technique [1].

Nonetheless, localization accuracy shall be degraded in very difficult propagation environment in particular due to outdoor Non Line-of-Sight (NLOS) propagation conditions. The aim of this phase-3 trials is to conduct advanced field tests in the city of Grenoble to be able to predict and evaluate the location algorithms accuracy.

For these tests, the propagation environment mainly focuses on urban environments with strong multipath components. During the field trials, Radio metrics (SNR, RSSI, ToA) have been collected through the CEA proprietary infrastructure deployed in the Grenoble Area. To achieve a large amount of data collection and as diverse as possible radio signals from an urban environment, crowdsourcing has been considered. The measurement database (more than 1 million samples) has been exploited in conjunction with a priori knowledge of LOS/NLOS propagation statistics to predict Localization Accuracy Map (LAM).

This LAM estimator might be exploited to identify and further improve localization approaches for 5G NB-IoT evolutions. Such information might be used to estimate performances for a given 5G Network base station deployment, or to find a suitable deployment to achieve given performances.

9.2 Final setup

This section details the main blocs used to provide achievable localization accuracy for 5G NB-IoT like network.

The aim of this phase 3 is to demonstrate the feasibility of localization over a 5G network from a low power wearable patch, transmitting “alert” messages at the network edge, and real-time position to the emergency unit (hospital, ambulance), all with acceptable or predictable accuracy.

Figure 68 presents the main Time of Arrival (ToA) measurement principle used for the multilateration localization algorithm. Each known-position NB-IoT 5G base station is time synchronized with GPS. On uplink message reception, estimated base station ToA can be used by the localization algorithm to compute tag position. Because uplink Time of Departure (ToD) is not known, the Time of Flight (ToF = ToA-ToD) cannot be directly computed by the localization algorithm. As a consequence, a minimum of 3 base stations are required to compute 2D position with Time Difference of Arrival (TDoA).
9.2.1 Network architecture

Because of frequency license regulations, a Commercial Off-The-Shelf (COTS) Low Power Wide Area (LPWA) network infrastructure (LoRa [33]) has been used to generate radio signals (same consideration done during phase 2, [1]). From Table 16, one may note the similar propagation characteristics between the LoRaWAN considered for testing and the 5G NB-IoT. Maximum Coupling Loss (MCL) has been chosen by 3GPP as the metric to evaluate coverage of a radio access technology.

Table 16 LPWAN LoRaWAN vs 5G NB-IoT comparison

<table>
<thead>
<tr>
<th>PHY Technology parameters</th>
<th>LoRaWAN</th>
<th>NB-IoT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>125 kHz</td>
<td>180 kHz</td>
</tr>
<tr>
<td>MCL</td>
<td>165 dB</td>
<td>164 dB</td>
</tr>
<tr>
<td>EU freq. band</td>
<td>ISM 868 MHz</td>
<td>Licensed B20 (800 MHz)</td>
</tr>
<tr>
<td>Datarate</td>
<td>~ kbps</td>
<td>~ kbps</td>
</tr>
</tbody>
</table>

The CEA proprietary network deployed in Grenoble - and in particular the base stations installation on 6 sites, is being used for the technical evaluation. Radio metrics (RSSI, SNR, ToA), required by the localization algorithm, have been collected in urban environment area through the infrastructure for more than 18 months (phase 3). Such amount of data collection, as diverse as possible, has been achieved via crowdsourcing.

Figure 69 shows an overall implementation of the network architecture used for the field trial.
Proprietary Tags, simulating the 5G wearable patch, periodically uplink signals received by known-position base stations. Base station metadata (radio metrics RSSI, SNR, ToA - position), collected by the Core Network, are stored in a database that is polled by real-time software, including the Localization Algorithm.

We assume that tag uplink messages can be received by more than one base station. This is mandatory for the multilateration localization algorithm to compute tag position from as diverse a possible Base Station #i metadata. This concept is furthermore sufficiently flexible to be advantageously adapted to 5G-NB-IoT evolutions.

### 9.2.2 User application architecture

This sub-section provides information regarding the four main network architecture blocks (See Figure 69) for the user application, being the tag localization.

These blocks are the following:

1. Tags
2. Base Station
3. Core Network and database collection
4. Localization Algorithm

**Tags**

Tags are battery powered LPWA devices (see Figure 70) that include a global navigation satellite system (GNSS) receiver used as ground truth and some keys UE id for authentication and ciphering. Their uplink packets are periodically broadcasted, including their ground truth position (latitude, longitude, altitude).

Note that these ground truth positions are thereafter compared with localization algorithm outputs to compute the project KPI: localization accuracy. This accuracy is analysed, depending on multiple urban environments (LOS, NLOS), assuming strong multipath components.
A tenth of devices have been distributed to volunteers to collect their locations, during their daily trips over a period of a little bit more than one year. These trips include mainly car, bicycle, and pedestrian displacements (See Figure 71).

**Base Station**

The network infrastructure is composed of 6 LPWA base stations deployed around the city of Grenoble, as described on Figure 72. To perform precise and accurate ToA measurements, COTS telecom grade base station iBTS from manufacturer Kerlink based on the LoRaWAN technology have been used.
The relative base station location from CEA and their altitude are given on Table 17.

Table 17 Base Station location

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Location</th>
<th>Altitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BCC@CEA</td>
<td>[-127m, -53m]</td>
<td>250m</td>
</tr>
<tr>
<td>2</td>
<td>Bastille</td>
<td>[1262m, 485m]</td>
<td>493m</td>
</tr>
<tr>
<td>3</td>
<td>Chamrousse</td>
<td>[15278m, -7601m]</td>
<td>2259m</td>
</tr>
<tr>
<td>4</td>
<td>B2I-testA@CEA</td>
<td>[-60m, 65m]</td>
<td>236m</td>
</tr>
<tr>
<td>5</td>
<td>B2I-testB@CEA</td>
<td>[-60m, 65m]</td>
<td>236m</td>
</tr>
<tr>
<td>6</td>
<td>B2I-testC@CEA</td>
<td>[-60m, 65m]</td>
<td>236m</td>
</tr>
<tr>
<td>7</td>
<td>Pont-De-Claix</td>
<td>[-872m, -8145m]</td>
<td>291m</td>
</tr>
<tr>
<td>8</td>
<td>Vouillants</td>
<td>[2686m, -2198m]</td>
<td>559m</td>
</tr>
<tr>
<td>9</td>
<td>Alpexpo</td>
<td>[2063m, -4294m]</td>
<td>232m</td>
</tr>
</tbody>
</table>

Base station deployment has been done in several steps to test and validate all the infrastructure. Consequently, from the beginning of the data collection process, three base stations (i.e. #4, #5 and #6, see in Table 17) have been temporarily installed on the roof of one building in CEA premises for testing only, and very close from #1. Such position results in a subset of 4 base stations with very poor geometric dilution of precision (GDOP). Their final locations have been fixed some months later (i.e. #7, #8 and #9).

Taking advantage of the fact the city of Grenoble is surrounded by mountains, some base stations have been installed on very high locations, especially base stations #2 and #3 which are both installed on the roof of cable car arrival stations at altitudes of 485m and 2259m respectively (Figure 73), offering a very wide radio coverage, even if this latter is located at more than 15km from the city.

Therefore, for the main subcase KPI analysis (localization accuracy) such deployment permits a high variety of propagation channel in urban environment, including LOS, NLOS and strong NLOS.

![Figure 73 Base stations installed on the roof of the cable-car arrival station at Bastille #2 (left) and Chamrousse #3 (right), overhanging the city of Grenoble, France](image)

From a network point of view, if covered by a base station *BaseStation*\#id, estimated radio metrics metadata such as *RSSI*, *SNR* and *ToA* are sent to the Core Network on each Tag uplink. One shall consider that in a star network such as LoRaWAN, tag to base station connectivity only depends on the radio coverage (radio sensitivity). In other words, if in their radio coverage, each base station send backhaul to the core network without tag to base station association. The core network shall be able to de-duplicate each message and is able to get all covered base station radio metadata.

One may consider that, using an unlicensed at 868 MHz (ISM band), on air collision may occur during the field trial. Consequently, particular radio analysis on each site may be performed to select the best
channel to use during the crowd sourcing period. A base station feature has been developed to send to the core Network real time spectral scan. Figure 74 shows a screenshot of the spectral scan for the BCC@CEA #1 site for a period of 30 days, for channels in the 867 MHz to 870 MHz frequency band.

![Figure 74 Base Station Heatmap Spectral Scan for different channel (Frequency in MHz) versus time, BCC@CEA #1, observation period = 30 days](image)

One can note some channels (e.g., freq = 867.5 MHz) are to be avoid during the test because of high usage. Other channels usage is quite stable over time if the channel list for uplink message does not need to be modified during the field test period.

**Core Network and database collection**

The Core Network mainly consists of different micro services connected with TCP/IP sockets (UDP, HTTP, MQTT and RPC).

A LoRaWAN Network Server (LNS) is used to collect information from base stations. Each base station metadata is received asynchronously on core network side because of jitter introduced by the base station to core network backhaul (3G/4G, Ethernet). The LNS is nevertheless able to compute the unique MIC (Message Integrity Check) from uplink packet and to de-duplicate this packet sent by each base station. Thanks to a unique message identifier look up table computed by the Network Server, python programming then:

- deciphers encrypted base station ToA
- decodes tag uplink packet (ground truth position including GNSS latitude, longitude, altitude)
- computes some stats (coverage, radio connectivity, radio performances, distances between tag and base station)
- merges in a single format de-duplicated / decoded uplink message, all base station radio metadata (frequency, DR, RSSI, SNR, ToA) and previously computed stats.
- writes this information to the InfluxdB Time Series DataBase (TSDB).

**Localization Algorithm**

The location algorithm is integrated into Matlab® software that:

- polls the TSDB
- compute tag position and extract statistics and generate maps from the algorithm
- generate a HTML page that includes algorithm outputs
Figure 75 shows some screenshot of the Localization webpage, accessible for all volunteers during the field test period.
9.2.3 Hardware components

The hardware system consists of the following parts:

1. A set of 12 LPWAN tags with a GNSS receiver, a low microcontroller, a battery and integrated antennas for GPS and LPWA (868 MHz).
2. A set of 6 LPWAN base station with GNSS receiver used as time reference and LPWA Time Of Arrival (ToA) capability.
3. A server with software requirement to deploy a Network Server and the proprietary Localization algorithm and Location Accuracy Map (LAM).
4. A laptop as the Web client

9.2.4 Software components

The software components are listed below:

On the LPWAN device:
- A low power embedded software deployed on LPWAN microcontroller that polls embedded GNSS position (latitude, longitude and altitude) and broadcasts position information on air periodically. External API is used to set Radio parameters (DR and channel)

On the Base station:
- Network tools for remote administration
- Script to generate LPWA base station statistics (heatmap for spectral scanning)

On the server:
- ChirpStack, an open-source LoRaWAN Network Server
- InfluxdB, an open-source TimeSeries Database server used to collect all LPWAN data and base station metadata radio metrics (ToA, RSS, SNR) and spectral scan
- Grafana, an open-source Web server and InfluxdB Client for metrics analysis
- Python programming to prepare data to be stored int Influxdb (connectivity, stats, …)
- MatLab ® and Open Street Map (OSM) access to compute devices position form algorithm and compute LAM and generate HTML pages (dashboard)
- Eclipse Paho MQTT (Message Queuing Telemetry Transport), an open-source broker to connect all services together
- NGINX, an open-source HTTP server, used as main proxy and Web server
- Thingsboard, an open-source IoT platform to display volunteers ground truth position
- Docker, a container to deploy all services on the server

9.3 Testing and verification

One objective for this wearable tag localization system is to optimize the 5G network infrastructure, such as the base station position, to get the highest location accuracy. In other words, the challenge is to be able to predict location precision considering both propagation environment and base station position with dedicated models. This section details methodology and algorithm used for the verification of the location accuracy.

9.3.1 Methodology

The accuracy is usually computed through the Cramer Rao Lower Bound considering various metrics such as ToA [34],[35], time-difference-of-arrival (TDoA) [36], angle-of-arrival (AoA) or fusion of several metrics [37]. However, few authors address the influence of hybrid propagation that mixes line-of-sight (LOS) and non-line-of-sight (NLOS) channels [35], and it is seldom that computed bounds are tested against real measurements or only with limited datasets and scenarios [38]. Large scale field tests using LPWA real measurements including both LOS/NLOS propagation has already been conducted [39], but it has not been used to build accuracy models.

Measurement model

In this sub case, location accuracy analysis based on the 5G base station ToA measurements is done by computing some performance bound for each possible patch position \( r_k = [x_g, y_g, z_g]^T \) belonging to a set with \( n \) positions arbitrarily defined, \( r_k \in \mathbb{M} = \{ r_1, r_2, ..., r_n \} \) which we call LAM.

Position estimation using ToA measurements

If we consider a single wearable patch uplink message from a position \( r_k \) to a base station with coordinates \( r_g = [x_g, y_g, z_g]^T \), the pseudo range \( z_{g,k} \) (i.e. ToA multiplied by the speed of light \( c \) measured by the base station \( g \)) can be modelled as:

\[
z_{g,k} = h_g(r_k, d_0) + \mu_g(r_k)
\]

where \( h_g \) is the measurement function associated to base station \( g \) defined by:

\[
h_g(r_k, d_0) = \sqrt{(x_k - x_g)^2 + (y_k - y_g)^2 + (z_k - z_g)^2} + d_0
\]

where \( d_0 = c \cdot t_0 \) is the distance offset, with \( t_0 \) the unknown departure time of the signal and \( \mu_g(r_k) \) a Gaussian distributed random variable that represents all the errors of the measurement process, including base station related errors (e.g. synchronization errors, time of arrival detection uncertainty) but also channel errors due to multipath and NLOS propagation.
Although channel is essentially static and thus, channel errors are highly correlated over time (i.e. channel errors are biased), here we are considering their spatial distribution that we assume to be independent from one position to another (i.e. $\mathbb{E}[\rho_{k,g}P_{j,g}] = 0$ for $j \neq k$). In the general case, we shall assume that the error distribution is different for every considered position $r_k$ because propagation conditions vary which makes the accuracy difficult to predict. In this study, we will assume that the ToA error distribution measured by a base station $g$ when wearable patch is located at position $r_k$ depends only on the channel conditions between these two points:

$$
\mu_g(p_k) = \begin{cases} 
\mathcal{N}(0, \sigma_{\text{LOS}}^2) & \text{if } \delta_{\text{LOS}}(r_k, r_g) = 1 \\
\mathcal{N}(0, \sigma_{\text{NLOS}}^2) & \text{if } \delta_{\text{LOS}}(r_k, r_g) = 0 
\end{cases}
$$

with $\delta_{\text{LOS}}(r_k, r_g)$ a function that is equal to 1 if the patch/base station position $r_k$ and $r_g$ are in LOS, and 0 otherwise.

Motivations for distinguishing between these two cases are that, in NLOS situations, the direct path of the LPWAN 5G NB-IoT radio signal can be highly attenuated (e.g. by a building) and only reflected signals are received by the very low sensitivity base station. Because reflected signals travel longer distances than the direct path, the effective time-of-flight (ToF) experiences an excess time delay resulting in a higher ToA error than for direct propagation (see Figure 76).

![Figure 76 LOS and NLOS propagation in urban condition](image)

Note that here ToF refers to the radio propagation time on air, contrary to the ToA which refers to the actual time of reception on base station side. In fact, time of departure is unknown on the wearable patch because ns scale synchronization cannot be guaranteed. Consequently, departure time $t_0$, or equivalently $d_0 = c \cdot t_0$, is unknown and must be estimated in addition to the position.

Assuming that the height (i.e. $z_k$) of the wearable patch is known with an accuracy of a few meters, the problem turns out to a parameter estimation:

$$ X_k = [x_k, y_k, z_k]^T $$

subject to the measurements:

$$ z_k = [z_{1,k}, z_{2,k}, \ldots, z_{n,k}]^T $$

which are distributed according to:

$$ z_k \sim \mathcal{N}(h(X_k), R_k) $$

with $h(X_k) = [h_1, h_2, \ldots, h_n]^T$ the vector of the measurement functions and $R_k$ the covariance matrix of the measurements which we assumed to be independent $(R_k)_{i \neq j} = 0$. 
We also assume that their variances, expressed by the diagonal coefficients \( R_k \), depend on the channel conditions between the base station \( g \) and the position \( r_g \) according to our model in Eq. (3).

The parameter \( X_k \) can be estimated using maximum likelihood estimation (MLE) given by

\[
\hat{X}_k = \arg \max_{\hat{X}_k} l(\rho_k, X_k)
\]

where \( l(\rho_k, X_k) \) is the logarithm of the likelihood function associated to the measurements’ distribution, which are in our case normally distributed, and is expressed as:

\[
\ln l(\rho_k, X_k) = \frac{1}{2} \ln \det R_k + n \ln (2\pi) + (\rho_k - h(X_k))^T R_k^{-1} (\rho_k - h(X_k))
\]

where \( \det R_k \) is the determinant of the matrix \( R_k \).

One sees easily that maximizing Eq. (8) is equivalent to minimizing:

\[
\hat{X}_k = \arg \min_{\hat{X}_k} \left( (\rho_k - h(X_k))^T R_k^{-1} (\rho_k - h(X_k)) \right)
\]

In practice, the solution of the Eq. (9) can be computed using an optimization algorithm (e.g., Leveberg-Marquardt, Steepest Descent [40], etc.).

When the position \( r_k \) is known and only the distance offset \( d_0 \) should be estimated (i.e., to compute ToA error for a given base station), Eq. (9) simplifies and the estimated solution is given by the weighted sum

\[
\hat{d}_0 = \frac{1}{s_\omega} \sum_{i=1}^{n} \omega_i \left( z_{i,k} - d_{i,k} \right)
\]

with \( d_{i,k} \) the (known) distance between the base station \( i \) and the considered position \( r_k \), the weight \( \omega_i = \frac{1}{(R_k)_{ii}} \) and \( s_\omega = \sum_{i=1}^{n} \omega_i \) the normalization coefficient.

**Cramér-Rao lower bound**

The location accuracy on a given point \( r_k \) of a map \( M \) will depend on the geometry of the deployment (i.e., position of each base station with respect to this point), the accuracy of each ToA measurement described by the distribution of \( \mu_g(r_k) \) and the efficiency of the algorithm used to compute the estimated location. Although the former two aspects have been modelled, it is very hard to describe the performance of an algorithm in the general case. A workaround to this problem is to compute the best achievable accuracy (i.e., instead of the actual accuracy), which cannot be beaten by any algorithm in a statistical sense and under certain conditions. The well-known Cramér-Rao lower Bound (CRLB) [36] provides such a lower bound for any unbiased estimator, and states that

\[
\mathbb{E} \left[ (\hat{X}_k - X)(\hat{X}_k - X)^T \right] \geq J_k^{-1}
\]

with \( J_k \) the Fisher information matrix (FIM) defined as

\[
J_k = \mathbb{E} [\nabla_X \ln l(\rho_k, X)]^T \mathbb{E} [\nabla_X \ln l(\rho_k, X)]
\]

with \( \ln l(\rho_k, X) \) the log-likelihood given in (8)

Taking the gradient descent and noting that

\[
\mathbb{E} \left[ (\rho_k - h(X_k))(\rho_k - h(X_k))^T \right] R_k^{-1} = I_3
\]

we obtain the following expression (Cf [36])
\[ J_k = \nabla_X h, R_k^{-1}, \nabla_X h^T \]  \hspace{1cm} (14) 

where  
\[
\nabla_X h = \begin{bmatrix}
    x_k - x_1 \\
    ||r_k - r_1|| \\
    ||r_k - r_1|| \\
    y_k - y_1 \\
    1 \\
    ||r_k - r_1|| \\
    ||r_k - r_1|| \\
    1 \\
    ... \\
    x_k - x_n \\
    ||r_k - r_1|| \\
    ||r_k - r_1|| \\
    y_k - y_2 \\
    1 \\
    ... \\
    ... \\
    y_k - y_n \\
    1 \\
    ... \\
\end{bmatrix} \hspace{1cm} (15)
\]

and \(\|\cdot\|\) denotes the norm operator. It should be noted that the FIM \(J_k\), and consequently the accuracy bound, depends on

- the considered position \(r_k\) on the map \(M\),
- the base stations position \(r_g\),
- the channel condition between these two points and finally both parameters \(\sigma_{LOS}\) and \(\sigma_{NLOS}\) that should be determined.

### 9.3.2 List of key performance indicators

**Localization Accuracy Map (LAM)**

**9.3.3 Measurement and testing tools**

Mainly two tools have been used for the application testing and functional verification:

- **Base Station Characterization testing tool**, and
- **Map model and map matching tool**

**The base station characterization testing tool**

The base station characterization testing tool is fully described in [3].

The conclusion is that, assuming a nominal GDOP of 1, in a pure AWGN channel with high SNR, localization accuracy might be less than few dozens of meters.

**Map model and map matching tool**

Over this phase 3 field trial, around 1.8 M uplink packets have been recorded into the database (which represent more than 640 000 tags transmission), which is nearly 3 times more than comparable studies [39].

The map \(M\) has been constructed on a grid of \(10 \times 10\) m cells over a square area of \(40 km \times 40 km\) centered on the site of CEA which is located close to downtown.

Hence, the map represents a potential of 16 million cells, but only those who collect at least one measurement are fully initialized, which represent a total of about 70 000 cells at the end of the project.

Because the tags embedded GNSS receiver present position accuracy that can be worse than the cell size especially in urban environment, a map matching is performed before assigning a packet to a cell.

This operation uses the road network from Open Street Map (OSM) database by pre-allocating any cell whose location matches certain type of "ways" (i.e., according to OSM representation) including the streets, highways, hiking trails etc. When a new packet is received, it will be assigned to the closest pre-allocated cell in a range of 50 m with respect to the GNSS position and, if no cell is found, it will be assigned to the closest cell of the grid.

Figure 77 presents each cell visited during the field trail (colour is assigned to each volunteer whose name has been removed for confidentiality reason).
LOS analysis using 3D buildings map

One central aspect of this study is to evaluate the benefits of prior knowledge about channel conditions between the transmitter and the receiver. Hence, for each cell, this information will be computed with respect to each base station using a ray-tracing tool included in Matlab Communication Toolbox®, which provides a binary information (LOS or NLOS) by detecting if the segment connecting two points intersect any building from the loaded OSM buildings map as illustrated on Figure 78.

Because this operation is computationally intensive, the full map is split into several "LOS Analysis Area" (LAA) of 2 km × 2 km (with 500 m of overlap between two adjacent LAA to avoid processing points on the LAA borders), and only buildings belonging to the considered LAA are loaded for processing. Using this technique, it reduces the processing time to approximately 10 s per cell for computing all base station visibility conditions and predicts the RSSI from a propagation model although this is not used in this study.
Cell structure and processing

Each cell gathers different types of information among which:

- static properties such as its position in a local frame of reference and channel conditions with respect to each base station but also the "full" CRLB corresponding to the accuracy for a packet received by all base stations,
- all measurements (e.g. ToA, RSSI etc.) assigned to the cell,
- results of cell computations which includes measurement statistics (mean value, standard deviations, etc.), various CRLB and MLE results (see Algorithm 1).

It should be noted that several configurations exist for a given cell, depending on the number of base stations receiving each packet. Indeed, some base stations will never receive a packet transmitted from a certain location because the path loss is too strong, but it may also happen that an interferer prevent a packet of being received even though the signal power is above the base station sensitivity. For this reason, if a cell $k$ receives $p_k$ packets, all results (MLE locations and location errors, CRLB, etc.) are computed independently for each packet and denoted as $v^{(i)} k$, the computed value $v$ from packet #i. Only the results corresponding to the best configuration are retained for analysis, where the best configuration is defined as the one that minimizes the CRLB, which can also be interpreted as the best GDOP.

**Algorithm 1 Cell k processing**

1: for all received packets $l \in \{1,2,...,p_k\}$ do
2:     select valid measurements $z^{(i)}_k$
3:     estimate $\delta d_0^{(i)}$ using Eq. (10)
4:     compute TOF error $\delta z_k^{(i)} = z_k^{(i)} - (d_k - \delta d_0)$
5:     compute Full solution $\hat{x}_k^{(i)}$ using Eq. (9)
6:     compute position error $\delta x_k^{(i)} = \hat{x}_k^{(i)} - x_k$
7:     compute CRLB $J_k^{(i)}$ using Eq. (12)
8: end for
9: Select best configuration: \( i_{\text{best}} = \arg\min_i \| J_k^{-1}(i) \| \)

10: Record best values \( J_k^{-1}(i_{\text{best}}) \), \( \delta x_k(i_{\text{best}}) \)

12: Compute TOF statistic (mean, std, rms, etc.)

### 9.3.4 Final results

#### ToA distribution

Besides the map preparation which includes the computation of channel conditions for each cell with all base stations, two important parameters are required by the model for the LAM prediction and localization algorithm which are \( \sigma_{\text{LOS}} \) and \( \sigma_{\text{NLOS}} \) (see Eq. (3)).

Base Station qualification (Cf section 9.3.3) gives a first idea of the value \( \sigma_{\text{LOS}} \) in an ideal case of conducted propagation. However, real propagation is much more complex, mainly because of multipath and signal obstruction. Hence, empirical cumulative distribution functions (CDF) of ToA errors have been computed as described in Algorithm 1, and the most relevant ones (i.e., based on a sufficient number of values) have been plotted according to their channel conditions as shown in Figure 79.

Interestingly, these CDF show that the two base stations (#2 and #3) with high altitudes (>500 m) are insensitive to their channel conditions as their LOS and NLOS curves nearly overlap, both having a P68 metric (i.e., 68% of the samples below the P68 value) equal to approx. 250 m. One possible explanation is that the high elevation base stations (i.e., elevation angle is a more relevant indication than altitude in that context), if not in LOS, are most often in near-LOS situation which is not detected by Matlab ray-tracing tool. Indeed, it can be assumed that in a NLOS condition a portion of the radio wave can reach the top of surroundings buildings by reflection with short extra path length, and from that point, the radio wave can reach a high elevation base station with LOS conditions. The same reasoning is not true for low elevation base stations because some additional obstacles (e.g., buildings) can be found on the wave path which propagates horizontally.

Another important result that can be highlighted from these CDF illustrated in Figure 79, at least for the low elevation base stations (i.e., #1 and #4), is the ability to predict the ToA accuracy using the 3D map information of the buildings. We clearly see that CDF of ToA errors is significantly degraded (P68=710 m and P68=1300 m) for cells predicted as NLOS compared to cells predicted as LOS (390 m and 380 m for base stations #1 and #4 respectively), which validates the approach. Nonetheless, we observe a difference for the \( \sigma_{\text{LOS}} \) values between high and low elevation base stations (i.e., 250 m and 380 m respectively) which is not fully understood but which probably originates from higher multipath level in case of low elevation base stations.

It should also be noticed that the ToA accuracy is significantly degraded in real conditions even in LOS conditions (i.e., 250 m and 380 m) with respect to conducted propagation where jitter (100 ns for DR0 mode, see section 9.3.3) would correspond to 33 m of TDoA errors, that is to say 23 m of ToA error.
Final results regarding localization accuracy are fully described in D6.4 [3].

### 9.4 Recommendations

This Phase 3 studied the feasibility of LAM prediction based on ToA measurements by using 3D buildings map and CRLB computation.

Our model assumes different ToA distributions according to the channel conditions of each considered base station-cell pair. The quality of this prediction has been assessed by comparison with a MLE-based location algorithm applied on a large data set of real measurements, and by analysis of ToA errors in conducted propagation as well as in radiated propagation conditions. This analysis clearly shows different distributions (i.e. \( \sigma_{\text{LOS}} = 230\text{m} \) vs \( \sigma_{\text{NLOS}} = 1000\text{m} \)) for low elevation base station according to their channel conditions with the considered point. One of the other notable results is the demonstration of a certain immunity to the effect of NLOS channels on ToA accuracy for base stations having high elevation, which seems to be always in LOS or in a near-LOS situation. This is not detected by Matlab® LOS detection function which only returns a binary information about the presence or not of an obstacle, but we believe that our approach would benefit from slightly richer information about the channel conditions, like the depth of the obstacles on the way of the direct path.

A first LAM has been generated according to the ToA error models found and the base stations position used during the data collection. This LAM has then been compared to the actual errors obtained from a MLE-based location algorithm, showing a very good match between the two (e.g. only 12 \% of difference for the P68 in the analysis zone). Location errors are significantly higher (e.g. P68 = 3000 m) than those usually reported in the literature but this is mainly explained by the poor GDOP of the initial base station deployment and perfectly predicted by our model (P68=3300 m). According to the same model, the accuracy of the new deployment should be close to 300m in the same zone, which can be verified as the measurement campaign is still going on.

Thanks to analysis, it is clear that localization accuracy almost depends on propagation characteristics (LOS/NLOS) and base station position (GDOP). The main recommendation is that one should consider increasing the density of base stations and place them at the highest point to increase LOS conditions.
10 SUBCASE H3C: AQUACULTURE REMOTE HEALTH MONITORING

10.1 Description and motivation

Health status and safety of workers within the aquaculture industry, particularly in low- and middle-income countries (LMICs), has not been given due importance to date. Farm hands and other workers in aquaculture are susceptible to many occupational diseases and injuries in the course of their work. Especially in Greece, where marine aquaculture ranks first in the aquaculture sector [41], the workplace hazards to which aquaculture workers are exposed, should be mitigated. The remote health monitoring system presented in this subcase is designed to serve among others this purpose, i.e., the mitigation of aquaculture workers’ hazards and risks.

The WINGS remote health monitoring system aims at providing real-time monitoring and constant situational awareness of the health status of the workers and/or vulnerable people in remote locations, such as aquaculture sites. Workers in the aquaculture industry will be equipped with wearable devices measuring major vital signs such as heart rate, sinus rhythm/electrocardiogram (ECG), oxygen saturation, etc. Supervisors of the aquaculture area and/or professional caretakers will be equipped with smart glasses to remotely provide additional insights to the professional caregiver/medical experts. They will be able to:

- be informed of the workers being monitored and appropriate alarms may be raised, if necessary,
- be on alert if the system identifies any abnormality in usual patterns of oxygen saturation or forecasts time periods with increased risk and
- facilitate and improve their work due to the automated decision making and the videocalls (from the dashboard and through the smart glasses) with the subjects, as well as the forecasting and visualization techniques.

10.2 Final setup

This section explains the final (phase-3) setup of the experimental use case, in accordance with previous descriptions in D3.2 [1] and D3.3 [2].

The test cases selected for the trials were defined as follows:

1. Remote monitoring of workers’ vital signs

   The focus of this test case is the collection of vital-signs measurements from wearable devices. The wearable devices transmit the data on workers vital signs to the Cloud over a mobile device. More specifically ECG, heart rate, oxygen saturation and geographical position/location sent by the wearable devices worn by the workers are analysed by the WINGS STARLIT platform [42]. All these data are represented in JavaScript Object Notation (JSON) format. Specifically, heart rate, oxygen saturation and geographical position are structured as key and value pairs in JSON objects while ECG is structured as a JSON array. The JSON objects also include metadata in JSON format that contain information for the accompanying device data, such as the unique wearable device id or the timestamp and unit of the measurement data. The size of a JSON object containing heart rate, oxygen saturation and location data is approximately 12 kB while the size of a single ECG JSON array is 30 kB. The analysis of the data provides additional insights into the health status of the workers in case of an emergency or potential future hazard. Vitals are displayed on interactive graphs in various time windows. This information can also be downloaded in common formats (.csv and .pdf). The dashboard offers the functionality of notification management (change criticality or status, add comments, etc.) and video-calls on demand. The solution will enhance the quality of life in aquaculture areas providing full confidence and peace of mind for the aquaculture workers, their relatives, and colleagues.
2. **Live video streaming with smart glasses**

The resolution of important issues is not always supported by indirect communication. This part of the subcase aims to connect supervisors or caretakers within the aquaculture areas to remote medical experts with “see-what-I-see” video collaboration. The supervisors/caretakers in aquaculture areas use smart glasses to enable live streaming in case direct communication with a remote medical expert if deemed necessary. The medical expert will be able to provide remotely the necessary guidelines and advice on worker care in case of emergencies or further examine them depending on the criticality of the event.

From the application side, the subcase will incorporate support for devices (smartwatches) measuring major vital signs, their visualization on devices easily portable on site (e.g., mobile, tablet) and, in a more specialized environment, an expert formatted dashboard that includes deeper analysis of the monitored parameters, as well as the efficient communication through smart glasses.

From the network perspective, it is critical to ensure the timely communication between the relevant actors (aquaculture workers and first responders).

**10.2.1 Network architecture**

This subcase is a concurrent aquaculture scenario and leverages the ICT-17 5G-EVE, (5G European Validation platform for Extensive trials) platform for the trials, while experiments using commercial 5G network have also been performed. More information about the architecture of 5G-EVE and regarding the aquaculture of Skironis can be found in D3.2 [1].

**10.2.2 User application architecture**

The Aquaculture Remote Health Monitoring solution (an instantiation of the WINGS STARLIT remote health monitoring system) comprises:

- Wearable devices for heart rate, oxygen saturation (SpO2) and sinus rhythm monitoring (ECG) and Smart Glasses as described in Section 10.2.3 below,
- intelligence for identification of current issues, forecasting of future issues and health emergencies and notification, and
- a dashboard for providing the health care professionals with visualization of health monitoring data, notifications, and alerts. Figure 80 illustrates the H3C overall concept of the Aquaculture Remote Health Monitoring.

![Figure 80 Aquaculture remote health monitoring subcase architecture for H3C](image)
10.2.3 Hardware components

The main hardware components are a) a Withings ScanWatch, b) Vuzix Blade Upgraded Smart Glasses, and c) a Samsung S10 5G phone acting as a gateway.

- Withings ScanWatch: the ScanWatch is a smartwatch that allows to continuously scan vital parameters to detect heart health conditions. It boasts a medical-grade ECG and an oximeter for SpO2 measurements. Overall, it provides tracking of the following metrics [43]:
  - Heartbeat notifications: high or low heart rate, irregular heartbeat.
  - Heart rate: beats per minute.
  - Breathing disturbances: detection via oxygen saturation.
  - Electrocardiogram: tracing of a 30-seconds ECG recording on a millimetric grid.
  - Oxygen saturation level (medical-grade SpO2).
- Vuzix Blade Upgraded Smart Glasses [44] are used as a live video streaming device from on premises personnel, delivering a view of the worker to the remote healthcare experts.
- Samsung S10 5G is used to make use of the 5G network available.

10.2.4 Software components

As introduced, data from the smart watch are transmitted and displayed continuously to remote medical experts via an appropriate dashboard. This dashboard is implemented as a Web-based User Interface (UI) and thus, is accessible through mobile devices as well as laptops, desktops, and tablets. In case the heart rate or oxygen saturation values are critical or out of range (based on predefined thresholds) or if a worker is outside a predefined geographical area, corresponding alarms are raised and the medical experts are notified (via pop-up windows in the dashboard, SMS and e-mail, depending on the selected settings).

Dashboard: User interface and data visualisation

To retrieve real-time data from the smartwatch the Withings REST API [45] is utilised. Whenever a new record is registered to a device, the Withings API notifies the STARLIT platform to retrieve the available data based on a specific timeframe. The communication is based on the OAuth2 protocol, that uses HTTPS requests and utilizes access and refresh token authentication to ensure security in all communications. Data are visualized on a user-friendly dashboard, screenshots of which can be found in the following figures. The user can log in to the web application through the login page (Figure 81). The authentication service sets the current user information object from the server response at the local storage of the client’s browser and the routing guard will check the user’s role and permissions and will redirect the user to the view requested.
The home page of STARLIT Remote Health Monitoring Platform is shown in Figure 82.

![Figure 81 WINGS STARLIT Remote Health Monitoring login page](image)

Figure 81 WINGS STARLIT Remote Health Monitoring login page

The vital parameters of the subscribed users are depicted on interactive graphs, while historical data and predictions can also be presented. The user can select the period of the data to be visualised and download the measurements in common formats (Figure 83).

![Figure 82 WINGS STARLIT Remote Health Monitoring patients view](image)

Figure 82 WINGS STARLIT Remote Health Monitoring patients view
Medical experts and/or supervisors of the aquaculture area can view the list of the recorded notifications, and respond with comments, change their status (from active, to pending, etc.), change their criticality (from high to low etc.), and so on. The notification list view of STARLIT Remote Health Monitoring Platform is shown in Figure 84.

The users of the system can view the subscribed workers on interactive map Figure 85.
Intelligence: identification of issues

A deep convolutional neural network is developed to analyse the ECG signal obtained from the smartwatch. For this purpose, the Pytorch python library is utilized to build the neural network since it provides tensor computing with strong acceleration via graphics processing units and a tape-based automatic differentiation system. Similarly, a deep convolutional neural network has been developed for the analysis of the Oxygen Saturation signal.

A deep convolutional neural network based on Residual connections (ResNet) for the analysis of the ECG signal has been developed and tested using open data repositories, specifically the MIT-BIH (Massachusetts Institute of Technology and former Boston's Beth Israel Hospital, now the Beth Israel Deaconess Medical Center) Arrhythmia Database [46]. This database is used as data source for labelled ECG recordings both from healthy subjects and subjects suffering from different kinds of arrhythmias. The training procedure is performed offline to make use of acceleration software (GPUs).

The deep convolutional neural network for the analysis of the Oxygen Saturation signal has been developed and tested using the UCD (St. Vincent's University Hospital / University College Dublin Sleep Apnoea Database) open data repository [47]. The database is used as a data source for labelled Oxygen Saturation recordings from adult subjects with respiratory events (obstructive, central apnoeas and hypopneas and periodic breathing episodes).

10.3 Testing and verification

This sub-section presents results from the testing and verification of H3C subcase.

10.3.1 Methodology

Testing has focused on the performance of the intelligence components as well as the overall system. WINGS personnel acted as aquaculture workers equipped with wearables (smartwatches) and smart glasses respectively. A medical expert was presented with a demonstration of the developed solution to provide feedback on the system usability.

10.3.2 List of key performance indicators

In terms of the performance of the intelligence components the following KPIs were measured:

- ECG classification accuracy (%) measured as the number of correctly classified heartbeats relative to the total number of heartbeats.
• ECG classification sensitivity (%) measured as the number of correctly identified abnormal heartbeats (i.e., heartbeats with arrhythmia) relative to the total number of ground truth abnormal heartbeats.

• ECG classification specificity (%) calculated as the number of correctly identified normal heartbeats relative to the total number of ground truth normal heartbeats.

• ECG classification precision (%) measured as the number of correctly identified abnormal heartbeats relative to the total number of predicted abnormal heartbeats.

• SpO2 classification accuracy (%) measured as the number of correctly classified segments (a segment is denoted as a part of the SpO2 signal of 1-minute duration) relative to the total number of segments.

• SpO2 classification sensitivity (%) measured as the number of correctly identified abnormal segments (i.e., segments with any respiratory event for at least 5 consecutive seconds inside the segment) relative to the total number of ground truth abnormal segments.

• SpO2 classification specificity (%) calculated as the number of correctly identified normal segments relative to the total number of ground truth normal segments.

• SpO2 classification precision (%) measured as the number of correctly identified abnormal segments relative to the total number of predicted abnormal segments.

In terms of the overall system performance the following KPIs were measured:

• RTT latency (ms) measured as the average of RTT latency recorded with the ping tool.

• Throughput (Mbps) calculated as the average of throughput recorded with Speedtest.

• Service availability (%) calculated based on packet losses recorded with ping.

• Service reliability (%) calculated based on packet losses and RTT latency recorded with ping.

• Usability calculated based on system usability questionnaire.

10.3.3 Measurement and testing tools

System performance metrics were calculated with the use of tools such as ping, Speedtest\(^5\) by Ookla and adding timestamps at various points of the system operation code.

10.3.4 Final results

Table 18 below presents the validation results on the performance of the intelligence components of the system. The overall system performance validation results are reported in D6.4 [3].

Table 18 Intelligence components performance validation results.

<table>
<thead>
<tr>
<th>KPI</th>
<th>Target values</th>
<th>Measurements (average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECG Accuracy</td>
<td>90.2%</td>
<td>98.31%</td>
</tr>
<tr>
<td>ECG Sensitivity</td>
<td>96.52%</td>
<td>92.11%</td>
</tr>
<tr>
<td>ECG Specificity</td>
<td>85.14%</td>
<td>99.12%</td>
</tr>
<tr>
<td>ECG Precision</td>
<td>83.84%</td>
<td>93.27%</td>
</tr>
<tr>
<td>SpO2 Accuracy</td>
<td>85.26%</td>
<td>82.03%</td>
</tr>
<tr>
<td>SpO2 Sensitivity</td>
<td>60.36%</td>
<td>61.4%</td>
</tr>
<tr>
<td>SpO2 Specificity</td>
<td>91.71%</td>
<td>89.5%</td>
</tr>
</tbody>
</table>

\(^5\) [https://www.speedtest.net/](https://www.speedtest.net/)
SpO2 Precision | N/A | 70%

The target values for the ECG signal are taken from a previous version of the ECG analytics system, which utilized a convolutional autoencoder and a One-class SVM classifier [48]. For the SpO2 case, the target values are extracted from a study which used a deep neural network (Deep Belief Network) to detect abnormal SpO2 segments [49].

### 10.4 Recommendations

As can be observed the validation for H3C Aquaculture Remote Health Monitoring has been successful both in terms of the intelligence for timely identification and prediction of issues, as well as the overall system performance, showcasing that healthcare, among other sectors, can be improved with the use of 5G.
11 CONCLUSION

This deliverable has documented the final (phase-3) technical solutions of the healthcare vertical use cases of 5G-HEART. The phase-3 trials have been performed per subcase, coordinated by the subcase owners, and using the 5GTN, 5G-VINNI, 5G-EVE, 5Groningen, Eindhoven and Grenoble platforms, as well as commercial networks.

Besides the results, observations and insights obtained from the phase-3 trials, recommendations have also been made addressing how the use cases can be further developed to become solutions for the future e-health paradigm. Table 19 contains an assessment of how well the project has addressed the different domains in developing testing the final solutions for the healthcare vertical.

Table 19 Assessment of total test emphasis of the different healthcare subcases in 5G-HEART.

<table>
<thead>
<tr>
<th>Subcase</th>
<th>Clinical tests</th>
<th>Application technology tests</th>
<th>5G technology tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1A</td>
<td>Some (1)</td>
<td>In-depth (3)</td>
<td>In-depth (3)</td>
</tr>
<tr>
<td>H1B CHD</td>
<td>In-depth (3)</td>
<td>In-depth (3)</td>
<td>Sufficient (2)</td>
</tr>
<tr>
<td>H1B Robot</td>
<td>In-depth (3)</td>
<td>In-depth (3)</td>
<td>Some (1)</td>
</tr>
<tr>
<td>H1C</td>
<td>In-depth (3)</td>
<td>In-depth (3)</td>
<td>In-depth (3)</td>
</tr>
<tr>
<td>H1D</td>
<td>In-depth (3)</td>
<td>In-depth (3)</td>
<td>Some (1)</td>
</tr>
<tr>
<td>H2A</td>
<td>Sufficient (2)</td>
<td>In-depth (3)</td>
<td>Some (1)</td>
</tr>
<tr>
<td>H3A</td>
<td>None (0)</td>
<td>In-depth (3)</td>
<td>Sufficient (2)\textsuperscript{1}</td>
</tr>
<tr>
<td>H3B</td>
<td>None (0)</td>
<td>In-depth (3)</td>
<td>Sufficient (2)\textsuperscript{2}</td>
</tr>
<tr>
<td>H3C</td>
<td>Some (1)</td>
<td>In-depth (3)</td>
<td>Sufficient (2)</td>
</tr>
</tbody>
</table>

1: The H3A subcase is mMTC, and has been tested on 4G/CIoT technologies LTE-M and NB-IoT
2: The H3B subcase has been tested using proprietary LPWA technology based on LoRa, with similar parameters as NB-IoT operating in the 800 MHz band.

Table 20 summarizes the motivation for each healthcare subcase, its phase-3 proposed trial setup, trial plan, performed trials, trial results and recommendations for improvement of future trials. The trial plans listed in the table are taken Section 5.2 of D3.1 [1], except for the subcase H3C that was added to the healthcare vertical during phase-2. For many of the other subcases, some deviations from the original plan have occurred, due to progressive insights and/or the COVID-19 pandemic related delays.
Table 20 Progress-to-plan for final (phase-3) trials

<table>
<thead>
<tr>
<th>Sub case</th>
<th>Motivation</th>
<th>Proposed setup</th>
<th>Trial plan (Phase-3)</th>
<th>Actual trials (Phase-3)</th>
<th>Trial results</th>
<th>Recommendations</th>
</tr>
</thead>
</table>
| **H1A** | To enhance participation to remote educational services and to improve situation awareness from the field. | 360° and single lens cameras 5G SA network RedZinc wearable video | • Using LTE and 5G  
• Introduce mobility at both sides  
• Considering 5Groningen for server side with the ambulance context (see subcase H1C)  
• Mobile clients at 5GTN, Oulu, 5Groningen  
• Measure e2e metrics with quality assessment | Testing on 5G SA  
Evaluating and configuring uplink channel Trials with RedZinc equipment at Oulu University Hospital | Verification of the proposed setup over 5G SA. Extensive network delay and end-to-end latency measurements with and without uplink congestion. E2E delay less than 180 ms observed, below the 200 ms baseline threshold. Oulu University Hospital trials show that 89% of the responders agreed that wearable video can be used for teaching or training, and 82% of the responders felt that RedZinc’s BluEye wearable video could replace traditional face-to-face teaching. | The use of 360 camera does not necessarily provide decent quality in case sharp video quality is needed and therefore the use of multiple single lens cameras may lead to better QoE providing also possibilities for lower E2E latency. Utilization of mobile networks and especially 5G can enable a flexible set-up for the live streaming application. |
| **H1B CHD** | To provide remote expert support when examining a baby suffering from Congenital Heart Disease using ultrasound. | EPIQ / Collaboration Live on 4G/5G network. DNL ultrasound / LiveSwitch on 5Groningen (and other facilities). AR/VR rendering of immersive cooperation setting for ultrasound examination on Ethernet. | • Improve the Skype-like and AR-based platforms based on the outcome of the evaluation experiments and perform further testing. | Tested usability of the EPIQ/Collaboration Live system with students and doctors. Tested usability of 3D capturing and visualization. | Remote guidance of novice echocardiographers with EPIQ/Collaboration Live could be useful to do crude evaluation of cardiac function. This approach is definitely worthy of further exploration. The remote guidance with EPIQ/Collaboration Live can differentiate between a critical case of CHD requiring transport to an expert center and cases that do not require transport. The evaluation 5G remains a promising, but not critical, enabler for remote ultrasound support. EPIQ/Collaboration Live does not need 5G support but can already be put to work on legacy 4G networks. Digital multi-stream ultrasound image and data sharing clearly requires a true 5G network configuration. The 3D telepresence experiments that use 3D |
## H1B Robot

**Motivation**
To provide a long-distance robotic teleoperation system over 5G network to enable the cardiologist in the centre hospital to capture ultrasound images from the remote hospital.

**Proposed setup**
- Master-Slave robotic system.
- 6DOF UR5 Manipulator.
- Phantom Omni haptic device.
- 3 Fixed cameras for video streaming.
- Gamma SI-65-5 force/torque sensor.

**Trial plan (Phase-3)**
- Improve the robotics framework based on the outcome of the experiments.
- Design an emergency safe mode for the robot considering location of patient, technician, and slave side devices.

**Actual trials (Phase-3)**
- Tested robot examination quality and time compared to reference examination with 23 healthy volunteers.
- Tested round-trip delay over the robot manipulation channel.

**Trial results**
- The robot examination time was almost 2 times longer than the manual echocardiography.
- Most patients reported slightly higher physical pressure from the robot probe.
- The image quality from the robot examinations was lower than for the reference.
- Round-trip latency average of 40.05 ms meets the target requirements.

**Recommendations**
- Further research and investigation is needed to bring the robotic system examinations up to the quality of manual reference.

## HIC

**Motivation**
To improve patient care quality and efficiency of ambulance services.

**Proposed setup**
- Audio-video headset.
- Vital parameters monitor device.
- 5G SA outdoor network.

**Trial plan (Phase-3)**
- Extending phase 2 with 5G using mobility.
- Further evaluation of ultrasound approach.

**Actual trials (Phase-3)**
- Tested guaranteed QoS using the outdoor 5G SA network slicing in a concurrent scenario with the transport vertical.

**Trial results**
- Validated the role of 5G in enabling the use of real-time video communication between paramedic and CMO.

**Recommendations**
- To continue similar trial-based investigation with a commercially deployed 5G network with more challenging network.
### Sub case | Motivation | Proposed setup | Trial plan (Phase-3) | Actual trials (Phase-3) | Trial results | Recommendations |
--- | --- | --- | --- | --- | --- | --- |
**H1D** | To evaluate the use of wearable video in a paramedic ‘buddy to buddy’ situation, with particular focus on urban search and rescue and incident commander support. | 4G and 5G network. BlueEye deployment on regional cloud. | • Evaluate results from phase 2. • Extend or repeat testing as applicable to gain more data. | Secure application architecture for MEC developed. Plans for an extended pilot between RedZinc, Telenor and OUS in progress (beyond 5G-HEART). | Significantly better video throughput for headcam when separating traffic using network slicing on 5G SA. | Investigate certain network features (e.g., slicing and relaying) which are expected to be able to improve network and resource availability and performance. Propose a trust-based approach to engage medical users. Recommendations for getting the approach to market. |
**H2A** | To increase the success rate of polyp detection in order to decrease cancer mortality while improving the capacity of clinicians to handle more patients in less time and with less burnout. | Simulator/Emulator of video capsule. Computer for the video streaming. Backscatter reader. Computer acting as a server. 5G Network. AI polyp Detector. | • Demonstrate a 5G interface and uplink and downlink latency assessments with AI model in python. • Demonstrate the concept in an animal experiment. • Prepare a prototype and documentation with an application to the Regional Ethics Committee to use AI for polyp detection. • Describe the regulatory steps and approval procedures as a roadmap to use capsule in human procedure. | End-to-end latency measurements using different protocols. AI-based model detection quality assessed. | E2E latency tests show that the UDP-RTP protocol exhibits the lowest latency of 46.74 ms, while TCP-based protocols need between 240 and 470 ms. AI-based detection gives 91.27% sensitivity and 67.21% precision. | The advent of deep learning in the field of capsule endoscopy, with its evolutionary character, could lead to a paradigm shift in clinical activity in this setting. |
<table>
<thead>
<tr>
<th>Subcase</th>
<th>Motivation</th>
<th>Proposed setup</th>
<th>Trial plan (Phase-3)</th>
<th>Actual trials (Phase-3)</th>
<th>Trial results</th>
<th>Recommendations</th>
</tr>
</thead>
</table>
| **H3A** | To keep a tab on patients, unobtrusively, and hassle-free, no matter where they are, by means of single-use, direct-to-cloud, vital-signs patches. | Automated test framework for NB-IoT/LTE-M energy, power and coverage measurements connecting to commercial networks in the Netherlands and on several international locations. | • Evaluation of selected, advanced features from Release 15-17.  
• Potentially broaden the application further into more demanding clinical cases.  
• Gap analysis as input to future 3GPP releases (i.e., requirements input to standardization). | Selection and optimization of upload protocols  
Feasibility of NB-IoT for improved coverage  
Feasibility of firmware over-the-air update | Under poor coverage, HTTPS uses 4.7x more energy than “plain UDP”, and 1.6x more energy than CoAP over DTLS. The improvements enabled by TLS session resumption have been quantified, 14-33% energy reduction is possible.  
A more lean-and-mean protocol based on CoAP is proposed for use with NB-IoT. Practical coverage improvement in the tests is about 6.5 dB, far from the 15-20 dB promise, probably due to network operator implementation limitations.  
The energy consumption for a single firmware upload is insignificant compared to the total energy budget, making a single firmware update over the operational life of the patch very well feasible. | Utilizing protocols that are supported by modern and cloud vendors, the stated KPI target of ~ 3mWh per vital-signs upload has been proven feasible without compromising cloud-nativeness. However, for future more demanding use cases monitoring developments in IETF protocol support by both cloud providers and modem vendors is recommended.  
Solutions for better indoor LTE-M coverage in hospitals seem to be a more feasible approach than the use of NB-IoT with its proprietary upload protocol.  
Mobile network operators should consider whether supporting RSRP values below -140 dBm for NB-IoT deployments would be worthwhile in order to better support small wearable and other battery-powered devices. |
| **H3B** | To provide a lightweight and accurate radio-localisation feature along with communication on wearable health monitoring patches. | Field measurements in the city of Grenoble using CEA experimental SDR platform.  
Six LPWA base stations deployed in Grenoble. | • Final field trials in the city of Grenoble  
• More advanced demonstrations  
• Final localization algorithm validation | Collecting radio metrics via crowdsourcing (SNR, RSSI, ToA).  
Analysis performed to predict a Localization Accuracy Map (LAM). | The resulting localization accuracy depends on the ToA measurements, and it is clear that the propagation conditions contribute significantly | Increased base station density and locations to highest points should be considered to improve the LOS conditions. |
<table>
<thead>
<tr>
<th>Sub case</th>
<th>Motivation</th>
<th>Proposed setup</th>
<th>Trial plan (Phase-3)</th>
<th>Actual trials (Phase-3)</th>
<th>Trial results</th>
<th>Recommendations</th>
</tr>
</thead>
</table>
| H3C      | To provide real-time monitoring and constant situational awareness of the health status of the workers and/or vulnerable people in remote locations, such as aquaculture sites. | WINGS STARLIT platform  
Remote monitoring of workers’ vital signs.  
Live video streaming with smart glasses.  
Wearable devices for heart rate, SpO2 and ECG monitoring.  
5G-EVE platform | ● Piloting activities with medical experts and volunteers  
● Completing integration of wearables with the WINGS Smart Gateway  
● Include RedZinc wearable video solution  
● 5G integration | Performance of AI-based analysis of vital-signs.  
Measured system performance KPIs | Measured ECG and SpO2 intelligence (AI/analytics) component performance:  
Accuracy, Sensitivity, Specificity and Precision.  
Most values meet the target performance.  
System performance measured to be 27 ms RTT latency, 134/37 Mb/s user experienced throughout (DL/UL)  
Usability score of 3.5 on a scale from 1 (worst) to 5 (best). | Successful validation shows that healthcare can be improved by using 5G. |
|          |            |                | (LOS vs NLOS) as well as the base station position. |           |               |                 |
REFERENCES


[47]. St. Vincent's University Hospital / University College Dublin Sleep Apnea Database (2007). https://physionet.org/content/ucdadb/1.0.0/
